FinTech Brings a Bias from Psychology Labs to a Two-trillion-dollar Market*

Hongjun Yan
DePaul University

Zhengwei Wang
Tsinghua University

Yulai Yuan
Xuanyuan Insurance Agency Co., Ltd

Xiaole Qiu
Tsinghua University

Jia Xiang
Tsinghua University

Yaqing Xiao
Capital University of Economics and Business

Erik Davidson
Baylor University

This Draft: November 10, 2025

*We are grateful for the helpful comments from James Choi, Erik Mayer, Lin Peng, Alberto Rossi, Giulio Trigilia, Baolian Wang, Michael Weber, Edward Van Wesep, Arthur Yan, Dexin Zhou, and seminar participants at the University of Rochester, UIC, University of Wisconsin Madison, George Mason University, and Baruch College. We thank Lili Huang, Chunyan Wang, Jun Xiao, Liang Zhang, and an anonymous robo-advisory company for the collaborations on the RCT in this study and Charles Rotblut and the AAII for the collaboration on the survey. The RCT in this study was pre-registered with AEA RCT Registry (RCT ID: AEARCTR-0014422) and received IRB approval from the DePaul University (Protocol ID: IRB-2024-1411). Please send all correspondence to Hongjun Yan, hongjun.yan.2011@gmail.com.

FinTech Brings a Bias from Psychology Labs to a Two-trillion-dollar Market

Abstract

Robo-advisors typically make investment recommendations based on surveys of clients' preferences. Consequently, biases induced by the surveys would be embedded in those recommendations and hence may influence investors' financial decisions. We examine this hypothesis through two studies. First, we administer a nationwide survey of experienced U.S. investors and find that the order of listed choices significantly influences their responses to risk-tolerance questions commonly used by robo-advisors. Second, we collaborate with a leading robo-advisor in China to conduct a preregistered RCT and demonstrate that the order effect significantly alters both the robo-advisor's risk assessments and its clients' investment decisions, raising new ethical concerns in the FinTech era.

JEL Classification Numbers: D1, D9, G10, G2, G23, G40, G41.

Keywords: FinTech, Robo-advisor, Order effect, RCT, Nudge, Survey, Interface.

1. Introduction

Personalized services have long been labor-intensive and resistant to automation. As pointed out in Baumol (1967), such sectors tend to exhibit limited productivity growth and experience rising relative costs over time. Consequently, personalized services have traditionally been affordable only to wealthy individuals. Modern technology, however, promises to change that. In recent years, numerous firms have started offering large-scale, low-cost, personalized services across a wide range of domains, including financial investments, insurance, retirement planning, housing decisions such as rent-versus-buy choices, career guidance, and even personalized nutrition.

While this development has transformed the service industry, we argue that it may also influence consumer decisions in subtle and often hidden ways. To deliver low-cost personalized services, providers cannot afford extensive personal interactions with each customer. Instead, they typically rely on brief surveys to gather essential information. As a result, any biases induced by these surveys are embeded in the service itself and may therefore shape customers' decisions.

Our paper provides one such example. Financial technology (FinTech) has fueled the rapid expansion of the robo-advisory market, which manages over $2 trillion in assets as of 2025. This innovation has the potential to democratize financial advice by delivering large-scale personalized advice to investors with modest net worth (D'Acunto and Rossi, 2023). However, our analysis shows that this development also brings an innocuous behavioral bias—long documented in psychology laboratories—into this two-trillion-dollar market.

We focus on the well-documented *response-order effect*: the order in which response options appear in a survey question can influence how respondents answer (Krosnick and Presser, 2010). We conjecture that this bias also operates in robo-advisory surveys, despite the large financial stakes involved. For instance, in a question assessing risk tolerance, presenting the most risk-tolerant option at the top rather than at the bottom of the list may elicit different responses from the same individual. Because robo-advisors use these survey responses as inputs to their algorithms, they may—perhaps unknowingly—embed the bias into their investment recommendations to clients. More importantly, we conjecture that this bias is not confined to the virtual world but extends to clients' real-world investment decisions.

We examine these conjectures in two studies. First, we administer a nationwide survey of experienced U.S. investors and document a substantial response-order effect in the risk-tolerance questions commonly used by robo-advisors. Second, motivated by this survey evidence, we collaborate with a leading robo-advisor in China to conduct a preregistered randomized controlled trial (RCT) that provides causal evidence that response order significantly affects both the robo-advisor's risk assessments and its clients' investment decisions.

Our first study examines whether the response-order effect documented in psychology also arises in a robo-advisory setting. Two features distinguish our design from typical psychology studies: it centers on the survey questions commonly used by robo-advisors, and it recruits experienced investors rather than students or the general public.

Specifically, we administered a survey of over 3,000 experienced investors across the U.S. between October 27 and November 2, 2020. To assess the response-order effect, we implemented two otherwise identical versions of the questionnaire that differed only in the order of response options. Each respondent was randomly assigned, with equal probability, to one of the two versions.

We find a strong response-order effect: respondents tend to select options near the top of the lists, leading to a systematic bias in survey results. For example, consider the following risk tolerance question, which is often used by robo-advisors: "Recognizing that investments can fluctuate in value, for a \$100,000 investment, how much of a decline in the value of your investment portfolio could you tolerate over the course of a year?" The response options were presented in the following two versions:

Original-order survey	*Reverse-order survey*
−10% (−\$10,000)	−40% (−\$40,000)
−20% (−\$20,000)	−30% (−\$30,000)
−30% (−\$30,000)	−20% (−\$20,000)
−40% (−\$40,000)	−10% (−\$10,000)

Because respondents tend to select options near the top of the list, those shown the reverse-order survey should, on average, report higher risk tolerance. Indeed, 39% of respondents to the reverse-order survey (611 out of 1,576) selected the most risk tolerant answer—"−40% (−\$40,000)". In contrast, when that option appeared at the bottom of the list in the original-order survey, only 28% of respondents (439 out of 1,550) selected it. This difference is highly statistically significant ($p = 6.29 \times 10^{-10}$). More broadly, we find a quantitatively large response-order effect across a variety of survey questions, including those on investment preferences, expectations on the macroeconomy, public health, and politics. By contrast, the response-order effect disappears for objective questions with clear, factual answers, such as demographics.

In the second study, we collaborate with a leading robo-advisor in China (which will be referred to as "the Advisor" hereafter) to conduct an RCT to examine whether the response-order effect persists when large financial stakes are involved, and, more importantly, whether it has a meaningful effect on investors' financial decisions.

When prospective clients open accounts with the Advisor, they must complete an 11-question survey covering investment experience, risk tolerance, and related topics. Existing clients are also prompted to retake the survey periodically, typically once every two years. For each question, the client receives a numerical score based on his response. The total score across all 11 questions determines the client's *risk score*, which is then used to classify clients into five *risk categories*—C1 through C5—with C5 indicating the highest risk tolerance.

During the trial, the Advisor administered two versions of the survey that differed only in the order of response options for each question. In the original-order version, the option corresponding to the lowest score appeared at the top of the list, whereas in the reverse-order version, the option corresponding to the highest score appeared at the top. Each client was randomly assigned, with equal probability, to one of the two versions. Randomization was implemented based on clients' account numbers: those with odd numbers received the reverse-order survey, and those with even numbers received the original-order version.

The treatment for new clients was administered between October 25 and December 27, 2024. For existing clients (i.e., those who had accounts before October 25, 2024), the treatment was rolled out in two waves. On November 1, 2024, the Advisor randomly selected 5,000 existing users and prompted them to retake the risk assessment survey. Similarly, on December 18, another 5,000 existing users were selected. As of April 1. 2025, 7,141 users participated in the trial. The original-order and reverse-order groups are balanced in both sample size and user characteristics.

To estimate the order effect, we regress users' risk scores on $Reverse$, which is one if a user's account number is odd (and hence the user received the reverse-order survey), and zero otherwise. The estimated coefficient on $Reverse$ is 3.35 (*s.e.* = 0.31), indicating that reversing the response order increases the user's risk score by an average of 3.35. We then perform a placebo test based on a sample of 3,142 new users who opened their accounts after our RCT concluded, and, as expected, find no difference in risk scores between those with odd account numbers ($Reverse = 1$) and those with even account numbers ($Reverse = 0$).

The response order also affects users' risk categories. Although the original-order and reverse-order groups are similar in size—3,542 and 3,599 users, respectively—the number of users in category C5 differs markedly: 1,697 in the original-order group versus 2,035 in the reverse-order group. The p-value for this difference is 2.83×10^{-13}. Thus, despite large financial stakes, response order continues to exert a substantial influence on users' survey responses.

Does this bias also affect users' actual investment decisions? There are at least two reasons to conjecture that it does. First, the order effect on risk assessments may influence users' investment decisions through a nudging channel. Since risk assessments are an input for the Advisor's algorithm, the order effect becomes embedded in those recommendations, potentially nudging users to act. Second, it may operate through a compliance channel. The Advisor classifies the financial products on its platform into five risk levels, R1 through R5 (riskiest). For regulatory considerations, a user is permitted to purchase only products with ratings at or below the user's assigned risk category. For example, a C4 user may purchase products rated R1 through R4. If the user attempts to purchase an R5 product, however, he would be reminded that the product's rating exceeds his risk category. To proceed with the purchase, the user must retake the risk assessment survey and obtain a C5 classification. Because of this restriction, we expect some of the response-order effect on risk categories to carry over to users' investment decisions.

To analyze the order effect on investment decisions, we focus on two primary types of products at the Advisor: robo-advised and self-directed. The Advisor provides three robo-advised investment strategies: "Aggressive," "Balanced," and "Stable" strategies, which have high, medium, and low risk levels, respectively. Each strategy aims to build and maintain a target portfolio tailored to a specific risk level. For example, the target allocation to equities is 47% under the Aggressive strategy but only 24% under the Stable strategy.

To invest in a strategy, a user needs to set up a "special purpose account" dedicated to the strategy. A user can choose any one or multiple strategies. However, his risk category may steer him toward a particular strategy. For example, a C5 user's assessment report begins with: "Your risk profile is Aggressive" before explaining how Aggressive investors typically invest. The phrase for the risk category—Aggressive—is the same as that for the strategy. Moreover, if a C4 user attempts to set up an account for the Aggressive strategy, he would be reminded that the strategy's risk level exceeds his category. Although the user can ignore this warning to set up the account, compliance restrictions may still limit the user's investment in the strategy: Whenever the strategy's recommendation includes purchasing an R5 product, the user would be reminded that he can only make the purchase after retaking the survey to obtain a C5 category. This impediment may limit the user's investment in the aggressive strategy.

The second type of investment is self-directed. It includes mutual funds, money market funds, and wealth management products offered by third parties, with the Advisor acting as an intermediary. In this case, users must actively make all decisions related to selecting and trading these products. In contrast, robo-advised investments are semi-automatic: users can accept or reject the Advisor's recommended transactions but cannot initiate alternatives. Because self-directed investing requires more active decision-making, the response-order effect is expected to be weaker for these investments than for robo-advised ones.

We first examine the order effect on robo-advised investments. To estimate the effect along the intensive margin, we obtain the sample of users who purchased robo-advised products during our sample period, which includes 1,746 users with a total purchase of RMB 48.59 million. We regress a user's total investment in the Aggressive strategy on $Reverse$. The estimated coefficient on $Reverse$ is 3.28 (*s.e.* = 1.66), indicating that users exposed to the reverse-order survey, on average, invest RMB3,280 more in the Aggressive strategy. In comparison, the average purchase size of robo-advised products is RMB27,840 in our sample period.

Our hypothesis is that the effect on investments operates through the response-order-induced variation in users' risk categories. To examine this mechanism, we use the survey response order as an instrument for each users' risk categories. Specifically, we regress a user's risk category on the indicator $Reverse$ to decompose it into two components: one capturing the user's intrinsic risk tolerance and the other reflecting the survey-induced bias from the response order. Consistent with our hypothesis, we find that a user's investment in the Aggressive strategy increase with both components of his risk category.

To estimate the effect along the extensive margin, we construct a dummy variable D_Robo_i that equals one if user i invested in any robo-advised product during our sample period and zero otherwise. We then regress this indicator on $Reverse$ using the overall RCT sample. The estimated coefficient on $Reverse$ is -0.0254 (s.e. = 0.0105), indicating that reverse-order users are 2.54 percentage points less likely to invest in robo-advised products. This result is unlikely to arise from the compliance channel because reverse-order users tend to have higher risk categories and therefore face fewer investment restrictions. A more plausible explanation is that the nudging effect backfired: inflated risk categories steered reverse-order users toward high-risk strategies, but many of them may have found these strategies too risky and opted out of investing altogether. Consistent with this interpretation, we find that a user is more likely to invest in robo-advised products when his intrinsic risk tolerance is higher, but less likely to do so when his survey-induced-bias component is higher.

