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: August 27, 2025

*We are grateful for the helpful comments from Michael Weber, Giulio Trigilia, Arthur Yan, and seminar participants at the University of Rochester and Arthur Yan for excellent research assistance. We thank Liang Zhang and LCMF 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 client 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 affluent and 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. These findings raise new ethical concerns in the FinTech era and carry important implications for decision-making theory.

JEL Classification Numbers: D1, D9, G10, G2A, G23, G40, G41.

Keywords: FinTech, Robo-advisor, Order effect, RCT, Survey, Interface, Ethical concerns, Expectations.

1. Introduction

Financial Technology (FinTech) has fundamentally transformed the financial service industry in recent decades. Many studies have examined its profound implications for consumers (Allen, Gu, and Jagtiani, 2021). An often-neglected aspect of the transformation is that, especially since the wide adoption of smartphones, FinTech has drastically changed the decision environment for investors. Instead of facing a human financial advisor, for example, an investor nowadays often makes decisions in front of a screen. In this paper, we argue that this change of environment may affect investors decisions in subtle and often hidden ways.

For instance, FinTech has led to rapid growth of the robo-advisory market, which has over $2 trillion of assets under management in 2025. This development has the potential to democratize financial advice, providing large-scale services to investors with modest net worth (D'Acunto and Rossi, 2025). However, our paper shows that this drastic change in investment practice also brings an innocuous psychological bias into the vast and fast-growing market.

Specifically, a robo-advisor provides individualized recommendations based on its clients' preferences, which are typically inferred from their responses to a survey. As a result, any biases introduced by the survey process can distort the robo-advisor's assessments and, in turn, its recommendations. We focus on the response-order effect, a bias that is well studied in psychology: the order of listed choices in a survey question may affect respondents' answers (Krosnick and Presser, 2010). We conjecture that this effect also applies to the survey questions of robo-advisors. In the survey question on risk tolerance, for example, presenting the most risk-tolerant response at the top or bottom of the list would elicit different answers from a given respondent. By using survey responses as inputs for constructing investment recommendations, a robo-advisor—perhaps unknowingly—brings the response-order bias onto its clients' screens. We conjecture that this bias would not be confined to the virtual world but would extend to influence clients' investment decisions.

We evaluate these conjectures in two studies. First, we field a nationwide survey of affluent and experienced U.S. investors and find a substantial order effect in their responses to risk-tolerance questions commonly employed by robo-advisors. Second, motivated by the survey evidence, we collaborate with a leading robo-advisor in China to conduct a preregistered randomized controlled trial (RCT) and provide causal evidence that the response order

systematically alters both the robo-advisor's risk assessments and its clients' investment decisions.

Specifically, the first study investigates whether the order effect applies to robo-advisors' survey questions on investment preferences. We conducted a nationwide survey of over 3,000 affluent and experienced U.S. investors from October 27 to November 2, 2020—during the height of the COVID pandemic and right before the U.S. presidential election. Our investment preference questions mirror those used by robo-advisors in the U.S. Our survey also includes "objective" questions with straightforward answers (such as age, and gender) as well as questions about macroeconomic expectations and current events, including the election and the COVID pandemic.

To assess the response-order effect, we administered two versions of the survey, differing only in that the response options for each question were listed in opposite orders. For convenience, we refer to them as the "original-order" and "reverse-order" surveys. 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 answers at or near the top of the lists, leading to a systematic bias in survey results. For example, consider the following risk tolerance question: "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	*Reverse order*
−10% (−$10,000)	−40% (−$40,000)
−20% (−$20,000)	−30% (−$30,000)
−30% (−$30,000)	−20% (−$20,000)
−40% (−$40,000)	−10% (−$10,000)

The tendency to select options near the top of the list implies that, for this question, respondents shown the reverse-order survey tend to report higher risk tolerance. Indeed, approximately 28% of respondents to the original-order survey (439 out of 1,550) selected the most risk tolerant answer—"−40% (−$40,000)"—which is at the bottom of the list. In contrast, when the same option appeared at the top of the list in the reverse-order survey, approximately 39% of respondents (611 out of 1,576) selected it. This difference is highly statistically significant ($p = 6.29 \times 10^{-10}$).

More broadly, we find a quantitatively large response-order effect for a variety of survey questions, including questions on economic forecasting, investment preferences, public health, and politics. A common feature is that these are subjective questions, for which respondents are typically uncertain about their answers. By contrast, the response-order effect vanishes for

objective questions with clear, factual answers, such as those on demographics.

In the second study, we collaborate with Licaimofang (LCMF hereafter), a leading robo-advisor in China, to conduct a preregistered RCT to examine whether the order effect persists if large financial stakes are involved, and, more importantly, whether it has a meaningful effect on investors' financial decisions.

When a new client opens an account at LCMF, he is required to complete an 11-question survey on his financial literacy, investment experience, income, risk tolerance, liabilities, and other relevant factors. For each survey question, the client is assigned a numerical score according to his answer, and the sum of the scores across all 11 questions determines the client's overall *risk score*, which, by design, ranges from 18 to 89. Clients are classified into five *risk categories* according to their risk scores: C1 (18–19), C2 (20–28), C3 (29–38), C4 (39–68), and C5 (69–89).

During the trail, LCMF administered two versions of the survey, differing only in that the response options for each question were listed in opposite orders. In one version of the survey (the original-order survey), the response that indicates the lowest score is listed at the top for each survey question. In the other version (the reverse-order survey), the response indicating the highest score is listed at the top. Each client was randomly assigned, with equal probability, one of the two versions. This randomization is implemented based on the client's account number. Clients with odd account numbers received the original-order survey, while those with even account numbers received the reverse-order version.

The treatment for new clients was administered from October 25 to December 27, 2024. For existing clients (i.e., those who opened their accounts before October 25, 2024), the treatment was rolled out in two waves. On October 25, 2024, LCMF 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. Our data was retrieved on April 1, 2025, and the overall sample comprises 8,245 users (1,323 new and 6,922 existing users). As expected, the original-order and reverse-order groups are balanced in terms of sample size and user characteristics, including demographics and investment holdings before the trial.

