Robo-Advising Meets Large Language Models: Educating Investors on Alpha and Beta of Mutual Funds and Stocks *

Fangzhou Lu
University of Hong Kong

Sixuan Li
University of Hong Kong

Yingyan Shi
University of Hong Kong

October 13, 2025

Abstract

We investigate the transformative potential of large language models (LLMs) when integrated into robo-advisory platforms, using financial literacy as a contextual setting. As a pilot program in partnership with a major brokerage firm, we enhance a robo-advisor with a back-end LLM (e.g., ChatGPT/DeepSeek) to provide personalized, conversational support to investors making decisions by differentiating between beta and alpha in mutual funds and stocks. Our study examines how the LLM-augmented robo-advisor influences investor behavior and portfolio selection compared with a traditional rule-based chatbot and standard non-premium human assistance. Results show that LLM support significantly increases investor engagement and understanding, narrowing the behavioral gap between automatic one-click enrollment and self-assembled portfolio construction. With continuous, interactive guidance reminiscent of high-touch wealth management services, the LLM-enabled robo-advisor fosters relational trust and encourages ETF adoption, particularly in trust-sensitive markets. Notably, LLM chatbots also generate positive spillovers: investors exposed to the LLM are more likely to pursue international diversification, even when such options are not actively promoted by the robo-advisor but are available. Overall, integrating LLMs into robo-advisors yields substantial benefits, with treated investors achieving monthly returns 40 basis points higher and realizing an average monthly gain of 750 RMB.

JEL Codes: D1, E62, H20

Keywords: Financial literacy, Robo-advising, LLMs, ChatGPT, Factor investment

*Fangzhou is grateful to Jennifer Carpenter, Jonathan Parker, Antoinette Schoar, and Robert Whitelaw for their guidance and advice. We also thank Kose John, Maarten Meeuwis, Christopher Palmer, Alan Kwan, Maureen O'Hara, Gregor Schubert, Alberto Rossi, and seminar participants at the 2025 INFORMS International Conference and the 2025 AI in Finance Conference, especially the discussant Nagpurnanand Prabhala, as well as colleagues at NYU, HKU, and MIT for their helpful comments and suggestions.

1 Introduction

The rapid advancement of artificial intelligence, particularly generative AI and large language models (LLMs), is transforming the landscape of investment advisory and portfolio management. As financial markets and products become increasingly complex, AI-powered systems are poised to play a pivotal role in supporting investors' decision-making by providing tailored guidance and real-time insights. Recent innovations in LLMs and machine learning have enabled these technologies to analyze vast amounts of financial data, assist in optimal portfolio construction, and deliver accessible explanations of key concepts such as beta and alpha, helping investors better understand risk and performance metrics. Despite persistent gaps in financial literacy among retail investors, AI-driven robo-advisors hold significant promise for democratizing financial knowledge and bridging these gaps, empowering a broader range of individuals to make informed investment choices (Tsoukalas, Wang and Giesecke, 2019; Horel, Giesecke, Storchan and Chittar, 2020; Cong, Tang, Wang and Zhang, 2019; D'Acunto and Rossi, 2021, 2023; D'Acunto and Rossi, 2023; Sheng, Sun, Yang and Zhang, 2024).[1] Building on the insights of D'Acunto, Prabhala and Rossi (2019) and Rossi and Utkus (2024), we examine the implications of incorporating LLMs as a core component of the next generation of robo-advisors.

In this study, we leverage a robo-advisory platform enhanced with a back-end LLM system to examine the broader impact of AI-driven investor education on behavior and performance. While we use beta and alpha, key metrics in mutual fund and stock analysis, as specific examples to evaluate the effectiveness of AI-powered explanations, the approach is designed to be extensible to a wide range of financial concepts. Through our partnership with a large brokerage firm in China, we are able to deploy a small-scale pilot program featuring an interactive AI-enabled system to a substantial and diverse investor base, providing insights into how generative AI can facilitate scalable, personalized financial education and

[1] Also see Chak, Croxson, D'Acunto, Reuter, Rossi and Shaw (2022); Gargano and Rossi (2024).

potentially improve decision-making across various investment contexts.[2]

To conduct the analysis, the robo-advisor systematically evaluates the investor's portfolio by calculating key metrics such as alpha and beta for each holding. It employs a multi-factor analysis, incorporating a comprehensive set of factors such as size, value, and momentum, and computes statistical measures such as t-statistics to determine the significance of the alpha. This helps distinguish genuine manager skill from random fluctuations or factor loadings. The robo-advisor then aggregates these metrics to provide an overall portfolio assessment, computing the portfolio level alpha and beta, and benchmarking them against market indices to contextualize performance and volatility. Based on this analysis, the robo-advisor generates personalized recommendations aimed at optimizing the portfolio. These recommendations may include reallocating funds from underperforming assets, maintaining positions in high-performing assets, and introducing low-cost, diversified investment vehicles such as ETFs or index funds. This data-driven approach helps investors achieve better returns, manage risk effectively, and minimize costs, thereby aligning their investment decisions with their financial goals and risk tolerance.

Specifically, to investigate the impact of LLMs, robo-advisors, and financial literacy on investor behavior and performance, we assisted this brokerage firm in conducting a Randomized Controlled Trial (RCT) involving 28,819 investors who expressed interest in participating. We randomly selected 14,409 of these investors to form the treatment group, while the remaining 14,410 constituted the control group. Both groups were similar in terms of investment portfolio size and personal characteristics, ensuring a balanced sample and comparable pre-trends.

In the treatment group, investors received targeted financial education on the concepts of beta and alpha through our robo-advising platform, which was enhanced with various

[2]Due to restrictions on accessing ChatGPT in mainland China, the ChatGPT chatbot was complemented by three other China-based LLMs: DeepSeek, Kimi, and Doubao. To enable the firm to use ChatGPT, a combination of an enterprise subscription, a VPN, and a server in Singapore was employed. Consequently, the China-based LLMs served as backup systems and were integrated and switched to periodically when ChatGPT was not available. Our data indicates no significant difference in investor performance when using either the ChatGPT or Chinese-based LLM such as Kimi, DeepSeek, or Doubao.

interactive learning tools. The treatment group was further divided into four subgroups to examine the heterogeneity of the intervention's effects. The first group is the LLM (ChatGPT/DeepSeek/Kimi/Doubao) Chatbot treatment group. This subgroup received financial education and personalized advice through a large language model (LLM)-powered chatbot, which provided detailed explanations and engaged in natural language interactions. The second is the without Chatbot group. This subgroup received financial education through static content of portfolio recommendation without interactive chatbot support. The third group is the Human Chat Assistant group. This subgroup had access to human financial advisors who provided real-time assistance and explanations via chat. The final group is the Normal Chatbot group. This subgroup used a standard rule-based chatbot that provided predefined responses and guidance based on a fixed set of rules.

While treated investors were informed that they had enrolled in the robo-advisory program, they also knew that some participants would receive chat-based assistance. However, they were not explicitly told which type of chat assistance they would receive, in order to minimize potential experimenter demand effects. In this pilot program, all three types of chat assistance (human advisor, rule-based chatbot, and LLM-based chatbot) were prohibited from revealing their true identity. Each assistant appeared under a neutral and human-like nickname,[3] and if asked whether they were human or AI, all systems were instructed to reply that they could not answer the question. None of the chat interfaces allowed users to transfer to human service, preventing any inadvertent revelation of which system was human-operated. During the sign-up process, investors were also not informed that some participants might interact with an LLM-based chatbot. These design choices ensured that any observed behavioral differences arose from the content and interaction style of the assistance itself, rather than from expectations about receiving a novel technology. Of course, investors could still form their own beliefs or guesses about the nature of their chat assistant.[4]

[3]Comparable to English-language chat agents named John from customer support or Emily, your assistant.

[4]One of the main reasons the brokerage firm did not openly advertise that its robo-advisory service might include an LLM-based chatbot was regulatory caution. Although the firm had received informal approval from China's regulatory committee to conduct a small-scale pilot program, the authorities advised against

In the overall sample, our analysis yields several key insights into how robo-advisors and financial education influence investor behavior. We find that, after gaining access to a robo-advisor and receiving education about beta, investors exhibit a marked preference for investing in ETFs and index funds, rather than directly in actively managed mutual funds or individual stocks. This shift suggests a growing appreciation for the benefits of diversification and the cost efficiencies associated with passive investment strategies. Most notably, our findings demonstrate that financial education has a significant positive impact on investment performance. Investors who receive education on beta and alpha experience an increase in their monthly returns by 40 basis points, alongside an average monthly gain of 750 RMB.[5] These results highlight the substantial financial benefits that can be realized through improved financial literacy. Moreover, this part of the results is largely consistent with the findings in D'Acunto, Prabhala and Rossi (2019) and Rossi and Utkus (2024), but significantly larger. This can be attributed in part to the high volatility of the Chinese stock market, as well as to the introduction of the new LLM-based chatbot.

More importantly, our findings indicate significant heterogeneity in the results across different subgroups within the treatment group, among which we compare investors using the LLM-enhanced robo-advisor with those using a traditional chatbot and basic human assistance. LLM Chatbot subgroup exhibited the highest advice-following rate and the most substantial improvements in investment outcomes. The interactive and natural language capabilities of the LLM-powered chatbot appeared to enhance investor engagement and comprehension, leading to better adherence to the robo-advisor's recommendations and superior portfolio performance. The CAPM alpha of clients using the pilot LLM-based chatbot is, on average, 29 basis points higher than that of clients using the rule-based chatbot, basic

publicly promoting the use of LLMs in financial advisory services. At the time, it remained unclear whether such disclosure might raise concerns about data transmission to the technology firms providing the LLMs. Nevertheless, all investors signed an online liability waiver agreement acknowledging full responsibility for their investment decisions and recognizing that the chat content might include AI-generated material. The pilot program was abruptly terminated at the end of 2024, in part due to ongoing regulatory concerns.

[5]We do not find evidence that investors using robo-advisors outperform the market. Instead, it is investors without robo-advisors who significantly underperform relative to the market.

human assistance, or no assistance. We conduct post usage survey and summarize the most frequently asked questions by investors to the Human/Robo advisors and their responses. The most common reason for human assistants to be ineffective is the slow response time and distrust. The most common reason for traditional robot assistants to be ineffective is the inaccurate response material that are "mechanical". This result sheds lights on the scenario when robo-advising is most effective.

Furthermore, Greig, Ramadorai, Rossi, Utkus and Walther (2022) find that human-robo hybrids optimize investor trust. Our results complement this finding by revealing that LLM-based chatbots uniquely emulate high-touch private wealth management, traditionally reserved for ultra-high-net-worth clients, through 24/7 availability and multi-topic engagement. Unlike standardized human advisors (akin to retail bank associates) or rule-based chatbots, LLMs may be able to provide relatively bias-free guidance (Gennaioli, Shleifer and Vishny, 2015; Egan, Matvos and Seru, 2019) while replicating relational depth via casual dialogue and off-hours accessibility. We find that investors engaging during non-trading hours or discussing non-investment topics achieved higher returns, mirroring the "client bonding" of private bankers who blend lifestyle and financial advice. Furthermore, we find that LLM scalability democratizes customization: high-interaction users acheive higher returns, resolving the cost-customization trade-off. Taken together, these results show that by simulating relational depth at near-zero marginal cost, particularly critical in distrust-prone markets, LLMs redefine robo-advising's value proposition: efficacy stems not merely from portfolio optimization, but from AI's capacity to deliver perceived exclusivity and accessibility.

We also investigate whether financial education has a lasting impact on investors' ability to manage their portfolios independently, particularly focusing on their understanding of alpha and beta. Through the randomized controlled trial, investors were further assigned either a one-click switch option to a recommended robo-advisor portfolio or a DIY (do-it-yourself) approach. The primary goal was to assess if investors could replicate the performance of the one-click portfolios in their DIY portfolios, indicating a solid grasp of alpha and beta. Our

findings reveal that the LLM (ChatGPT/DeepSeek/Kimi/Doubao) chatbot significantly enhances investors' understanding and performance, narrowing the gap between one-click and DIY portfolios. While initial financial education is beneficial, the effectiveness of follow-up assistance, particularly from the LLM chatbot, plays a crucial role in helping investors apply their knowledge effectively. This suggests that advanced AI-driven advisory tools can empower investors with the skills needed for long-term financial independence and literacy.

One interesting puzzle we find is that investors in the treatment group also began diversifying their portfolios to include S&P 500 ETFs and other foreign ETFs, despite the robo-advisors initially recommending only reallocations within the Chinese A-share market ETFs. Further analysis shows significant increases in allocations to S&P500, US Factor Style Investment, and Chinese ADRs, suggesting that the robo-advising service effectively promotes diversification. However, this unexpected foreign diversification appears to be driven by LLM-based chatbot assistance, as this effect is only evident in the treatment subgroup with LLM-based chatbot. This indicates the potential of AI-driven assistance to extend and amplify traditional robo-advisor recommendations, leading to more diversified investment portfolios.

In addition to the primary findings, we observed significant heterogeneity in the effectiveness of our interventions across different investor characteristics. Younger investors, who are typically more tech-savvy, exhibited higher engagement and better investment outcomes when interacting with the robo-advisor, as evidenced by their increased advice-following rates and superior portfolio performance. Moreover, investors with larger portfolios demonstrated greater benefits from the detailed financial education, leveraging advanced strategies to enhance their investment returns more effectively. These variations underscore the importance of tailoring financial literacy programs to address the specific needs and preferences of different investor segments, thereby optimizing the impact of robo-advisory services.

Our research primarily contributes to the literature on robo-advising, as ChatGPT, DeepSeek, Doubao, Kimi, or modified versions of LLM models can potentially serve as inte-

grated components of robo-advisors. D'Acunto, Prabhala and Rossi (2019) show that robo-advising can increase investors' portfolio diversification, and Rossi and Utkus (2024) finds that Vanguard's robo-advisors lead to greater diversification. Our results extend Greig, Ramadorai, Rossi, Utkus and Walther (2022) by demonstrating that LLM-based chatbots can replicate the exclusive, high-touch engagement of traditional wealth management through 24/7 availability, fostering relational bonding and improving returns in markets prone to distrust. While our findings are consistent with previous literature, we further show that robo-advisors enhanced with LLM tools can help retail investors better understand complex financial concepts and select superior portfolios, compared to robo-advisors paired with non-premium human advisors or standard chatbots.

Our research is related to the emerging literature on large language models (LLMs) and artificial intelligence-generated content (AIGC) more broadly. Sheng, Sun, Yang and Zhang (2024) show that ChatGPT can help hedge fund managers increase their returns and outperform those who do not use ChatGPT. Cong, Liang, Zhang and Zhu (2024) develop a general approach to analyzing large-scale text-based data using neural network language processing and generative statistical models. Our paper demonstrates that LLMs can have a lasting impact on investor behavior, and can encourage rational choices, such as international diversification, even without explicit instructions.