We also conduct a parallel analysis for self-directed investments and find that all estimated effects have the same sign as their counterparts for robo-advised investments but are weaker in both statistical and economic terms. A likely explanation is that users play a more active role in decision-making for self-directed investments, which attenuates the response-order effect. This intuition also suggests that the response-order effect is likely to be even stronger in the U.S. robo-advisory market. In the U.S. setting, investors typically play a more passive role in the decision-making process. Once the risk-assessment survey is completed, investment decisions are executed automatically based on the survey results (Reher and Sokolinski, 2024). Consequently, the order effect on risk assessments is transmitted directly to actual investments.

Our findings raise an important ethical concern. A robo-advisor may profit by steering its clients toward products that are suboptimal from the clients' perspective. Because survey-elicited preferences are susceptible to manipulation, the robo-advisor can influence investors' purchases while preserving the appearance that these choices reflect investors' own preferences. More broadly, the evidence indicates that preference constructs such as risk aversion are not fully determinate: their estimated values—whether inferred from survey responses or investment decisions—systematically vary with exogenous factors such as the order of survey responses.

Our paper contributes the literature on technology and the democratization of professional services. Autor, Levy, and Murnane (2003) show that the computerization in the second half of the 20th century cannot reduce the cost for nonroutine cognitive tasks, which are often required for personalized service tasks. The digitalization of the 21st century, however, raised the hope to deliver low-cost customized services broadly (Varian, 2010; Goldfarb and Tucker, 2019). Philippon (2020) shows that financial services have been surprisingly expensive over centuries but argues that technologies such as robo-advising has the potential to reduce the cost of financial advice for a broader investor base. Prior work on the robo-advisory market has emphasized its potential to help retail investors overcome behavioral biases (D'Acunto, Prabhala, and Rossi, 2019; D'Acunto and Rossi, 2023; Reher and Sokolinski, 2024). Our paper complements this literature by showing that when the personalization requires inputs from surveys, it may embed biases into the service and influence customers decisions.

Our paper also adds to the emerging literature on how the decision environment shapes financial investments. For example, Liao et al. (2021) document the effects of user interface on online financial investments. We show that the business model of robo-advising creates an environment in which a survey-induced bias creeps in the investment process.

The response-order effect has been extensively studied in psychology (Krosnick and Presser, 2010). Our paper extends it to economic surveys and analyzes its effects on high-stake real-world financial decisions. Prior studies have also documented the order effects on auction revenue (Hong et al., 2015) and attention to academic papers (Feenberg, et al., 2017). There is a similar phenomenon for the order effect (or ranking effect) on investor attention (e.g., Hartzmark, 2015, Frydman and Wang, 2020, Jiang et al, 2022). More generally, order effects have also been documented in citations (Huang, 2015; Brogaard et al, 2023), school admissions (Jurajda and Munich, 2010), and election outcomes (Chen, et al., 2014).

A growing literature uses surveys to measure subjective beliefs (Manski, 2004) to study various topics in belief formation, such as experience effects (Malmendier and Nagel, 2011), partisan biases (Coibion, Gorodnichenko, Weber, 2020), personality traits (Jiang, Peng, and Yan, 2024), and memory (Jiang, Liu, Peng, and Yan, 2025). There has been renewed interest in survey-based beliefs in finance since Greenwood and Shleifer (2014), as they offer unique insights for distinguishing asset pricing theories (see, for example, Choi and Robertson, 2020; Giglio, Maggiori, Stroebel, and Utkus, 2020, 2021; De la O and Myers, 2021; Chinco, Hartzmark, and Sussman, 2022; and Bordalo, Gennaioli, LaPorta, and Shleifer, 2025). Our paper adds to the literature by evaluating survey methodologies, highlighting the large response-order bias on survey expectations and preferences.

The rest of the paper is as follows. Section 2 reports a study based on a survey of U.S. investors. Section 3 presents an RCT at a leading robo-advisor in China. Section 4 concludes.

2. A Nationwide Survey

We administered a nationwide survey through the American Association of Individual Investors (AAII), a nonprofit organization with approximately 160,000 members. According to the AAII website, its main purpose is to help "individuals become effective managers of their own assets through programs of education, information, and research." A typical AAII member is described as a male in his mid-60s with a bachelor's or graduate degree, and members tend to be affluent, with a median portfolio size exceeding $1 million. Surveys of AAII members have been used in prior studies. For example, Greenwood and Shleifer (2014) show that expectations elicited from AAII surveys are highly correlated with those from other investor surveys. Jiang, Peng, and Yan (2024) also surveyed AAII members to examine the relationship between personality traits and investment decisions.

To assess the response-order effect, we created two versions of the survey. Both contained the same 17 questions in the same order; the only difference was the order of the responses for each question. Each respondent was randomly assigned to one version with equal probability. One version, shown in Appendix A1, is referred to as the "original-order survey." In the other version, "reverse-order survey," the order of response options was reversed for each question. For example, the first question is "How old are you?" The response options for the two versions are as follows:

Original order	*Reverse order*
18 – 39 years	70 years or older
40 – 49 years	60 – 69 years
50 – 59 years	50 – 59 years
60 – 69 years	40 – 49 years
70 years or older	18 – 39 years

Our survey was distributed via an email by the AAII to its members on October 27, 2020, and data collection concluded on November 2, 2020—the day before the U.S. presidential election. We received responses from 3,146 individuals: 1,561 from the original-order survey and 1,585 from the reverse-order one.

Respondent i's answer to survey question j is denoted as $Answer_{ij}$ and coded as

$$Answer_{ij} = k, \tag{1}$$

if the selected response occupies the k-th position in the original-order list. For example, if respondent i receives the original-order survey and selects the fourth response to the first question (i.e., "60–69 years"), the answer is coded as $Answer_{i1} = 4$. If respondent i receives the reverse-order survey and selects the second response (which is also "60–69 years"), the answer is likewise coded as $Answer_{i1} = 4$. That is, the answer "60–69 years" is coded as 4 in both survey versions.

Table 1 reports the summary statistics. The first three questions gather basic information. As expected, our sample is skewed towards seniors and males. Around 90% of the respondents are over 60 years old and 93% are male. The respondents are experienced investors; around 92% had more than 20 years of experience.

Questions 4 and 5 are about investment horizon and risk tolerance and closely resemble the survey questions used by robo-advisors. Most respondents have long investment horizons. Around 60% select “More than 20 years” or “Beyond my lifespan (estate).” Around 63% indicate that they can tolerate a 30% or 40% loss within a year.

Questions 6 and 7 are about forecasting the GDP growth and S&P 500 stock index return, respectively, for the next 12 months. The median forecast for GDP growth is -2% to 2% and for the S&P 500 index return is -5% to 5%. Questions 8 and 9 are about the respondent’s experiences during the COVID pandemic. Question 8 is “Do you directly know someone who has been infected by COVID-19 with mild or no symptoms?” Around 50% replied yes. Responses to Question 9 suggest that 27% of our respondents “directly know someone who has been severely ill due to COVID-19 infection.” Questions 10 through 12 focus on their expectations on the pandemic, such as the vaccine development. Questions 13 through 15 are about the U.S presidential election. For example, in their responses to Question 13, 42% selected “Definitely Republican” or “Usually Republican” and 30% selected “Definitely Democrat” or “Usually Democrat.”

Finally, the responses to Question 16 show that 91% of our respondents have a bachelor’s degree or above. The responses to Question 17 indicate that 69% live in suburbs and that rural and urban areas account for around 16 percent each. More details of this survey can be found in Xiao and Yan (2023), which examines how information from social network disproportionately shapes expectations.

2.1 The Response-order Effect

A typical form of response-order effect is the tendency that survey respondents are more likely to select a choice found near the top of the list (Krosnick and Presser, 2010). This is consistent with the anchor-and-adjust heuristic (Tversky and Kahneman, 1974): when making choices, people tend to use the quantities they first encounter as the anchors and make adjustment afterwards. Because adjustments tend to be insufficient, choices tend to be biased toward the anchors. In our context, this heuristic implies that respondents are biased towards choices near the top of the lists, which, according to the definition in (1), have larger values for $Answer$ for the reverse-order survey. Hence, our first hypothesis is that the respondents to the reverse-order survey should have higher values of $Answer$ than those to the original-order survey.

We conjecture that the effect should be stronger if respondents are more uncertain about their answers. This is motivated by the evidence that the anchoring effect is stronger when respondents are less confident about their answers (Jacowitz and Kahneman, 1995).

Generally, people are less certain about the answers to subjective questions, such as macro expectations, than those to objective questions, such as demographic questions. For example, our first three survey questions are about age, gender, and investment experience, respectively. Questions 8 and 9 are about recent personal experience during the pandemic. Finally, Questions 16 and 17 are about education and residence. We expect respondents to have clear answers to these "objective questions" and hence expect little or no order effect. The other questions call for subjective opinions; we refer to them as "subjective questions." Many may not have clear answers to these questions. Hence, our second hypothesis is that the response-order effect should be stronger for subjective questions than for objective ones.

2.2 Estimation

We first estimate the response-order effect across all survey questions. Note that, according to the definition in (1), the value of $Answer$ ranges from 1 to 5 for Question 1, but only from 1 to 2 for Question 2. Hence, we standardize this variable to make it comparable across questions: the standardized version of respondent *i*'s answer to Question *j* is defined as

$$SAnswer_{ij} = (Answer_{ij} - 1)\frac{100}{n-1}, \tag{2}$$

where $Answer_{ij}$ is defined in (1) and *n* is the number of choices for Question *j*. That is, for each question, its responses are assigned values between [0,100]. For example, if a survey question has five responses, $SAnswer$ is 0, 25, 50, 75, and 100, if one selects response 1 through 5, respectively.

We then run the following panel regression:

$$SAnswer_{ij} = \alpha + \beta\ Reverse_i + \varepsilon_{ij}, \tag{3}$$

where $Reverse_i$ is a dummy variable that is 1 if respondent *i* faces a reverse-order survey and 0 otherwise. The regression results are reported in the first two columns of Panel A of Table 2. Column (1) shows that the coefficient on *Reverse* is 2.54 (*s.e.* = 0.29). That is, on average, reversing the response order increases the standardized answer, $SAnswer$, by 2.54. In column (2), we include survey-question fixed effects. The coefficient on *Reverse* remains similar in both magnitude and statistical significance.

To examine the cross-sectional variation, we add an interaction term to regression (3):

$$SAnswer_{ij} = \alpha + \ \beta \ Reverse_i + \ \gamma \ Reverse_i \times Sub_j + \delta Sub_j + \varepsilon_{ij}, \qquad (4)$$

where Sub_j is a dummy variable that is 1 if Question j is a "subjective question" and 0 otherwise. The regression results in column (3) shows that the interaction coefficient is 3.80 (*s.e.* = 0.58). The results remain similar after we include survey-question fixed effects in column (4). These results show that the response-order effect is limited to subjective questions. For example, column (4) shows that the coefficient on *Reverse* is merely 0.36 (*s.e.* = 0.47), i.e., the response-order effect is statistically insignificant for objective questions. In contrast, the effect for subjective questions is substantially stronger, with a point estimate of 4.11 (=3.75+0.36).

We also run regression (3) for each survey question separately and summarize the results in Panel B of Table 2. For six of the seven objective questions, the estimated coefficient on $Reverse$ is statistically indistinguishable from zero. The exception is Question 3, on investment experiences, for which the estimated coefficient is 1.16 (*s.e.* = 0.56). Although a respondent's investment experience is an objective fact, it requires the respondent to recall the distant past; 92% of our respondents report more than 20 years of experience. Some may not be able to recall precisely when they started investing, making a response-order effect possible. In contrast, there is strong evidence of the response-order effect for subjective questions. For nine of the ten questions, the coefficient on $Reverse$ is positive and highly statistically significant.

2.3 Quantify the Effect on Survey-based Expectations

In this section, we quantify the response-order effect on survey expectations and contrast it with the effects of two salient features at the time of our survey. The first is the effect of the COVID pandemic. Our survey was conducted from October 27 to November 2, 2020. This was before Pfizer's November 9 announcement of the first major breakthrough in COVID vaccine development. There was substantial uncertainty regarding vaccine development. Xiao and Yan (2023) analyze the data from this AAII survey and show that personal experience of the COVID pandemic plays a disproportionate role in shaping expectations. To measure a respondent's personal experience during the pandemic, we construct two dummy variables: $Mild_i$ is 1 if respondent i, according to his answer to Question 8, directly knows someone who has COVID infection with mild or no symptoms but, according to Question 9, does not directly know anyone who has been severely ill due to COVID. $Severe_i$ is 1 if respondent i, according to his answer to Question 9, directly knows someone who has been severely ill due to COVID.