To estimate the order effect, we regress a user's risk score on $Reverse$, a dummy variable equal to one if the user's account number is odd (and hence the user received the reverse-order survey), and zero otherwise. For our overall sample, the estimated coefficient on $Reverse$ is 2.83 ($s.e.$ = 0.27), indicating that reversing the response order increases the user's risk score by an

average of 2.83. We also estimate this regression separately for new and existing users and find similar results.

We also perform a placebo test. Specifically, we obtained records for 4,464 users who opened accounts between October 25, 2024 and March 31, 2025. Because the treatment for new users concluded on December 27, 2024, only 1,323 of these users participated in the RCT. The remaining 3,141 users, who opened accounts after the RCT period, form the sample for our placebo test. For this sample, we regress users' risk scores on $Reverse$. Since all of these users received the same version of the survey, there should be no difference in risk scores between those with odd account numbers ($Reverse = 1$) and those with even account numbers ($Reverse = 0$). Indeed, the estimated coefficient on $Reverse$ is -0.17 with a standard error of 0.63, indicating no meaningful difference.

The response order also has a significant effect on users' assigned risk categories. For example, although the original-order and reverse-order groups are similar in size—4,091 and 4,154 users, respectively—the number of users in category C5 differs markedly: 1,973 users in the original-order group versus 2,341 in the reverse-order group—an increase of 19%. The difference is highly statistically significant ($p = 1.49 \times 10^{-13}$).

So far, our evidence demonstrates that, despite the large financial stakes, the response order continues to exert a significant influence on the robo-advisor's assessments of users' risk scores and categories. Do these effects on risk assessments translate into changes in users' actual investment decisions? There are at least two reasons to conjecture that they do.

First, the order effect on risk assessments may influence users' investment decisions through a nudging mechanism. Since the risk assessments are used as an input for the robo-advisor's investment recommendations, the order effect becomes embedded in those recommendations, potentially nudging users to act.

Second, the order effect on risk assessments may also influence users' investment decisions through a compliance mechanism. LCMF classifies the financial products on its platform into five risk levels, R1 through R5, with R5 being the riskiest. For regulatory considerations, a user is permitted to purchase only products whose risk ratings are 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, he would be reminded that the product's risk rating exceeds his assigned 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 financial decisions, we focus on two primary types of investments at LCMF: robo-advised and self-directed. LCMF 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 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 is free to choose 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 strategy'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 strategy's recommendations to make purchases as long as no R5 products are involved. If the recommended transaction includes purchasing an R5 product, however, 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 a C5 category. This impediment may limit the user's investment size.

The second type of investment in our analysis is self-directed. It includes mutual funds, money market funds, and wealth management products offered by third parties, with LCMF serving only as the intermediary. Similar to investing through an online brokerage account, the user must actively make all decisions related to identifying and trading these products. In contrast, the previously mentioned robo-advised strategies are designed by LCMF: users can only accept or reject the recommended transactions and cannot initiate alternatives. Because self-directed investing requires more active decision-making by users, the order effect is expected to be weaker for these investments than for robo-advised ones.

To estimate the order effect on robo-advised investments along the intensive margin, we obtain the sample of users who made purchases of robo-advised products during our sample period.

This sample includes 2,225 users, and their total purchase during our sample period is RMB 72,497,175. We first regress LnA_i, the logarithm of user i's investment in the Aggressive strategy (in RMB) plus one, on $Reverse_i$. The estimated coefficient on $Reverse$ is 0.41 ($s.e.$ = 0.20), suggesting that users exposed to the reverse-order survey invest more in the Aggressive strategy. On average, a reverse-order user invests RMB3,912 more in this strategy. We then regress αA_i, user i's investment in the Aggressive strategy normalized by his total robo-advised investment, on $Reverse_i$. The estimate suggests that on average, reverse-order users allocate 3.50% ($s.e.$=2.02%) more to the Aggressive strategy than original-order users. Finally, consistent with our conjecture that the order effect operates through users' risk categories, we find that once we include category fixed effects, the coefficient of $Reverse$ becomes statistically insignificant while the coefficients of the category dummies are highly significant.

To estimate the effect along the extensive margin, we construct a dummy variable D_Robo_i equal to one if user i invested in any robo-advised product during the sample period, and zero otherwise. We then regress it on $Reverse$ using our overall sample. The estimated coefficient of $Reverse$ is -0.03 ($s.e.$ = 0.01). This implies that reverse-order users are three percentage points less likely to invest in robo-advised products. This result is unlikely to be driven by the compliance mechanism, as reverse-order users tend to have higher risk profiles and therefore face fewer investment restrictions. One possible explanation is that the nudging effect backfired: reverse-order users received inflated risk profiles, which nudged them toward the high-risk strategy. Some users, however, ultimately chose not to invest. Consistent with this explanation, we find that the effect is driven mostly by users in the highest risk category, namely C5. Specifically, reverse-order users were significantly more likely to be classified into the C5 category, and those reverse-order C5 users were substantially less likely to invest than their original-order counterparts.

We also conduct a similar analysis for self-directed investments and find that most results are statistically insignificant. This lack of statistical significance may be due to the smaller sample size for self-directed investments. Moreover, users must play a much more active role in the decision process for self-directed investments, which likely attenuates the impact of response order.

Our evidence is based on a robo-advisor in China. What is its implication on the U.S. robo-advisory market? We expect the order effect to be even stronger because investors there play a more passive role in the decision process. In the U.S., once a user completes the survey and sets up an account, investment decisions are executed automatically based on the risk assessment

results. As a result, the order effect on risk assessments translates more directly into actual investment decisions. By contrast, investors at LCMF play a more active role even in their robo-advised investments, as each recommended transaction must be individually authorized.

Our findings raise a potential ethical concern. For example, a robo-advisor may benefit by steering clients toward financial products that are not optimal from the clients' perspective. Since responses to investor preference surveys can be manipulated, a robo-advisor can potentially guide investors to purchase those products while maintaining the appearance that investors "prefer" them.