This study contributes to the literature on financial education and investor behavior by providing empirical evidence of how integrating advanced technologies, such as AI-driven educational tools and LLM-based robo-advisory platforms, can enhance investor learning and performance. The importance of financial literacy in shaping investor behavior and improving financial outcomes has been widely documented. For example, Campbell (2006) demonstrates that financial literacy plays a major role in shaping household financial outcomes, while Van Rooij, Lusardi and Alessie (2011) finds that households typically lack knowledge of basic financial concepts beyond the time value of money and inflation. More specifically, Calvet, Campbell and Sodini (2009) shows that the disposition effect is the most

dominant behavioral bias when trading stocks, with investors often making "fight or flight" decisions based on mutual fund performance. Our findings show that after investors learn about the difference between alpha and beta, they exhibit less disposition effect and are more likely to "fight" by reallocating their funds based on factor-adjusted alpha. By equipping investors with a deeper understanding of key financial concepts through tailored financial education powered by LLMs, robo-advisors can help investors make more informed decisions and achieve better financial outcomes.

The remainder of the paper is organized as follows. Section II reviews the relevance of the LLM used in our studies, the role of robo advisor, as well as beta and alpha for investment. Section III describes the data and methodology used in our study. Section IV presents our empirical results. Section V concludes.

2 LLM and Financial Literacy

2.1 LLM

During the intervention period beginning in late 2023, the brokerage firm utilized four large language models (LLMs) to support its operations. The primary model, ChatGPT-4 (gpt-4-1106-preview), developed by OpenAI, served as a multimodal foundation for text generation, contextual analysis, and cross-market reasoning, leveraging its 128k-token context window to synthesize global financial trends.[6] Alongside it, two China-focused LLMs operated in parallel: Kimi (Moonshot-v1-200k), developed by Moonshot AI, specialized in processing complex Chinese financial documents and regulatory texts with its 200k-token context capacity, while Doubao (v2.3.0), ByteDance's multimodal model, integrated real-time data from platforms like Douyin (TikTok) to automate client-facing tasks. These models collectively form the AI chatbot supplementing the firm's robo-advisor.

A late-2024 adjustment saw DeepSeek (R1) replace Doubao for technical workflows. De-

[6]ChatGPT4 is also used in parallel with ChatGPT 3.5.

veloped by DeepSeek Inc., this model brought quant-oriented precision to derivatives pricing and risk modeling, achieving 81.7% accuracy on coding benchmarks (HumanEval) and integrating natively with tools like Python for financial simulations. While ChatGPT-4 and Kimi remained central to general analysis and regulatory tasks, DeepSeek's domain-specific optimization for STEM and quantitative finance filled a critical gap in high-complexity technical outputs. All models operated with bilingual (Mandarin/English) compliance guardrails aligned with China's Algorithmic Recommendation Management Provisions, ensuring adherence to local regulations. This phased deployment balanced real-time retail engagement with escalating demands for technical rigor as market conditions evolved.

To ensure seamless integration with existing systems, the brokerage firm partnered with a third-party IT vendor to preconfigure the LLMs with contextual knowledge of the robo-advisor's ETF-centric strategies. The vendor developed a middleware layer that synchronized real-time robo-advisor outputs, such as ETF selection criteria (e.g., liquidity thresholds, expense ratios) and sector/thematic recommendations (e.g., AI semiconductors, renewable energy ETFs), with the LLMs' knowledge bases. Structured data from historical ETF portfolio recommendations were integrated, including rationale such as tracking error parameters, sector concentration rules, and details about the brokerage firm's current ETF product offerings. The LLMs were further trained on annotated regulatory guidelines governing ETF disclosures, particularly China's regulatory requirements for thematic fund exposures and investor risk categorization. This preparatory work enabled the chatbots to reference real-time ETF metrics during interactions, such as explaining, "Your recommended switch to the XXX ETF](https://examplelink.com), which is available for purchase on the brokerage firm's platform, might be a suitable adjustment given the China Tech Index's 35% year-to-date performance," while cross-referencing live net asset values and creation/redemption activity. By grounding responses in the firm's predefined ETF frameworks, while retaining the ability to generate compliant, generalized advice when needed, the system ensured alignment with the robo-advisor's methodology and regulatory protocols.

2.2 Financial Literacy: Beta and Alpha

Educating investors about the fundamental concepts of beta and alpha is crucial because it helps them discern the true sources of their investment returns. Beta measures a security's sensitivity to market movements and various factors, indicating how much of the investment's performance is due to general market risk and factor risks. In contrast, alpha represents the excess return of an investment relative to a benchmark, reflecting the portion of performance attributed to the mutual fund manager or investor's skill or strategy. By understanding these concepts, investors can more accurately assess whether their returns are due to market movements (beta) or the mutual fund managers and their own investment acumen (alpha), particularly when evaluating their fund managers' ability when purchasing mutual funds or their own skill when buying individual stocks.

This distinction is vital for improving financial literacy and investor performance. Often, investors may attribute their gains to personal skill when, in reality, these returns are primarily driven by market factors. By understanding that beta explains a significant portion of their returns, investors are less likely to overestimate their abilities and take undue risks. This awareness directly addresses overconfidence biases, particularly when it comes to stock investments. Recognizing that market exposure, rather than individual stock-picking prowess, drives their returns can lead to more prudent investment decisions, such as preferring diversified, low-cost ETFs or index funds that capture market returns efficiently.

In addition to addressing overconfidence biases, educating investors about beta and alpha also helps them correctly evaluate mutual fund managers. Investors often struggle to distinguish between a manager's genuine skill and performance driven by market movements. By understanding alpha, investors can better assess whether a mutual fund manager is delivering true value beyond what can be attributed to market exposure. This knowledge empowers investors to make more informed decisions when selecting mutual funds, ensuring they are not overpaying for active management that does not deliver corresponding alpha. In essence, it enables investors to accurately evaluate their fund managers' ability when purchasing mutual

funds or their own investment acumen (alpha) when it comes to buying their own stocks.

In comparison to typical robo-advisors that generate portfolio that rely on making tailor made suggestions based investors' age and risk preference, educating investors about beta and alpha might have a more profound and lasting impact. Behavioral biases often lead to suboptimal trading behaviors and require continuous effort to correct. However, by understanding beta and alpha, investors gain a foundational framework that inherently guides them toward more rational and data-driven decisions. This knowledge helps them recognize that outperforming the market consistently is challenging and often not due to skill but to exposure to market factors.

We use education on alpha and beta as an illustrative example to demonstrate the application of an LLM-enhanced robo-advisor. The insights derived from this setting may have external validity and could potentially be generalized to other types of robo-advisors. It is important to note that any measure of alpha and beta is subject to considerable measurement error; nevertheless, brokerage firms often regard these metrics as effective tools for encouraging investors to switch ETFs.

3 Institutional Details

3.1 The Brokerage Firm

The brokerage firm we partner with has extensive market coverage and a strong reputation within the financial industry. It is one of the largest and most established brokerage firms in China, with over 200 branches spread across the country. This extensive network ensures that our study could reach a diverse and broad spectrum of investors, providing a solid foundation for our research on the impact of financial literacy.

The firm manages assets worth over 1 trillion RMB and serving more than 10 million clients, it stands as one of the top brokerage firms in China. This significant market share and large client base are essential for obtaining a representative sample for our study. The

broad reach allows us to analyze the behaviors and outcomes of a varied group of investors, enhancing the generalizability and validity of our findings. Additionally, the firm's strong reputation and credibility in the financial industry lend weight to our study, ensuring that our research is conducted within a reliable and trustworthy framework.

Furthermore, the firm's commitment to innovation and technology aligns well with our research objectives. By integrating a back-end LLM (ChatGPT/DeepSeek/Kimi/Doubao) system with their robo-advisor, we were able to provide an advanced and interactive educational experience for investors. This firm's willingness to adopt and implement cutting-edge technology demonstrated their dedication to improving client services and financial literacy. Their support and collaboration were instrumental in facilitating the seamless integration of our educational tools, thereby enhancing the overall effectiveness of our study.

3.2 RCT designs

To investigate the impact of LLM, robo-advisor, and financial literacy on investor behavior and performance, we conducted a Randomized Controlled Trial (RCT) involving 28,819 investors who expressed interest in participating.[7] We randomly selected 14,409 of these investors to form the treatment group, while the remaining 14,410 investors constituted the control group. Both groups were similar in terms of investment portfolio size and personal characteristics, ensuring a balanced and fair comparison. Our sample period begins in December 2022, one year prior to the implementation of the robo-advisor. The pilot program of the robo-advisor was launched at full scale in December 2023 and concluded in November 2024.[8]

[7]The brokerage firm sent a push notification through its app to slightly fewer than 5% of all its customers. Ultimately, 28,819 investors expressed interest in participating.

[8]For the randomization process, The brokerage firm use the Stata code automatically to re-randomized as follows. They first allocated investors to the treatment and control groups on the basis of a randomly generated number. Using this allocation, they then calculated the maximum and the average t-statistics on the differences in averages across the treatment and control groups for the following variables: investors' age, gender, years of schooling, total asset value, P/L in terms of percentage in the past year, If the maximum t-statistic for these variables was higher than 1.25 or the average t-statistic was higher than 0.35, we drew a new random number and allocated investors to the treatment and control groups on the basis of this new number. We repeated this process until the maximum t-statistic was 1.25 or lower and the average t-statistic

In the treatment group, investors received targeted financial education on the concepts of beta and alpha through our robo-advising platform. The education module provided detailed explanations, examples, and personalized insights into how these metrics influence investment risk and performance. Investors in the control group did not receive the robo-advisory services.[9] Moreover, a subsample of the treatment group, was further enhanced with a back-end LLM system or other follow-up assistance for interactive learning.

Throughout the intervention period, we closely monitored the investment decisions and performance outcomes of both groups. The robo-advisor in the treatment group not only provided the usual portfolio analysis and recommendations but also incorporated educational content to help investors understand the rationale behind each suggestion. For instance, if an investor in the treatment group held mutual funds with varied alpha and beta values, the robo-advisor would explain the implications of these metrics and recommend adjustments based on their statistical significance and performance potential.

By comparing the outcomes of the treatment and control groups, we aimed to isolate the effect of the robo-advisor on investment behavior. Specifically, we measured changes in portfolio composition, trading activity, and overall investment performance. This RCT design allowed us to rigorously evaluate the causal impact of robo-advisor on investor behavior and performance, providing robust evidence to support the integration of educational tools within robo-advisory platforms. The insights gained from this study highlight the potential benefits of enhancing financial literacy to help investors make more informed decisions and achieve superior financial outcomes.

To conduct the analysis, the robo-advisor begins by evaluating the investor's current portfolio holdings. It systematically assesses each mutual fund or stock within the portfolio by calculating key metrics such as alpha and beta, which are essential for understanding the performance and risk characteristics relative to the market. The robo-advisor employs

was 0.35 or lower.

[9] In the Appendix, we show how exactly our robo-advisor function to help investor understand the difference between alpha and beta.

a comprehensive set of factors, specifically those that have been shown to be effective in the Chinese stock market, to conduct a multi-factor analysis.[10] This analysis includes commonly used factors such as size (SMB), value (HML), and momentum (MOM), among others. It also computes statistical measures, such as t-statistics, to determine the significance of the alpha, thereby distinguishing genuine manager skill from random fluctuations. In the Appendix section, we provide a detailed example of how robo-advisors work.

Once the individual metrics for each holding are calculated, the robo-advisor aggregates these metrics to provide an overall assessment of the portfolio. This involves computing an overall average alpha and beta for the entire portfolio, taking into account the proportion of each asset within the portfolio. The robo-advisor then benchmarks these aggregated metrics against market indices to provide context on the portfolio's performance and volatility.

Based on this detailed analysis, the robo-advisor generates personalized recommendations for the investor. These recommendations are aimed at optimizing the portfolio by suggesting actions such as reallocating funds from underperforming assets, maintaining positions in high-performing assets, and introducing low-cost, diversified investment vehicles like ETFs or index funds. The ultimate goal is to help the investor achieve better returns, manage risk more effectively, and minimize investment costs. Through this methodical and data-driven approach, the robo-advisor empowers investors to make informed decisions that align with their financial goals and risk tolerance.

3.3 LLMs and Robo-advisors

In the treatment group, investors received targeted financial education on the concepts of beta and alpha through our robo-advising platform. The treatment group was further divided into four subgroups to examine the heterogeneity of the intervention's effects:

LLM Chatbot: This subgroup received financial education and personalized advice through a large language model (LLM)-powered chatbot, which provided detailed explanations and

[10]The Market and SIZE factors are the most important considerations for adjustment.

engaged in natural language interactions.[11] Without Chatbot: This subgroup received financial education through static content such as articles and videos, without interactive chatbot support. Human Chat Assistant: This subgroup had access to human financial advisors who provided real-time assistance and explanations via chat. Normal Chatbot: This subgroup used a standard rule-based chatbot that provided predefined responses and guidance based on a fixed set of rules. We also further divide each subgroup into two groups. The first group receive the benefit of automatic execution or one-click execution. The second group can only execute the recommended portfolio by doing it by themselves (DIY). This setup helps us to distinguish whether investors actually understand the logic behind the recommended portfolios. Throughout the intervention period, we closely monitored the investment decisions and performance outcomes of all subgroups in the treatment group, as well as the control group.

While treated investors were informed that they had enrolled in the robo-advisory program, they also knew that some participants would receive chat-based assistance. However, they were not explicitly told which type of chat assistance they would receive, in order to minimize potential experimenter demand effects. In this pilot program, all three types of chat assistance (human advisor, rule-based chatbot, and LLM-based chatbot) were prohibited from revealing their true identity. Each assistant appeared under a neutral and human-like nickname,[12] and if asked whether they were human or AI, all systems were instructed to reply that they could not answer the question. None of the chat interfaces allowed users to transfer to human service, preventing any inadvertent revelation of which system was human-operated. During the sign-up process, investors were also not informed that some participants might interact with an LLM-based chatbot. These design choices en-

[11] Due to restrictions on accessing ChatGPT in mainland China, the ChatGPT chatbot was complemented by three other China-based LLMs: DeepSeek, Kimi, and Doubao. To enable the firm to use ChatGPT, a combination of an enterprise subscription, a VPN, and a server in Singapore was employed. Consequently, the China-based LLMs served as backup systems and were integrated and switched to periodically when ChatGPT was not available. Our data indicate no significant difference in investor performance when using either ChatGPT or other Chinese-based large language models (LLMs), such as Doubao or Kimi.

[12] Comparable to English-language chat agents named John from customer support or Emily, your assistant.

sured that any observed behavioral differences arose from the content and interaction style of the assistance itself, rather than from expectations about receiving a novel technology. Of course, investors could still form their own beliefs or guesses about the nature of their chat assistant.[13]

By comparing the outcomes of the treatment and control groups, as well as the different subgroups within the treatment group, we aimed to isolate the effect of LLM chatbot, robo-advisor and financial literacy on investment behavior. Specifically, we measured changes in portfolio composition, trading activity, and overall investment performance. In conclusion, this RCT design allowed us to rigorously evaluate the causal impact of robo-advisor, LLMs, and financial literacy on investor behavior and performance.