The second feature is that our survey was conducted right before the 2020 presidential election. The political division was salient and there was considerable uncertainty about the outcome. A recent literature has shown that an individual's expectations are shaped by her political beliefs (Kempf and Tsoutsoura, 2021; Meeuwis, Parker, Schoar, and Simester, 2022). We construct two dummy variables: R_i is 1 if respondent *i*'s answer to Question 13 on political leaning is "Definitely Republican" or "Usually Republican" and D_i is 1 if respondent *i*'s answer is "Definitely Democrat" or "Usually Democrat." As noted in Table 2, the stated political leanings are subject to the order effect; that is, R_i and D_i are correlated to $Reverse_i$. To isolate the effect from political beliefs, we orthogonalize R_i and D_i against $Reverse_i$ by regressing them on $Reverse_i$ to obtain the residuals, which are denoted R'_i and D'_i, respectively.

To compare the magnitude of the response-order effect with that of the effects of COVID experiences and political leanings, we run the following regression:

$$Y_i = \alpha + \beta_1 Reverse_i + \beta_2 Mild_i + \beta_3 Severe_i + \beta_4 D'_i + \beta_5 R'_i + \varepsilon_i, \tag{5}$$

where Y_i is respondent *i*'s expectation. For ease of interpretation, we set Y_i to be the midpoint value of the respondent's choice range. That is, for the GDP forecast in Question 6, Y_i is -7%, -3.5%, 0%, 3.5%, and 7%, if respondent *i* selected choices 1 through 5, respectively. For the S&P500 return forecast in Question 7, Y_i is -20%, -10%, 0%, 10%, and 20%, if respondent *i* selected choices 1 through 5, respectively. The regression results are reported in Table 3.

In column (1), the coefficient on *Reverse* is 0.95 (*s.e.* = 0.26). That is, reversing the response order increases the S&P500 return forecast by 95 basis points. This is comparable to the effect of COVID experiences. The coefficient on *Mild* is -0.85 (*s.e.* = -0.31), implying that a respondent reduces his return forecast by 85 basis points if he personally knows someone with COVID with mild symptoms. The coefficient on *Severe* is -0.90 (*s.e.* = 0.31). That is, knowing someone with severe illness from COVID reduces a respondent's forecast of the stock market index return by 90 basis points. The order effect is also comparable to the effect of political leanings. The coeffects of D' and R' are -0.63 (*s.e.* = 0.34) and 1.36 (*s.e.* = 0.31), respectively. That is, on average, relative to the forecasts by Independents, Democrats' forecasts are 63 basis points lower while Republicans' are 136 basis points higher. Column (2) shows that the effects on GDP forecast are similar. Reversing the response order increases the GDP growth forecast by 0.45%

(*s.e.* = 0.10%). Directly knowing someone who is severely ill from COVID reduces a respondent's forecast by 0.26% (*s.e.* = 0.12%). On average, relative to the forecast by Independents, Democrats' forecast is 0.62% (*s.e.* = 0.13%) lower while Republicans' is 0.86% (*s.e.* = 0.12%) higher.

The above evidence shows that the response-order effect is substantial for survey-based expectations on the macroeconomy. In Appendix A3, we find that the response order also has significant effects on the expectations in other domains such as public health and politics. During our survey period, the COVID-19 pandemic and the presidential election were top of mind and had substantial effects on expectations. Our evidence suggests that the response-order effect is on par with the effect of the dramatic experience during a global pandemic and the effect of political leanings during one of the most divisive presidential elections.

Our survey is conducted only once and hence our direct evidence is on the *levels* rather than *changes* of the economic variables. However, our evidence does have implications on the potential effect on changes. If a survey is repeated over time, how does the response-order effect influence the changes of a survey-based variable? It is easy to see that if the response order is not consistent across surveys over time, the order effect would make the survey variable more volatile. Suppose the response order stays the same across surveys over time; how does the response-order effect affect the changes in the variables? Let's use the survey of stock return forecast as an example. Suppose we conduct two surveys on the stock return forecast at two points in time. Both have the same list of responses, say, in descending order. The response-order effect implies a positive bias in the forecast in both surveys. When we compute the change in forecast across these two surveys, the two biases would cancel each other out *only if* they have the same size. However, our evidence suggests that the size of the bias is likely increasing in uncertainty. Suppose, for example, the stock market becomes more uncertain at the time of the second survey. The response-order effect should be stronger, leading to a larger positive bias. Therefore, the change in return forecast is not only biased, but also positively correlated with the market uncertainty, a property that often has important implications to studies in finance.

2.4 Implications for robo-advisor surveys

A typical robo-advisor recommends financial products to its clients based on their risk tolerance and investment horizons, which are estimated from surveys (D'Acunto and Rossi, 2023). Indeed, Questions 4 and 5 in our survey are often used by robo-advisors to infer an investor's planned investment horizon and risk tolerance, respectively.

Table 2 has shown substantial order effects for both. In Question 4, for example, in the reverse-order survey (where longest investment horizon is listed at the top), 50% of the respondents choose the longest investment horizon (i.e., "Beyond my lifespan (estate)"). In contrast, when this answer is listed at the bottom in the original-order survey, only 35% of the respondents choose it. For Question 5, in the original-order survey, where the most risk tolerant answer is listed at the bottom, 28% of its respondents choose this answer. In contrast, this answer was chosen by 39% of the respondents to the reverse-order survey. If a robo-advisor use these inferred investment horizon and risk tolerance to design its recommendations, the response-order effect would directly affect its recommendations.

The respondents in our survey do not expect any consequences from their answers. In an actual robo-advisory market, knowing that their survey answers may have important consequences, would investors behave differently? Would the order effect still hold? More importantly, would the order effect influence investors' financial decisions? We address these questions next.

3. A Pre-registered RCT

Motivated by the above survey evidence, we collaborate with a leading robo-advisor in China (which will be referred to as "the Advisor") to conduct an RCT to examine the response-order effect on real-world financial decisions. This RCT was preregistered with the AEA RCT Registry (RCT ID: AEARCTR-0014422).

3.1 Background

Established in 2014, the Advisor is one of the first robo-advisors in China. As of 2024, it serves approximately 150,000 clients and manages RMB 20 billion in assets. Similar to robo-advisors in the U.S., the Advisor assesses clients' investment preferences through a survey. When new clients open accounts with the Advisor, they are required to complete a risk assessment survey (see Appendix A2) consisting of 11 questions covering financial literacy, investment experience, income, risk tolerance, liabilities, and related factors. Typically, existing clients are prompted to retake the survey every two years or when attempting to purchase products that exceed their assessed risk levels.

For each survey question, a client is assigned a numerical score based on his answer. The sum of these scores across all 11 questions constitutes the client's overall *risk score*, which ranges from 18 to 89. Clients are assigned to one of five *risk categories* based on their risk scores: C1 (18–19), C2 (20–28), C3 (29, 38), C4 (39–68), and C5 (69–89). There is one exception to this classification rule: Question 8 of the survey asks, "What is your investment preference?", and the listed responses in the original-order survey are as follows:

4% return with no principal loss
6% return with a maximum 2% potential principal loss
10% return with a maximum 7% potential principal loss
15% return with a maximum 15% potential principal loss
30% return with a maximum 40% potential principal loss

If a client selects the first choice—"4% return with no principal loss"—his risk score will be capped at 19, and he will be assigned to category C1.

The Advisor offers a range of financial products, and our analysis focuses on two main types. We refer to the first type as "self-directed" investment products to highlight users' active role in the investment process. This category includes mutual funds, money market funds, and wealth management products offered by third parties, with the Advisor serving as the intermediary. The Advisor also offers portfolios composed of these third-party products, allowing clients to invest directly in those portfolios rather than individual components.

The Advisor classifies these products into five risk categories, R1 through R5, with R5 being the highest risk level. For compliance reasons, a user is only allowed to purchase products whose risk levels are at or below his assigned risk category. For example, a client in category C4 can invest in products rated R1 through R4. If the client attempts to purchase an R5 product, he would be reminded that the product's risk level is beyond his category and prompted to retake the risk survey before proceeding.[1]

The main page of the Advisor's app displays a curated set of products based on a proprietary algorithm that incorporates a client's risk score and investment history. According to the management at the Advisor, while these prominently displayed products tend to receive more attention, users can easily navigate away to explore and purchase other offerings.

The second type is "robo-advised" investment products, which are similar to the robo-

[1] In principle, a user can contact customer service to declare that he understands the risks associated with an R5 product and insist on purchasing it without retaking the risk-assessment survey. In practice, however, this is extremely rare. In our sample, only 19 users purchased R5 products while classified in category C4.

advisory services in the U.S. market. The Advisor provides three robo-advised investment strategies: “Aggressive,” “Balanced,” and “Stable” strategies, which have high, medium, and low risk levels. Each strategy aims to build and maintain a target portfolio tailored to a specific risk level. For example, as shown in Panel A of Table 4, the target allocation to Chinese and U.S. equities is 47% under the Aggressive strategy but only 24% under the Stable strategy.

To invest in a strategy, a user needs to set up a special purpose account dedicated to the strategy. A user can open accounts for any one or multiple strategies. However, the user’s risk assessment may steer him toward a particular strategy. For example, a C5 user's risk assessment report begins with: “Your risk profile is Aggressive.” It then explains how Aggressive investors typically invest. Note that the phrase for the risk profile—Aggressive—is the same as that for the strategy. Moreover, if a C4 user attempts to set up an account for the Aggressive strategy, the user is reminded that the product’s risk level exceeds his risk category. Although the user can ignore this warning to set up the account, compliance restrictions may still limit the user’s investment size in the strategy: The C4 user can follow the recommendations to make purchases as long as no R5 products are involved. If the recommended transaction includes purchasing an R5 product, the user would be reminded that the product’s risk level is beyond his category, he can only make the purchase after retaking the survey to obtain the C5 category.

We note that users play a more active role in self-directed investments than in robo-advised investments. A user needs to actively make all decisions in his self-directed investments. In contrast, for robo-advised investments, the trading strategies are designed by the Advisor. Although a user needs to authorize each recommended transaction, he can only accept or reject a recommended transaction, but cannot make an alternative transaction.

It is also noteworthy that users in the U.S. robo-advisory market play an even more passive role. Typically, once a user completes the risk-assessment survey and establishes an account with a robo-advisor, subsequent investment decisions are executed automatically based on the user’s risk-assessment results.

3.2 The RCT Procedure

As in the AAII survey, treatment is implemented by randomizing the response order for each survey question. The risk assessment survey was presented in two versions that differ only in the order of listed responses. In one version—referred to as the original-order survey and shown in

Appendix A2—the responses for each question are ordered from lowest to highest risk tolerance. In the other version—referred to as the reverse-order survey—the responses are listed from highest to lowest risk-tolerance. Each user was randomly assigned to one of the two versions with equal probability.

Our RCT started on October 25, 2024. We will refer to users with accounts before this date as "existing users," and those who opened accounts after this date as "new users." The treatment for new users was administered from October 25 to December 27, 2024. During this period, a new user received the original-order survey if his account number was even and the reverse-order survey if it was odd. The treatment for existing users was rolled out in two waves. On November 1, 2024, 5,000 existing users with holdings of robo-advised products were randomly selected and added to the RCT. Upon logging into their accounts, they were prompted to retake their risk assessment survey. If a user proceeded to take the survey, the randomization is the same as for new users. He would receive the original-order survey if his account number was even and the reverse-order version otherwise. A user could ignore the prompt and would no longer see it after declining three times. The second wave followed the same design and was rolled out on December 18, 2024, to another 5,000 randomly selected users with holdings of robo-advised investments. There was no overlap between the two groups.

Our data were retrieved on April 1, 2025. As shown in Panel B of Table 4, our RCT sample comprises 7,141 users. The vast majority fall into categories C4 and C5, with only 260 users classified into categories C1 through C3. The second and third rows indicate that 3,542 users received the original-order survey, while 3,599 received the reverse-order version. Note that our RCT for new users ended on December 27, 2024. When retrieving data on April 1, 2025, we also obtained records for 3,142 new users who opened accounts between December 28, 2024, and March 31, 2025. The response order in their surveys was not randomized, and hence these users form our placebo sample.

Panel C shows that the average risk score in our RCT sample is 67. The average age is 46 years, and 60% of clients are male. During the sample period, 1,746 users made robo-advised investments, with an average purchase of RMB 27,830, and 1,419 users made self-directed investments, with an average purchase of RMB 54,000.

Panel D reports the summary statistics for existing and new users separately. It shows that our RCT sample consists of 5,818 existing users and 1,323 new users. Before the launch of the

RCT, existing users in our sample had an average total asset holding of RMB 315,830. Finally, Panel E shows that the original- and reverse-order groups are comparable across all characteristics: gender, age, new user status, and pre-treatment risk tolerance and asset holdings.

