

Online Appendix for “Taming the Bias Zoo”

Hongqi Liu, Cameron Peng, Wei A. Xiong, and Wei Xiong*

This appendix contains the following materials:

1. Original survey questions in Chinese
2. Survey questions translated to English
3. Distribution of respondents across brokers and provinces
4. Relationship between the financial literacy score and the time taken to complete the survey
5. Construction of each variable used in the empirical analysis
6. A comparison between Chinese investors and U.S. investors
7. A comparison between the pre-merged sample and the merged sample
8. Additional validation exercises
9. Robustness checks for the horse-race analysis
10. Gambling preferences and characteristics of stocks purchased
11. A “nudge” experiment to examine the effect of awareness of trading costs on turnover

*Hongqi Liu: Chinese University of Hong Kong, Shenzhen (email: hongqiliu@cuhk.edu.cn); Cameron Peng: London School of Economics and Political Science (email: c.peng9@lse.ac.uk); Wei A. Xiong: Shenzhen Stock Exchange (email: wxiong@szse.cn); Wei Xiong: Princeton University and NBER (email: wxiong@princeton.edu).

1. Survey Questions (Chinese Version)

This section presents the Chinese version of our full survey. Figure A1 shows the cover page of the survey, and the full text of the survey follows.

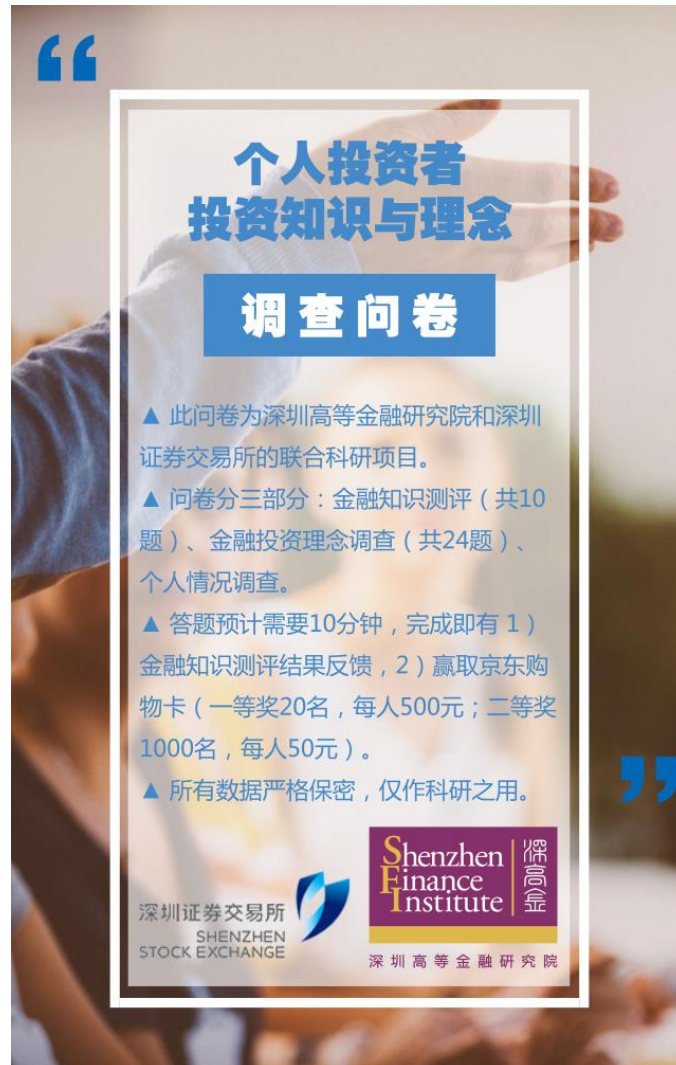


Figure A1: Cover Page of the Survey and Confidentiality Agreement

个人投资者投资知识与理念调查

1. 以1到7来打分, 1分代表非常差, 7分代表非常好, 请问您会给自己的金融知识水平打几分?

- A. 1-非常差
- B. 2-比较差
- C. 3-有点差
- D. 4-一般
- E. 5-有点好
- F. 6-比较好
- G. 7-非常好
- H. 拒绝回答

接下来的八道题, 我们会对您的金融知识进行一个小测试, 快来看看您能得几分吧(完成问卷即有结果反馈)!

2. 假设您有100元的银行存款, 银行的利息是每年2%。如果您将钱一直存在银行里, 5年以后, 您的银行账户里面会有多少钱?

- A. 多于102元
- B. 正好102元
- C. 少于102元
- D. 不知道
- E. 拒绝回答

3. 假设您把钱存入银行, 银行存款利息是每年1%, 物价水平预计每年上涨2%。一年以后您能够买到的物品预计会

- A. 比现在多
- B. 和现在相等
- C. 比现在少
- D. 不知道
- E. 拒绝回答

4. 请问您是否同意以下说法: 购买单支股票通常比购买股票型基金的风险更小

- A. 同意
- B. 不同意
- C. 不知道
- D. 拒绝回答

5. 通常来说, 以下哪种投资产品的收益波动最大?

- A. 银行存款
- B. 债券
- C. 股票

- D. 不知道
E. 拒绝回答
6. 如果一个人购买了 B 公司的股票，以下正确的描述是
A. 他拥有 B 公司的一部分
B. 他借钱给 B 公司
C. 他对 B 公司债务承担无限责任
D. 不知道
E. 拒绝回答
7. 通常来说，当市场利率下降时，已发行债券的价格会
A. 上升
B. 不变
C. 下降
D. 不知道
E. 拒绝回答
8. 市盈率的定义是
A. 每股股价/每股收益
B. 每股股价/每股股息
C. 每股股价/每股账面价值
D. 不知道
E. 拒绝回答
9. 以下哪一项关于基金的描述是正确的？
A. 所有的基金都投资股票
B. 基金可以投资多种资产，例如股票和债券
C. 基金根据过往的业绩支付固定的收益
D. 不知道
E. 拒绝回答
10. 第 2 题至第 9 题这八道题中，您认为自己答对了几题？
A. 0 道
B. 1 道
C. 2 道
D. 3 道
E. 4 道
F. 5 道
G. 6 道
H. 7 道
I. 8 道

- J. 不知道
- K. 拒绝回答

11. 请问您认为有多少比例的股民在 2017 年的股市收益比您的高？

- A. 低于 10%
- B. 10% 到 20%
- C. 20% 到 30%
- D. 30% 到 40%
- E. 40% 到 50%
- F. 50% 到 60%
- G. 60% 到 70%
- H. 70% 到 80%
- I. 80% 到 90%
- J. 高于 90%
- K. 不知道
- L. 拒绝回答