3.4 Summary Statistics

Table 1 presents summary statistics for investor and portfolio characteristics, comparing the treatment and control groups immediately prior to the implementation of the robo-advisor in 2023. For each characteristic, the table reports the mean and standard deviation (in parentheses), along with the calculated difference between the treatment and control groups. The differences observed are not statistically significant. The average age of investors in the treatment group is 45.6 years, with a standard deviation of 10.2 years, compared to the control group, which has an average age of 46.1 years and a standard deviation of 10.5 years. This results in a difference of -0.5 years, which is not statistically significant. The average account age is 8.5 years in the treatment group, with a standard deviation of 5.2 years, while the control group has an average account age of 8.7 years and a standard deviation of 5.0 years. The difference here is -0.15 years, again, not statistically significant. Initial portfolio

[13]One of the main reasons the brokerage firm did not openly advertise that its robo-advisory service might include an LLM-based chatbot was regulatory caution. Although the firm had received informal approval from China's regulatory committee to conduct a small-scale pilot program, the authorities advised against publicly promoting the use of LLMs in financial advisory services. At the time, it remained unclear whether such disclosure might raise concerns about data transmission to the technology firms providing the LLMs. Nevertheless, all investors signed an online liability waiver agreement acknowledging full responsibility for their investment decisions and recognizing that the chat content might include AI-generated material. The pilot program was abruptly terminated at the end of 2024, in part due to ongoing regulatory concerns.

values are quite similar between the groups. The treatment group has an average initial portfolio value of $184.4 thousand with a standard deviation of $50.3 thousand, whereas the control group has an average value of $182.8 thousand with a standard deviation of $51.2 thousand, leading to a difference of $1.5 thousand, which is not statistically significant.

The gender distribution, measured as the male ratio, shows that males constitute 74% of the treatment group (standard deviation 0.44) and 76% of the control group (standard deviation 0.43), with a difference of -0.02, which is not statistically significant. Investors in the treatment group hold an average of 15.3 stocks (standard deviation 10.4), slightly higher than the control group, which holds an average of 14.8 stocks (standard deviation 10.7), resulting in a difference of 0.5 stocks, which is not statistically significant. The average number of mutual funds held is 5.6 (standard deviation 3.2) in the treatment group and 5.8 (standard deviation 3.1) in the control group, leading to a difference of -0.2, which is not statistically significant. Returns three years before the intervention show a negative average return of -8.4% (standard deviation 2.5) for the treatment group and -8.2% (standard deviation 2.6) for the control group, with a difference of -0.2%, which is not statistically significant. Finally, the average number of trades per month is 10.8 (standard deviation 24.1) in the treatment group, compared to 10.5 (standard deviation 26.2) in the control group, resulting in a difference of 0.3 trades per month, which is not statistically significant.

These summary statistics provide a comprehensive overview of the comparability between the treatment and control groups, highlighting the similarities and differences across various investor characteristics and portfolio attributes. The lack of statistically significant differences suggests that the two groups are comparable across these dimensions.

4 Results

We next present our baseline results. Table 2 shows the relationship between robo-advising and investment outcomes, categorizing the results for mutual fund investors, stock investors,

and investors participating in both markets since the implementation of the robo-advisor in December 2023. The table includes key metrics such as Average Annual Return, Annualized CAPM Alpha, Portfolio Diversification (measured by the Herfindahl Index), and Investor Satisfaction (measured on a scale of 1 to 5). Panel A of Table 2 shows the results for mutual fund investors. The treatment group, which received robo-advising, achieved an Average Annual Return of 14.8%, compared to 10.2% for the control group, resulting in a statistically significant difference of 4.6 percentage points ($p < 0.01$). The Alpha for the treatment group was 0.1% while the control group had a CAPM Alpha of -2.3%, leading to a significant difference of 2.4 percentage points ($p < 0.01$). Furthermore, Investor Satisfaction was significantly higher in the treatment group, with an average score of 4.5 compared to 3.8 in the control group (difference = 0.7, $p < 0.05$).

For mutual fund investors, Figure 1 shows that there is a significant decrease of active mutual fund investment for the treatment group. While for the control group during the same time period, there is no difference in the share of active equity mutual fund investment. Figure 1 also shows that there is a significant increase in passive index fund and ETFs investments for the treatment group. For the control group, again, there is no change in the investment in passive index fund and ETFs investments. Figure 2 shows that there is a significant increase in Sharpe ratio and investor gain for investors that receive the treatment. But there is not much Sharpe ratio change and investor gain for the control group.

Panel B highlights the outcomes for stock investors. The treatment group showed an Average Annual Return of 14.5%, which is 4.1 percentage points higher than the control group's 10.4% return ($p < 0.05$). The Alpha for the treatment group was 0.16%, whereas the control group had an Alpha of -2.24%, resulting in a significant difference of 2.4 percentage points ($p < 0.05$). Similar to mutual fund investors, the treatment group exhibited better Portfolio Diversification with a Herfindahl Index of 0.19 as opposed to 0.30 in the control group (difference = -0.11, $p < 0.01$). Investor Satisfaction also improved significantly, with the treatment group scoring 4.4 compared to 3.7 for the control group (difference = 0.7, p

< 0.05).

For individual stock investors, Figure 4 shows that there is a significant decrease of active common equity investment for the treatment group. While for the control group during the same time period, there is no difference in the share of active common equity investment. Figure 4 also shows that there is a significant increase in passive index fund and ETFs investments for the treatment group, similar to the results of active mutual fund. For the control group, again, there is no change in the investment in passive index fund and ETFs investments. Figure 5 shows that there is a significant increase in Sharpe ratio and investor gain for common equity investors that receive the treatment. But there is not much Sharpe ratio change and investor gain for the control group.

Panel C provides the results for investors participating in both mutual funds and stocks. The treatment group outperformed the control group with an Average Annual Return of 14.7%, which is 4.4 percentage points higher than the 10.3% return of the control group ($p < 0.01$). The Alpha for the treatment group was 0.1%, compared to -2.3% for the control group, resulting in a significant difference of 2.4 percentage points ($p < 0.01$). Finally, Investor Satisfaction in the treatment group was significantly higher, with an average score of 4.5 compared to 3.7 in the control group (difference = 0.7, $p < 0.05$).

The results across all panels consistently show that robo-advising has a positive impact on investment outcomes. The treatment groups, which received robo-advising, exhibited higher Average Annual Returns and Alpha, better Portfolio Diversification, and greater Investor Satisfaction compared to the control groups. These findings suggest that robo-advising can significantly enhance investment performance and investor satisfaction, making it a valuable tool for both mutual fund and stock investors. We do not find evidence that investors using robo-advisors outperform the market. Instead, it is investors without robo-advisors who significantly underperform relative to the market. Our findings reveal that after learning about beta with the help of the robo-advisor, investors tend to favor ETFs or index funds over mutual funds or individual stocks. This shift indicates a recognition that

market exposure, rather than individual security selection, is the primary driver of their returns.

4.1 LLM-based robo-advisors

We next show our main results by investigating the heterogeneous effect of the treatment based on different follow-up assistance after giving the initial recommendation, especially the LLM-based follow-up assistance. the Figure 3 shows the heterogeneous effect of the treatment based on different follow-up assistance after giving the initial recommendation for mutual fund investors. LLMs-based chatbot assistance induces the most economic significant increase in shift to passive index fund and ETFs. The results of without assistance at all is much weaker than LLMs-based chatbot but slightly stronger than normal chatbot and human based assistance. Figure 6 shows a similar result for equity investors. Panel A of Table 3 presents the difference-in-differences regression coefficients and t-statistics (in parenthesis) for various performance indicators, including Monthly Return, Monthly CAPM Alpha, and Portfolio Diversification (measured by the Herfindahl Index).[14]

$$Y_{it} = \beta_1 + \beta_2 Robo_i * Post_t + \gamma^{Charact} Charact_{it} + \alpha_i + \lambda_t + \epsilon_{it} \tag{1}$$

The treatment group (*Robo*, indicator variable) is defined as investors who intended to and actually received robo-advising services, while the control group includes those who intended to but did not receive such services. The post-treatment period (Post) is defined as months after December 2023. $Charact_{it}$ are various characteristics of the investors. We also include individual investor fixed effects and month fixed effects in our specifications.

Overall, investors receiving robo-advising services showed significantly better performance outcomes. The treatment group achieved a monthly raw return of 39 bps higher compared to the control group, with a t-statistic of 4.45. The Alpha for the treatment group

[14]Panel A, row one of Table 3 essentially shows a finding that mirrors the results reported in Table 2, but the analysis is conducted at the monthly level and is more rigorously examined using regression analysis. Starting from the second row of Panel A in Table 3, we present results pertaining to LLM-backed chatbots.

was 0.21% higher, significantly better than the control group, with a t-statistic of 4.98. Portfolio Diversification improved significantly in the treatment group, as evidenced by a lower Herfindahl Index of -0.11, with a t-statistic of -3.03.

To test the parallel trends assumption, we define January to June 2023 as the pre-treatment period and July to October 2023 as the pseudo-treatment period, and find no evidence of treatment effects. Alternatively, when we define January to December 2022 as the pre-treatment period and January to October 2023 as the pseudo-treatment period, we similarly observe no treatment effects. These results suggest that our difference-in-differences regressions are likely to satisfy the parallel trends assumption.

Next, we conduct the regression analysis in four subsamples of treatment group where we keep the control group the same.[15] Further breakdown (rows two to five of Table 3) shows that the LLM (ChatGPT/DeepSeek/Kimi/Doubao) Chatbot subgroup outperformed other groups with a raw return of 0.64% (t-statistic 6.14) and an Alpha of 0.43% (t-statistic 6.20), along with a significant improvement in Portfolio Diversification (-0.17, t-statistic -3.61). In contrast, groups without chatbot support, human chat assistants, and normal chatbots showed lower but still significant improvements in performance metrics compared to the control group.

The Human Chat Assistant subgroup also showed positive results, with high advice-following rates and improved investment outcomes, though not as pronounced as those in the LLM/ChatGPT subgroup. The personalized and real-time assistance provided by human advisors contributed to better understanding and execution of investment strategies. The Normal Chatbot subgroup demonstrated moderate improvements, with advice-following rates and investment outcomes better than the control group but not as significant as those seen in the LLM/ChatGPT or Human Chat Assistant subgroups. The predefined responses of the rule-based chatbot were helpful but lacked the dynamic interaction and depth provided by the LLM-powered chatbot and human advisors. Lastly, the Without Chatbot subgroup

[15]The results are qualitatively similar if we use matched control group for each subsamples of the treatment group.

showed the least improvement among the treatment subgroups. While the static educational content did enhance financial literacy to some extent, the lack of interactive support limited the investors' ability to fully understand and apply the concepts to their investment decisions.

Alternatively, we also use a triple difference-in-differences to test that the treatment effect from the LLM subgroup is larger than those in other treatment subgroups.

$$Y_{it} = \beta_1 + \beta_2 Robo_i * Post_t + \beta_3 Robo_i * Post_t * LLM_i + \gamma^{Charact} Charact_{it} + \alpha_i + \lambda_t + \epsilon_{it} \quad (2)$$

LLM_i is an indicator that equals one if the investors belongs to the LLM treatment group. The coefficient of interest is β_3, which measures the additional treatment effect of the LLM subgroup relatively to other treatment subgroups. Panel B of Table 3 shows that the additional treatment effect from the LLM treatment group is both statistically and economically significant for raw return, alpha, and the Herfindahl index. The positive effect of the LLM chatbot on raw return and alpha is approximately three times higher than that observed in the rest of the treatment group who do not use the LLM chatbot.[16]

Panel C of Table 3 examines investor reactions, including Average Response Time, Interaction Times, and Investor Sentiment. The LLM (ChatGPT/DeepSeek/Kimi/Doubao) Chatbot subgroup demonstrated the most efficient response time of 3.7 seconds on average (standard deviation of 1.4), with the highest interaction frequency of 14.3 times per month (standard deviation of 28.9) and a positive investor sentiment score of 0.25 (standard deviation of 0.37). Conversely, the human chat assistant subgroup had a much longer average response time of 64.8 seconds (standard deviation of 264.7), fewer interaction times (5.1, standard deviation of 12.4), and a negative investor sentiment score (-0.17, standard deviation of 0.46). The normal chatbot subgroup showed the quickest average response time of 1.2 seconds (standard deviation of 0.3) but had the lowest interaction frequency and investor sentiment

[16]In the Appendix Table A.3, we present the full set of coefficients from these regressions.

score. In summary, the results consistently indicate that robo-advising, particularly when integrated with advanced LLM-based chatbots such as DeepSeek or ChatGPT, significantly enhances investment performance and investor satisfaction. The treatment groups exhibited higher average returns, better portfolio diversification, and greater investor sentiment, demonstrating the substantial benefits of financial education and advanced robo-advising tools.

Our findings indicated significant heterogeneity in the results across the different subgroups within the treatment group. Investors in the LLM (ChatGPT/DeepSeek/Kimi/Doubao) Chatbot subgroup exhibited the highest advice-following rate and the most substantial improvements in investment outcomes. The interactive and natural language capabilities of the LLM-powered chatbot appeared to enhance investor engagement and comprehension, leading to better adherence to the robo-advisor's recommendations and superior portfolio performance.

4.2 Potential Mechanisms

Greig, Ramadorai, Rossi, Utkus and Walther (2022) demonstrate that hybrid human-robo advisory models optimize investor trust. Therefore, we hypothesize that the superior performance from LLM-based chatbots could be attributed to their capacity to emulate high-touch private wealth management traditionally reserved for ultra-high-net-worth clients. Unlike the standardized, office-hour-bound human advisors in our experiment (who mirror retail bank associates with fixed consultation protocols), LLM chatbots provide relatively bias-free (Gennaioli, Shleifer and Vishny, 2015; Egan, Matvos and Seru, 2019) financial guidance while simulating the relational depth of elite wealth managers. Crucially, private bankers cultivate personal relationships through value-added services beyond portfolio advice, such as lifestyle engagements. We thus posit that LLM chatbots replicate this premium experience through two channels: (1) 24/7 availability that accommodates off-hour investor needs, and (2) multi-topic dialogue that builds client relationship through casual interactions.