3.3 Order Effects on Risk Assessments

Does the order effect persist when large financial stakes are involved? To examine this, we estimate the following regression using our RCT sample:

$$Y_i = \alpha + \beta Reverse_i + \gamma X_i + \varepsilon_i, \tag{6}$$

where Y_i is user i's risk score or category, $Reverse_i$ is a dummy variable equal to one if user i's account number is odd (and hence the user received the reverse-order survey), and zero otherwise. X includes control variables: the user's age, gender, and initial asset holding at the Advisor.

The regression results, reported in Panel A of Table 5, indicate a strong order effect. In column (1), the coefficient on $Reverse$ is 3.35 (s.e. = 0.31), suggesting that reversing the response order increases the average risk score by 3.35. Column (2) adds control variables, and the coefficient on $Reverse$ remains mostly unchanged.

Does the response order influence users' risk categories? Panel B of Table 4 provides a clear answer. In the original-order group, 161 users are classified in category C1. In contrast, only 78 users fall into category C1 in the reverse-order group. The p-value for the null hypothesis that the probability of being in C1 is the same across the two groups is 2.29×10^{-8}. The response order also significantly affects the distribution between categories C4 and C5. In the original-order group, the number of users in C4 and C5 is roughly equal (1,673 vs. 1,697). In the reveres-order group, however, category C5 contains 2,036 users—38% more than the 1,476 users in category C4. The p-value for the null hypothesis that the likelihood of being in C5 is the same across the original- and reverse-order groups is 2.83×10^{-13}. In summary, reversing the response order substantially increases the number of users in the highest risk category (C5) while decreasing the number in the lowest (C1).

We also estimate regression (6) by setting Y_i to 1 through 5 if respondent i's risk category is C1 through C5, respectively. Column (3) of Panel A shows that the coefficient on $Reverse$ is 0.16 (s.e. = 0.02), indicating that—consistent with the results based on risk scores—reversing the response order increases a user's risk category. Column (4) adds control variables, and the coefficient on $Reverse$ remains mostly unchanged.

Finally, we conduct a placebo test using the placebo sample. Recall that the users in this sample answered the same set of survey questions. However, but the response order was no longer randomized because they opened their accounts after our RCT concluded. We run regression (6) on this sample as a placebo test. Specifically, the dummy variable $Reverse_i$ has the same definition as before: it is one if user *i*'s account number is odd and zero otherwise. Since there was no randomization, we should observe no systematic difference in risk assessments between users with $Reverse_i = 1$ and those with $Reverse_i = 0$. Indeed, as shown in Panel B, the estimated coefficients on *Reverse* are close to zero and statistically insignificant in all columns. This null result is not due to a lack of statistical power: the standard errors are quite small compared to the estimated coefficients on *Reverse* in Panel A.

3.4 Why Should the Response Order Affect Investment Decisions?

The previous subsection shows that the response order influences users' risk scores and risk categories. But does it also affect their investment decisions? There are at least two reasons to conjecture that it does. First, the order effect on risk assessments may influence users' investment decisions through a nudging channel. Since the risk assessments are used as an input for a robo-advisor's investment recommendations, the order effect becomes embedded in those recommendations, potentially nudging users to act. In the U.S., for example, once a user completes his risk assessment survey and sets up his account with a robo-advisor, the rest of the investment decisions are executed automatically based on the user's risk assessment results (Reher and Sokolinski, 2024). Hence, the order effect on risk assessments automatically turns into an effect on users' investment decisions. As noted in Section 3.1, investors play a more active role in our setting. Hence, one would expect the order effect on investment decisions to be weaker in our setting than in the U.S. robo-advisory market.

Second, the order effect on risk assessments may also influence users' investment decisions through a compliance channel. For regulatory reasons, a user at the Advisor can only purchase assets whose risk rating is at or below his assigned risk category. For example, a user classified as C4 may purchase products rated R1 through R4. If the user attempts to purchase an R5 product, he would be reminded that the product's risk rating exceeds his assigned risk category. To proceed with the purchase, the user needs to retake his risk assessment survey to obtain a C5 assessment.[2] As a result of this restriction, we expect the response order to influence users' investment choices through its effects on their risk categories.

Our analysis focuses on robo-advised and self-directed investments. As noted in Section 3.1, self-directed investments require more active involvement from users, whereas robo-advised investments resemble the U.S. robo-advisory market, where investors play a relatively passive role. Accordingly, we expect the order effect on investment decisions to be stronger for robo-advised than for self-directed investments.

3.5 Order Effects on Robo-advised Investments

This section examines the order effect on robo-advised investments. We first analyze the effects along the intensive margin. That is, we examine how the response order influences users' investment allocations, conditional on their participation. As shown in Table 5, the reverse-order survey increases users' assigned risk categories. How might this shift affect their investment decisions? Intuitively, the order effect can operate through both a nudging mechanism and a compliance mechanism.

First, the reverse order users' higher risk categories may nudge them towards riskier products. This is especially the case for robo-advised investment products. For example, a user classified to the C5 category in the risk assessment survey receives the following summary:

> *Your risk profile is "Aggressive." Aggressive investors typically seek to achieve above-average returns and have a strong tolerance for market volatility and potential capital losses. Their investment objective is to earn high returns while accepting high levels of risk.*
>
> *Most of the assets for aggressive investors are allocated to domestic and international equities and other similar products. A very small portion is allocated to cash and fixed-income products, adjusted based on market conditions, to help diversify risk.*

Notably, the term "Aggressive" is also used to label the investment strategy, reinforcing the alignment between the assigned risk profile and the recommended strategy.

The operational procedure may also guide users toward investment strategies aligned with

[2] Consistent with this stated policy, Table A4 in the appendix shows that users' asset allocations exhibit a discrete jump at a risk score of 69, the cutoff between the C4 and C5 categories.

their assigned risk categories. Specifically, when a user opens a special purpose account for robo-advised investments, the assigned risk category can influence the selection of investment strategies. For instance, the Aggressive strategy is designed for C5 users. If a C4 user attempts to enroll in this strategy, a warning message pops up and indicates that the strategy's risk level exceeds the user's risk category. Although the user can override the warning and proceed, such prompts are likely to deter some C4 users from pursuing the Aggressive strategy.

Second, the compliance mechanism may further steer reverse-order users—who are more likely to receive higher risk categories—toward riskier strategies. As noted earlier, a C4 user is permitted to invest in the Aggressive strategy. However, compliance constraints may limit the size of that investment. For example, if the Aggressive strategy recommends purchasing an R5 product, the user will be prevented from completing the transaction unless he retakes the risk assessment survey and obtains a C5 risk category. This restriction effectively reduces a C4 user's investment size in the Aggressive strategy.

To estimate the order effect, we run regression (6) using the sample of users who purchased robo-advised products during our sample period and report the results in Panel A of Table 6. In column (1), the dependent variable is a user's total investment in the Aggressive strategy during our sample period. The coefficient on Reverse is 3.28 (s.e. = 1.66), suggesting that, on average, facing the reverse-order survey increases users' investment in the Aggressive strategy by approximately RMB 3,280 during our sample period. This effect corresponds to about 12% of the average purchase size of robo-advised products (RMB 27,830).

Our hypothesis posits that these response-order effects operate through users' risk categories: the response order influences the assigned risk category, which in turn steers the user toward certain products. To shed further light on this hypothesis, we decompose a user's risk assessment into two components. The first is determined by the response order in the survey, while the second reflects the user's characteristic, which is likely related to the user's innate risk tolerance. Specifically, we run regression (6) with the dependent variable $C5_i$, which is a dummy variable that is one if user i's risk category is C5, and zero otherwise.[3] From this regression, we decompose a user's risk assessment into the predicted value $\widehat{C5}_i$ and the regression residual $C5Residual_i$. The former is the user's risk assessment that is determined by the response order in

[3] All users in this sample are in the C4 and C5 categories.

the survey, while the latter reflects the user's innate risk tolerance. We include both components in the regression in column (3). The coefficient of $\widehat{C5}_i$ is 44.26 (s.e. = 22.28). That is, the risk assessment result that is determined by an external factor—the response order in the survey—affects the user's investment decision. Moreover, the coefficient of $C5Residual_i$ is 8.90 (s.e. = 1.35). As expected, if a user is innately more risk tolerant, he would invest more in the Aggressive strategy.

We now turn to the extensive margin, examining whether the response order affects users' participation in robo-advised investments. Specifically, we run regression (6) based on our entire RCT sample, using D_Robo_i as the dependent variable, which equals 1 if user i purchased robo-advised products during the sample period and zero otherwise. The results are reported in Panel B of Table 6. Column (1) shows that the coefficient on $Reverse$ is –0.0254 (*s.e.* = 0.0105), indicating that reverse-order users are 2.54 percentage points less likely to make robo-advised investment.

This result cannot be explained by the compliance mechanism. If compliance were the primary driver, reverse-order users—who tend to receive higher risk categories—should face fewer investment restrictions and therefore be more likely to invest.

One possible explanation is that nudging backfired. With inflated risk profiles due to the order effect, reverse-order users are steered toward riskier strategies. However, many of them may have found those strategies too risky for them and opted out of investing altogether. Our evidence in columns (2) and (3) is consistent with this interpretation. Similar to the decomposition analysis in Panel A, the regression in column (2) decomposes the risk assessment $C5_i$ into a component that is due to the response-order effect $\widehat{C5}_i$ and a component that reflects the user's innate risk tolerance $C5Residual_i$. The regression in column (3) shows that the coefficient of $C5Residual_i$ is 0.10 (s.e. = 0.01). That is innately more risk tolerant users are more likely to have invested in robo-advised products. In contrast, the coefficient of $\widehat{C5}_i$ is -0.30 (s.e. = 0.12), suggesting that if user get an inflated risk category due to the order effect, he is less likely to make an investment.

In summary, our evidence suggests that reversing the response order reduces users' likelihood of investing in robo-advised products. However, conditional on investing, reverse-order users tend to allocate more to the Aggressive strategy.

3.6 Order Effects on Self-directed Investments

This section examines the order effects on self-directed investments by repeating the the analysis in Table 6 on the sample of self-directed investments. To estimate the effect along the intensive margin, we run regression (6) using the sample of users who purchased self-directed products during our sample period and report the results in Panel A of Table 7. In column (1), the dependent variable is a user's total investment in self-directed products in the category of R5 during our sample period. The coefficient on Reverse is 3.09 (s.e. = 1.84), suggesting that, on average, facing the reverse-order survey increases users' investment in the R5 category by approximately RMB 3,090 during our sample period. We note that this is similar to but somewhat weaker than the results on robo-advised products, both statistically and in economic magnitudes. This effect corresponds to about 6% of the average purchase size of self-directed products (RMB 54,000). Columns (2) and (3) conduct similar decomposition analysis of users' risk assessments. The regression in column (2) decomposes a user's risk assessment $C5_i$ into his innate risk tolerance $C5Residual_i$ and the component that is due to the response-order effect $\widehat{C5}_i$. Columns (3) shows that a user's total purchase of R5 products is increasing in both components.

To estimate the effect along the extensive margin, we estimate regression (6) based on our entire RCT sample and use D_Direct_i as the dependent variable, which equals 1 if user *i* purchased any self-directed products during the sample period and zero otherwise. The results are reported in column (1) in Panel B of Table 7. The coefficient on $Reverse$ is –0.0178 (*s.e.* = 0.0090), indicating that reverse-order users are 1.56 percentage points less likely to make self-directed investment. Again, this result is weaker, both statistically and in economic magnitude, than its counterpart for robo-advised investments in the previous section. We also conduct similar decomposition analysis. The results in columns (2) and (3) show that while innately more risk tolerant users are more likely to have invested in self-directed products, and that if a user get an inflated risk category due to the response-order effect, he is less likely to make an investment.

In summary, the analysis in this section suggests that reversing the response order reduces users' likelihood of investing in self-directed products. However, conditional on investing, reverse-order users tend to allocate more to products in the riskiest category R5. We note that the results long both the intensive and extensive margins are weaker than their counterparts for robo-advised investments. This is consistent with our earlier observation that the order effect is likely weaker for self-directed investments, possibly because users play a more active role in their investment decisions.

3.7 Implications

Our findings raise a potential ethical concern. Suppose there exists a conflict of interest between a robo-advisor and its clients—for instance, the robo-advisor may profit from steering users toward financial products that are not optimal from the clients' perspective. Because investor preferences are inferred from survey responses that can be manipulated through design choices such as response order, a robo-advisor could exploit this mechanism to nudge users toward products that serve its own interests, all while maintaining the appearance that users "prefer" those products.