More broadly, our evidence has implications on constructs such as "risk aversion," which is typically treated as a primitive. Our findings suggest that such constructs are not fully determined in the sense that their estimated values, whether based on survey or even revealed preference approaches, systematically vary with external factors.

Our paper contributes to the growing economics literature that relies on surveys. This literature primarily uses surveys to measure subjective beliefs (Manski, 2004) to study various topics in belief formation, such as experience effects (Malmendier and Nagel, 2011), partisan bias (Coibion, Gorodnichenko, Weber, 2020), personality traits (Jiang, Peng, and Yan, 2023), 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, Chinco, Hartzmark, and Sussman, 2020; Choi and Robertson, 2020; and Giglio, Maggiori, Stroebel, and Utkus, 2020, 2021; De la O and Myers, 2021; 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. Moreover, it shows that this bias affects investment decisions in a large and growing financial market.

Our paper also adds to the emerging literature on how the decision environment shapes financial investments. For example, Liao et al. (2021) document how user interfaces affect 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. 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, 2025). Our paper complements this literature by showing that it can also introduce biases into investors' financial decisions.

The response-order effect has been extensively studied in psychology (Krosnick and

Presser, 2010). Our paper extends this literature by examining its implications on economic variables such as expectations, preferences, and, more importantly, its effects on real-world financial decisions. More broadly, our study contributes to the broad literature on the effects of order. For instance, Hong, Kremer, Kubik, Mei, and Moses (2015) show that auction order affects total revenue. Feenberg, Ganguli, Gaule, and Gruber (2017) find that NBER working papers listed at the top or bottom of announcement emails receive more attention. Order effects have also been documented in citations (Huang, 2015; Brogaard, Engelberg, Eswar, and Van Wesep, 2020), school admissions (Jurajda and Munich, 2010), and election outcomes (Chen, Simonovits, Krosnick, and Pasek, 2014).

The rest of the paper is organized as follows. Section 2 reports a study based on a survey of U.S. investors. Section 3 presents a preregistered 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.

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. 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 response options 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 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% state that they can tolerate a 30% 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 opinions on the pandemic. Question 10 is "Do you believe that wearing a mask reduces *your* risk of COVID-19 infection?" Around 84% chose "Probably yes" or "Definitely yes"; only 3% chose "Definitely no." Responses to Question 11 indicate that the respondents' median forecast is that a COVID-19 vaccine will be widely available in 6 to 12 months. The responses to Question 12 suggest that 84% of respondents would "Definitely" or "Probably" take the vaccine.

Questions 13 through 15 are about political leanings: 42% selected "Definitely Republican" or "Usually Republican" and 30% selected "Definitely Democrat" or "Usually Democrat." Around 53% would "Definitely" or "Likely" vote for Biden while 39% would "Definitely" or "Likely" vote for Trump. Around 56% predict that Biden will "Definitely win" or is "Likely to win" and 17% predict that Trump will "Definitely win" or is "Likely to win."

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 at or 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 at 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.

Moreover, we expect the response-order effect to vary across survey questions. In particular, we conjecture that the effect should be stronger if respondents are more uncertain about the value of the surveyed variable. 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 much 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 respondents 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 for each question. Note that the domain of $Answer_{ij}$ varies across survey questions. For example, the value of $Answer$ is either 1 or 2 in Question 2, but ranges from 1 to 6 in Question 4. Hence, we standardize $Answer$ to make it comparable across questions: the standardized answer of respondent *i*'s choice in Question *j* is given by

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

where $Answer_{ij}$ is defined in (1) and *n* is the number of responses for Question *j*. That is, for the responses on the original-order list, we set the first response to 0 and the last to 100. When there are more than two responses, we assign values to those responses such that they divide the interval [0,100] evenly. Suppose, for example, a survey question has five responses. The value of $SAnswer$ is 0, 25, 50, 75, and 100 when the value of $Answer$ is 1 through 5, respectively.

We then run the following panel regression:

$$SAnswer_{ij} = \alpha + \beta\, Reverse_i + \varepsilon_{ij}, \quad (3)$$

where $SAnswer_{ij}$ is respondent *i*'s standardized answer to Question *j* and is defined in (2) and $Reverse_i$ 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 of *Reverse* is 2.54 (*s.e.*=0.29). That is, on average, reversing the response order increases the standardized answer, $SAnswer$, by 2.54. Hence, the average response-order effect across all questions is 2.54% of the difference between the top and bottom responses. In column (2), we include survey-question fixed effects. The coefficient of *Reverse* remains similar

in both magnitude and statistical significance.

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

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

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). The response-order effect is limited to subjective questions. For example, column (4) shows that the coefficient of *Reverse* is merely 0.36 (*s.e.*=0.47), i.e., the response-order effect is statistically insignificant for objective questions. 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 estimate of β is statistically indistinguishable from zero. The exception is Question 3, on investment experiences, for which the coefficient of *Reverse* 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 of $Reverse$ is positive and highly statistically significant.

2.3 Quantify the Effect on Survey-based Expectations

In this section, we assess the magnitude of 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. A large literature has shown that expectations are shaped by people's experience (Malmendier and Nagel, 2011). Moreover, 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 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 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 (Meeuwis, Parker, Schoar, and Simester, 2019; Kempf and Tsoutsoura, 2020). 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 for each survey question:

$$Y_i = \alpha + \beta_1 Reverse_i + \beta_2 Mild_i + \beta_3 Severe_i + \beta_4 D'_i + \beta_5 R'_i + \varepsilon_i, \qquad (5)$$

where Y_i is respondent i's expectation. The coefficient β_1 captures the response order effect while the coefficients β_2 through β_5 reflect the effects of COVID experiences and political leanings.

Macroeconomic expectations. We first run regression (5) for Questions 6 and 7, the forecast of the GDP growth and the S&P 500 index return over the next 12 months, respectively. 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 of *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 of *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 of *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.