We examine heterogeneous treatment effects by comparing different subgroups within the treatment group to the control group. Table 4 shows these results, the differential efficacy across usage contexts reveals several distinct advantages of LLM-based chatbots in robo-advising. First, their 24/7 availability, evidenced by superior outcomes for non-trading/non-working hour users, demonstrates their capacity to deliver just-in-time financial education when investors are cognitively available to process complex concepts like alphas and betas. Panel A shows that users who primarily use the LLM chatbot during non-trading hours achieve a 0.43% higher raw return, compared to a 0.36% higher raw return for those who mainly use the chatbot during trading hours, both relative to the control group. Panel B shows that users who primarily use the LLM chatbot during non-working hours achieve a 0.47% higher raw return, compared to a 0.32% higher raw return for those who mainly use the chatbot during working hours, both relative to the control group. We distinguish user groups based on the median number of tokens spent during each time period. Unlike human advisors constrained by office hours or rule-based chatbots limited to scripted interactions, LLMs adaptively support reflective learning during evenings or weekends, aligning with evidence that off-hours financial planning improves decision quality. Table 7 subfigures C and D also show that investors spend close to 50% of their time using LLM during the weekends, and they mostly use the LLM after work instead of during trading hours.

Second, Panel C shows that users who engage in mixed-topic conversations with the LLM chatbot achieve a 0.45% higher raw return, compared to a 0.34% higher raw return for users who focus exclusively on investment-related topics, both relative to the control group. We again classify users based on the median number of LLM token used, using this cutoff to distinguish between mixed-topic and investment-focused interactions. The higher alpha observed among users who engage in mixed-topic conversations suggests that LLMs foster trust through holistic interactions; discussing non-investment topics such as tax planning and life goals may help replicate the relationship-building typical of high-touch private banking. By contextualizing portfolio advice within personal circumstances, LLMs replicate the "client

bonding" traditionally exclusive to human wealth managers, but at near-zero marginal cost. This is particularly consequential in emerging markets like China, where distrust in human advisors (cited by 62% in post-surveys) stems from perceived conflicts of interest, whereas LLMs' neutrality enhances credibility.

The broader implications might be scalable personalization. Panel D shows that users with above-median interaction time with the LLM chatbot achieve a 0.44% higher raw return, compared to a 0.37% higher raw return for users with below-median interaction time, both relative to the control group. The intensity results further underscore LLMs' ability to deliver tiered personalization. Casual users benefit from baseline automation, while more engaged investors receive increasingly personalized guidance, all within a single platform. This scalability resolves the longstanding trade-off between cost and customization in financial advice, democratizing services once reserved for high-net-worth individuals. Critically, LLMs achieve this not merely through algorithmic precision but by creating a premium: investors perceive the chatbot as a dedicated advisor rather than a tool, despite the latter's operational reality. While the differences between all four groups lack statistical significance, the directional consistency supports our overall hypothesis. Our findings thus redefine robo-advising (combined with LLM)'s value proposition—it is not just portfolio optimization, but AI's capacity to simulate relational depth while retaining low-cost efficiency, that drives adoption and efficacy.

4.3 One Click automated Robo-advisors and LLM chatbot

Robo-advisors often offer a one-click switch option, enabling investors to transition to a recommended portfolio without further reliance on human judgment. A key question we seek to investigate is whether continuous assistance from an LLM chatbot can produce longer-term improvements in investors' ability to independently manage their portfolios—even in the absence of the one-click switch feature. To address this question, in the randomized controlled trial in which investors were randomly assigned to receive one of several types of

advisory assistance, half of the investors receive a one-click switch option to a recommended robo-advisor portfolio and the other half need to manage their portfolios themselves (DIY). The one-click option allows investors to seamlessly transition to a professionally managed portfolio by simply clicking an agreed button, while the DIY option requires investors to implement the recommended changes on their own.

We test these hypothesis and show the results in Table 5. Table 5 shows the effectiveness of different advisory methods on investor portfolio performance, specifically comparing the performance of one-click robo-advisor portfolios with DIY (do-it-yourself) portfolios. The table presents the difference-in-differences regression coefficients and t-statistics for CAPM alpha, which serves as a measure of risk-adjusted return.

Across all groups, we observe that the one-click alpha tends to be higher than the DIY alpha. This indicates that the simplicity and ease of switching to a robo-advisor portfolio with a single click generally lead to better performance outcomes. Moreover, the key finding in this table is the significant impact of LLM (ChatGPT/DeepSeek/Kimi/Doubao) follow-up assistance on investor understanding and performance. For the LLM chatbot group, the one-click alpha is 0.59%, and the DIY alpha is 0.58%, with an insignificant difference of just 1 bps per month. This suggests that investors who received follow-up assistance from the ChatGPT chatbot were able to effectively manage their portfolios themselves, achieving performance nearly identical to the one-click option. The high t-statistics (5.92 and 5.79) further underscore the robustness of these results.

In stark contrast, other types of follow-up assistance show a substantial gap between one-click and DIY alphas. For example, the group without any chatbot assistance shows a one-click alpha of 0.33% compared to a DIY alpha of just 0.03%, resulting in a significant difference of 30 bps. Similar patterns are observed for the Human Chat Assistant and Normal Chatbot groups, with differences of 25 and 18 bps, respectively.

These findings suggest that while initial financial education is beneficial, not all investors fully grasp the distinction between alpha and beta. As a result, they may rely more heavily

on the straightforward one-click option, which leads to higher performance. However, the LLM chatbot's ability to provide clear and effective follow-up assistance enables investors to apply their knowledge more effectively, thereby narrowing the performance gap between one-click and DIY portfolios.

In summary, the table demonstrates that the LLM (ChatGPT/DeepSeek/Kimi/Doubao) chatbot not only enhances initial financial literacy but also significantly improves investors' ability to manage their portfolios independently, achieving performance levels comparable to the more automated one-click options. This underscores the potential of advanced AI-driven advisory tools to empower investors with both knowledge and practical portfolio management skills. The results from our tests also suggest that while initial financial education positively impacts investor behavior, the effectiveness of follow-up assistance plays a critical role in ensuring that investors internalize and apply their knowledge. Specifically, the LLM (ChatGPT/DeepSeek/Kimi/Doubao) chatbot demonstrates a significant capability in bridging the gap between education and practical application, enabling investors to achieve comparable performance in their DIY portfolios as in the one-click portfolios. This has profound implications for investors' long-term financial independence and literacy, as access to LLM chatbots enables them to manage their own portfolios.

4.4 Heterogeneity Based On Investor Characteristics

Table 6 presents the difference-in-differences regression coefficients and t-statistics (in parentheses) for various performance indicators, including Raw Return, Alpha, and Portfolio Diversification (measured by the Herfindahl Index), across different investor characteristics. The subgroups analyzed are based on age, account age, account size, and education level.

Overall, the results indicate significant heterogeneity in investment performance across different investor segments. Younger investors aged 18-35 achieved 0.48% higher Raw Returns, with a t-statistic of 5.12, and 0.29% higher monthly Alpha (t-statistic 5.45). This group also exhibited a notable improvement in Portfolio Diversification, as indicated by a

Herfindahl Index of 0.13 lower (t-statistic -4.53). Investors aged 36-50 showed 0.38% higher Raw Returns (t-statistic 3.85) and 0.20% higher monthly Alpha (t-statistic 3.91), with a Herfindahl Index of 0.10 lower (t-statistic -2.82). The 51+ age group demonstrated 0.34% higher Raw Returns (t-statistic 3.20) and 0.14% higher monthly Alpha (t-statistic 3.18), along with a Herfindahl Index of 0.08 lower (t-statistic -2.13).

Examining account age, investors with accounts aged 0-5 years exhibited the highest performance with 0.46% higher Raw Returns (t-statistic 5.01) and 0.27% higher monthly Alpha (t-statistic 5.22), with a Herfindahl Index of 0.12 lower (t-statistic -4.23). Those with 6-10 year-old accounts had 0.39% higher Raw Returns (t-statistic 3.91) and 0.21% higher monthly Alpha (t-statistic 4.13), with a Herfindahl Index of 0.10 lower (t-statistic -2.74). Investors with accounts older than 11 years showed 0.32% higher Raw Returns (t-statistic 3.02) and 0.14% higher monthly Alpha (t-statistic 3.11), with a Herfindahl Index of 0.08 lower (t-statistic -1.92).

Regarding account size, investors with low account sizes (<300k RMB) achieved 0.35% higher Raw Returns (t-statistic 3.42) and 0.18% higher monthly Alpha (t-statistic 3.53), with a Herfindahl Index of 0.08 lower (t-statistic -2.01). Medium-sized accounts (300k-1M RMB) exhibited 0.41% higher Raw Returns (t-statistic 4.52) and 0.22% higher monthly Alpha (t-statistic 4.63), with a Herfindahl Index of 0.11 lower (t-statistic -3.02). High account sizes (>1M RMB) showed 0.44% higher Raw Returns (t-statistic 4.82) and 0.24% higher monthly Alpha (t-statistic 4.92), with a Herfindahl Index of 0.12 lower (t-statistic -3.51).

Finally, analyzing education levels, investors with a high school education achieved 0.33% higher Raw Returns (t-statistic 3.12) and 0.16% higher monthly Alpha (t-statistic 3.22), with a Herfindahl Index of 0.07 lower (t-statistic -1.93). Those with a Bachelor's Degree exhibited 0.39% higher Raw Returns (t-statistic 4.01) and 0.21% higher monthly Alpha (t-statistic 4.12), with a Herfindahl Index of 0.10 lower (t-statistic -2.61). Investors with a postgraduate degree showed 0.42% higher Raw Returns (t-statistic 4.32) and 0.23% higher monthly Alpha (t-statistic 4.43), with a Herfindahl Index of 0.11 lower (t-statistic -2.93). Together, these

results suggests that younger investors with larger account size and higher education level benefit more from the treatment.

4.5 Investor Behavior

We hypothesize that, after learning about beta and alpha and gaining access to LLM-backed chatbots and robo-advisors, investors will exhibit changes in their behavior. Specifically, we hypothesize that investors would spend less time trading after they switch to passive investment. Table 7 panel A presents the regression results for the impact of treatment on three variables: monitoring time, login counts, and trading frequency, across three different time periods: one month, three months, and six months. The coefficients, along with their corresponding t-statistics, are reported for each variable and time period. In the one-month period after robo-advisor implementation, the coefficients for the treatment group are 0.05, 0.22, and 0.17 for monitoring time, login counts, and trading frequency, respectively. These coefficients are positive but statistically insignificant, as indicated by the t-statistics (0.13, 0.09, and 0.11, respectively).

For the three-month period after robo-advisor implementation, the treatment group coefficients show a significant reduction in all three variables. The coefficients are -1.53, -5.65, and -4.50 for monitoring time, login counts, and trading frequency, respectively, with corresponding t-statistics of -5.33, -4.38, and -3.51. These results indicate a significant negative impact of the treatment over the three-month period as the mean value of these variables before treatment are 5.63, 29.93, and 10.87 correspondingly. In the six-month period after robo-advisor implementation, the results are similar to those observed in the three-month period, suggesting a sustained negative impact of the treatment. The treatment group coefficients are -1.59, -6.70, and -4.27 for monitoring time, login counts, and trading frequency, with t-statistics of -5.31, -4.59, and -3.67, respectively.

In summary, while the one-month results indicate no significant effect of the treatment, the three-month and six-month results reveal a significant and sustained reduction in moni-

toring time, login counts, and trading frequency. This suggests that the treatment effectively reduces these activities over longer periods, with the magnitude of the effect remaining consistent between the three-month and six-month intervals. Moreover, the results in Panel A are highly consistent with Rossi and Utkus (2024) that robo-advisors "untie" investors from trading.

Panel B shows the results when we estimate a triple difference-in-differences regression where we add an indicator LLM, that equals one if the treatment group has LLM-based follow-up assistance. The results show that when investors have LLM-based follow-up assistance, there is a significant marginal decrease in monitoring time, login time, and trading frequency. This suggests that LLM-based follow-up assistance indeed helps investors manage their portfolios more efficiently by reducing the need for constant oversight and potentially improving their decision-making process.

4.6 Fund Characteristics

We also investigate the changes in investor's holding in terms of fund characteristics. Figure 2 shows that there is a significant decrease of average mutual fund management fee being charged for the treatment group but not for the control group. Table 8 presents the difference-in-differences regression coefficients and t-statistics (in parentheses) for various characteristics of mutual funds that investors purchase, including Expense Ratio, Turnover Ratio, and Beta. The analysis focuses on the impact of the robo-advising service on these mutual fund characteristics. The average expense ratio before the treatment was 1.4%. The coefficient for the Expense Ratio in the treatment group is -0.35 (35 bps lower), with a t-statistic of -4.84, indicating a statistically significant reduction at the 1% significance level. This suggests that investors using the robo-advising service tend to choose funds with lower expense ratios. The average annual turnover ratio before the treatment was 24.6. The Turnover Ratio for the treatment group is -4.18, with a t-statistic of -3.21, also significant at the 1% level, indicating that these investors prefer funds with lower turnover

ratios after received robo advising, and these funds are ETFs that mechanically have lower turnovers. However, the Beta coefficient for the treatment group is 0.03, with a t-statistic of 0.07, showing no significant change in the risk profile of the funds selected by investors using the robo-advising service. Overall, these results indicate that the robo-advising service significantly influences investors' preferences for mutual fund characteristics, particularly in reducing expense and turnover ratios, which can lead to better investment outcomes.

Table 9 presents the results of the difference-in-differences regression analysis examining the impact of the robo-advising service on investor account net flows, specifically focusing on overall net flow and ETF net flow in percentage terms. The table reports the coefficients and corresponding t-statistics (in parentheses) for the interaction term (Robo*Post). The results suggest that there is a significant increase (0.637%) in overall inflow into the treatment group's accounts compared to the control group's accounts, which implies a moderate increase in stock market participation among the treatment group.[17] We also show that this inflow mainly goes to ETF investment, indicating a significant increase in allocation to ETF investment.

4.7 Oversea diversification and factor investment

One other interesting result we find is that, investors in the treatment group actually also start to diversify to S&P 500 ETFs and other foreign ETFs, even though the initial recommendation only consists of reallocating the Chinese A-share market ETFs.[18] Table 10 presents the results of the difference-in-differences regression analysis examining the impact of the robo-advising service on investors' allocation to S&P500, US Factor Style Investment, and Chinese ADRs. The table reports the coefficients and corresponding t-statistics (in parentheses) for the interaction term (Robo*Post).

[17]This implies that each month, an investor who initially has \$150,000 would deposit an additional $150{,}000 \times 0.637\% = 955.50$ at the end of the first month. In the second month, the investor would deposit another $150{,}000 \times (1 + 0.637\%) \times 0.637\% \approx 961.56$, with the pattern continuing in subsequent months. Over a one year period after treatment, the inflow of capital is roughly 7.7%.

[18]Although these are not directly recommended from the robo-advisors, the choices are available through investor's investment menu.