More broadly, our findings have implications for constructs such as risk aversion and preference. These concepts are so deeply embedded in economic analysis that they are often treated as primitive variables. Our evidence suggests that they are not fully determined in the sense that they cannot be fully inferred from either surveys or even the revealed preference approach. This is because that the response order not only affects the preference parameters implied by survey data, but also affects investment decisions. Hence, the response order—an external factor—can directly affect the inferred parameters even based on the revealed preference approach, challenging the notion that such parameters are intrinsic and context independent.

How do we mitigate response-order bias? The solution depends on the goal. For example, if the goal is to estimate the average opinion across a population, we can mitigate the bias by randomizing the response order such that one half of the respondents face the choices in the original order while the other half face them in the reverse order. The average across all respondents is then a more accurate estimate because the biases of the two subsamples would largely offset each other. The solution is different if the goal is to estimate the opinion of an individual. Suppose, for example, the goal of a robo-advisor is to estimate a customer's risk tolerance. One solution is to include two different questions on risk tolerance in the survey, with the most risk-tolerant response listed at the top for one question and at the bottom for the other. The average of the inferred risk tolerances across these two questions is likely to be a more accurate estimate as the biases of these two questions offset each other.

4. Conclusions

Our evidence shows that FinTech has transferred a seemingly innocuous psychological bias into the two-trillion-dollar robo-advisory market. We begin with a nationwide survey of more than 3,000 experienced U.S. investors and document a substantial response-order effect in the risk-tolerance questions commonly used by robo-advisors. Building on these findings, we partnered with a leading robo-advisor in China to conduct a preregistered RCT involving more than 10,000 investors. We find that even when substantial financial stakes are involved, the order of survey responses continues to exert a significant influence on clients' risk profiles, which directly affect the robo-advisor's product recommendations. More importantly, the response order also has significant effects on users' actual investment behavior.

While our evidence is derived from an RCT in China, the response-order effect is expected to be even stronger in the U.S. context, where investors generally play a more passive role. Once users establish and fund their accounts with a robo-advisor, subsequent investment decisions are typically executed automatically based on the survey results. Consequently, any bias introduced by the response order is transmitted directly to investment decisions.

Our findings raise an important ethical concern: a robo-advisor may steer users toward products that advance its own interests while preserving the appearance that these products reflect users' preferences. More broadly, the evidence indicates that preference constructs such as risk aversion are not fully determinate: their estimated values—whether inferred from surveys or investment decisions—systematically vary with exogenous factors, such as the response order.

Finally, our analysis focuses on short-term effects. A long-term RCT would be required to evaluate the persistence of these effects on investment behavior. In addition, long-term trials are essential for achieving sufficient statistical power to reliably assess investment performance. We leave these important questions to future research.

References

Autor, David, Frank Levy, and Richard Murnane, 2003, The Skill Content of Recent Technological Change: An Empirical Exploration, *Quarterly Journal of Economics* 118, 1279–1333.

Baumol, William, 1967, Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis." *American Economic Review* 57, 415–426.

Bordalo, Pedro, Nicola Gennaioli, Rafael LaPorta, and Andrei Shleifer, 2025, Finance Without Exotic Risk, *Journal of Financial Economics*, forthcoming.

Brogaard, Jonathan, Joseph Engelberg, Sapnoti Eswar, and Edward Van Wesep, 2023, On the Causal Effect of Fame on Citations, *Management Science* 70, 7187–7214.

Chen, Eric, Gabor Simonovits, Jon A. Krosnick, and Josh Pasek, 2014, The impact of candidate name order on election outcomes in North Dakota, *Electoral Studies* 35, 115–122.

Choi, James and Adriana Z. Robertson, 2020, What Matters to Individual Investors? Evidence from the Horse's Mouth, *Journal of Finance* 75, 1965–2020.

Chinco, Alex, Samuel Hartzmark, and Abigail Sussman, 2022, A New Test of Risk Factor Relevance, *Journal of Finance* 77, 2183–2238.

Coibion, Olivier, Yuriy Gorodnichenko, and Michael Weber, 2020, Political Polarization and Expected Economic Outcomes, Working paper.

D'Acunto, Francesco and Alberto Rossi, 2023, Robo-Advice: Transforming Households into Rational Economic Agents, *Annual Review of Financial Economics* 15, 543-563.

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

D'Acunto, Francesco and Michael Weber, 2024, Why Survey-Based Subjective Expectations are Meaningful and Important, *Annual Review of Economics* 16, 329–357..

De la O, Ricardo, and Sean Myers, 2021, Subjective cash flow and discount rate expectations. *Journal of Finance* 76, 1339–1387.

Feenberg, Daniel R., Ina Ganguli, Patrick Gaule and Jonathan Gruber, 2017, It's Good to Be First: Order Bias in Reading and Citing NBER Working Papers, *Review of Economics and Statistics* 99, 32–39.

Frydman, Cary and Baolian Wang, 2020, The Impact of Salience on Investor Behavior: Evidence from a Natural Experiment, *Journal of Finance* 75, 229–276.

Goldfarb, Avi, and Catherine Tucker, 2019, Digital Economics, *Journal of Economic Literature* 57, 3–43.

Greenwood, Robin, and Andrei Shleifer, 2014, Expectations of Returns and Expected Returns, *Review of Financial Studies* 27, 714–746.

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

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus, 2021, The joint dynamics of investor beliefs and trading during the COVID-19 crash, *PNAS* 118 (4) e2010316118.

Hartzmark, Samuel, 2015, The Worst, The Best, Ignoring All the Rest: The Rank Effect and Trading Behavior, *Review of Financial Studies* 28, 1024–1059.

Hong, Harrison, Ilan Kremer, Jeffrey Kubik, Jianping Mei and Michael Moses, 2015, Ordering, Revenue and Anchoring in Art Auctions, *RAND Journal of Economics* 46, 186–216.

Huang, Wei, 2015, Do ABCs Get More Citations than XYZs? *Economic Inquiry* 53, 773–789.

Jacowitz, Karen E. and Daniel Kahneman, 1995, Measures of Anchoring in Estimation Tasks, *Personality and Social Psychology Bulletin* 21, 1161–1166.

Jiang, Lei, Jinyu Liu, Lin Peng, and Baolian Wang, 2022, Investor Attention and Asset Pricing Anomalies, *Review of Finance* 26, 563–593.

Jiang, Zhengyang, Hongqi Liu, Cameron Peng and Hongjun Yan, 2025, Investor Memory and Biased Beliefs: Evidence from the Field, *Quarterly Journal of Economics*, forthcoming.

Jiang, Zhengyang, Cameron Peng and Hongjun Yan, 2024, Personality Differences and Investment Decision-Making, *Journal of Financial Economics* 153, 103776.

Jurajda, Stepan, and Daniel Munich, 2010, Admission to Selective Schools, Alphabetically, *Economics of Education Review* 29, 1100–1109.

Kempf, Elisabeth and Margarita Tsoutsoura, 2021, Partisan Professionals: Evidence from Credit Rating Analysts, *Journal of Finance* 76, 2805–2856.

Krosnick, Jon and Presser, Stanley, 2010, Question and Questionnaire Design, in *Handbook of Survey Research* (2nd Edition) James D. Wright and Peter V. Marsden (Eds). San Diego, CA: Elsevier.

Liao, Li, Jia Xiang, Zhengwei Wang, Hongjun Yan, and Jun Yang, 2021, User Interface and First-hand Experience in Retail Investing, *Review of Financial Studies* 34, 4486–4523.

Malmendier, Ulrike, and Stefan Nagel, 2011, Depression babies: Do Macroeconomic Experiences Affect Risk-taking? *Quarterly Journal of Economics* 126, 373–416.

Manski, Charles, 2004, Measuring Expectations, *Econometrica*, 72, 1329–1376.

Meeuwis, Maarten, Jonathan A. Parker, Antoinette Schoar, and Duncan I. Simester, 2022, Belief Disagreement and Portfolio Choice, *Journal of Finance* 77, 3191–3247.

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

Tversky, Amos and Daniel Kahneman, 1974, Judgment under Uncertainty: Heuristics and Biases, *Science* 185, 1124–1131.

Varian, Hal, 2010, Computer Mediated Transactions, *American Economic Review: Papers and Proceedings* 100, 1–10.

Xiao, Yaqing and Hongjun Yan, 2023, Relatable Information Disproportionately Shapes Expectations, Working paper.

Table 1. Summary Statistics for the AAII Survey

This table reports the summary statistics for each survey question. The survey was sent out by the American Association of Individual Investors (AAII) to its members on October 27, 2020; data collection stopped on November 2, 2020. The survey received responses from 3,146 individuals. The survey is presented in two different versions. Each respondent is randomly assigned one version with a 50% chance. Both versions have the same set of questions presented in the same order. The only difference is the order of response choices. 1,561 respondents faced the original-order survey (shown in the Appendix A1) and 1,585 faced the survey with the choices for each question in the reverse order. The first column lists the survey questions. Those with asterisks are classified as "subjective questions" and all others are "objective questions." The second column reports the number of choices (#choice) for each question. The next three columns report the mean, standard deviation of *Answer* (defined in equation (1)), and number of respondents (N) for each question for the original-order subsample. Similarly, the last three columns report the results for the reverse-order subsample.

		Original order (1,561 respondents)			Reverse order (1,585 respondents)		
Question	#choice	Mean	S.D.	N	Mean	S.D.	N
1 Age	5	4.43	0.77	1549	4.45	0.74	1567
2 Gender	2	1.93	0.26	1529	1.94	0.24	1544
3 Experience	4	3.86	0.49	1557	3.90	0.44	1576
4* Horizon	6	4.47	1.40	1553	5.01	1.21	1579
5* Risk tolerance	4	2.75	0.98	1550	3.02	0.93	1576
6* GDP	5	3.29	0.89	1548	3.44	0.78	1570
7* Stock return	5	3.32	0.71	1536	3.41	0.69	1564
8 COVID mild	2	1.50	0.50	1551	1.49	0.50	1577
9 COVID serious	2	1.27	0.45	1551	1.27	0.45	1575
10* Facemask	5	4.26	1.04	1555	4.24	1.08	1577
11* Vaccine forecast	5	2.91	0.79	1546	3.09	0.83	1572
12* Vaccine take	5	4.17	0.93	1551	4.24	0.92	1570
13* Political leaning	5	3.10	1.32	1545	3.23	1.34	1565
14* Vote	5	2.64	1.79	1509	2.78	1.83	1521
15* Election forecast	5	2.52	0.98	1533	2.64	0.99	1556
16 Education	6	5.44	0.84	1553	5.43	0.86	1576
17 Residence	3	1.99	0.55	1551	2.01	0.57	1576

Table 2. Response-order Effects

Panel A reports the results from the following regression:

$$SAnswer_{ij} = \alpha + \beta Reverse_i \times Sub_j + \delta\, Reverse_i + \gamma\, Sub_j + \varepsilon_{ij},$$

where $SAnswer_{ij}$ is respondent i's standardized answer to Question j and is defined in (3); $Reverse_i$ is 1 if respondent i faces a reverse-order survey and 0 otherwise; and Sub_j is 1 if Question j is a subjective question and 0 otherwise. Standard errors, in parentheses, are clustered by respondent. Then, for each question j, we regress $SAnswer_{ij}$ on $Reverse_i$, and Panel B reports the estimated coefficient on $Reverse$, its standard error, and the number of observations for each regression. ***, **, and * indicate significance levels of 1%, 5%, and 10%, respectively.