COVID pandemic. We run Regression (5) with the data from Question 11 ("When do you expect a COVID-19 vaccine to become widely available?"). 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 of *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.

The presidential election outcome. Question 15 is "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 of *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 of *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.

2.4 Implications

First, the above evidence shows that the response-order effect is substantial for survey-based expectations on the macroeconomy, as well as on variables 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. As expected, COVID experiences and political leanings have substantial effects public opinions on issues in all three domains. 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.

Second, although our survey is conducted only once and hence our direct evidence is on the *levels* rather than *changes* of the economic variables, 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; 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.

Last but not least, if survey data are used as inputs, the response-order effect would spread over to other domains and have unexpected consequences. For example, 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, 2025). 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.

Note that in the survey above, our respondents 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? Would the order effect influence investors' financial decisions? We attempt to address these questions next.

3. A Pre-registered RCT

Motivated by the above evidence from the AAII survey, we collaborate with LCMF—a leading robo-advisor in China—to conduct an RCT to examine the 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, LCMF is the first robo-advisor 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., LCMF assesses clients' investment preferences through a survey and recommends financial products accordingly.

The risk assessment process operates as follows: when new clients open accounts on the platform, 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. 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.

LCMF 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 LCMF serving as the intermediary. LCMF also offers portfolios composed of these third-party products, allowing clients to invest directly in portfolios rather than individual components.

These products are classified into five risk categories, R1 through R5. For compliance purposes, users may only purchase products whose risk categories are at or below their assigned risk categories. 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.

The main page of the LCMF 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 LCMF, 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-advisory services in the U.S. market. LCMF 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, 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 may 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 investment strategies are designed by LCMF. 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 interesting to note that users in the U.S. robo-advisory market play an even more passive role. Typically, once a user finishes his risk assessment survey and sets up his account with a robo-advisor in the U.S., the rest of the investment decisions are automatic and are 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 (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-bearing capacity. Each user was randomly assigned to one of the two versions with equal probability.

The timeline of the RCT is as follows. 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 odd and the reverse-order survey if it was even. The treatment for existing users (i.e., those who opened their accounts before October 25, 2024) was rolled out

in two waves. On October 25, 2024, 5,000 existing users with holdings of robo-advised investments 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, he would receive the original-order survey if his account number was odd 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, the overall sample comprises 8,245 users. The vast majority fall into categories C4 and C5, with only 274 users classified into categories C1 through C3. The second and third rows indicate that 4,091 users received the original-order survey, while 4,154 received the reverse-order version. Finally, the sample consists of 1,323 new users and 6,922 existing users.

Note that our RCT for new users ended on December 27, 2024. When retrieving data on April 1, 2025, we also obtained records for 3,141 new users who opened accounts between December 28, 2024, and March 31, 2025. The response order in their surveys was not randomized, and hence we refer to this group as the "placebo sample."

Panel C shows that the average risk score in our overall sample is 67. The average age is 46 years, and approximately 60% of clients are male. During the sample period, 1,739 users made self-directed investments, with an average purchase of RMB 61,234, and 2,225 users made robo-advised investments, with an average purchase of RMB 32,682. Before the launch of the RCT, existing users in our sample had an average total asset holding of RMB 308,111. Finally, Panel D shows that the treatment and control groups are comparable across all observed characteristics: gender, age, new user status, and initial asset holdings.

3.3 Order Effect on Risk Assessments

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

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

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 LCMF.

The regression results, reported in Panel A of Table 5, indicate a strong order effect. In column (1), the coefficient on $Reverse$ is 2.96 (s.e. = 0.28), suggesting that reversing the response order increases the average risk score by 2.96. Column (2) adds control variables, and the coefficient on $Reverse$ remains mostly unchanged. We also estimate the regression separately for the subsamples of new and existing users. The results, reported in columns (3) through (6), are similar.

Does the response order influence users' risk categories? Panel B of Table 4 provides a clear answer. In the control group, where users received the original-order survey, 166 users are classified in category C1. In contrast, only 85 users fall into category C1 in the treatment group. The p-value for the null hypothesis that the probability of being in C1 is the same across the two groups is 1.06×10^{-7}. The response order also significantly affects the distribution between categories C4 and C5. In the control group, the number of users in C4 and C5 is roughly equal (1,941 vs. 1,973). In the treatment group, however, category C5 contains 2,341 users—36% more than the 1,716 users in category C4. The p-value for the null hypothesis that the likelihood of being in C5 is the same across the two groups is 1.49×10^{-13}. In short, reversing the response order substantially increases the number of users in the highest risk category (C5) while reducing 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, and report the results in Panel B. Columns (1) and (2) use the overall sample. In column (1), for example, the coefficient on $Reverse$ is 0.14 (s.e. = 0.02), indicating that—consistent with the results based on risk scores—reversing the response order increases a user's risk category. Columns (3) through (6) present similar findings for the subsamples of new and existing users.

Finally, we conduct a placebo test using the sample of new users who opened accounts after December 27, 2024, when our RCT for new users concluded. As noted in the previous subsection, our placebo sample includes 3,141 new users who opened accounts between December 28, 2024, and March 31, 2025. These users answered the same set of survey questions, but the response order was no longer randomized. We run regression (6) on this sample as a placebo test. Specifically, the dummy variable $Reverse_i$ is one if user i's account number is an odd number and is 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 C, the coefficient on *Reverse* is statistically indistinguishable from zero in all columns. This null result is not due to a lack of statistical power: the standard errors in this placebo sample are smaller than those in the new-user subsample reported in Panels A and B.

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 mechanism. 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 finishes his risk assessment survey and sets up his account with a robo-advisor, the rest of the investment decisions are automatic and are based on the user's risk assessment results. 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 at LCMF. Hence, one would expect the order effect on investment decisions to be weaker in our analysis than in the U.S. robo-advisory market.