Before the treatment, the mean percentage that investors allocate to S&P 500 ETFs, US Factor Style Investments, and Chinese ADRs is below 1% of the overall portfolio. After the treatment, for the allocation to S&P500, the coefficient for the interaction term (Robo*Post) is 2.04, with a t-statistic of 2.24, indicating a statistically significant increase at the 5% significance level. This suggests that the robo-advising service significantly increases the allocation to S&P500 by 2.04 percentage points. Similarly, After the treatment, for the allocation to US Factor Style Investment, the coefficient for the interaction term (Robo*Post) is 1.52, with a t-statistic of 1.99, indicating a statistically significant increase at the 5% significance level. This implies that the robo-advising service significantly increases the allocation to US Factor Style Investment by 1.52 percentage points. For the allocation to Chinese ADRs, the coefficient for the interaction term (Robo*Post) is 1.81, with a t-statistic of 2.01, indicating a statistically significant increase at the 5% significance level. This suggests that the robo-advising service significantly increases the allocation to Chinese ADRs by 1.81 percentage points.

This result is puzzling given that the robo-advisor did not recommend foreign diversification directly. Therefore, we hypothesize that LLM-based chatbot may have recommended foreign diversification beyond what the brokerage firms initially recommended to them. Indeed, Panel B of Table 10 presents the results of the triple difference-in-differences regression, which includes an additional indicator for LLM (AI) interaction effects as in regression specification (2). The coefficient for the Robo*Post*LLM interaction term is 3.48 for S&P500, 3.07 for US Factor Style Investment, and 4.71 for Chinese ADRs, all with t-statistics greater than 4, indicating statistically significant increases at the 1% significance level. This further suggests that LLM-based follow-up assistance significantly enhances the impact of the robo-advising service on these allocations. At the same time, Robo*Post coefficients are not significant when controlling for Robo*Post*LLM, suggesting that the effect is concentrated in the subgroup that use LLM chatbots.[19]

[19]There is also additional evidence from the textual conversations extracted from the LLM indicating that these models tend to recommend diversification in overseas markets. For brevity, we have not presented

Overall, the results suggest that the introduction of LLM-based follow-up assistance significantly amplifies the impact of robo-advising services on foreign diversification, guiding investors towards greater allocations in S&P 500 ETFs, US Factor Style Investments, and Chinese ADRs. This highlights the potential of AI-driven assistance to enhance and extend the recommendations of traditional robo-advisors, leading to more diversified investment portfolios.

5 Robustness Checks

5.1 Beyond CAPM Alpha

In this robustness check, we examine the impact of different advisory methods on financial literacy using various factor models to compute alphas. Table A.1 presents the difference-in-differences regression coefficients along with their respective t-statistics for the Four Factor Alpha configurations. The alphas are calculated using the Fama-French Three Factor Model (Fama and French, 1992), the Three Factor Model with Momentum (Carhart, 1997), the Fama-French Five Factor Model (Fama and French, 2015), and the Five Factor Model with Momentum (FF Five + MOM Alpha).

The table highlights the comparative effectiveness of different robo-advisory approaches: a LLM-based robo-advisory chatbot, a traditional rule-based chatbot, a non-premium human chat assistant, and a control group of investors who intended to but did not receive the advisory service. The results across all alpha definitions consistently show that the LLM-based chatbot provides the most significant positive impact on investors' performance and financial literacy.

The LLM (ChatGPT/DeepSeek/Kimi/Doubao) Chatbot group generates alphas of 0.44%, 0.43%, 0.41%, and 0.40% across the FF Three Alpha, FF Three + MOM Alpha, FF Five Alpha, and FF Five + MOM Alpha, respectively, with highly significant t-statistics ranging

these results here, but they are available upon request.

from 5.70 to 6.27. These results clearly demonstrate the superior effectiveness of the LLM-based chatbot in enhancing financial outcomes compared to both traditional methods and other chatbots.

Conversely, the group without any chatbot support shows lower, yet still significant, alphas ranging from 0.13% to 0.10%, with t-statistics between 1.99 and 2.25. This suggests that while traditional methods do have a positive impact, their effectiveness is markedly less than that of the LLM-based advisory.

The Human Chat Assistant and Normal Chatbot groups also show significant positive alphas, but these are lower than those of the LLM Chatbot group. The Human Chat Assistant group shows alphas ranging from 0.14% to 0.11% with t-statistics between 2.24 and 2.41, while the Normal Chatbot group shows alphas ranging from 0.12% to 0.09% with t-statistics between 2.03 and 2.29.

In summary, the robustness check confirms that the LLM-based robo-advisor significantly outperforms both human and traditional chatbot advisors in improving financial outcomes, as indicated by the higher and more significant alphas across multiple factor models.

5.2 Alternative robo-advising model

We also conduct robustness checks based on the alternative robo-advising models that the brokerage firm applied during the pilot period. Our current empirical analysis is based on the sample where treated investors are advised to retain mutual funds or stocks that are performing well, using a simple rule of thumb to keep only those with positive alpha (after fees) and a t-statistic above 3. However, extensive literature has documented that past performance of stocks and mutual funds cannot reliably predict future performance, raising questions about whether the current model is the most beneficial for investors. Despite this, it is noteworthy that only a small percentage of stocks or mutual funds are suggested to be kept by investors under this model.

Additionally, the brokerage firm experimented with alternative robo-advising models.

One model applied a "one-size-fits-all" approach, suggesting that the entire portfolio be converted to passive ETFs. Another model involved a complex algorithm incorporating machine learning techniques and predictors such as past mutual fund performance, fund size, expense ratio, turnover ratios, manager tenure, active share, fund flows (for mutual funds), and neural network analysis for stocks to make keep-or-switch recommendations. Interestingly, the results from the "one-size-fits-all" model, the machine learning model, and the simple rule of thumb based on retaining mutual funds and stocks with positive alpha (after fees) and a t-statistic above 3 were very similar.

6 Conclusion

In this study, we investigate the impact of large language models (LLMs) and generative AI on investor behavior and decision-making within a robo-advisory context. Specifically, we integrated a back-end LLM system with a robo-advisor to serve as an interactive tool for investors. This AI-powered system enabled users to engage in natural language conversations, ask complex questions about investment strategies, and receive tailored, real-time explanations and guidance.

The LLM also facilitated financial education by helping investors better understand key concepts such as beta and alpha in mutual funds and stocks. Our partnership with a large brokerage firm enabled us to deploy this AI-enhanced robo-advisory platform to a broad user base and analyze its effects on both investor engagement and knowledge acquisition. Our findings reveal several significant outcomes resulting from the integration of LLM technology into the investment advisory process.

First, investors who learned about beta demonstrated a noticeable shift in their investment preferences. They showed a preference for investing in ETFs or index funds rather than directly in mutual funds or individual stocks. This shift suggests that understanding beta helps investors appreciate the benefits of diversification and the lower costs associated

with passive investment vehicles.

Second, and most importantly, the LLM chatbot system greatly enhances the traditional robo-advisor system. It significantly amplifies the economic benefits documented with traditional robo-advisors and also seems to educate investors on key financial concepts such as beta and alpha. Additionally, it helps investors achieve global diversification, a feature that our pre-programmed robo-advisors did not include in their recommendations.

Third our study highlights the tangible financial benefits of financial education. Investors who received education on beta and alpha experienced a substantial improvement in their investment performance, with their monthly returns increasing by 40 basis points. Additionally, these investors achieved an average monthly gain of 750 RMB. These results underscore the value of financial literacy in enhancing investor outcomes.

Our research contributes to the growing literature on the application of artificial intelligence and large language models in the investment advisory space. By integrating AI-powered tools such as ChatGPT within robo-advisory platforms, we demonstrate how these technologies can enhance investor engagement, support interactive learning, and improve investment outcomes. Leveraging big data and advanced analytics, AI-driven systems empower investors to better understand complex financial concepts and make more informed decisions, ultimately fostering a more resilient and financially literate investment community (D'Acunto, Prabhala and Rossi, 2019; Rossi and Utkus, 2024).

References

Calvet, Laurent E, John Y Campbell, and Paolo Sodini, "Fight or flight? Portfolio rebalancing by individual investors," *The Quarterly Journal of Economics*, 2009, *124* (1), 301–348.

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

Carhart, Mark M, "On persistence in mutual fund performance," *The Journal of Finance*, 1997, *52* (1), 57–82.

Chak, Ida, Karen Croxson, Francesco D'Acunto, Jonathan Reuter, Alberto G Rossi, and Jonathan M Shaw, "Improving household debt management with robo-advice," *NBER Working Paper*, 2022.

Cong, Lin William, Ke Tang, Jingyuan Wang, and Yang Zhang, "AlphaPortfolio: Direct construction through deep reinforcement learning and interpretable AI," *SSRN Working Paper*, 2019.

_ , Tengyuan Liang, Xiao Zhang, and Wu Zhu, "Textual factors: A scalable, interpretable, and data-driven approach to analyzing unstructured information," *NBER Working Paper*, 2024.

D'Acunto, Francesco and Alberto G Rossi, "Robo-advice: Transforming households into rational economic agents," *Annual Review of Financial Economics*, 2023, *15* (1), 543–563.

_ , Nagpurnanand Prabhala, and Alberto G Rossi, "The promises and pitfalls of robo-advising," *The Review of Financial Studies*, 2019, *32* (5), 1983–2020.

D'Acunto, Francesco and Alberto G Rossi, "New frontiers of robo-advising: Consumption, saving, debt management, and taxes," *Machine Learning in Finance: A Guide to Contemporary Practice, forthcoming*, 2021.

_ and _ , "IT meets finance: financial decision-making in the digital era," *Handbook of Financial Decision Making*, 2023, pp. 336–354.

Egan, Mark, Gregor Matvos, and Amit Seru, "The market for financial adviser misconduct," *Journal of Political Economy*, 2019, *127* (1), 233–295.

Fama, Eugene F and Kenneth R French, "The cross-section of expected stock returns," *the Journal of Finance*, 1992, *47* (2), 427–465.

_ and _ , "A five-factor asset pricing model," *Journal of Financial Economics*, 2015, *116* (1), 1–22.

Gargano, Antonio and Alberto G Rossi, "Goal setting and saving in the FinTech era," *The Journal of Finance*, 2024, *79* (3), 1931–1976.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny, "Money doctors," *The Journal of Finance*, 2015, *70* (1), 91–114.

Greig, Fiona, Tarun Ramadorai, Alberto G Rossi, Stephen P Utkus, and Ansgar Walther, "Algorithm aversion: Theory and evidence from robo-advice," *SSRN Working Paper*, 2022.

Horel, Enguerrand, Kay Giesecke, Victor Storchan, and Naren Chittar, "Explainable clustering and application to wealth management compliance," *Proceedings of the First ACM International Conference on AI in Finance*, 2020, pp. 1–6.

Lewellen, Jonathan, "Institutional investors and the limits of arbitrage," *Journal of Financial Economics*, 2011, *102* (1), 62–80.

_ , "How Many Factors?," *Working Paper*, 2023.

_ **and Katharina Lewellen**, "Institutional investors and corporate governance: The incentive to be engaged," *The Journal of Finance*, 2022, *77* (1), 213–264.

Liberman, Joseph, Clemens Sialm, Nathan Sosner, and Lixin Wang, "The Tax Benefits of Separating Alpha from Beta," *Financial Analysts Journal*, 2020, *76* (1), 38–61.

Rooij, Maarten Van, Annamaria Lusardi, and Rob Alessie, "Financial literacy and stock market participation," *Journal of Financial Economics*, 2011, *101* (2), 449–472.

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

Sheng, Jinfei, Zheng Sun, Baozhong Yang, and Alan L Zhang, "Generative AI and Asset Management," *SSRN Working Paper*, 2024.

Sialm, Clemens and David Xiaoyu Xu, "Information Acquisition By Mutual Fund Investors: Evidence From Stock Trading Suspensions," *Working Paper*, 2024.

_ **and T Mandy Tham**, "Spillover effects in mutual fund companies," *Management Science*, 2016, *62* (5), 1472–1486.

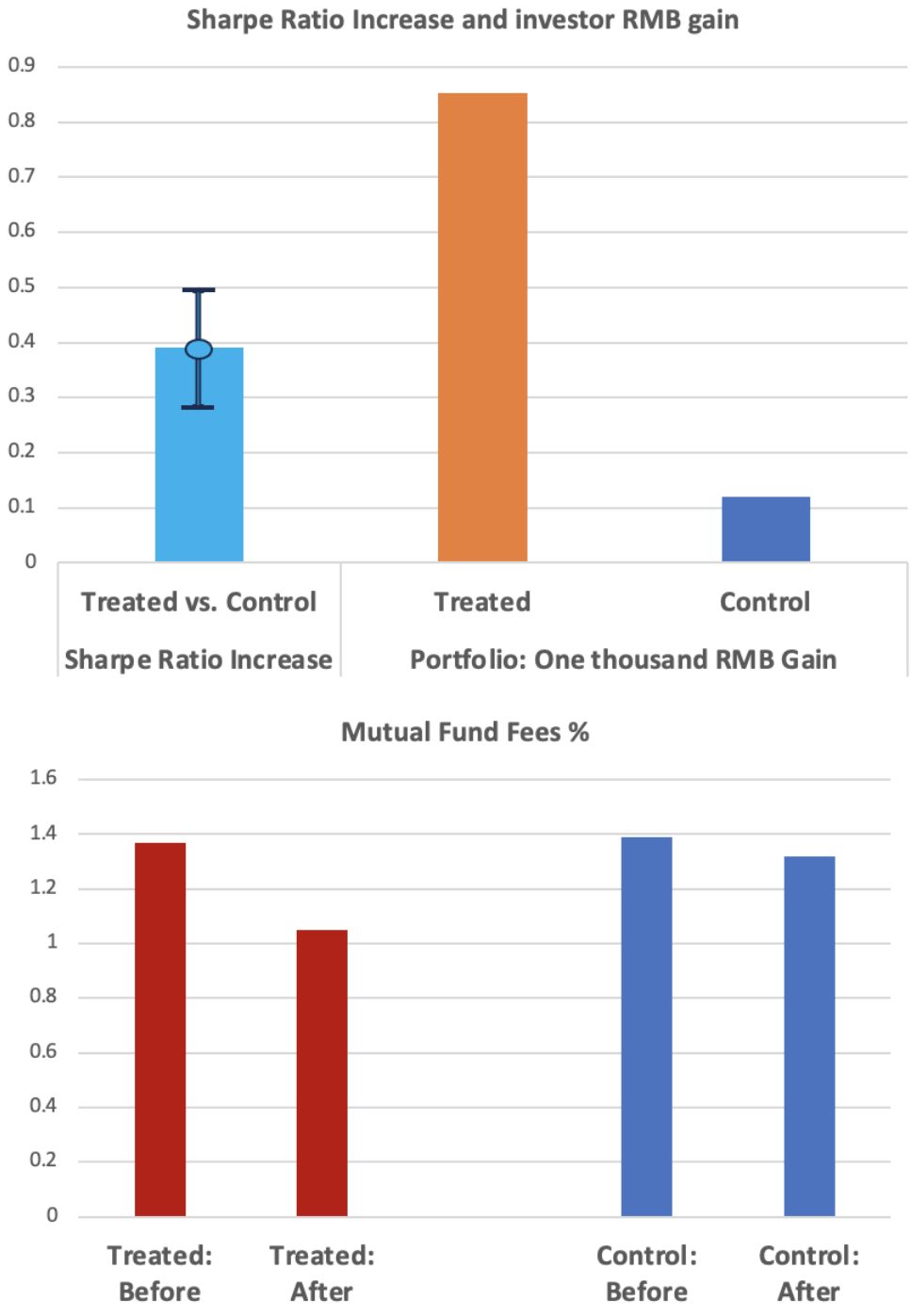
Tsoukalas, Gerry, Jiang Wang, and Kay Giesecke, "Dynamic portfolio execution," *Management Science*, 2019, *65* (5), 2015–2040.