12. 接下来，我们想请您对未来一年中国股市的走势做出预测。具体来说，在现在开始计算的**未来一年内**，您认为上证指数的累积收益会是多少？

为了方便您答题，以下我们给出了以 5% 为区间的多个选项，请您作出预测。

【下拉框中给出-50%到 50%的选项，以 5% 递增】

低于-50%;-50% 到 -45%;-45% 到 -40%;-40% 到 -35%;-35% 到 -30%;-30% 到 -25%;-25% 到 -20%;-20% 到 -15%;-15% 到 -10%;-10% 到 -5%;-5% 到 0%;0% 到 5%;5% 到 10%;10% 到 15%;15% 到 20%;20% 到 25%;25% 到 30%;30% 到 35%;35% 到 40%;40% 到 45%;45% 到 50%;高于 50%;不知道;拒绝回答

13. 接上题，我们请您预测上证指数上涨的空间。假设上证指数未来一年的累积收益有**一成**的概率能**高于**某个水平，您认为这个水平会是多少？

为了方便您答题，以下我们给出了以 5% 为区间的多个选项，请您作出预测。

【下拉框中给出-50%到 50%的选项，以 5% 递增】

低于-50%;-50% 到 -45%;-45% 到 -40%;-40% 到 -35%;-35% 到 -30%;-30% 到 -25%;-25% 到 -20%;-20% 到 -15%;-15% 到 -10%;-10% 到 -5%;-5% 到 0%;0% 到 5%;5% 到 10%;10% 到 15%;15% 到 20%;20% 到 25%;25% 到 30%;30% 到 35%;35% 到 40%;40% 到 45%;45% 到 50%;高于 50%;不知道;拒绝回答

14. 接上题，我们请您预测上证指数下跌的空间。假设上证指数未来一年的累积收益有**一成**的概率会**低于**某个水平，您认为这个水平会是多少？

为了方便您答题，以下我们给出了以 5% 为区间的多个选项，请您作出预测。

【下拉框中给出-50%到 50%的选项，以 5%递增】

低于-50%;-50% 到 -45%;-45% 到 -40%;-40% 到 -35%;-35% 到 -30%;-30% 到 -25%;-25% 到 -20%;-20% 到 -15%;-15% 到 -10%;-10% 到 -5%;-5% 到 0%;0% 到 5%;5% 到 10%;10% 到 15%;15% 到 20%;20% 到 25%;25% 到 30%;30% 到 35%;35% 到 40%;40% 到 45%;45% 到 50%;高于 50%;不知道;拒绝回答

15. 在实际交易中，假设您以每股 100 元的价格购入了 100 股 A 股票，然后再以同样的价格全部卖出。请问这两笔交易您需要支付的交易费用（包括券商佣金和相关税费）总共是多少？为了方便您答题，下面我们给出了以 5 元为区间的多个选项，请选择您认为正确的区间。

【下拉框中给出 0 到 50 元的选项，以 5 元递增】

0 - 5 元; 5 - 10 元; 10 - 15 元; 15 - 20 元; 20 - 25 元; 25 - 30 元; 30 - 35 元; 35 - 40 元; 40 - 45 元; 45 - 50 元; 50 元以上; 不知道; 拒绝回答