Panel A. Objective vs subjective questions

	(1)	(2)	(3)	(4)
Reverse	2.54***	2.56***	0.31	0.36
	(0.29)	(0.29)	(0.47)	(0.47)
Reverse × Sub			3.80***	3.75***
			(0.58)	(0.58)
Sub			-11.60***	-12.36***
			(0.42)	(0.70)
Constant	63.16***	48.69***	69.97***	49.80***
	(0.21)	(0.52)	(0.33)	(0.55)
Question FEs	No	Yes	No	Yes
N	52908	52908	52908	52908
R-squared	0.001	0.317	0.020	0.318

Panel B. Order effect for each question

		Reverse	s.e.	N
Objective questions	1 Age	0.71	0.68	3116
	2 Gender	0.91	0.90	3073
	3 Experience	1.16**	0.56	3133
	8 COVID mild	-0.83	1.79	3128
	9 COVID severe	-0.04	1.60	3126
	16 Education	-0.21	0.61	3129
	17 Residence	0.80	1.00	3127
Subjective questions	4 Horizon	10.67***	0.93	3132
	5 Risk tolerance	9.08***	1.14	3126
	6 GDP	3.69***	0.75	3118
	7 Stock return	2.29***	0.63	3100
	10 Facemask	-0.50	0.95	3132
	11 Vaccine forecast	4.31***	0.73	3118
	12 Vaccine take	1.65**	0.83	3121
	13 Political leaning	3.34***	1.19	3110
	14 Vote	3.52**	1.65	3030
	15 Election forecast	2.94***	0.89	3089

Table 3. Quantify the Response-order Effect on Expectations

This table reports the estimated results from the following regression:

$$Y_i = \alpha + \beta_1 Reverse_i + \beta_2 Mild_i + \beta_3 Severe_i + \beta_4 D_i' + \beta_5 R_i' + \varepsilon_i,$$

where Y_i is respondent i's answer to a survey question; $Reverse_i$ is a dummy variable that is 1 if respondent i's survey is in the reverse order and 0 otherwise; $Mild_i$ is a dummy variable that is 1 if respondent i answered "Yes" to Question 9 ("Do you directly know someone who has been infected by COVID-19 with mild or no symptoms?") and "No" to Question 10 ("Do you directly know someone who has been severely ill due to COVID-19 infection?"); $Severe_i$ is a dummy variable that is 1 if respondent *i* answered "Yes" to Question 10; D_i' is the residual from regressing D_i on $Reverse_i$; R_i' is the residual from regressing R_i on $Reverse_i$; D_i is a dummy variable that is 1 if respondent *i* answered "Definitely Democrat" or "Usually Democrat" to Question 13 ("What is your political leaning?"); and R_i is a dummy variable that is 1 if respondent i answered "Definitely Republican" or "Usually Republican" to Question 13. In column (1), Y_i is the forecast of the S&P 500 stock index return and is set to -20%, -10%, 0%, 10%, and 20% if respondent *i* selected choices 1 through 5, respectively, in their responses Question 7 ("What do you think GDP growth in the U.S. will be over the next 12 months?"). In column (2), Y_i is the forecast of the GDP growth rate and is set to -7%, -3.5%, 0%, 3.5%, and 7% if respondent *i* selected choices 1 through 5, respectively, in their responses Question 6 ("What do you think the total return of the S&P 500 stock market index will be over the next 12 months?"). All regressions include age, gender, education, and residence fixed effects. Standard errors are reported in parentheses. ***, **, and * indicate significance levels of 1%, 5%, and 10%, respectively.

	Stock return (%)	GDP growth (%)
	(1)	(2)
Reverse	0.95***	0.45***
	(0.26)	(0.10)
Mild	-0.85***	-0.13
	(0.31)	(0.12)
Severe	-0.90***	-0.26**
	(0.31)	(0.12)
D'	-0.63*	-0.62***
	(0.34)	(0.13)
R'	1.36***	0.86***
	(0.31)	(0.12)
Constant	-1.60	-1.97*
	(2.91)	(1.15)
N	3,009	3,027
R-squared	0.035	0.071

Table 4. Summary Statistics for the RCT

Panel A reports the composites of the target portfolios of the Aggressive-, Balanced-, and Stable-Strategies. Panel B reports the number of users in each risk category for our RCT Sample, its two subsamples (the Original-order Subsample and the Reverse-order Subsample), and the Placebo Sample. Panels C through E report the summary statistics for various subsamples. *Score* is a user's risk score as of April 1, 2025. *Category* is a user's risk category as of April 1, 2025, and is 1 through 5 if the user's is in category C1 through C5, respectively. *Reverse* is one if a user's account number is odd (and hence if the user is in the RCT, he receives the reverse-order survey), and zero otherwise. *Age* is a user's age, denominated in years. *Male* is a dummy variable that is one is a user is a male and zero otherwise. *BuyRobo* and *BuyDirect* denote a user's total purchases of robo-advised and self-directed investment products, respectively, during the RCT period. *InitialAll* is a user's total asset holding before entering the RCT. *InitialRobo* and *InitialDirect* denote a user's total holdings of robo-advised and self-directed investment products, respectively, prior to the RCT. *InitialScore* and *InitialCategory* are a user's risk score and category, respectively, before entering the RCT. Columns labeled by "N" report the number of observations. The last two columns report the difference and its standard error. *NewUser* is a dummy variable that is one is a user's account was opened before October 25, 2024, and zero otherwise.

Panel A. Target portfolios

	Aggressive	Balanced	Stable
Chinese Stocks (%)	35	25	20
U.S. Stocks (%)	12	8	4
Alternative Investments (%)	25	20	10
Money Market Funds (%)	18	12	6
Chinese Government Bonds (%)	10	35	60

Panel B. Sample Size

	C1	C2	C3	C4	C5	Total
RCT Sample	239	3	18	3149	3732	7141
Original-order	161	1	10	1673	1697	3542
Reverse-order	78	2	8	1476	2035	3599
Placebo Sample	256	0	28	1653	1205	3142

Panel C. RCT and Placebo Samples

	RCT Sample			Placebo Sample		
	Mean	S.D.	N	Mean	S.D.	*N*
Score	66.88	13.32	7141	61.25	16.59	3142
Category	4.42	0.81	7141	4.13	1.05	3142
Reverse	0.50	0.50	7141	0.49	0.50	3142
Age (years)	45.66	10.28	6440	39.69	11.12	1640
Male	0.60	0.49	6440	0.47	0.50	1640
BuyRobo (1000 RMB)	27.83	64.21	1746	19.56	18.33	21
BuyDirect (1000 RMB)	54.00	142.08	1419	23.39	46.87	1067

Panel D. Existing- and New-user Samples

	Existing Users			New Users		
	Mean	S.D.	*N*	Mean	S.D.	*N*
Score	67.95	12.15	5818	62.16	16.78	1323
Category	4.47	0.73	5818	4.18	1.05	1323
Reverse	0.50	0.50	5818	0.52	0.50	1323
Age (years)	46.47	9.99	5818	38.09	9.84	622
Male	0.61	0.49	5818	0.48	0.50	622
BuyRobo (1000 RMB)	27.84	64.40	1734	26.06	25.24	12
BuyDirect (1000 RMB)	51.41	146.26	959	59.41	132.93	460
InitialAll (1000 RMB)	315.83	658.54	5818			
InitialRobo (1000 RMB)	158.01	414.60	5818			
InitialDirect (1000 RMB)	87.65	293.16	5818			

Panel E. Original- and Reverse-Order Samples

	Original Order		Reverse Order			
	Mean	N	Mean	N	Difference	*s.e.*
*Male*100*	58.86	3189	60.47	3251	1.62	1.22
Age (year)	45.68	3189	45.65	3251	-0.04	0.26
NewUser	0.18	3542	0.19	3599	0.01	0.01
InitialScore	49.89	2903	49.47	2915	-0.43	0.54
InitialCategory	3.68	2903	3.65	2915	0.03	0.03
InitialAll (1000 RMB)	308.69	2903	322.94	2915	14.25	17.27
InitialRobo (1000 RMB)	155.16	2903	160.85	2915	5.69	10.87
InitialDirect (1000 RMB)	83.14	2903	92.15	2915	9.01	7.69
Score (post-treatment)	65.19	3542	68.53	3599	3.35	0.31
Category (post-treatment)	4.34	3542	4.50	3599	0.16	0.02
BuyRobo (1000 RMB)	25.70	910	30.14	836	4.44	3.08
BuyDirect (1000 RMB)	49.32	715	58.76	704	9.44	7.54

Table 5. Order Effects on Risk Assessment

This table reports the results from the following regression:

$$Y_i = \alpha + \beta Reverse_i + X_i + \varepsilon_i,$$

where $Reverse_i$ is a dummy variable equal to one if user i's account number is odd (and hence if the user is in the RCT, he receives the reverse-order survey), and zero otherwise. Panels A and B are based on the RCT and Placebo Samples, respectively. For both panels, Y_i is user i's risk score in columns (1) and (2); and is 1 through 5 if user i's risk category is C1 through C5, respectively, for columns (3) and (4). X_i refers to all control variables, which are defined in Table 4. Standard errors are reported in parentheses. ***, **, and * indicate significance levels of 1%, 5%, and 10%, respectively.

Panel A. RCT Sample

	Risk Score		Risk category	
	(1)	(2)	(3)	(4)
Reverse	3.35***	3.12***	0.16***	0.14***
	(0.31)	(0.28)	(0.02)	(0.02)
Male		5.00***		0.23***
		(0.30)		(0.02)
Age		-0.06***		-0.00***
		(0.02)		(0.00)
LnInitialAll		1.91***		0.10***
		(0.11)		(0.01)
NewUser		7.34***		0.48***
		(0.74)		(0.04)
Constant	65.19***	56.92***	4.34***	3.95***
	(0.24)	(0.88)	(0.01)	(0.05)
N	7,141	6,440	7,141	6,440
R-squared	0.02	0.11	0.01	0.08

Panel B. Placebo sample

	Risk Score		Risk Category	
	(1)	(2)	(3)	(4)
Reverse	0.41	-0.27	0.02	-0.03
	(0.59)	(0.62)	(0.04)	(0.04)
Male		3.52***		0.13***
		(0.63)		(0.04)
Age		0.17***		0.01***
		(0.03)		(0.00)
Constant	61.05***	56.32***	4.12***	4.09***
	(0.42)	(1.28)	(0.03)	(0.07)
N	3,142	1,640	3,142	1,640
R-squared	0.00	0.04	0.00	0.01

Table 6. Order Effects on Robo-Advised Investments

Panel A reports the analysis on the intensive margin. The sample includes users who made robo-advised investments during our sample period. Columns (1) and (2) report the results from the following regression:

$$Y_i = \alpha + \beta Reverse_i + X_i + \varepsilon_i,$$

where $Reverse_i$ is a dummy variable equal to one if user i's account number is odd (and hence the user receives the reverse-order survey), and zero otherwise. X_i refers to all control variables, which are defined in Table 4. The dependent variable Y_i is A_i, user i's total purchase of the *Aggressive* portfolio, denominated in 1000 RMB, in column (1), and is $C5_i$, which is a dummy variable that is one if user i's risk category is C5, and zero otherwise, in column (2). The regression in column (3) the same as in column (1) with $Reverse_i$ replaced by $\widehat{C5}_i$ and $C5Residual_i$, which are the predicted value of $C5_i$ and residual based on the regression in column (2). Panel B reports similar analysis on the extensive margin and is based on the entire RCT Sample. In columns (1) and (3), the dependent variable Y_i is a dummy variable D_Robo_i, which is one if user i invested in robo-advised investments during our sample period, and zero otherwise. Robust standard errors are reported in parentheses. ***, **, and * indicate significance levels of 1%, 5%, and 10%, respectively.

Panel A. Intensive Margin

	A	*C5*	*A*
	(1)	(2)	(3)
Reverse	3.28**	0.07***	
	(1.66)	(0.02)	
Male	3.36*	0.18***	-4.82
	(1.80)	(0.02)	(4.41)
Age	0.11	-0.00	0.14
	(0.09)	(0.00)	(0.09)
InitialRobo	0.02***	0.00***	0.01**
	(0.00)	(0.00)	(0.00)
NewUser	8.92	-0.18	17.02**
	(8.22)	(0.15)	(8.30)
$\widehat{C5}$			44.26**
			(22.28)
C5Residual			8.90***
			(1.35)
Constant	-0.26	0.52***	-23.42*
	(4.68)	(0.06)	(12.39)
Observations	1,746	1,746	1,746
R-squared	0.05	0.05	0.06

Panel B. Extensive Margin

	D_Robo_i	C5	D_Robo_i
	(1)	(2)	(3)
Reverse	-0.0254**	0.09***	
	(0.0105)	(0.01)	
Male	0.01	0.17***	0.06**
	(0.01)	(0.01)	(0.02)
Age	-0.00***	-0.00*	-0.00***
	(0.00)	(0.00)	(0.00)
LnInitial	0.06***	0.03***	0.07***
	(0.00)	(0.00)	(0.00)
NewUser	-0.09***	0.06***	-0.08***
	(0.01)	(0.02)	(0.01)
$\widehat{C5}$			-0.30**
			(0.12)
C5Residual			0.10***
			(0.01)
Constant	0.19***	0.36***	0.29***
	(0.03)	(0.03)	(0.06)
Observations	6,440	6,440	6,440
R-squared	0.11	0.05	0.12

Table 7. Order Effects on Self-Directed Investments

Panel A reports the analysis on the intensive margin. The sample includes users who made self-directed investments during our sample period. Columns (1) and (2) report the results from the following regression:

$$Y_i = \alpha + \beta Reverse_i + X_i + \varepsilon_i,$$

where $Reverse_i$ is a dummy variable equal to one if user *i*'s account number is odd (and hence the user receives the reverse-order survey), and zero otherwise. X_i refers to all control variables, which are defined in Table 4. The dependent variable Y_i is $R5_i$ (user i's total purchase of self-directed investment products in the risk category of R5, denominated in 1000 RMB) in column (1), and is $C5_i$, which is a dummy variable that is one if user i's risk category is C5, and zero otherwise, in column (2). The regression in column (3) the same as in column (1) with $Reverse_i$ replaced by $\widehat{C5}_i$ and $C5Residual_i$, which are the predicted value of $C5_i$ and residual based on the regression in column (2). Panel B reports similar analysis on the extensive margin and is based on the entire RCT Sample. In columns (1) and (3), the dependent variable Y_i is a dummy variable D_Direct_i, which is one if user i invested in self-directed investments during our sample period, and zero otherwise. Robust standard errors are reported in parentheses. ***, **, and * indicate significance levels of 1%, 5%, and 10%, respectively.