Figure 1: Robo-advising and portfolio reallocation for mutual fund investors



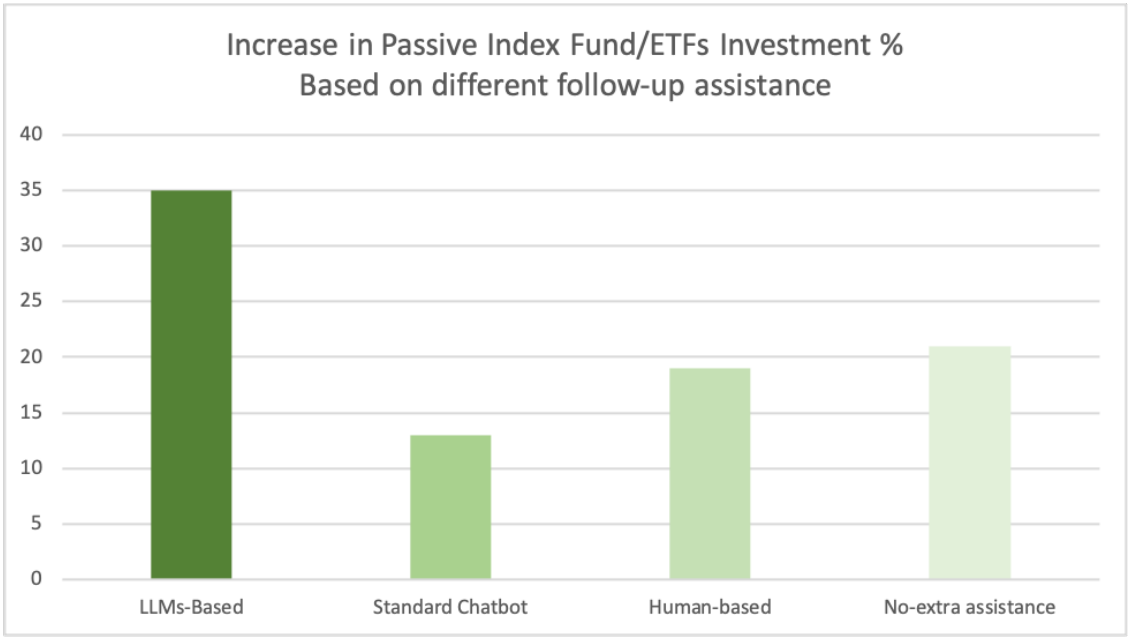
This Figure shows the portfolio reallocation after robo-advising. The "Treated: Recommended Level" refers to the percentage allocation recommended by the robo-advisor. "Before" and "After" for Treated and Control indicate the actual allocation before and after the implementation of the robo-advisor, shown separately for the Treatment group and the Control group.

Figure 2: Robo-advising and investors' gain for mutual fund investors



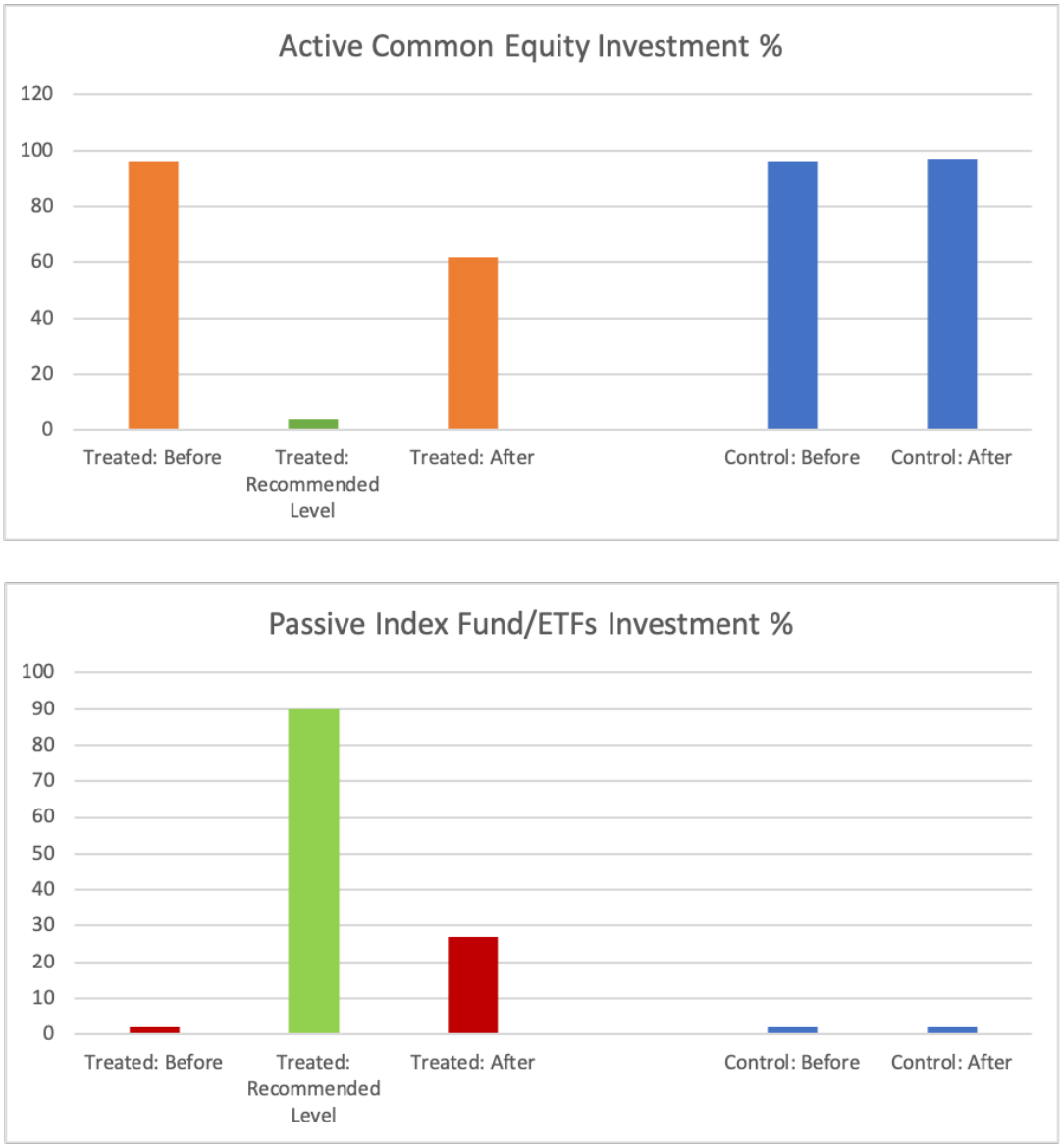
This Figure shows the Sharpe ratio increase and monetary gain after robo-advising.

Figure 3: Heterogeneous effect based on different follow-up assistances for mutual fund investors



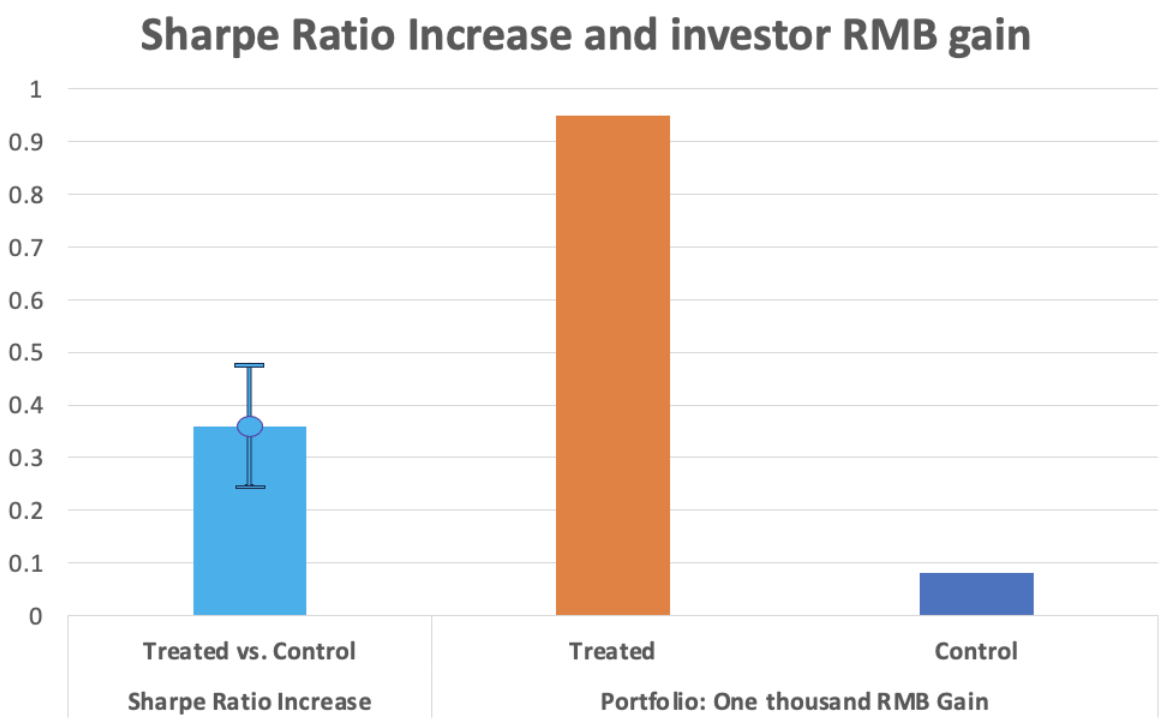
This Figure shows the heterogeneous treatment effect based on different follow-up assistances.

Figure 4: Robo-advising and portfolio reallocation for stock investors



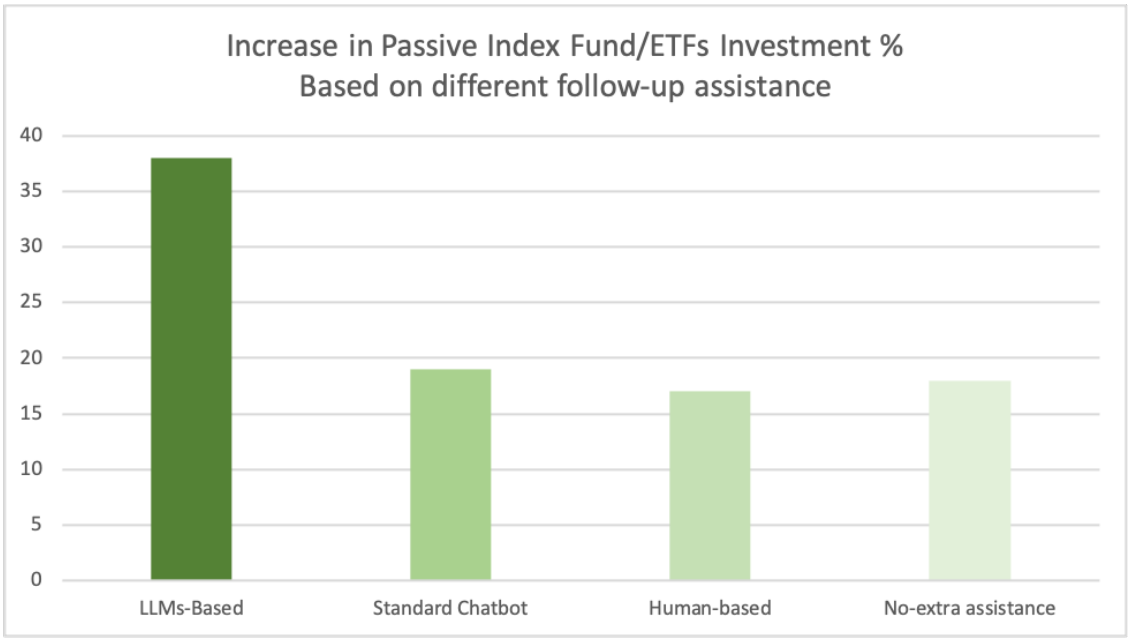
This Figure shows the portfolio reallocation after robo-advising. The "Treated: Recommended Level" refers to the percentage allocation recommended by the robo-advisor. "Before" and "After" for Treated and Control indicate the actual allocation before and after the implementation of the robo-advisor, shown separately for the Treatment group and the Control group.