16. 请问您在做出买卖股票的决定时，有多少时候考虑过交易费用？

- A. 从来没有
- B. 很少
- C. 有时候
- D. 经常
- E. 总是
- F. 不知道
- G. 拒绝回答

17. 请问您是否同意股票市场上的买卖价差是一种交易成本？买卖价差是指某一个时间点上，市场上限价委托单的最低卖价和最高买价之间的差值。

- A. 同意
- B. 不同意
- C. 不明白买卖价差的含义
- D. 不知道
- E. 拒绝回答

18. 请问您是否同意以下说法？我平时买卖股票时，寄希望于能通过买到短期内连续上涨的少数个股实现快速致富。

- A. 非常同意
- B. 同意
- C. 中立
- D. 不同意
- E. 非常不同意

- F. 不知道
- G. 拒绝回答

19. 请问您是否同意以下说法？我平时买卖股票时，抱着一种买彩票的心态，愿意承担一定的亏损以换取小概率的高收益。

- A. 非常同意
- B. 同意
- C. 中立
- D. 不同意
- E. 非常不同意
- F. 不知道
- G. 拒绝回答

20. 通常情况下，当您持有的一支股票市场价显著高于您的平均购买价时，卖出或继续持有该股票这两种情形中，哪一种让您感到更快乐？

- A. 卖出该股票
- B. 一样快乐
- C. 继续持有该股票
- D. 没什么感觉
- E. 不知道
- F. 拒绝回答

21. 通常情况下，当您持有的一支股票市场价显著低于您的平均购买价时，卖出或继续持有该股票这两种情形中，哪一种让您感到更痛苦？

- A. 卖出该股票
- B. 一样痛苦
- C. 继续持有该股票
- D. 没什么感觉
- E. 不知道
- F. 拒绝回答

22. 请问您是否同意以下说法？接触新的股票和公司让我感到兴奋。

- A. 非常同意
- B. 同意
- C. 中立
- D. 不同意
- E. 非常不同意
- F. 不知道

G. 拒绝回答

23. 请问您是否同意以下说法？我享受股票价格的上涨和下跌带来的刺激感。

- A. 非常同意
- B. 同意
- C. 中立
- D. 不同意
- E. 非常不同意
- F. 不知道
- G. 拒绝回答

24. 请问您在做出买卖股票的决定时，有多少时候是因为您认为自己比其他投资者更了解这些股票？

- A. 从来没有
- B. 很少
- C. 有时候
- D. 经常
- E. 总是
- F. 不知道
- G. 拒绝回答

25. 请问您在买卖股票的时候，有多少时候担心其他投资者知道一些您不知道的消息？

- A. 从来没有
- B. 很少
- C. 有时
- D. 经常
- E. 总是
- F. 不知道
- G. 拒绝回答

26. 请问您是否同意以下叙述？当某支股票价格持续上涨一段时间后，我通常认为其价格在短期内还会继续上涨。

- A. 非常同意
- B. 同意
- C. 中立
- D. 不同意
- E. 非常不同意
- F. 不知道

G. 拒绝回答

27. 请问您是否同意以下叙述？当某支股票价格持续下跌一段时间后，我通常认为其价格在短期内还会继续下跌。

- A. 非常同意
- B. 同意
- C. 中立
- D. 不同意
- E. 非常不同意
- F. 不知道
- G. 拒绝回答

28. 请问您在做出买卖股票的决定时，有多少时候是因为亲人，朋友或熟人的推荐或影响？

- A. 从来没有
- B. 很少
- C. 有时候
- D. 经常
- E. 总是
- F. 不知道
- G. 拒绝回答

29. 请问您在做出买卖股票的决定时，有多少时候是因为投资顾问（如证券公司客户经理，银行理财经理，炒股专家等）的推荐或影响？

- A. 从来没有
- B. 很少
- C. 有时候
- D. 经常
- E. 总是
- F. 不知道
- G. 拒绝回答

30. 请问您在做出买卖股票的决定时，有多少时候是因为需要调整投资组合？

- A. 从来没有
- B. 很少
- C. 有时候
- D. 经常
- E. 总是

- F. 不知道
- G. 拒绝回答

31. 请问您在做出卖出股票的决定时，有多少时候是因为其他地方需要用钱？

- A. 从来没有
- B. 很少
- C. 有时候
- D. 经常
- E. 总是
- F. 不知道
- G. 拒绝回答

32. 假设您是家庭中唯一的收入来源，您有一个很好的工作可以保证今后每年给您稳定不变的收入。现在您有机会换一个同样好的新工作，这份工作有 50% 的可能将您的收入增加一倍，同时也有 50% 的可能将收入减少五分之一。您会换新工作吗？

- A. 会
- B. 不会
- C. 不知道
- D. 拒绝回答

33. 接上题，假设新工作有 50% 的可能将您的收入增加一倍，50% 的可能将您的收入减少三分之一。您会换新工作吗？

- A. 会
- B. 不会
- C. 不知道
- D. 拒绝回答

34. 接上题，假设新工作有 50% 的可能将您的收入增加一倍，50% 的可能将您的收入减少一半。您会换新工作吗？

- A. 会
- B. 不会
- C. 不知道
- D. 拒绝回答

【针对实验组】

35. 非常感谢您完成答卷！快来看看巴菲特给个人投资者的建议：

数据显示，A股市场中百分之八十以上的交易量来自个人投资者。大量个人投资者存在频繁交易的行为。然而研究发现：交易频率越高的投资者，收益往往越差。

为什么会这样呢？

那是因为频繁交易
有两大负面影响：
1. 频繁交易会产生产高额的交易费用。
2. 股市的短期走势是无法预测的。
个人投资者由于缺乏经验和信息，
经常不好买入卖出的时机，
容易高买低卖，造成亏损。

交易费用详解：

假设：证券公司收取的佣金是万三

如果你以相同的价格买卖股票一次，那么：

则交易费用占初始投资额的百分比=
买入佣金0.03%+卖出佣金0.03%
+印花税0.1%=0.16%

我明白了！

如果我每月进行一次买入再卖出的操作，交易费用就占到我初始投资的1.92%。如果每周一次，则交易费用高达8.32%！

对的！你知道吗？过去15年间，沪深300指数的年平均收益为8.96%。

这样呀！这意味着什么呢？

这意味着，
一个和市场收益持平的投资者，
如果每月交易一次，
则交易费用占总收益的五分之一。
如果每周交易一次，
则交易费用几乎和总收益持平。
而且！很多投资者的收益是低于市场收益的，那么频繁交易的后果很有可能就是亏损。

股神巴菲特给个人投资者的建议

获得更好投资回报的秘诀就是：
买进一只股票，然后忘掉它。
如果试图频繁交易股票，
那么将赚不到钱。
如果你经常买卖股票，
你将以交易佣金和税收的形式
拿走很大一部分回报。所以，
最好是买好股票并持有很久。

咱们来看看您是否理解了图片中的内容吧！在上图的例子中，请问如果证券公司的佣金为万三，您每周进行一次买入股票，再以相同价格全部卖出的交易的话，一年之后，交易费用会占到您的初始投资额的多少呢？

- A. 0%
- B. 0.16%
- C. 1.92%
- D. 8.32%
- E. 不知道

36. 上一题的答案是 8.32%！交易费用不可忽视，请尽量避免频繁交易。

感谢您完成关于投资知识及理念的题目，下面请帮助我们了解一下您的个人情况吧！完成这一部分之后，会有您的答题成绩，更可参与抽奖。一等奖 20 名，每人 500 元，二等奖 1000 名，每人 50 元。若中奖，我们会将相应金额的京东电子购物卡通过短信发送至您的手机，请注意查收。请继续吧！若此时放弃，则答卷作废。

请问您开户的券商是？

华泰证券;中信证券;国泰君安;广发证券;银河证券;海通证券;招商证券;申万宏源;国信证券;中信建投

37. 请问您开户的营业部是？

38. 请问您的出生于哪一年？

39. 请问您的出生于几月？

40. 请问您的出生于几号？

41. 请问您的性别是

- A. 男
- B. 女

42. 请问您的最高学历是？

- A. 没有正式入学
- B. 小学
- C. 初中
- D. 高中/中专/技校
- E. 大专/成人大专/夜大/电大
- F. 大学本科
- G. 研究生及以上

43. 请问您的家庭净资产是（总资产-总负债）多少？

- A. 2 万元以下
- B. 2 万-10 万元（不含 10 万）

- C. 10 万元-50 万元（不含 50 万）
- D. 50 万元-100 万元（不含 100 万）
- E. 100 万元-500 万元（不含 500 万）
- F. 500 万-1000 万元（不含 1000 万）
- G. 1000 万元及以上

44. 请问您的家庭年收入是多少？

- A. 2 万元以下
- B. 2 万-10 万元（不含 10 万）
- C. 10 万-20 万（不含 20 万）
- D. 20 万-50 万（不含 50 万）
- E. 50 万-100 万（不含 100 万）
- F. 100 万-200 万（不含 200 万）
- G. 200 万-1000 万（不含 1000 万）
- H. 1000 万元及以上

45. 请问您的出生省份是？（注：在选择出生省市时，以现有行政划分为准。如直辖市重庆市在 1997 年以前隶属四川省。若张三在 1988 年出生于该市，则应按照现有行政区划选择重庆市，而非四川省）

46. 请问您的出生城市是？（注：在选择出生省市时，以现有行政划分为准。如直辖市重庆市在 1997 年以前隶属四川省。若张三在 1988 年出生于该市，则应按照现有行政区划选择重庆市，而非四川省）

这是最后一道题啦，非常感谢您完成问卷。请在下一页中填入您的姓名和手机号码参与抽奖以及查看答卷成绩。若中奖，我们会将京东电子购物卡通过短信发送至您的手机，请注意查收。若此时放弃，则答卷作废

47. 请问您的姓名是？

48. 请问您的手机号是？

2. Survey Questions (English Version)

We translated the original Chinese survey into English, with slight modifications made to accommodate the different institutional details of the U.S. market. This translated version was used for a smaller survey of U.S. investors on Amazon Mechanical Turk. The English version of the survey is attached below.