Second, the order effect on risk assessments may also influence users' investment decisions through a compliance mechanism. For regulatory reasons, a user at LCMF can only purchase assets whose risk rating is at or below his assigned risk category. For example, a user classified as C4 may purchase assets rated R1 through R4. If the user attempts to purchase an R5 asset, he would be reminded that its 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. 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 two types of investments: robo-advised and self-directed. As noted in Section 3.1, the latter requires more active involvement from users. Similar to the common practice in the U.S., the robo-advisor—LCMF—plays a significant role in robo-advised investments, whereas users must make all decisions in self-directed investments. 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 focus on the intensive margin—that is, it analyzes how the response order influences users' investment allocations, conditional on their participation. We then turn to the extensive margin, evaluating how the response order affects the likelihood that a user chooses to invest.

Intensive margin. 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 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 as C5 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 portfolio.

The operational procedure may also guide users toward investment strategies aligned with their assigned risk categories. Specifically, when a user opens an 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 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 LnA_i, the natural logarithm of one plus user i's total investment (in RMB) in the Aggressive strategy. The coefficient on *Reverse* is 0.41 (*s.e.* = 0.20), indicating that facing the reverse-order survey makes users invest more in the high-risk strategy. In monetary terms, reverse-order users invest RMB 4,985 more in the Aggressive strategy on average.

We next regression (6) with the dependent variable αA_i, user i's investment in the Aggressive strategy normalized by his total robo-advised investment. Column (3) shows that the coefficient on $Reverse$ is 3.50 (*s.e.* = 2.02), indicating that, on average, reverse-order users allocate 3.50 percentage points more to the high-risk strategy than original-order users.

Our hypothesis posits that these response-order effects operate through users' risk categories: the order in which survey options are presented influences the assigned risk category, which in turn steers the user toward certain products. This hypothesis implies that the estimated treatment effect of $Reverse$ should vanish once we control for users' risk categories. Consistent with this prediction, columns (2) and (4) of Table 6 show that including risk category fixed effects renders the coefficient on $Reverse$ statistically insignificant. Moreover, the risk category dummies themselves are highly statistically significant, supporting the view that risk category is the primary channel through which the response order affects robo-advised investments.

Extensive margin. We now examine whether the response order affects users' participation in robo-advised investments. Specifically, we run regression (6) based on our overall 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 –0.03 (*s.e.* = 0.01), indicating that reverse-

order users are 3 percentage points less likely to make robo-advised investment. We add control variables in Column (2), and the estimated treatment effect remains similar.

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, reverse-order users are nudged toward the high-risk strategy. However, many of them may have opted out of investing altogether. For example, within the C5 group, the number of reverse-order users (2,341) significantly exceeds that of original-order users (1,973). Yet, the number of those who actually made investments to any robo-advised products is only slightly higher among reverse-order users: 708 versus 692. In the C4 group, not only are there fewer reverse-order users, but their investment rate is also lower, resulting in a notably smaller number of investors—357 compared to 464 among original-order users.

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. In contrast, the evidence on the order effect for self-directed investments is inconclusive, possibly due to the smaller sample size and the more active role users play in those investment decisions.

3.6 Order Effects on Self-directed Investments

This section examines the order effects on self-directed investments. We first estimate the effect along the intensive margin. Since nearly 97% of users in our sample fall into categories C4 and C5, we estimate regression (6) based on the sample of users who made investments in R4 or R5 products and use $\alpha R5_i$ as the dependent variable, which user i's investment in R5 products divided by his total investment to R4 and R5 products. The results are reported in Panel A of Table 7. In column (1), the coefficient on $Reverse$ is 2.83 ($s.e.$=1.59). Column (2) shows that, after including control variables, the coefficient of $Reverse$ falls to 2.09 ($s.e.$ = 1.51).

In both specifications, the estimated treatment effects have the predicted sign, but are not

statistically significant.[1] This lack of statistical power may be related to the smaller sample size: the sample for self-directed investments consists of 1,406 users, compared to 2,225 for robo-advised investments. Moreover, although both the nudging and compliance mechanisms are expected to be present, in principle, they are expected to be less effective because users play a much more active role in self-directed investments.

To estimate the effect along the extensive margin, we estimate run regression (6) based on our overall 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 Panel B of Table 7. In both specifications, the coefficient on $Reverse$ is statistically insignificant. Given the large sample sizes, this lack of statistical significance cannot be attributed to the statistical power. The evidence 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 foundational concepts such as risk aversion and preference. These constructs 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 an external factor such as the response order not only affects the preference parameters implied by survey data, but also affects respondents' decisions. Hence, the

[1] Even if one takes the more "charitable" interpretation of our evidence based on the one-sided test, the treatment effect is only marginally significant. Specifically, suppose one takes the stance that given our prediction of $\beta > 0$, the appropriate null hypothesis should be $\beta \leq 0$. The *p*-value for rejecting this null hypothesis based on a one-sided *t*-test is 8.33% for the specification in column (2).

response order 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 a single subject. 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 bias from psychology labs into the two-trillion-dollar robo-advising market. Specifically, we conduct two studies to examine the effects of survey response order. The first is a nationwide survey of over 3,000 affluent and experienced U.S. investors. We document a pervasive response-order effect: respondents disproportionately select options appearing at or near the top of the list, introducing systematic bias into survey responses. This effect is particularly pronounced for questions related to risk preferences and expectations, but it vanishes for objective questions with clear answers.

Building on these findings, we partnered with a leading robo-advisor in China to conduct a preregistered RCT involving over 8,000 investors. We find that even when substantial financial stakes are involved, the order of survey responses significantly influences clients' risk profile—which directly affects the robo-advisor's product recommendations. More importantly, the response order also has a significant impact on clients' actual investment decisions. 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 high-risk strategy. In contrast, the evidence on the order effect for self-directed investments is inconclusive, possibly due to the more active role users play in those investment decisions.

While our evidence is based on an RCT in China, the response-order effect is expected to be even stronger in the U.S. context, where clients tend to play a more passive role. Once a U.S. client transfers funds to a robo-advisor, subsequent investment decisions are typically executed automatically based on the survey responses. As a result, any bias introduced by response order is more likely to directly translate into differences in clients' investment outcomes.