Figure 5: Robo-advising and investors' gain for stock investors



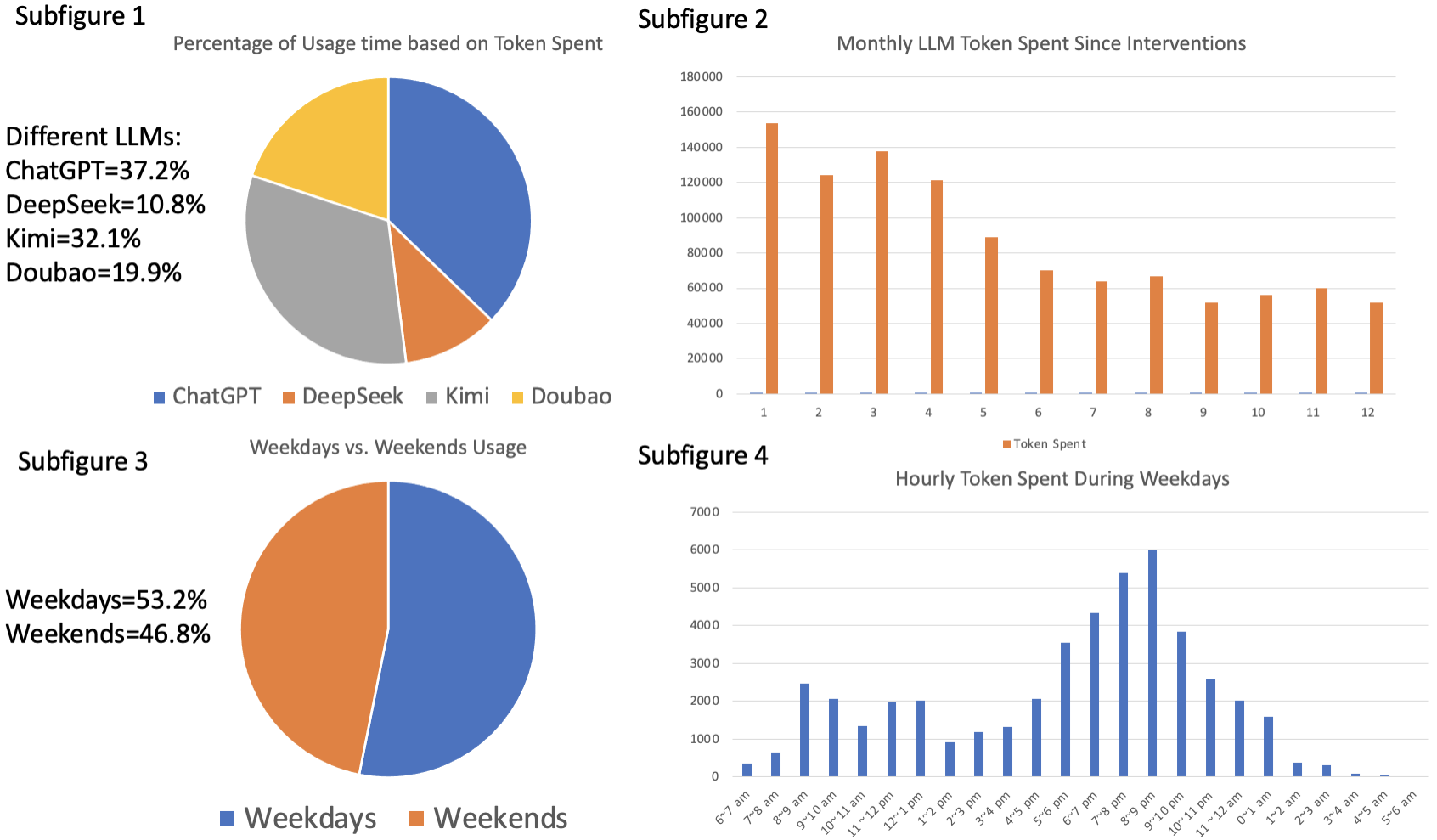
This Figure shows the Sharpe ratio increase and monetary gain after robo-advising.

Figure 6: Heterogeneous effect based on different follow-up assistances for stock investors



This Figure shows the heterogeneous effect based on different follow-up assistances.

Figure 7: LLM usage breakdown



This Figure shows the LLM usage breakdown by LLM types (subfigure 1), usage by month since the treatment (subfigure 2), weekdays vs. weekends usage (subfigure 3), usage by hour during weekdays (subfigure 4).

Table 1: Summary Statistics

This table presents summary statistics for investor and portfolio characteristics immediately prior to the implementation of the robo-advisor in 2023. For each characteristic, the mean and standard deviation (in parentheses) are reported for the period before the treatment. The variables include investor age, account age, initial portfolio value, gender, number of stocks held, number of mutual funds held, average three-year raw returns, monitoring time, login counts, and trading frequency, all measured in the month prior to treatment. Monitoring time is defined as the number of hours logged in per month; login count refers to the number of logins per month; and trading frequency is the number of trades per month. Additionally, the table reports the mean differences between the treatment and control groups. T-tests for these mean differences are performed, and the results are marked with *, **, and *** to signify statistical significance at different levels. Specifically, * indicates significance at the 10% level, ** indicates significance at the 5% level, and *** indicates significance at the 1% level.

Characteristics	**Treatment Group**	**Control Group**	**Difference**
Age (years)	45.63 (10.21)	46.19 (10.54)	-0.56
Account age (years)	8.56 (5.20)	8.71 (5.08)	-0.15
Initial Portfolio Value ($1000)	184.4 (50.39)	182.84 (51.21)	1.57
Gender (Male ratio)	0.74 (0.44)	0.76 (0.43)	-0.02
Number of Stocks held	15.38 (10.42)	14.82 (10.75)	0.56
Number of Mutual Funds held	5.63 (3.21)	5.89 (3.17)	-0.26
Returns Three Years before intervention (%)	-8.45 (2.58)	-8.27 (2.64)	-0.18
N	14,409	14,410	

Trading Behavior	**Treatment Group**	**Control Group**	**Difference**
Monitoring Time	5.63 (12.48)	5.51 (14.29)	0.12
Login Counts	29.93 (31.49)	29.52 (33.84)	0.42
Trading Frequency	10.87 (24.15)	10.54 (26.23)	0.33
N	14,409	14,410	

Table 2: Robo-advising and Investment Outcomes

This Table shows the relationship between robo-advising and investment outcomes separately for mutual fund investors only, stock investors, and investors in both markets after the implementation of the robo-advisor. The table shows the metrics for the control and treatment groups, along with the calculated differences. The differences are marked with *, **, and *** to signify statistical significance at different levels. Specifically, * indicates significance at the 10% level, ** indicates significance at the 5% level, and *** indicates significance at the 1% level. The metrics include Average Annual Return, Alpha (CAPM Alpha), Portfolio Diversification (measured by the Herfindahl Index), and Investor Satisfaction (measured on a scale of 1 to 5 from in-app rating). Panel A presents results for Mutual Fund Investors, Panel B for Stock Investors, and Panel C for Investors in both markets.

Panel A: Mutual Fund Investors

Metric	**Control Group**	**Treatment Group**	**Difference**
Average Annual Return (%)	10.24	14.89	4.65***
Annualized Alpha (%)	-2.33	0.11	2.44***
Investor Satisfaction (out of 5)	3.81	4.57	0.76**
N	2,849	2,849	

Panel B: Stock Investors

Metric	**Control Group**	**Treatment Group**	**Difference**
Average Annual Return (%)	10.41	14.53	4.12**
Annualized Alpha (%)	-2.24	0.16	2.40**
Portfolio Diversification (Herfindahl Index)	0.30	0.19	-0.11***
Investor Satisfaction (out of 5)	3.72	4.46	0.74**
N	6,381	6,381	

Panel C: Investors in both markets

Metric	**Control Group**	**Treatment Group**	**Difference**
Average Annual Return (%)	10.32	14.71	4.39***
Annualized Alpha (%)	-2.33	0.09	2.42***
Investor Satisfaction (out of 5)	3.77	4.51	0.74**
N	5,179	5,180	

Table 3: Investor Performance Outcomes: Comparing Human Assistance, LLM Chatbots, and Rule-Based Chatbots

Panel A shows the difference-in-differences regression coefficients β_2 in regression specification (1) and t-statistics (in parentheses). Treatment equals one if the investors intended to receive the robo-advising service and actually received the service. Control equals one if the investors intended to receive the robo-advising service but were not selected to receive the service. The overall sample result, as well as the results for four subtreatment groups, are reported. The four subtreatment groups are: LLM-based Chatbot, Human Chat Assistance, Normal Chatbot, and No Chatbot. The control group remains unchanged in these regressions. The sample period is from December 2022 to November 2024. Post equals one for observations (months) after December 2023. Raw Return is the monthly raw return. Alpha is the monthly CAPM alpha. Panel B shows the triple DiD regression coefficient β_3 in regression specification (2) for the whole sample. Panel C reports summary statistics on the heterogeneous effects of various follow-up services following the initial treatment. Mean and standard deviations (in parentheses) are reported. Response Time is the average response time in seconds. Interaction Times is the number of times investors interact with the chatbot or human live chat service per month. Investor sentiment is derived solely from the textual dialogue generated using the Chinese LLMs Kimi and Doubao. The sentiment value is scaled from -1 (lowest) to 1 (highest) using the VADER natural language processing package. Individual fixed effects and month fixed effects are included. Standard errors are clustered at the investor level.

Panel A: Performance (Monthly)	**Raw Return (%)**	**Alpha (%)**	**Herfindahl Index**	**N**
Overall	0.39***	0.21***	-0.11***	691,656
	(4.45)	(4.98)	(-3.03)	
ChatGPT (LLM) Chatbot	0.65***	0.43***	-0.17***	172,914
	(6.14)	(6.20)	(-3.61)	
Without Chatbot	0.22**	0.13**	-0.09**	172,914
	(2.20)	(2.38)	(-2.10)	
Human Chat Assistant	0.25**	0.14**	-0.04*	172,914
	(2.16)	(2.47)	(-1.73)	
Normal Chatbot	0.21**	0.12**	-0.05*	172,914
	(2.17)	(2.29)	(-1.84)	
Panel B: Performance (LLM)	**Raw Return (%)**	**Alpha (%)**	**Herfindahl Index**	**N**
LLM*Robo*Post	0.41***	0.29***	-0.08**	691,656
	(2.83)	(2.60)	(-2.35)	
Panel C: Investor Reaction	**Response time**	**Interaction times**	**Investor sentiment**	
ChatGPT (LLM) Chatbot	3.71	14.36	0.25	
	(1.42)	(28.92)	(0.37)	
Human Chat Assistant	64.89	5.14	-0.17	
	(264.7)	(12.4)	(0.46)	
Normal Chatbot	1.24	1.31	-0.24	
	(0.31)	(10.62)	(0.27)	

Table 4: Heterogeneity Tests based on LLM User Preference and Actual Usage

Panels A through D present the triple difference-in-differences (DiD) regression coefficient β_3 from specification (2), estimated for distinct subsamples of LLM chatbot users. Panel A categorizes investors based on chatbot usage timing: those who predominantly used the LLM-based chatbot during trading hours (9:25 AM – 3:00 PM) versus during non-trading hours. Weekends are also counted as non-trading hours. Panel B distinguishes between investors who primarily engaged with the LLM-based chatbot during working hours (8:00 AM – 8:00 PM) versus outside working hours. Weekends are also counted as non-working hours. Panel C separates users by chat content: those who discussed exclusively investment-related topics versus those who engaged in mixed discussions (investment and non-investment topics). Panel D groups investors by usage intensity, splitting the sample at the median total LLM tokens used in chatbot interactions. Treatment equals one if the investors intended to receive the robo-advising service and actually received the service. Control equals one if the investors intended to receive the robo-advising service but were not selected to receive the service. The sample period is from December 2022 to November 2024. Post equals one for observations (months) after December 2023. Raw Return is the monthly raw return. Alpha is the monthly CAPM alpha. Individual fixed effects and month fixed effects are included. Standard errors are clustered at the investor level.

Panel A: Trading hours vs. Non-trading hours	**Raw Return (%)**	**Alpha (%)**
Non Trading Hour: LLM*Robo*Post	0.43**	0.31**
	(2.24)	(2.40)
Trading Hour: LLM*Robo*Post	0.36*	0.22*
	(1.90)	(1.83)
Differences	0.07	0.09
t-statistics	(0.26)	(0.51)
Panel B: 8am to 8pm (Working) vs. 8pm to 8am (Non)	**Raw Return (%)**	**Alpha (%)**
Non-Working Hours: LLM*Robo*Post	0.47**	0.35***
	(2.45)	(2.79)
Work Hours: LLM*Robo*Post	0.32*	0.23*
	(1.81)	(1.60)
Differences	0.15	0.12
t-statistics	(0.57)	(0.63)
Panel C: Investment Topics Only vs. Mixed Topics	**Raw Return (%)**	**Alpha (%)**
Mixed Topics: LLM*Robo*Post	0.45**	0.33***
	(2.37)	(2.38)
Investment Topics Only: LLM*Robo*Post	0.34*	0.24***
	(1.94)	(1.86)
Differences	0.11	0.09
t-statistics	(0.42)	(0.48)
Panel D: High Interaction Time vs. Low Interaction Time	**Raw Return (%)**	**Alpha (%)**
High Interaction Time: LLM*Robo*Post	0.44***	0.32**
	(2.61)	(2.31)
Low Interaction Time: LLM*Robo*Post	0.37**	0.26*
	(2.13)	(1.89)
Differences	0.07	0.06
t-statistics	(0.28)	(0.30)

Table 5: One-Click Robot Advisor vs. DIY portfolio

This table presents the difference-in-differences regression coefficients and t-statistics (in parentheses) for various treatment group subsamples, as specified in regression specification (1). The Treatment equals one if the investors intend to receive the robo-advising service and actually received the robo-advising service. The Control equals one if the investor intend to receive the robo-adivisng service but was not selected to receive the service. The sample period is from December 2022 to November 2024. Post equals one for months after 2023 December. Alpha is the CAPM alpha. The treatment group is further sorted into two groups. One-click alpha refers to the treatment group's alpha for investors who can switch to a robo-advisor portfolio by simply clicking an agreed-upon button. DIY alpha refers to the treatment group's alpha for investors who must switch portfolios themselves. The control group remains unchanged in these regressions. Individual fixed effects and month fixed effects are included. Standard errors are clustered at the investor level.

Performance (Monthly)	**One-click Alpha (%)**	**DIY Alpha (%)**	**Difference**
Overall	0.38***	0.19*	0.19
	(3.14)	(1.73)	(1.16)
ChatGPT (LLM) Chatbot	0.59***	0.58***	0.01
	(5.92)	(5.79)	(0.08)
Without Chatbot	0.33***	0.03	0.30**
	(2.67)	(0.32)	(2.07)
Human Chat Assistant	0.34***	0.09	0.25
	(2.82)	(0.79)	(1.55)
Normal Chatbot	0.28**	0.10	0.18
	(2.31)	(0.64)	(0.90)

Table 6: Heterogeneity in terms of Investor Characteristics

This table presents the difference-in-differences regression coefficients and their corresponding t-statistics (in parentheses) for various subgroups based on different investor characteristics, as specified in regression specification (1). These characteristics include age, account age, account size, and education level. The table highlights the heterogeneity in investment performance, measured by monthly raw return, alpha, and the Herfindahl Index, across these different investor segments. Individual fixed effects and month fixed effects are included. Standard errors are clustered at the investor level. Significance levels are indicated by asterisks, where *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Investor Segment	**Raw Return (%)**	**Alpha (%)**	**Herfindahl Index**
Age			
18-35	0.48***	0.29***	-0.13***
	(5.12)	(5.45)	(-4.53)
36-50	0.38**	0.20**	-0.10**
	(3.85)	(3.91)	(-2.82)
51+	0.34**	0.14**	-0.08*
	(3.20)	(3.18)	(-2.13)
Account Age			
0-5 years	0.46***	0.27***	-0.12***
	(5.01)	(5.22)	(-4.23)
6-10 years	0.39**	0.21**	-0.10**
	(3.91)	(4.13)	(-2.74)
11+ years	0.32**	0.14**	-0.08*
	(3.02)	(3.11)	(-1.92)
Account Size			
Low (<300k RMB)	0.35**	0.18**	-0.08
	(3.42)	(3.53)	(-2.01)
Medium (300k-1M RMB)	0.41***	0.22***	-0.11**
	(4.52)	(4.63)	(-3.02)
High (>1M RMB)	0.44***	0.24***	-0.12***
	(4.82)	(4.92)	(-3.51)
Education Level			
High School	0.33**	0.16**	-0.07*
	(3.12)	(3.22)	(-1.93)
Bachelor's Degree	0.39**	0.21**	-0.10**
	(4.02)	(4.23)	(-2.72)
Master's/PhD	0.43***	0.25***	-0.12***
	(4.72)	(4.83)	(-3.42)