1. How do you assess the level of your financial literacy?
 - A. Very low
 - B. Low
 - C. Slightly low
 - D. Fair
 - E. Slightly high
 - F. High
 - G. Very high
 - H. Decline to answer
2. Normally, which of the following assets displays the highest fluctuation over time?
 - A. Savings accounts
 - B. Bonds
 - C. Stocks
 - D. Do not know
 - E. Decline to answer
3. Which of the following statements is correct? If somebody buys a stock of firm B in the stock market:
 - A. He owns a part of firm B
 - B. He has lent money to firm B
 - C. He is fully liable for firm B's total debts
 - D. Do not know
 - E. Decline to answer
4. Normally, when the market interest rate falls, the price of an existing bond will:
 - A. Go up
 - B. Remain the same
 - C. Go down
 - D. Do not know
 - E. Decline to answer

5. What is the P/E ratio?
- A. Price per share/Earnings per share
 - B. Price per share/Dividend per share
 - C. Price per share/Book value per share
 - D. Do not know
 - E. Decline to answer
6. Which of the following statements about mutual funds is correct?
- A. All mutual funds invest in stocks
 - B. Mutual funds can invest in multiple asset classes, such as stocks and bonds
 - C. Mutual funds pay fixed returns based on their past performance
 - D. Do not know
 - E. Decline to answer
7. From Question 2 to 9, how many questions do you think you have answered correctly?
- A. 0
 - B. 1
 - C. 2
 - D. 3
 - E. 4
 - F. 5
 - G. 6
 - H. 7
 - I. 8
 - J. Do not know
 - K. Decline to answer
8. What fraction of retail investors do you think earned higher returns than you in 2018?
- A. Lower than 10%
 - B. 10% to 20%
 - C. 20% to 30%
 - D. 30% to 40%
 - E. 40% to 50%
 - F. 50% to 60%
 - G. 60% to 70%
 - H. 70% to 80%
 - I. 80% to 90%
 - J. Higher than 90%
 - K. Do not know
 - L. Decline to answer
9. We now ask you to make predictions about the U.S. stock market over the next year.
What do you think is the expected return for the S&P 500 Index over the next year?

10. As a follow-up question, we ask you to predict the upside potential for the S&P 500 Index. Assume that with 10% chance the return of the S&P 500 Index over the next year will be higher than a certain level. What do you think this level will be?
11. As a follow-up question, we ask you to predict the downside potential for the S&P 500 index. Assume that with 10% chance the return of the S&P 500 Index over the next year will be lower than a certain level. What do you think this level will be?
12. Suppose you purchase 100 shares of stock A at \$100 per share and then sell all the shares at the same price. Based on your trading experience, what's your estimated total transaction cost (including the commission fees) in \$?
- A. 0-5
 - B. 5-10
 - C. 10-15
 - D. 15-20
 - E. 20-25
 - F. 25-30
 - G. 30-35
 - H. 35-40
 - I. 40-45
 - J. 45-50
 - K. 50 or more
 - L. Do not know
 - M. Decline to answer
13. How often do you consider transaction costs when you trade stocks?
- A. Never
 - B. Rarely
 - C. Sometimes
 - D. Often
 - E. Always
 - F. Do not know
 - G. Decline to answer
14. Do you agree the following statement? The bid-ask spread is one form of transaction cost. (The bid-ask spread is the difference between the lowest ask price and the highest bid price.)
- A. Agree
 - B. Disagree
 - C. Do not understand what the bid-ask spread is
 - D. Do not know
 - E. Decline to answer

15. Do you agree with the following statement? When I trade stocks, I often wish to select those stocks whose price would rise sharply in a short period time so that I can make a lot of money quickly.
- A. Strongly agree
 - B. Agree
 - C. Neutral
 - D. Disagree
 - E. Strongly disagree
 - F. Do not know
 - G. Decline to answer
16. Do you agree with the following statement? When I trade stocks, I often think of them as lotteries: I am willing to accept small losses in exchange for the possibility of a big upside.
- A. Strongly agree
 - B. Agree
 - C. Neutral
 - D. Disagree
 - E. Strongly disagree
 - F. Do not know
 - G. Decline to answer
17. Normally, if the price of a stock in your portfolio rose substantially since you bought it, which of these two actions would make you feel happier: holding on to the stock, or selling that stock?
- A. Holding on to the stock
 - B. The same
 - C. Selling the stock
 - D. I am indifferent to price changes
 - E. Do not know
 - F. Decline to answer
18. Normally, if the price of a stock in your portfolio dropped substantially since you bought it, which of these two actions would make you feel more painful: holding on to the stock, or selling that stock?
- A. Holding on to the stock
 - B. The same
 - C. Selling the stock
 - D. I am indifferent to price changes
 - E. Do not know
 - F. Decline to answer
19. Do you agree with the following statement? I feel excited about getting to know new stocks and new firms.
- A. Strongly agree
 - B. Agree

- C. Neutral
- D. Disagree
- E. Strongly disagree
- F. Do not know
- G. Decline to answer

20. Do you agree with the following statement? I feel excited about the stock market moving up and down.

- A. Strongly agree
- B. Agree
- C. Neutral
- D. Disagree
- E. Strongly disagree
- F. Do not know
- G. Decline to answer

21. When you decide to trade a stock, how often do you believe that you know the stock better than others?

- A. Never
- B. Rarely
- C. Sometimes
- D. Often
- E. Always
- F. Do not know
- G. Decline to answer

22. When you decide to trade a stock, how often do you worry that other investors know about the stock better than you do?

- A. Never
- B. Rarely
- C. Sometimes
- D. Often
- E. Always
- F. Do not know
- G. Decline to answer

23. Do you agree with the following statement? After a stock's price keeps rising for a while, I usually believe that the price will rise even further in the future.