Our findings raise a potential ethical concern: a robo-advisor may steer users toward products that serve its own interests while maintaining the appearance that these products align with users' preferences. More broadly, our evidence suggests that constructs such as risk aversion are not fully determined in the sense that their estimated values, whether based on survey or revealed preference approaches, systematically vary with external factors.

Finally, our analysis focuses on short-term effects. A long-term randomized controlled trial would be needed to evaluate the persistence of these effects on investment decisions. Moreover, long-term studies are essential to achieving sufficient statistical power to reliably assess investment performance. We leave this to future research.

References

Allen, Franklin, Xian Gu, and Julapa Jagtiani, 2021, A Survey of Fintech Research and Policy Discussion, *Review of Corporate Finance* 1, 259–339.

Bertrand, Marianne and Sendhi Mullainathan, 2001, Do People Mean What They Say? Implications for Subjective Survey Data, *American Economic Review: Papers & Proceedings* 91, 67–72.

Bloom, Nicholas and John Van Reenen, 2010, New Approaches to Surveying Organizations, *American Economic Review: Papers & Proceedings* 100, 105–109.

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, 2020, On the Causal Effect of Fame on Citations, Working paper.

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 J. 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 M. Hartzmark, and Abigail Sussman, 2020, Necessary Evidence for a Risk Factor's Relevance, Working paper.

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

D'Acunto, Francesco and Alberto Rossi, 2025, Robo-Advice: Transforming Households into Rational Economic Agents, *Annual Review of Financial Economics*, Forthcoming.

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, forthcoming.

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.

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, 2020, Five facts about

beliefs and portfolios, *American Economic Review*, forthcoming.

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.

Gorodnichenko, Yuriy 2024, Subjective Expectations and Decisions, AFA Lecture, https://www.youtube.com/watch?v=PaIHuRqRhMo.

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, 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, 2020, Partisan Professionals: Evidence from Credit Rating Analysts, Working paper.

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, 2019, Belief Disagreement and Portfolio Choice, Working paper.

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

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. Order effect

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 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. Then, for each question *j*, we run the following regression:

$$Score_{ij} = \alpha_j + \beta Reverse_i + \varepsilon_{ij},$$

Panel B reports the estimated coefficient of $Reverse$, its standard error, and the number of observations for each regression. Standard errors, in parentheses, are clustered by respondent. ***, **, 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 order effect on expectations

This table 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), 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 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. For column (3), 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. For column (4), 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.

	Stock return (%)	GDP growth (%)	Vaccine dev. (m)	Election outcome
	(1)	(2)	(3)	(4)
Reverse	0.95***	0.45***	1.89***	0.14***
	(0.26)	(0.10)	(0.30)	(0.03)
Mild	-0.85***	-0.13	-0.63*	-0.002
	(0.31)	(0.12)	(0.35)	(0.04)
Severe	-0.90***	-0.26**	-0.05	-0.10**
	(0.31)	(0.12)	(0.36)	(0.04)
D'	-0.63*	-0.62***	0.78**	-0.48***
	(0.34)	(0.13)	(0.39)	(0.04)
R'	1.36***	0.86***	-3.61***	0.55***
	(0.31)	(0.12)	(0.36)	(0.04)
Constant	-1.60	-1.97*	18.70***	3.40***
	(2.91)	(1.15)	(3.36)	(0.37)
N	3,009	3,027	3,026	3,007
R-squared	0.035	0.071	0.091	0.226

Table 4. Summary statistics for the RCT

Panel A reports the composites of the target portfolios with High-, Medium-, and Low-risk strategies in columns (1) through (3), respectively. Panel B reports the number of users in each risk category, for the overall sample, the control and treatment groups, and the subsamples of new users and existing users, in rows 1 through 5, respectively. Row 6 reports the statistics for the placebo sample. Panel C reports the summary statistics for the main variables in this study. *Age* is the user's age in years. *Male* is a dummy variable that is one is the user is a male and zero otherwise. *Inverse* is a dummy variable that is one if the user's survey is in the reverse order and 0 otherwise. *Score1* is the user's total score when the user took the survey for the first time after our RCT started. *Score* is the user's latest total score from the survey at the Robo-advisor as of April 1, 2025. Duration is the time it took for the user to finish his survey at the Robo-advisor, denoted in seconds, during the latest take of the survey by April 1, 2025. *BuyDirect* is a user's total direct purchase, denominated in RMB, during the RCT period: from the time the user took the survey to April 1, 2025. *BuyRobo* is a user's total purchase in the user's robo-advisory account, denominated in RMB, during the RCT period. *InitialAll* is a user's total asset holding before entering the RCT, denominated in RMB. *InitialRobo* is a user's total asset holding in the robo-advisory account before entering the RCT, denominated in RMB. Panel D reports the average of the characteristics of the users who faced the reverse-order survey (the Treatment group) and those who faced the original-order survey (the Control group). *NewUser* is a dummy variable that is one is the user is a new user when our RCT was rolled out on October 25, 2024, and zero otherwise. *InitialAll* is a user's total asset holding before entering the RCT, denominated in RMB. *InitialRobo* is a user's total asset holding in the robo-advisory account before entering the RCT, denominated in RMB. Columns labeled by "N" report the number of observations. The last two columns report the difference and its standard error.