Table 7: Login Count and Monitoring Time

Panel A presents the difference-in-differences regression coefficients and their corresponding t-statistics (in parentheses) for different variables that measure investors' involvement in portfolio management, as specified in regression specification (1). Monitoring time is the number of hours logged in per month. Login count is the number of logins per month. Trading frequency is the number of trades per month. Panel B presents the triple difference-in-differences results, as specified in regression specification (2). Individual fixed effects and month fixed effects are included. Standard errors are clustered at the investor level. Significance levels are indicated by asterisks, where *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A			
Variables:	**Monitoring Time**	**Login Count**	**Trading Frequency**
One Month:			
Robo*Post	0.05	0.22	0.17
	(0.13)	(0.09)	(0.11)
Three Month:			
Robo*Post	-1.53***	-5.65***	-4.50***
	(-5.33)	(-4.38)	(-3.51)
Six Month:			
Robo*Post	-1.59***	-6.70***	-4.27***
	(-5.31)	(-4.59)	(-3.67)
Panel B (Six Month Only)			
LLM (AI) Interaction effects			
Robo*Post*LLM	-2.62***	-4.92***	-3.80***
	(-4.82)	(-4.21)	(-4.27)
Robo*Post	-1.23**	-4.48***	-3.31**
	(-2.39)	(-2.85)	(-2.18)

Table 8: Robo-advising and the Mutual fund Characteristics

Panel A presents the difference-in-differences regression coefficients along with their corresponding t-statistics (in parentheses) based on different characteristics of the funds that investors purchase, using regression specification (1). Panel B presents the triple difference-in-differences results using regression specification (2). The Turnover Ratio is defined as the annualized monthly turnover ratio of the mutual fund. Beta refers to the rolling monthly CAPM beta, estimated over the five years prior to the treatment. Individual fixed effects and month fixed effects are included. Standard errors are clustered at the investor level. Significance levels are indicated by asterisks, where *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A			
	Expense Ratio (%)	Turnover Ratio (%)	Beta
Robo*Post	-0.35***	-4.18***	0.03
	(-4.84)	(-3.21)	(0.07)
Panel B			
LLM (AI) Interaction effects			
Robo*Post*LLM	-0.51***	-4.63***	0.04
	(-3.69)	(-3.58)	(0.11)
Robo*Post	-0.13**	-3.38**	0.02
	(-2.41)	(-2.35)	(0.05)

Table 9: Individual Investor Account Netflow

This table reports both the net flow of deposits into or out of investors' accounts and the corresponding net flow into or out of their ETF holdings. The "Overall Funding Netflow" represents the percentage of net inflow (positive) or net outflow (negative) into an investor's account, calculated as the RMB amount of total deposits or withdrawals at the end of the month divided by the account balance at the beginning of the month. The "ETF Netflow" is defined as the net inflow (positive) or net outflow (negative) into ETF holdings at the end of the month, also expressed as a percentage of the account balance at the beginning of the month. Panel A shows the results using regression specification (1), and Panel B shows the results using regression specification (2). Individual fixed effects and month fixed effects are included. Standard errors are clustered at the investor level. Significance levels are indicated by asterisks, where *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A:		
	Overall Funding Netflow (%)	ETF Netflow (%)
Robo*Post	0.637***	0.419***
	(5.12)	(4.75)
Panel B		
LLM (AI) Interaction effects		
Robo*Post*LLM	0.849***	0.631***
	(4.25)	(4.19)
Robo*Post	0.342***	0.102***
	(3.48)	(3.51)

Table 10: Spillover Effect: Oversea Diversification

Panel A presents the difference-in-differences regression coefficients along with their corresponding t-statistics (in parentheses) for the impact of the robo-advising service on the allocation to the S&P 500, US Factor Style Investments, and Chinese ADRs in the US and HK, using regression specification (1). These three dependent variables are measured as the corresponding percentage weight in the entire portfolio. Panel B presents the results using regression specification (2). The coefficients are expressed as percentages (%). Individual fixed effects and month fixed effects are included. Standard errors are clustered at the investor level. Statistical significance is denoted by *, **, and *** for the 10%, 5%, and 1% significance levels, respectively.

Panel A			
	Allocation to S&P500	Allocation to US Factor Style Investment	Allocation to Chinese ADRs
Robo*Post	2.04**	1.52**	1.81**
	(2.24)	(1.99)	(2.01)
Panel B			
LLM (AI) Interaction effects			
Robo*Post*LLM	3.48***	3.07***	4.71***
	(4.84)	(5.35)	(4.90)
Robo*Post	-0.06	-0.25	-0.09
	(-0.11)	(-0.42)	(-0.23)

G Online Appendix

G.1 How robo-advising function

G.1.1 Example 1: Mutual funds

Consider an investor, Jane, who currently holds three mutual funds: Mutual Fund A, Mutual Fund B, and Mutual Fund C. Jane is curious about the performance of these funds and whether she should make any changes to her portfolio. She decides to use a robo-advisor for a detailed analysis and personalized recommendations.

Upon analyzing Jane's holdings, the robo-advisor provides the following insights (All beta values are calculated on a monthly or quarterly basis, depending on the availability of higher frequency data, using a five-year rolling window and a t-statistic cutoff of 3 for statistical significance.):

Mutual Fund A: Alpha = 3%, Beta = 1.2, HML (High Minus Low) = 0.4, SMB (Small Minus Big) = 0.5, t-statistic = 1.8. Although Mutual Fund A has a positive alpha, the t-statistic below 3 indicates that the alpha is not statistically significant, suggesting that the outperformance may not be due to skill. Mutual Fund B: Alpha = -1%, Beta = 0.8, HML = -0.2, SMB = -0.1, t-statistic = -0.5. Mutual Fund B has a negative alpha and an insignificant t-statistic, confirming underperformance relative to its benchmark. Mutual Fund C: Alpha = 2%, Beta = 1.0, HML = 0.3, SMB = 0.2, t-statistic = 3.1. Mutual Fund C has a positive and statistically significant alpha, indicating that the fund manager has likely added value beyond market returns.

The robo-advisor calculates the overall portfolio's alpha and beta, adjusting for factor returns such as HML and SMB:

Overall Portfolio Alpha: 1.33% Overall Portfolio Beta: 1.0 (indicating market-level volatility) Overall Portfolio HML: 0.16 Overall Portfolio SMB: 0.2

Based on these insights, the robo-advisor makes the following recommendations to Jane:

Reevaluate Mutual Fund A: Despite its positive alpha, the t-statistic of 1.8 suggests that

the alpha is not statistically significant. The robo-advisor recommends considering reallocating funds from Mutual Fund A to investments with more reliable performance metrics. Sell Mutual Fund B: Given its negative alpha and insignificant t-statistic, Mutual Fund B is clearly underperforming. The robo-advisor advises selling this fund and reallocating the capital. Maintain Position in Mutual Fund C: With a positive and statistically significant alpha, Mutual Fund C demonstrates strong management performance. The robo-advisor suggests keeping this fund. Introduce Low-Cost ETFs or Index Funds: To enhance cost efficiency and maintain diversification, the robo-advisor recommends reallocating the funds from Mutual Fund B and potentially Mutual Fund A into low-cost ETFs or index funds that track broad market indices. For example, it suggests ETF X (Usually Hushen 300 ETFs, similar to S&P 500 ETFs), with a beta of 1.0, minimal exposure to HML and SMB, and very low fees, to capture market returns efficiently.

By following these recommendations, Jane can optimize her portfolio, ensuring she benefits from statistically significant high-alpha mutual funds while reducing costs and maintaining market exposure through ETFs or index funds. This balanced approach helps her achieve better returns while managing risk and minimizing fees.

G.1.2 Example 2: Equity

Consider an investor, John, who currently holds three individual stocks: Stock X, Stock Y, and Stock Z. John is interested in understanding the performance of these stocks and whether he should make any adjustments to his portfolio. He decides to use a robo-advisor for a detailed analysis and personalized recommendations.

Upon analyzing John's holdings, the robo-advisor provides the following insights (All beta values are calculated on a monthly or quarterly basis, depending on the availability of higher frequency data, using a five-year rolling window and a t-statistic cutoff of 3 for statistical significance.):

Stock X:

Alpha = 4% Beta = 1.3 t-statistic = 1.9 Stock X has a positive alpha, indicating outperformance relative to its benchmark. However, the t-statistic is well below 3, suggesting that the alpha is not statistically significant, and the outperformance may not be due to skill. Stock Y:

Alpha = -2% Beta = 0.7 t-statistic = -1.2 Stock Y has a negative alpha, indicating underperformance relative to its benchmark. The t-statistic is also negative and not significant, confirming that the stock is likely underperforming. Stock Z:

Alpha = 3% Beta = 1.0 t-statistic = 3.4 Stock Z has a positive and statistically significant alpha, indicating that the stock has likely provided returns above the market due to some stock-specific factors. The robo-advisor calculates the overall portfolio's alpha and beta based on the weighted average of the individual stocks:

Overall Portfolio Alpha: 1.67% Overall Portfolio Beta: 1.0 (indicating market-level volatility)

Based on these insights, the robo-advisor makes the following recommendations to John:

Reevaluate Stock X: Despite its positive alpha, the t-statistic of 1.9 suggests that the alpha is not statistically significant. The robo-advisor recommends monitoring Stock X closely and considering reallocating funds if better investment opportunities arise.

Sell Stock Y: Given its negative alpha and insignificant t-statistic, Stock Y is underperforming. The robo-advisor advises selling this stock and reallocating the capital to more promising investments.

Maintain Position in Stock Z: With a positive and statistically significant alpha, Stock Z demonstrates strong performance. The robo-advisor suggests keeping this stock in the portfolio.

Introduce Diversification Through ETFs: To enhance diversification and reduce individual stock risk, the robo-advisor recommends reallocating the funds from Stock Y into low-cost ETFs or index funds that track broad market indices. For example, it suggests ETF A (Usually Hushen 300 ETF, similar to S&P 500 ETFs), which has a beta of 1.0 and

very low fees, to efficiently capture market returns.

By following these recommendations, John can optimize his stock portfolio by retaining stocks with statistically significant positive alpha, reducing exposure to underperforming stocks, and enhancing diversification with low-cost ETFs. This approach helps him achieve better returns while managing risk and minimizing fees.

Table A.1: Investor Performance Outcomes: Comparing Human Assistance, LLM Chatbots, and Rule-Based Chatbots (Robustness Checks)

This Table shows the difference-in-differences regressions coefficients and t-statistics (in parenthesis) as in regression specification (1). The Treatment equals one if the investors intend to receive the robo-advising service and actually received the robo-advising service. The Control equals one if the investor intend to receive the robo-advising service but was not selected to receive the service. Post equals one for months after 2023 December. Alphas are the Three Factor Alpha, Three Factor + Momentum Alpha, Five Factor Alpha, and Five Factor + Momentum Alpha. Standard errors are clustered at the investor level.

Group	**FF Three Alpha (%)**	**FF Three + MOM Alpha (%)**	**FF Five Alpha (%)**	**FF Five + MOM Alpha (%)**
Overall	0.20***	0.19***	0.18***	0.17***
	(4.72)	(4.62)	(4.51)	(4.44)
ChatGPT (LLM) Chatbot	0.44***	0.43***	0.41***	0.40***
	(6.27)	(6.03)	(5.89)	(5.70)
Without Chatbot	0.13**	0.12**	0.12**	0.10**
	(2.25)	(2.21)	(2.04)	(1.99)
Human Chat Assistant	0.14**	0.13**	0.12**	0.11**
	(2.41)	(2.35)	(2.29)	(2.24)
Normal Chatbot	0.12**	0.12**	0.10**	0.09**
	(2.29)	(2.24)	(2.19)	(2.03)

Table A.2: Functionality Comparison of LLM Chatbot, Human Chat Assistant, and Normal Chatbot

Functionality Comparison of LLM Chatbot, Human Chat Assistant, and Normal Chatbot. Other Y variables: Number of question asked, positive and negative sentiment from the the investors' questions

Functionality	**LLM Chatbot**	**Human Chat Assistant**	**Normal Chatbot**
Explaining Beta	Comprehensive, contextual understanding (Uses advanced language models)	Good, but may lack consistency (Relies on human expertise)	Limited, basic explanations (Relies on predefined scripts)
Explaining Alpha	Detailed, data-driven insights (Access to large datasets)	Good, with practical examples (Uses personal experience)	Basic, often superficial (Limited to script accuracy)
Answering Financial Queries	Accurate, quick responses (High processing power)	Reliable, but slower (Depends on human availability)	Variable accuracy, slower (Limited by scripting)
User Interaction	High engagement, adaptive (Adapts to user input)	Personalized, but variable (Depends on human mood)	Standard responses, less engaging (Predefined interactions)
Learning and Improvement	Continuous learning, updates (AI model updates)	Static, based on human learning (Depends on individual growth)	Limited, no self-improvement (Requires manual updates)

Table A.3: Baseline Robot Advising Treatment Effect and the LLM Chatbot Effect

Columns (1) and (2) show the triple DiD regression coefficient β_3 in regression specification (2) with different fixed effects and control variables. Columns (3) and (4) show the difference-in-differences regression coefficients β_2 in regression specification (1) and t-statistics (in parentheses) with different fixed effects and control variables. Treatment equals one if the investors intended to receive the robo-advising service and actually received the service. Control equals one if the investors intended to receive the robo-advising service but were not selected to receive the service. The sample period is from December 2022 to November 2024. Post equals one for observations (months) after December 2023. The dependent variable is Raw Return, the monthly raw return. Standard errors are clustered at the investor level.

Performance (Monthly)	(1)	(2)	(3)	(4)
LLM*Robo*Post	0.43***	0.41***		
	(3.06)	(2.83)		
Robo*Post	0.28**	0.25**	0.43***	0.39***
	(2.42)	(2.18)	(4.73)	(4.45)
Robo	-0.03		-0.05	
	(-0.51)		(-0.62)	
Age	0.01		-0.02	
	(0.36)		(-0.42)	
Account Age	0.02		-0.01	
	(0.32)		(-0.16)	
Initial Portfolio Value	0.01		0.01	
	(0.14)		(0.17)	
Gender	-0.02		-0.01	
	(-0.35)		(-0.42)	
Monitoring Time	-0.01		-0.01	
	(-0.08)		(-0.04)	
Login Counts	0.00		0.00	
	(0.09)		(0.07)	
Trading Frequency	0.01		0.01	
	(0.11)		(0.09)	
N	691,656	691,656	691,656	691,656
Adj. R^2	0.052	0.146	0.049	0.142
City FEs	Yes	No	Yes	No
Month FEs	Yes	Yes	Yes	Yes
Individual FEs	No	Yes	No	Yes