- A. Strongly agree
- B. Agree
- C. Neutral
- D. Disagree
- E. Strongly disagree
- F. Do not know
- G. Decline to answer

24. Do you agree with the following statement? After a stock's price keeps dropping for a while, I usually believe that the price will drop even further in the future.
- A. Strongly agree
 - B. Agree
 - C. Neutral
 - D. Disagree
 - E. Strongly disagree
 - F. Do not know
 - G. Decline to answer
25. When you decide to trade a stock, how often are you influenced by your family members, friends, or other acquaintances?
- A. Never
 - B. Rarely
 - C. Sometimes
 - D. Often
 - E. Always
 - F. Do not know
 - G. Decline to answer
26. When you decide to trade a stock, how often are you influenced by your investment advisors?
- A. Never
 - B. Rarely
 - C. Sometimes
 - D. Often
 - E. Always
 - F. Do not know
 - G. Decline to answer
27. When you decide to trade a stock, how often is it that you need to rebalance your portfolio?
- A. Never
 - B. Rarely
 - C. Sometimes
 - D. Often
 - E. Always
 - F. Do not know
 - G. Decline to answer
28. When you decide to trade a stock, how often is it that you need money somewhere else?
- A. Never
 - B. Rarely
 - C. Sometimes
 - D. Often
 - E. Always

- F. Do not know
- G. Decline to answer

29. Suppose you are the only income earner in the family, and you have a good job guaranteed to give you your current income every year for life. You are given the opportunity to take a new, equally good job. With a 50% chance it will double your income, and with a 50% chance, it will cut your income by 20%. Would you take the new job?

- A. Yes
- B. No
- C. Do not know
- D. Decline to answer

30. Suppose the chances were 50% that it would double your income and 50% that it would cut it by $1/3$. Would you take the new job?

- A. Yes
- B. No
- C. Do not know
- D. Decline to answer

31. Suppose the chances were 50% that it would double your income and 50% that it would cut it by $1/2$. Would you take the new job?

- A. Yes
- B. No
- C. Do not know
- D. Decline to answer

3. Distribution of respondents across brokers and provinces

Table A1 reports the distribution of respondents across brokers and provinces. By design, respondents were evenly distributed across the ten brokers, with only slight variation. In terms of geographic variation, areas that are more financially developed (e.g., Guangdong, Zhejiang, Jiangsu, and Shanghai) are more represented in our sample.

Panel A: By Broker	Respondents	Percentage
Guotai Junan Securities	1,519	11.8%
CITIC Securities	1,410	11.0%
Haitong Securities	1,390	10.8%
China Merchants Securities	1,372	10.7%
Huatai Securities	1,350	10.5%
Guosen Securities	1,252	9.8%
China Securities	1,203	9.4%
Shenwan Hongyuan Securities	1,169	9.1%
GF Securities	1,111	8.7%
China Galaxy Securities	1,051	8.2%
Panel B: By Province/Region		
Guangdong	1,674	13.1%
Zhejiang	1,201	9.4%
Jiangsu	1,138	8.9%
Shanghai	1,135	8.9%
Hubei	629	4.9%
Beijing	622	4.9%
Fujian	600	4.7%
Hunan	572	4.5%
Shandong	542	4.2%
Henan	531	4.1%
Sichuan	530	4.1%
Anhui	463	3.6%
Jiangxi	388	3.0%
Hebei	385	3.0%
Liaoning	331	2.6%
Chongqing	284	2.2%
Heilongjiang	250	2.0%
Guangxi	230	1.8%
Shanxi	222	1.7%
Shaanxi	198	1.5%
Others	931	7.2%
Total	12,856	100%

Table A1 Distribution of Survey Respondents Across Brokers and Provinces

4. Relationship Between the Financial Literacy Score and the Time Taken to Complete the Survey

Figure A2 plots the distribution of the total amount of time spent on the survey: it took a median respondent about eight minutes to complete the survey, and 95% of respondents finished within 20 minutes. However, we find that respondents who spent less than three minutes on the survey experienced a sharp drop in their financial literacy score, suggesting that they may have shirked during the survey. In the main analysis of the paper, we dropped these observations, which reduced our sample size to 11,268.

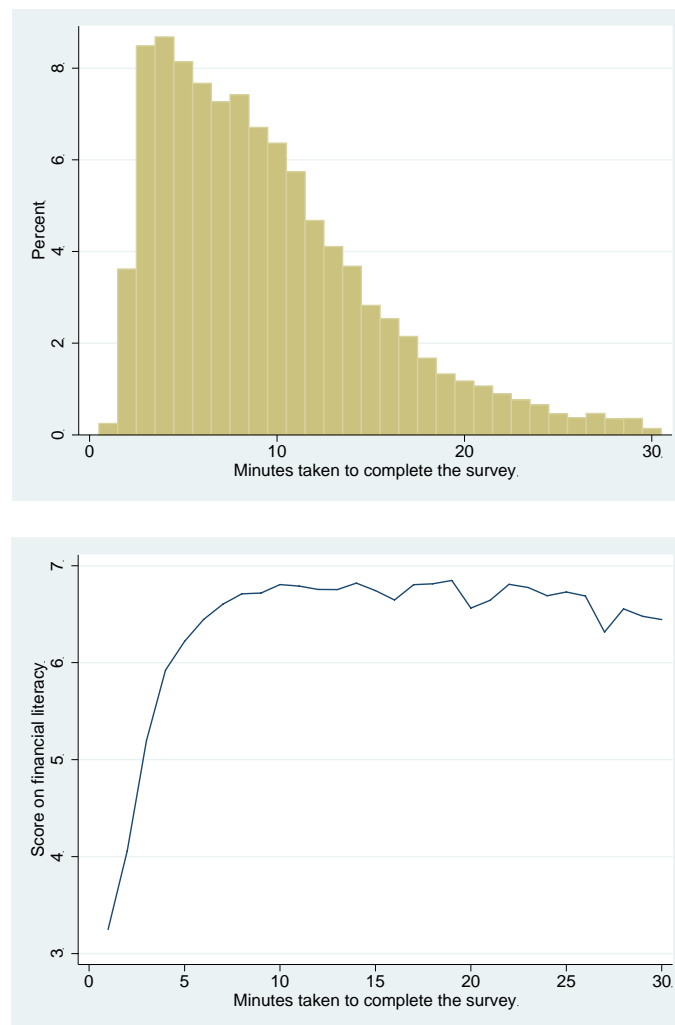


Figure A2 Relationship Between Financial Literacy Score and the Time Taken to Complete the Survey

5. Construction of Variables

Table A2 provides a detailed description about the construction of each variable we use in the paper.