Panel A. Target portfolios

	(1)	(2)	(3)
	Hight risk	Medium risk	Low risk
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 and Risk Category Distribution

	C1	C2	C3	C4	C5	Total
Overall Sample	251	3	20	3657	4314	8245
Control	166	1	10	1941	1973	4091
Treatment	85	2	10	1716	2341	4154
New users	101	3	10	651	558	1323
Existing users	150	0	10	3006	3756	6922
Placebo Sample	256	0	28	1653	1204	3141

Panel C. Summary Statistics

	Overall sample			New users			Existing users			Placebo sample		
	Mean	S.D.	N	Mean	S.D.	*N*	Mean	S.D.	*N*	Mean	S.D.	*N*
Score	66.98	13.00	8245	62.16	16.78	1323	67.90	11.93	6922	62.13	12.63	3141
Reverse	0.50	0.50	8245	0.52	0.50	1323	0.50	0.50	6922	0.49	0.50	3141
Duration (seconds)	72.79	395.29	8245	63.92	331.93	1323	74.49	406.27	6922	53.67	66.96	3141
Age (years)	45.73	10.22	7404	38.29	10.00	482	46.24	10.03	6922	39.66	11.29	1115
Male	0.60	0.49	7404	0.44	0.50	482	0.61	0.49	6922	0.43	0.50	1115
BuyDirect (RMB)	61234	156080	1739	69640	155856	478	58047	156108	1261	28310	59189	1097
BuyRobo (RMB)	32682	78301	2225	35106	44402	12	32669	78451	2213	322487	23996	22
InitialAll (RMB)							308111	637111	6922			
InitialRobo (RMB)							145392	394740	6922			
InitialDirect (RMB)							91814	290528	6922			

Panel D. Treatment and Control Groups

	Treatment		Control			
	Mean	N	Mean	N	Difference	*s.e.*
Male	0.61	3724	0.59	3680	0.01	0.01
Age (year)	45.80	3724	45.66	3680	0.14	0.24
NewUser	0.16	4154	0.16	4091	0.01	0.01
InitialAll (RMB)	316096	3470	300084	3452	16013	15315
InitialRobo (RMB)	147736	3470	143036	3452	4700	9490
InitialDirect (RMB)	97083	3470	86518	3452	10566	6983

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 that is 1 if respondent i's survey is in the reverse order and 0 otherwise. In Panel A and the first two columns of Panel C, Y_i is user i's risk score. In Panel B and Columns (3) and (4) of Panel C, Y_i is 1 through 5 if user i's risk category is C1 through C5, respectively. All control variables are defined in Table 4. Standard errors are reported in parentheses. ***, **, and * indicate significance levels of 1%, 5%, and 10%, respectively.

Panel A. Risk Score

	Overall Sample		New Users		Existing Users	
	(1)	(2)	(3)	(4)	(5)	(6)
Reverse	2.96***	2.83***	3.55***	2.01**	2.91***	2.90***
	(0.28)	(0.27)	(0.92)	(0.99)	(0.28)	(0.27)
Male		4.57***		2.05**		5.11***
		(0.27)		(1.00)		(0.28)
Age		-0.04***		0.13***		-0.06***
		(0.01)		(0.05)		(0.01)
LnInitialAll		0.42***				1.50***
		(0.04)				(0.08)
Constant	65.49***	61.05***	60.33***	60.16***	66.44***	48.99***
	(0.20)	(0.71)	(0.66)	(2.11)	(0.20)	(1.04)
N	8,245	7,404	1,323	482	6,922	6,922
R-squared	0.01	0.07	0.01	0.03	0.01	0.11

Panel B. Risk Category

	Overall Sample		New User Sample		Existing User Sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Reverse	0.14***	0.13***	0.18***	0.08*	0.14***	0.14***
	(0.02)	(0.02)	(0.06)	(0.05)	(0.02)	(0.02)
Male		0.22***		0.08*		0.25***
		(0.02)		(0.05)		(0.02)
Age		-0.00***		0.00*		-0.00***
		(0.00)		(0.00)		(0.00)
LnInitialAll		0.02***				0.08***
		(0.00)				(0.00)
Constant	4.36***	4.24***	4.09***	4.32***	4.41***	3.51***
	(0.01)	(0.04)	(0.04)	(0.10)	(0.01)	(0.06)
N	8,245	7,404	1,323	482	6,922	6,922
R-squared	0.01	0.04	0.01	0.02	0.01	0.08

Panel C. Placebo sample

	Risk Score		Risk Category	
	(1)	(2)	(3)	(4)
Reverse	0.39	-0.17	0.02	-0.01
	(0.59)	(0.63)	(0.04)	(0.03)
Male		3.79***		0.13***
		(0.64)		(0.03)
Age		0.15***		0.00***
		(0.03)		(0.00)
Constant	61.05***	58.68***	4.12***	4.26***
	(0.42)	(1.25)	(0.03)	(0.06)
N	3,141	1,115	3,141	1,115
R-squared	0.00	0.05	0.00	0.02

Table 6. Order Effects on Robo-Advised Investments

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 that is 1 if respondent i's survey is in the reverse order and 0 otherwise. Panel A is based on the sample of users who made robo-advised investments during our sample period. In Columns (1) and (2), Y_i is LnA_i, the logarithm of 1 plus user i's total purchase of the *Aggressive* portfolio, in RMB. In Columns (3) and (4), Y_i is αA_i, user i's total purchase of the *Aggressive* portfolio, normalized by the user's total purchase of the *Aggressive*, *Balanced*, and *Conservative* portfolios. Panel B is based on the Overall Sample, and 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. All control variables are defined in Table 4. Standard errors are reported in parentheses. ***, **, and * indicate significance levels of 1%, 5%, and 10%, respectively.

Panel A. Intensive Margin

	LnA		αA (%)	
	(1)	(2)	(3)	(4)
Reverse	0.41**	0.22	3.50*	1.67
	(0.20)	(0.19)	(2.02)	(1.95)
Male	1.40***	0.88***	14.39***	9.27***
	(0.21)	(0.20)	(2.07)	(2.04)
Age	0.00	0.00	-0.01	-0.03
	(0.01)	(0.01)	(0.10)	(0.10)
LnInitialRobo	0.04	0.01	0.17	-0.09
	(0.03)	(0.03)	(0.33)	(0.32)
NewUser	2.76*	2.97**	20.79	22.93*
	(1.41)	(1.35)	(14.18)	(13.68)
Constant	2.16***	3.69	26.11***	47.09**
	(0.59)	(2.35)	(5.98)	(23.71)
Category FEs	No	Yes	No	Yes
Observations	2,225	2,225	2,225	2,225
R-squared	0.02	0.10	0.02	0.09