Panel A. Intensive Margin

	R5	*C5*	R5
	(1)	(2)	(3)
Reverse	3.09*	0.08***	
	(1.84)	(0.02)	
Male	2.04	0.15***	-4.09
	(1.58)	(0.03)	(4.32)
Age	0.15	-0.00	0.17
	(0.11)	(0.00)	(0.11)
InitialDirect	0.02	0.00***	0.02
	(0.02)	(0.00)	(0.02)
NewUser	15.35***	-0.11***	19.68***
	(3.70)	(0.03)	(5.78)
$\widehat{C5}$			41.22*
			(24.39)
C5Residual			9.41***
			(1.03)
Constant	-10.70	0.58***	-34.46*
	(7.40)	(0.06)	(20.00)
Observations	1,419	1,419	1,419
R-squared	0.08	0.06	0.09

Panel B. Extensive Margin

	D_Direct_i	C5	D_Direct_i
	(1)	(2)	(3)
Reverse	-0.0178**	0.08***	
	(0.0090)	(0.01)	
Male	-0.03***	0.18***	0.01
	(0.01)	(0.01)	(0.02)
Age	-0.00	-0.00	-0.00*
	(0.00)	(0.00)	(0.00)
LnInitial	0.05***	0.03***	0.06***
	(0.00)	(0.00)	(0.00)
NewUser	0.69***	0.04*	0.70***
	(0.02)	(0.02)	(0.02)
$\widehat{C5}$			-0.21**
			(0.11)
C5Residual			0.09***
			(0.01)
Constant	0.10***	0.36***	0.17***
	(0.02)	(0.03)	(0.05)
Observations	6,440	6,440	6,440
R-squared	0.24	0.05	0.25

Appendix A1. AAII Survey Questions

Question 1: How old are you?
18 – 39 years
40 – 49 years
50 – 59 years
60 – 69 years
70 years or older

Question 2: What is your gender?
Female
Male

Question 3: How long have you been investing?
0 – 9 years
10 – 15 years
16 – 20 years
More than 20 years

Question 4: For the goal(s) of your overall investment portfolio, what is your approximate time horizon?
Less than 1 year
1 - 5 years
6 - 10 years
11 - 20 years
More than 20 years
Beyond my lifespan (estate)

How certain do you feel about your answer to this question?
Very certain: I have a clear idea about my investment horizon
Somewhat certain: I have a rough idea about my investment horizon
Not certain: I have barely considered my investment horizon

Question 5: Recognizing that investments can fluctuate in value, for a $100,000 investment, how much of a decline in the value of your investment portfolio could you tolerate over the course of a year?
-10% (-$10,000)
-20% (-$20,000)
-30% (-$30,000)
-40% (-$40,000)

Question 6 What do you think GDP growth in the U.S. will be over the next 12 months?
Less than -5%
-5% to -2%
-2% to 2%
2% to 5%
Greater than 5%

How certain are you about your forecast?
I am very certain about my forecast
I am somewhat certain about my forecast
I am not certain about my forecast.

Question 7 What do you think the total return of the S&P 500 stock market index will be over the next 12 months?

Less than -15%
-15% to -5%
-5% to 5%
5% to 15%
Greater than 15%

How certain are you about your forecast?

I am very certain about my forecast
I am somewhat certain about my forecast
I am not certain about my forecast

Question 8 Do you directly know someone who has been infected by COVID-19 with mild or no symptoms?

No
Yes

Question 9 Do you directly know someone who has been severely ill due to COVID-19 infection?

No
Yes

Question 10 Do you believe that wearing a mask reduces *your* risk of COVID-19 infection?

Definitely not
Probably not
Not sure
Probably yes
Definitely yes

How certain are you about your answer?

I am very certain about my answer
I am somewhat certain about my answer
I am not certain about my answer

Question 11 When do you expect a COVID-19 vaccine to become widely available?

Within 3 months
3 months to 6 months
6 months to 12 months
1 year to 3 years
More than 3 years

How certain are you about your forecast?

I am very certain about my forecast
I am somewhat certain about my forecast
I am not certain about my forecast

Question 12 When a COVID-19 vaccine becomes available, do you plan to take it?

Definitely not
Probably not
Not sure
Probably yes
Definitely yes

How certain are you about your answer?

I am very certain about my answer
I am somewhat certain about my answer
I am not certain about my answer

Question 13 What is your political leaning?
Definitely Democrat
Usually Democrat
Independent or Other
Usually Republican
Definitely Republican

Question 14 Who will you vote for in the upcoming U.S. Presidential election?
Definitely Joe Biden
Likely Joe Biden
Undecided, Someone else, or Not voting
Likely Donald Trump
Definitely Donald Trump

Question 15 Who do you think will win the upcoming U.S. Presidential election?
Definitely Joe Biden
Likely Joe Biden
Hard to say
Likely Donald Trump
Definitely Donald Trump

How certain are you about your forecast?
I am very certain about my forecast
I am somewhat certain about my forecast
I am not certain about my forecast

Question 16: What is the highest level of school you have completed or the highest degree you have received?
Less than high school degree
High school degree or equivalent (e.g., GED)
Some college but no degree
Associate degree
Bachelor's degree
Graduate degree

Question 17: Which of the following best describes the area you live in?
Urban
Suburban
Rural

Is there any ambiguity in determining your area to be Urban/Suburban/Rural?
Yes
No

Appendix A2: Risk Assessment Survey Questions at the Advisor

Question 1: How would you describe your investment knowledge?

Limited: Virtually no knowledge of financial products
Moderate: Basic knowledge and understanding of financial products and their associated risks
Extensive: Comprehensive knowledge and understanding of financial products and their associated risks

Question 2: What is your expected investment duration?

Less than half a year
Less than 1 year
1 - 3 years
3 - 5 years
Over 5 years

Question 3: How old are you?

Under 25 years old
25 – 35 years
35 – 50 years
Over 50 years old

Question 4: What is your family's annual after-tax income?

Less than 100,000 yuan
100,000 - 300,000 yuan
300,000 - 500,000 yuan
500,000 - 1,000,000 yuan
Over 1,000,000 yuan

Question 5: What is your family's investable assets?

Less than 100,000 yuan
100,000 - 500,000 yuan
500,000 - 2,000,000 yuan
Over 2,000,000 yuan

Question 6: How long have you been investing?

Less than 1 year
1 - 3 years
3 - 5 years
Over 5 years

Question 7: Please select your primary investment category?

Money market funds
Bank wealth management products
Bonds
Stocks

Question 8: What is your investment preference?

4% return with no principal loss
6% return with a maximum 2% potential principal loss
10% return with a maximum 7% potential principal loss
15% return with a maximum 15% potential principal loss
30% return with a maximum 40% potential principal loss

Question 9: What is your primary source of income?

Wages and labor remuneration
Income from business operations
Income from financial assets such as interest, dividends, and asset transfers
Income from non-financial assets such as rental and sale of real estate
No steady income

Question 10: Do you have any large outstanding debts?

No
Yes, long-term debt such as mortgage
Yes, short-term debt such as credit card debts and consumer credit
Yes, loans from relatives and friends

Question 11: Do you have any negative credit records?

Yes
No

Appendix A3. Additional Analysis

1. Probit Analysis

Using the AAII survey data in Section 2, we run the following Probit regression

$$\Pr(Answer_{ij} = k) = \Phi(\alpha + \beta Reverse_i \times Sub_j + \gamma Reverse_i + \delta Sub_j + \varepsilon_{ij}), \qquad (A1)$$

for $k = 1, 2, 3, 4, 5$, where $\Phi(\cdot)$ is the CDF of a standard normal distribution. For the case of $k = 1$, the above regression estimates the effect of reversing the response order on the probability of selecting the first choice on the original-order list. Similarly, in the cases of $k = 2$ and $k = 3$, Regression (5) estimates the effects on the probability of selecting the second and third choices. In the cases of $k = 4$ and $k = 5$, Regression (A1) estimates the effect on the probability of selecting the penultimate and last choices.[4] Regression results are reported in Table A1.

The first column of Panel A shows that the coefficient on *Reverse* is -0.00 (*s.e.* = 0.03) and the coefficient on the interaction term is -0.11 (*s.e.* = 0.04). This implies that for objective questions, reversing the response order does not affect the probability of selecting the first choice. For subjective questions, however, the first choice is significantly less likely to be selected when it is presented at the bottom of the list in the reverse-order survey. The implied marginal effects are reported in the first column of Panel B. For subjective questions, reversing the response order reduces the selection probability of the first choice by 1.95 (*s.e.* = 0.44) percentage points, while, for objective questions, the marginal effect of *Reverse* is merely 0.03 (*s.e.* = 0.24) percentage points.

The rest of the table shows that for objective questions, reversing the order has no significant effect on the selection probabilities of all other choices. For subjective questions, however, reversing the order decreases the selection probability by 3.66 (*s.e.* = 0.45) percentage points for the second choice, but increases the selection probability by 1.57 (*s.e.* =0.44) and 3.66 (*s.e.* = 0.42) percentage points for the penultimate and last choices, respectively. For the middle choice ($k = 3$), reversing order appears to have an insignificant effect.

The above results show that respondents have a bias towards the choices at or near the top

[4] Note that the number of choices for a survey question ranges from 2 to 6. In the regressions for the cases of $k = 1$ and $k = 5$, the regression is based on the data from all survey questions. In the regression for the cases of $k = 2$ and $k = 4$, the data from survey questions with 2 or 3 choices are excluded. In the regression for the case of $k = 3$, only the data from questions with 5 choices are included.

of their lists. When the response order is reversed, the last and penultimate choices on the original-order list are now the first and second choices and hence more likely to be selected. The first and second choices on the original-order list are now at the bottom and less likely to be selected. The third choice stays in the middle and reversing order has an insignificant effect. The overall effect is strong for subjective questions and disappears for objective questions to which respondents are likely to have clear answers (such as their age and residence).

2. Decision Time

Our analysis in Section 2 was motivated by the interpretation based on anchoring. For a given survey question, the top choice serves as the anchor and hence respondents tend to select choices close to it, leading to response order bias. For subjective questions, respondents are more uncertain about their answers. Since the anchoring effect tends to be stronger in such circumstances (Jacowitz and Kahneman, 1995), the response-order effect should be stronger. An alternative interpretation of the response order bias is based on the notion of "satisficing behavior" (Krosnick and Presser, 2010). Respondents start from the top of the list and go down until reaching a choice they are satisfied with, leading to an order effect.

Under the satisficing-behavior interpretation, respondents who make decisions more quickly are more likely to be subject to the satisficing behavior; they stop reading the choices as soon as they reach an "acceptable" one and therefore tend to finish their surveys more quickly. Hence, the response-order effect should be stronger among respondents with shorter decision times. The anchoring-based interpretation implies the opposite. If a respondent is more confident about her answers and hence less subject to the order effect, she would also spend less time on her responses. Therefore, the anchoring-based interpretation implies that the order effect should be weaker among respondents with shorter decision times.

To examine these implications, we extend Regression (4) to include a triple interaction term $Reverse_i \times Sub_j \times Fast_i$, where $Fast_i$ is a dummy variable that is 1 if respondent i's survey completion time is among the shortest α ($0<\alpha<1$) of our sample and 0 otherwise. The regression results are reported in Table 5. Column (1) is for the case of α=1/2. Consistent with the anchoring-based interpretation, the triple interaction coefficient is -3.45 (*s.e.*=1.15). That is, the order effect is weaker for respondents with shorter decision times. We also repeat the regression with α=1/3 and α=1/4. The results, reported in Columns (2) and (3), are similar.

3. Order Effect on Survey Expectations

This section reports additional analysis that quantifies the order effect on survey expectations based on the AAII survey data. We first analyze the order effect on respondents' forecast of the vaccine development. Question 11 of the survey is "When do you expect a COVID-19 vaccine to become widely available?" We run Regression (5) based on the data from this question. For ease of interpretation of regression coefficients, we set the dependent variable Y_i as 1.5, 4.5, 9, 24, and 48 months, if respondent *i* selected choices 1 through 5, respectively. In column (3), the coefficient on *Reverse* is 1.89 (*s.e.*=0.30). That is, when responses are listed in the reverse order, the average forecast for a COVID vaccine is delayed by 1.89 months. The coeffects of D' and R' are 0.78 (*s.e.*=0.39) and -3.61 (*s.e.*=0.36). That is, relative to Independents, Democrats expect a 0.78-month longer wait for a COVID vaccine while Republicans expect a 3.61-month shorter wait.