Variable	Definition
Over-placement, performance	1: self-assessed performance rank among the investor population in 2017 is higher than her actual rank 0: otherwise
Over-placement, literacy	1: self-assessed score in the financial literacy test is higher than her actual score 0: otherwise
Miscalibration of uncertainty	Investors are asked to estimate how much the stock market will go up (down) with 10% probability within the next year. The difference between these two estimates gives the 80% confidence interval of an investor's forecast for the market's next year return. 1: confidence interval is narrower than 76%, the estimated confidence interval from the historical data 0: otherwise
Underestimation of transaction cost	Investors are asked to estimate the total transaction cost from buying 10,000 RMB worth of stock and then selling everything at the same value. 1: answer falls below the estimated true cost (15 to 26 RMB) 0: otherwise
Do not consider transaction cost	1: answers "Rarely" or "Never" when asked how often she considers transaction costs when trading a stock 0: otherwise
Do not think bid-ask spread is a cost	1: answers "Disagree", "Do not understand what the bid-ask spread is", "Do not know" or "Decline to answer" when asked if she agrees that the bid-ask spread is one form of transaction cost 0: otherwise
Extrapolation, up	1: answers "Strongly agree" or "Agree" when asked if she believes stock price will rise even further in the future after it keeps rising for a while 0: otherwise
Extrapolation, down	1: answers "Strongly agree" or "Agree" when asked if she believes the stock price will drop even further in the future after it keeps dropping for a while 0: otherwise
Gambling preference, blockbusters	1: answers "Strongly agree" or "Agree" when asked if she aims to make a lot of money quickly through stock investment

	0: otherwise
Gambling preference, lotteries	1: answers “Strongly agree” or “Agree” when asked if she often think of stocks as lotteries 0: otherwise
Realization utility, winners	1: answers “Selling the stock” when asked if the price of a stock in her portfolio rose substantially since she bought it, which of these two actions make her fell happier: holding on to the stock, or selling that stock? 0: otherwise
Realization utility, losers	1: answers “Selling the stock” when asked if the price of a stock in her portfolio dropped substantially since she bought it, which of these two actions would make her feel more painful: holding on to the stock, or selling that stock? 0: otherwise
Sensation seeking, novelty	1: answers “Strongly agree” or “Agree” when asked if she feels excited about getting to know new stocks and new firms 0: otherwise
Sensation seeking, volatility	1: answers “Strongly agree” or “Agree” when asked if she feels excited about the stock market moving up and down 0: otherwise
Perceived information advantage	1: answers “Often” or “Always” when asked how often she believes that she knows the stock better than others when she decides to trade a stock 0: otherwise
Dismissive of others’ information	1: answers “Rarely” or “Never” when asked how often she worries that other investors know about the stock better than she does when she decides to trade a stock 0: otherwise
Social influence	1: answers “Often” or “Always” when asked how often she was influenced by her family members, friends, or other acquaintances when she decides to trade a stock 0: otherwise
Advisor influence	1: answers “Often” or “Always” when asked how often she was influenced by her investment advisors when she decides to trade a stock 0: otherwise
Portfolio rebalancing needs	1: answers “Often” or “Always” when asked how often it is because she need to rebalance her portfolio when she decides to trade a stock 0: otherwise

Liquidity needs	1: answers “Often” or “Always” when asked how often it is because she need to rebalance her portfolio when she decides to trade a stock 0: otherwise
Risk aversion	Investors are asked if they would be willing to give up their current stable jobs for other jobs with higher expected income but also higher uncertainty in three hypothetical scenarios. 1: answers “No” in all three questions 0: otherwise
Expected future one-year market return	Investors' prediction for the future one-year market returns, ranging from 1 (lower than –50%) to 22 (higher than 50%).
Financial literacy, dummy	1: scores 7 or higher out 8 questions designed to test her financial literacy 0: otherwise

Other variables	Definition
Turnover	For each month, we calculate an investor's turnover ratio as the total value of buy and sell orders divided by the maximum value of investment in that month. We then take the time series average of monthly turnover ratios weighted by monthly maximum value of investment.
Raw return	For each month, we calculate an investor's raw return as raw profit divided by the maximum value of investment in that month. We then take the time series average of monthly raw returns weighted by monthly maximum value of investment. Monthly raw profit = ending portfolio balance + value of sell orders + dividends – beginning portfolio balance – value of buy orders
Net return	For each month, we calculate an investor's net return as net profit divided by the maximum value of investment in that month. We then take the time series average of monthly net returns weighted by monthly maximum value of investment. Monthly net profit = ending portfolio balance + value of sell orders + dividends – beginning portfolio balance – value of buy orders – comission fees – stamp duty
Volume-weighted Past One-month Count of Up-limit Hits Based on Initial Buys	For each month, we first calculate the past one-month # of up-limit hits of the stock for each initial buy transaction and then take the transaction value weighted average across all initial buy orders. We then take the time-series average value weighted by monthly initial buy values. A stock purchase is considered as an initial buy if the investor holds zero share of the stock before the purchase.

Volume-weighted Past One-quarter Count of Up-limit Hits Based on Initial Buys	For each month, we first calculate the past one-quarter # of up-limit hits of the stock for each initial buy transaction and then take the transaction value weighted average across all initial buy orders. We then take the time-series average value weighted by monthly initial buy values. A stock purchase is considered as an initial buy if the investor holds zero share of the stock before the purchase.
Volume-weighted Past One-month Return Based on Initial Buys	We first calculate the past one-month return of the stock for each initial buy order and then take the average for all initial buy orders in a month weighted by the transaction value. We then calculate the time-series average weighted by the total value of monthly initial buy orders. A stock purchase is considered as an initial buy if the investor holds zero share of the stock before the purchase.
Volume-weighted Past One-quarter Return Based on Initial Buys	We first calculate the past one-quarter return of the stock for each initial buy order and then take the average for all initial buy orders in a month weighted by the transaction value. We then calculate the time-series average weighted by the total value of monthly initial buy orders. A stock purchase is considered as an initial buy if the investor holds zero share of the stock before the purchase.
Volume-weighted Past One-month Return Volatility Based on Initial Buys	We first calculate the past one-month return volatility of the stock for each initial buy order and then take the average for all initial buy orders in a month weighted by the transaction value. We then calculate the time-series average weighted by the total value of monthly initial buy orders. A stock purchase is considered as an initial buy if the investor holds zero share of the stock before the purchase.
Volume-weighted Past One-quarter Return Volatility Based on Initial Buys	We first calculate the past one-quarter return volatility of the stock for each initial buy order and then take the average for all initial buy orders in a month weighted by the transaction value. We then calculate the time-series average weighted by the total value of monthly initial buy orders. A stock purchase is considered as an initial buy if the investor holds zero share of the stock before the purchase.
<i>After</i>	The dummy equals one for the period from October 2018 onward and zero otherwise.
<i>Treated</i>	We randomly assign 500 targeted branches of brokerage firms into treated and control groups. Investors in the two groups receive questionnaires that are otherwise identical except for one difference: the questionnaire for the treated group include educational content about how frequent trading negatively affect their investment performance. The dummy equals one if an investor is in the treated group and correctly answered the follow-up question designed to test if the respondent understands the content of the message. The dummy equals zero if an investor is in the control group.