Panel B. Extensive Margin

	D_Robo_i	
	(1)	(2)
Reverse	-0.03***	-0.03***
	(0.01)	(0.01)
Male		-0.01
		(0.01)
Age		-0.00**
		(0.00)
LnInitial		0.03***
		(0.00)
NewUser		-0.07***
		(0.02)
Constant	0.28***	0.15***
	(0.01)	(0.03)
Observations	8,245	7,404
R-squared	0.00	0.11

Table 7. Order Effects on Self-Directed Investments

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 that is 1 if user i's survey is in the reverse order and 0 otherwise. Panel A is based on the sample of users who made self-directed investments during our sample period, and Y_i is $\alpha R5_i$, the ratio of user i's purchase of R5 products to his total purchase of R4 and R5 products and is denominated in percentage points. Panel B is based on the Overall Sample, and 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. All control variables are defined in Table 4. Standard errors are reported in parentheses. ***, **, and * indicate significance levels of 1%, 5%, and 10%, respectively.

Panel A. Intensive Margin

	$\alpha R5$ (%)			
	(1)	(2)	(3)	(4)
Reverse	2.83*	2.03	1.42	0.36
	(1.59)	(1.51)	(1.54)	(1.44)
Male		0.79		-1.68
		(1.54)		(1.47)
Age		-0.14*		-0.13*
		(0.07)		(0.07)
LnInitialDirect		0.37		0.11
		(0.26)		(0.25)
NewUser		23.12***		22.87***
		(3.16)		(3.00)
Constant	13.39***	9.54**	-1.42	5.82
	(1.12)	(4.48)	(28.80)	(26.93)
Category FEs	No	No	Yes	Yes
Observations	1,406	1,406	1,406	1,406
R-squared	0.00	0.11	0.07	0.20

Panel B. Extensive Margin

	D_Direct_i	
	(1)	(2)
Reverse	-0.01	-0.01
	(0.01)	(0.01)
Male		-0.02**
		(0.01)
Age		-0.00
		(0.00)
LnInitial		0.02***
		(0.00)
NewUser		0.97***
		(0.02)
Constant	0.21***	0.06***
	(0.01)	(0.02)
Observations	8,245	7,404
R-squared	0.00	0.29

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: LCMF Risk Assessment Survey Questions

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

To sheds further light on the response-order effect, we run the following Probit regression

$$\Pr\left(Answer_{ij} = k\right) = \Phi\left(\alpha + \beta Reverse_i \times Sub_j + \gamma Reverse_i + \delta Sub_j + \varepsilon_{ij}\right), \quad (5)$$

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 (5) estimates the effect on the probability of selecting the penultimate and last choices.[2] Regression results are reported in Table 4.

The first column of Panel A shows that the coefficient of *Reverse* is -0.00 (*t*=-0.12) and the coefficient of the interaction term is -0.11 (*t*=-3.02). 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% (*t*=4.45), while, for objective questions, the marginal effect of *Reverse* is merely -0.03% (*t*=-0.12).

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% (*t*=8.21) for the second choice, but increases the selection probability by 1.57% (*t*=3.57) and 3.66% (*t*=8.68) 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 of their lists. When the response order is reversed, the last and penultimate choices on the original-

[2] 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.

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 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 $\alpha=1/2$. Consistent with the anchoring-based interpretation, the triple interaction coefficient is -3.45 (t=-3.00). That is, the order effect is weaker for respondents with shorter decision times. We also repeat the regression with $\alpha=1/3$ and $\alpha=1/4$. The results, reported in Columns (2) and (3), are similar.

3. 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 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 whose risk scores lie within a narrow bandwidth around a cutoff should exhibit markedly different investment behaviors depending on which side of the threshold they fall. 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:

$$\alpha R5_i = \alpha + \beta 1_{Score_i \geq 69} + \gamma X_i + \varepsilon_i, \quad (7)$$

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 8 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 15.31 (*s.e.* = 3.07), indicating that users whose scores are just above the threshold (and are therefore classified as C5) allocate 15.31 percentage points more to 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 similar in magnitude. In summary, there is a substantial discontinuity in R5 allocation 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 R5 allocation across the artificial threshold, consistent with the absence of a risk category change.

It is important to note that the estimated effect in Panel A should not be interpreted as causal. 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

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*. T-statistics, 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 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. Discontinuity in Self-Directed Investments

This table shows discontinuity of investors' purchases of R5 products at risk assessment score 69. Specifically, run the following regression:

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

where $\alpha R5_i$ is the ratio of respondent i's purchase of R5 products to his total purchase of R4 and R5 products and is denominated in percentage points, $1_{Score_i \geq K}$ 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. Regressions are based on the original-order sample of investors whose final 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., $K - 3 \leq Score_i \leq K + 2$). $K = 69,\ 66$, and 72 in Panels A, B, and C, respectively. Standard errors are reported in parentheses. ***, **, and * indicate significance levels of 1%, 5%, and 10%, respectively.

Panel A. $K = 69$

	$\alpha R5$ (%)		
	$n = 3$	$n = 2$	$n = 1$
$1_{Score_i \geq K}$	15.31***	17.79***	18.68***
	(3.07)	(4.18)	(5.52)
Controls	Y	Y	Y
N	354	223	113
R-squared	0.27	0.27	0.38

Panel B. $K = 66$

	$\alpha R5$ (%)		
	$n = 3$	$n = 2$	$n = 1$
$1_{Score_i \geq K}$	1.51	2.08	-5.96
	(1.61)	(2.09)	(5.22)
Controls	Y	Y	Y
N	207	153	68
R-squared	0.02	0.02	0.09

Panel C. $K = 72$

	$\alpha R5$ (%)		
	$n = 3$	$n = 2$	$n = 1$
$1_{Score_i \geq K}$	-1.59	0.29	-1.19
	(2.70)	(3.06)	(4.78)
Controls	Y	Y	Y
N	440	317	150
R-squared	0.31	0.32	0.31