We then estimate the order effect on respondents' forecast of the election outcome. Questions 15 is about "Who do you think will win the upcoming U.S. Presidential election?" We run Regression (5) using the responses to this question as the dependent variable. That is, Y_i is 1 through 5 if respondent i's answer is "Definitely Joe Biden," "Probably Joe Biden," "Hard to say," "Probably Donald Trump," and "Definitely Donald Trump," respectively. Column (4) shows that the coefficient on *Reverse* is 0.14 (*s.e.*=0.03), i.e., the reverse-order list (with Trump at the top) leads to a higher predicted chance of a Trump win. The coefficient on *Severe* is -0.10 (*s.e.*=0.04), i.e., respondents who personally know someone seriously ill from COVID are less likely to predict a Trump win. Interestingly, the coefficients of D' and R' are -0.48 (*s.e.*=0.04) and 0.55 (*s.e.*=0.04), respectively. That is, relative Independents, both Democrats and Republicans predict that their own side has a higher chance to win.

In summary, in both cases, the response-order effect is sizeable and comparable to the effects of COVID experiences and political leanings.

4. Discontinuity in Self-Directed Investments

The estimated effects in Panel A of Table 7 capture the average treatment effect across all reverse-order users. However, the treatment effect should manifest only when the response order leads to a change in a user's risk category. If the response order alters the user's risk score without crossing a threshold into a different category, no change in investment behavior is expected. In other words, the response-order effect operates primarily through category shifts—when a user's score is pushed across a threshold into a new risk category. Therefore, even a small change in a user's risk score that crosses a category threshold can result in a substantial change in their self-directed investment. This suggests that users with risk scores within a narrow bandwidth around a cutoff should display markedly different investment behaviors depending on whether their scores are above or below the threshold. To test this discontinuity, we focus on users whose risk scores are close to the cutoff between the C4 and C5 categories—the risk score of 69—and estimate the following regression:

$$R5_i = \alpha + \beta 1_{Score_i \geq 69} + \gamma X_i + \varepsilon_i,$$

where $1_{Score_i \geq 69}$ is a dummy variable that is 1 if respondent i's score, $Score_i$, is greater than or equal to 69, and 0 otherwise. X_i denotes a vector of control variables.

Panel A of Table A4 presents the regression results. In column (1), the bandwidth is 3—that is, the regression includes users with risk scores between 66 and 71 (i.e., within 3 points of the cutoff at 69). The estimate of β is 1.42 (*s.e.* = 3.08), indicating that users whose scores are just above the threshold (and are therefore classified as C5) invest RMB 1420 more R5 products than users just below the cutoff. Columns (2) and (3) narrow the bandwidth to 2 and 1, respectively, and the estimates of β remain qualitatively similar. In summary, there is a substantial discontinuity in the investment in R5 products at the score threshold of 69, the boundary between risk categories C4 and C5.

We conduct two placebo tests. Panel B reports the results using a fictitious cutoff point of 66. In this case, all users in the sample fall within the C4 category. As expected, the estimates of β are close to zero with small standard errors, indicating that users just above and below this artificial threshold allocate similar amounts to R5 products. Likewise, Panel C presents results based on a fictitious cutoff point of 72, where all users are in the C5 category. Again, the estimates show no meaningful difference in the investment in R5 products across the artificial threshold, consistent with the absence of a risk category change.

We note that the estimated effect in Panel A should not be interpreted as causal. In principle, users on either side of the cutoff may have self-selected into their positions, and such selection could partly explain the large difference in portfolio allocations. Nonetheless, the evidence lends support to the compliance mechanism by demonstrating a strong association between risk category and investment behavior—even when the underlying risk score changes only marginally.

Table A1. Probit Regressions

This table is based on the AAII survey data. Panel A reports the results of the following Probit regression for *k*=1, 2, 3, 4, 5:

$$\Pr(Choice_{ij} = n) = \Phi(\alpha + \beta Reverse_i \times Sub_j + \delta Reverse_i + \gamma Sub_j + \varepsilon_{ij}),$$

where $Choice_{ij}$ is respondent i's choice for Question *j*; *k*=1, 2, 3, 4, 5 refer to the first, second, third, penultimate, and last choices, respectively; $\Phi(\cdot)$ is the CDF of a standard normal distribution; $Reverse_i$ is a dummy variable that is 1 if respondent *i* faces a reverse-order survey and 0 otherwise; and Sub_j is a dummy variable that is 1 if Question *j* is a subjective question and 0 otherwise. The classification of subjective and objective questions is reported in Table 1. For *k*=2 and *k*=4, the data from survey questions with 2 or 3 responses are excluded. For *k*=3, the regression is based on the data from only the Questions with 5 responses. Survey-question fixed effects are included in all specifications. Panel B reports the implied marginal effect of *Reverse*. Standard errors, in parentheses, are based on standard errors that are clustered by respondent. ***, **, and * indicate significance levels of 1%, 5%, and 10%, respectively.

Panel A. Probit regressions

	First (*k*=1)	Second (*k*=2)	Middle (*k*=3)	Penultimate (*k*=4)	Last (*k*=5)
Reverse	-0.00	-0.08	-0.03	-0.02	0.04
	(0.03)	(0.07)	(0.05)	(0.03)	(0.02)
Reverse × Sub	-0.11***	-0.07	-0.04	0.08**	0.13***
	(0.04)	(0.07)	(0.05)	(0.03)	(0.03)
Sub	1.33***	2.05***	0.81***	-0.82	-1.83***
	(0.10)	(0.07)	(0.05)	(0.04)	(0.05)
Constant	-2.64***	-2.04***	-1.38***	-0.45***	0.16***
	(0.10)	(-33.06)	(0.04	(0.03)	(0.03)
Question FEs	Yes	Yes	Yes	Yes	Yes
N	49779	40454	34195	40454	49779
Pseudo R-squared	0.352	0.190	0.162	0.074	0.297

Panel B. Marginal effects of *Reverse* (%)

	First (*k*=1)	Second (*k*=2)	Middle (*k*=3)	Penultimate (*k*=4)	Last (*k*=5)
Subjective questions	-1.95***	-3.66***	-1.78***	1.57***	3.66***
	(0.44)	(0.45)	(0.55)	(0.44)	(0.42)
Objective questions	-0.03	-0.12	-3.89	-0.71	0.88
	(0.24)	(0.10)	(0.69)	(1.07)	(0.56)

Table A2. Variations across Decision Time

This table reports the results from the regression that expands the regression in Panel A of Table 2 with a triple interaction term $Reverse_i \times Sub_j \times Fast_i$, where $Fast_i$ is a dummy variable that is 1 if respondent i's survey completion time is among the shortest α (=1/2, 1/3, or 1/4) of the respondents' completion times. Survey-question fixed effects are included in all specifications. T-statistics, in parentheses, are based on standard errors that are clustered by respondent. ***, **, and * indicate significance levels of 1%, 5%, and 10%, respectively.

Dep. var. = $Score$	(1) α=1/2	(2) α=1/3	(3) α=1/4
Reverse × Sub × Fast	-3.45***	-3.48***	-2.59**
	(1.15)	(1.22)	(1.32)
Reverse × Sub	5.48***	4.88***	4.37***
	(0.81)	(0.71)	(0.67)
Sub × Fast	0.74	0.76	0.18
	(0.84)	(0.87)	(0.94)
Reverse × Fast	2.91***	3.17***	2.11*
	(0.93)	(0.99)	(1.09)
Reverse	-1.10	-0.68	-0.15
	(0.65)	(0.57)	(0.54)
Sub	-12.73***	-12.62***	-12.41***
	(0.81)	(0.76)	(0.74)
Fast	-0.81	-0.86	-0.41
	(0.67)	(0.71)	(0.78)
Constant	50.21***	50.10***	49.91***
	(0.64)	(0.60)	(0.58)
Question FEs	Yes	Yes	Yes
N	52908	52908	52908
R-squared	0.318	0.318	0.318

Table A3. Quantify the Order Effect on Expectations

This table extends the analysis in Table 3, and reports the results from the following regression:

$$Y_i = \alpha + \beta_1 Reverse_i + \beta_2 Mild_i + \beta_3 Severe_i + \beta_4 D_i' + \beta_5 R_i' + \varepsilon_i,$$

where Y_i is respondent i's answer to a survey question; $Reverse_i$ is a dummy variable that is 1 if respondent i's survey is in the reverse order and 0 otherwise; $Mild_i$ is a dummy variable that is 1 if respondent i answered "Yes" to Question 9 ("Do you directly know someone who has been infected by COVID-19 with mild or no symptoms?") and "No" to Question 10 ("Do you directly know someone who has been severely ill due to COVID-19 infection?"); $Severe_i$ is a dummy variable that is 1 if respondent *i* answered "Yes" to Question 10; D_i' is the residual from regressing D_i on $Reverse_i$; R_i' is the residual from regressing R_i on $Reverse_i$; D_i is a dummy variable that is 1 if respondent *i* answered "Definitely Democrat" or "Usually Democrat" to Question 13 ("What is your political leaning?"); and R_i is a dummy variable that is 1 if respondent i answered "Definitely Republican" or "Usually Republican" to Question 13. In column (1), the survey question is "When do you expect a COVID-19 vaccine to become widely available?" For these two regressions, Y_i is 1.5, 4.5, 9, 24, and 48 months if respondent *i* selected choices 1 through 5, respectively. In column (2), the survey question is "Who do you think will win the upcoming U.S. Presidential election?" Y_i is 1 through 5 if respondent i's answer is "Definitely Joe Biden," "Probably Joe Biden," "Hard to say," "Probably Donald Trump," and "Definitely Donald Trump," respectively. All regressions include age, gender, education, and residence fixed effects. Standard errors are reported in parentheses. ***, **, and * indicate significance levels of 1%, 5%, and 10%, respectively.

	Vaccine development (m)	Election outcome
	(1)	(2)
Reverse	1.89***	0.14***
	(0.30)	(0.03)
Mild	-0.63*	-0.002
	(0.35)	(0.04)
Severe	-0.05	-0.10**
	(0.36)	(0.04)
D'	0.78**	-0.48***
	(0.39)	(0.04)
R'	-3.61***	0.55***
	(0.36)	(0.04)
Constant	18.70***	3.40***
	(3.36)	(0.37)
N	3,026	3,007
R-squared	0.091	0.226

Table A4. Discontinuity in Self-Directed Investments

This table, based on the data from the RCT in Section 3, shows the discrete jump in investors' purchases of R5 products at the risk score 69. Specifically, we run the following regression:

$$R5_i = \alpha + \beta 1_{Score_i \geq K} + X_i + \varepsilon_i,$$

where $R5_i$ is user i's purchase of R5 products and is denominated in RMB 1000, $1_{Score_i \geq K}$ is a dummy variable that is 1 if respondent i's risk assessment score, $Score_i$, is greater than or equal to K, and 0 otherwise. Control variables includes *Male, Age, LnInitialDirect, NewUser*, all of which are defined in Table 4. The regressions are based on the users whose scores are within n points from the cutoff point K. For example, in column "$n = 3$", investors' final scores are within 3 points of the cutoff point K (i.e., $K - 3 \leq Score_i \leq K + 2$). The table only reports the estimated coefficients on $1_{Score_i \geq K}$. $K = 69,\ 66$, and 72 in Panels A, B, and C, respectively. Robust standard errors are reported in parentheses. ***, **, and * indicate significance levels of 1%, 5%, and 10%, respectively.

Panel A. $K = 69$

	$R5$		
	$n = 3$	$n = 2$	$n = 1$
$1_{Score_i \geq K}$	1.42***	1.60***	0.84*
	(0.35)	(0.56)	(0.49)
Controls	Y	Y	Y
N	1,525	950	505
R-squared	0.10	0.09	0.11

Panel B. $K = 66$

	$R5$		
	$n = 3$	$n = 2$	$n = 1$
$1_{Score_i \geq K}$	0.14	0.07	-0.04
	(0.11)	(0.06)	(0.13)
Controls	Y	Y	Y
N	1,193	832	402
R-squared	0.00	0.02	0.04

Panel C. $K = 72$

	$R5$		
	$n = 3$	$n = 2$	$n = 1$
$1_{Score_i \geq K}$	0.00	-0.21	-0.06
	(0.48)	(0.57)	(0.77)
Controls	Y	Y	Y
N	1,534	1,157	500
R-squared	0.17	0.18	0.24