Table A2 Definition of Variables

6. Comparison Between Chinese and U.S. investors

To study the differences between Chinese and US investors, we run the English version of our survey on Mechanical Turk with a small sample of 400 investors. The results are reported in Table A3. American investors are more “sophisticated” in three aspects: they consider transaction cost more often when they trade, they are less dismissive of others’ information advantage, and they have a closer relationship with their advisors. At the same time, on several fronts, American retail investors are more prone to behavioral biases than their Chinese counterparts: they exhibit stronger gambling preferences, are more prone to realization utility and sensation seeking, and display slightly stronger extrapolative beliefs.

Survey question	China sample	US sample	Difference
# of responses	12,856	400	
Over-placement, literacy	0.25	0.20	0.04**
Miscalibration of uncertainty	0.67	0.88	-0.21***
Do not consider transaction cost	0.51	0.25	0.26***
Do not think bid-ask spread is a cost	0.42	0.40	0.02
Extrapolation, up	0.32	0.40	-0.08***
Extrapolation, down	0.33	0.35	-0.02
Gambling preference, blockbusters	0.36	0.52	-0.16***
Gambling preference, lotteries	0.30	0.52	-0.22***
Realization utility, winners	0.35	0.55	-0.20***
Realization utility, losers	0.23	0.41	-0.18***
Sensation seeking, novelty	0.26	0.74	-0.48***
Sensation seeking, volatility	0.28	0.65	-0.37***
Perceived information advantage	0.18	0.15	0.03
Dismissive of others’ information	0.45	0.25	0.20***
Social influence	0.14	0.12	.02
Advisor influence	0.08	0.31	-0.22***
Portfolio rebalancing needs	0.18	0.21	-0.03
Liquidity	0.12	0.19	-0.06***
Risk aversion	0.19	0.29	-0.09***

Table A3 Differences in Survey Responses between Chinese and American Investors

7. Comparison Between the Pre-merged Sample and the Merged Sample

Table A4 reports the comparison between the pre-merged sample and the merged sample. The merged sample represents the 6,013 investors we are able to locate in the Shenzhen Stock Exchange's centralized database. Overall, the merged sample and the pre-merged sample are very similar in their survey responses. For certain responses such as risk aversion and return expectations, the difference is significant, but the magnitude is generally small. Overall, we don't find strong evidence that our merging process is biased towards certain investor groups.

Survey question	Merged sample	Pre-merged sample	Difference
Over-placement, literacy	0.24	0.26	-0.02**
Miscalibration of uncertainty	0.67	0.66	0.01
Underestimation of transaction cost	0.68	0.67	0.00
Do not consider transaction cost	0.52	0.50	0.02*
Do not think bid-ask spread is a cost	0.41	0.43	-0.01
Extrapolation, up	0.32	0.31	0.01
Extrapolation, down	0.34	0.32	0.02*
Gambling preference, with prob. weighting	0.37	0.36	0.00
Gambling preference, without prob. weighting	0.31	0.30	0.01
Realization utility, winner	0.36	0.35	0.01
Realization utility, loser	0.23	0.22	0.01
Sensation seeking, novelty	0.25	0.26	-0.01
Sensation seeking, volatility	0.28	0.28	0.01
Perceived information advantage	0.17	0.18	-0.01
Dismissive of others' information	0.46	0.45	0.01
Social influence	0.14	0.14	-0.01
Advisor influence	0.08	0.09	-0.01**
Portfolio rebalancing needs	0.17	0.19	-0.02***
Liquidity	0.11	0.14	-0.03***
Risk aversion	0.20	0.18	0.02***
Expected future one-year market return	9.92	9.25	0.67***

Table A4 Comparison between the Pre-merged Sample and the Merged Sample

8. Additional Validation Exercises

In the main text of the paper, we report evidence from the validation exercises for gambling preferences. In this section, we present evidence from three other validation exercises: one for extrapolative beliefs, one for risk aversion, and one for return expectations.

We measure extrapolative behavior by the volume-weighted past return among all the stocks bought by an investor. Table A5 reports the results when regressing transaction-based extrapolative behavior on survey-based extrapolative beliefs, where, in measuring extrapolative behavior, Panel A uses the past one-month return and Panel B uses the past one-quarter return. Indeed, investors who report having extrapolative beliefs exhibit stronger extrapolative behavior: on average, the stocks they purchase experience 1% higher returns in the preceding month and more than 2% higher returns in the preceding quarter, and this holds in both pre-survey and post-survey samples. The two measures of extrapolation have equally strong explanatory power for extrapolative behavior.

Table A6 reports the relationship between risk aversion and the average return volatility of stocks in the portfolio. Overall, investors with a higher risk aversion tend to hold stocks with less volatility. This further validates that survey-based risk aversion is useful for explaining portfolio-level characteristics. The results are robust to alternative specifications of return volatility and different sample periods.

Table A7 reports the relationship between return expectations and changes in stock holdings. Consistent with the evidence from Giglio et al. (2020), we find that higher return expectations indeed lead to an increase in stock holdings. However, we also reaffirm their finding that the sensitivity of holding changes to return expectations is small in magnitude. Note that, unlike Giglio et al. (2020), which studies holdings of passive funds at Vanguard, we study holdings of individual stocks, which are arguably traded more actively.

Panel A: Volume-Weighted Past One-Month Return Based on Initial Buys												
	Full sample (2018:01 – 2019:06)				Pre-survey (2018:01 – 2018:09)				Post-survey (2018:10 – 2019:06)			
Extrapolation, up	0.011** (2.170)	0.011** (2.134)			0.012*** (2.689)	0.013*** (2.902)			0.011* (1.668)	0.011* (1.704)		
Extrapolation, down			0.014*** (2.751)	0.013*** (2.640)			0.012*** (2.655)	0.012*** (2.691)			0.014** -2.142	0.014** -2.142
Male		-0.014*** (-2.854)		-0.014*** (-2.816)		-0.012*** (-2.740)		-0.012*** (-2.697)		-0.014** (-2.284)		-0.014** (-2.237)
Controls	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
R2	0.001	0.017	0.002	0.018	0.002	0.016	0.002	0.016	0.001	0.017	0.001	0.017
N	4,142	4,142	4,142	4,142	3,432	3,432	3,432	3,432	3,550	3,550	3,550	3,550

Panel B: Volume-Weighted Past One-Quarter Return Based on Initial Buys												
	Full sample (2018:01 – 2019:06)				Pre-survey (2018:01 – 2018:09)				Post-survey (2018:10 – 2019:06)			
Extrapolation, up	0.020** (2.406)	0.020** (2.419)			0.019*** (2.999)	0.022*** (3.446)			0.026** (2.451)	0.028*** (2.597)		
Extrapolation, down			0.021*** (2.615)	0.020** (2.532)			0.020*** (3.112)	0.021*** (3.316)			0.021** (2.032)	0.021** (2.091)
Male		-0.028*** (-3.685)		-0.028*** (-3.638)		-0.037*** (-5.848)		-0.036*** (-5.801)		-0.030*** (-3.113)		-0.029*** (-3.031)
Controls	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
R2	0.001	0.023	0.002	0.023	0.003	0.033	0.003	0.033	0.002	0.021	0.001	0.02
N	4,136	4,136	4,136	4,136	3,428	3,428	3,428	3,428	3,544	3,544	3,544	3,544

t-statistics in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A5 Validating Extrapolative Belief Using Trend-Chasing Behavior

Note: This table studies the relationship between survey-based extrapolative beliefs and transaction-based trend-chasing behavior. Trend-chasing behavior is measured as the buy-volume (in RMB) weighted average of past one-month (Panel A) or one-quarter (Panel B) returns of the stocks an investor purchases in a given sample period. A stock purchase is considered as an initial buy if the investor holds zero share of the stock before the purchase. Each panel presents OLS regression results based on three sample periods: full (January 2018 through June 2019), pre-survey (January 2018 through September 2018), and post-survey (October 2018 through June 2019). Extrapolation-up (Extrapolation-down) equals one if an investor answers “Strongly agree” or “Agree” when asked if they believe stock price will rise (drop) even further in the future after it has risen (dropped) for a while. Otherwise, extrapolation-up (extrapolation-down) equals zero. Control variables include age, gender, wealth, income, trading experience, account size, and education. *T*-statistics are based on robust standard errors are reported in parentheses.

Panel A: Average Past One-month Volatility Based on Stock Holdings						
	Full sample (2018:01 – 2019:06)		Pre-survey (2018:01 – 2018:09)		Post-survey (2018:10 – 2019:06)	
Risk aversion	-0.068*** (-2.932)	-0.054** (-2.327)	-0.041 (-1.342)	-0.033 (-1.075)	-0.084*** (-3.071)	-0.069** (-2.515)
Male		0.043** (2.185)		0.052** (2.125)		0.029 (1.035)
Controls	NO	YES	NO	YES	NO	YES
R2	0.002	0.027	0.001	0.016	0.002	0.016
N	4,381	4,381	3,807	3,807	3,956	3,956
Panel B: Average Past One-quarter Volatility Based on Stock Holdings						
	Full sample (2018:01 – 2019:06)		Pre-survey (2018:01 – 2018:09)		Post-survey (2018:10 – 2019:06)	
Risk aversion	-0.065*** (-2.614)	-0.050** (-2.001)	-0.059* (-1.771)	-0.048 (-1.433)	-0.077*** (-2.662)	-0.062** (-2.125)
Male		0.046** (2.200)		0.062** (2.306)		0.021 (0.674)
Controls	NO	YES	NO	YES	NO	YES
R2	0.002	0.029	0.001	0.018	0.001	0.014
N	4,381	4,381	3,807	3,807	3,956	3,956

t-statistics in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A6 Validation of Risk Aversion

Note: This table studies the relationship between investors' risk attitude and return volatilities of the stocks they buy. The dependent variables are buy-volume (in RMB) weighted average of past one-month (Panel A) or one-quarter (Panel B) return volatilities of stocks an investor purchases during various sample periods. The return volatility is defined as the standard deviation of daily returns. A stock purchase is considered as an initial buy if the investor holds zero share of the stock before the purchase. Each panel presents results of an OLS regression based on three sample periods: full (January 2018 through June 2019), pre-survey (January 2018 through September 2018), and post-survey (October 2018 through June 2019). The key independent variables are dummies that indicate if an investor is strongly risk averse or not. It equals one if an investor answers "No" when asked if she would be willing to quit her current stable job for a higher expected pay but higher risk one in all three hypothetical scenarios. Otherwise, the dummy equals zero. Control variables include age, net worth, income, trading experience, account size, and education. *T*-statistics are based on robust standard errors are reported in parentheses.

Panel A: Change in Stock Holdings During the 6-month Window Around the Survey						
	Full sample		Pre-survey		Post-survey	
	(2018:07 – 2018:12)		(2018:07 – 2018:09)		(2018:10 – 2018:12)	
Return Expectation	0.006*	0.007*	0.005**	0.005***	0.003	0.003
	(1.781)	(1.836)	(2.363)	(2.603)	(1.496)	(1.391)
Male		0.113***		0.032		0.037*
		(3.209)		(1.596)		(3.209)
Controls	NO	YES	NO	YES	NO	YES
R2	0.001	0.053	0.002	0.042	0.001	0.021
N	3,122	3,122	3,174	3,174	3,159	3,159

Panel B: Change in Stock Holdings During the 18-month Window Around the Survey						
	Full sample		Pre-survey		Post-survey	
	(2018:01 – 2019:06)		(2018:01 – 2018:09)		(2018:10 – 2019:06)	
Return Expectation	0.030**	0.029**	0.015***	0.012**	-0.005	-0.005
	(2.084)	(2.106)	(2.614)	(2.186)	(-0.631)	(-0.572)
Male		0.359***		0.046		0.162**
		(2.760)		(0.820)		(1.969)
Controls	NO	YES	NO	YES	NO	YES
R2	0.001	0.099	0.002	0.077	0.000	0.041
N	3,187	3,187	3,143	3,143	3,225	3,225

t-statistics in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A7 Validation of Survey Expectations

Note: This table studies the relationship between investors' one-year market return forecasts and portfolio changes. The dependent variables are percentage changes in investors' total stock holdings during various sample periods. Panel A looks at investors' portfolio changes in the 6-month window around the survey, and Panel B examines a longer 18-month window around the survey. The key independent variables are investors' forecasts of one-year ahead market return. Its discrete values range from 1 (corresponds to less than -50%) to 21 (corresponds to 45% to 50%). Each panel presents results for an OLS regression based on three sample periods: full (January 2018 through June 2019), pre-survey (January 2018 through September 2018), and post-survey (October 2018 through June 2019). Control variables include age, net worth, income, trading experience, account size, and education. *T*-statistics are based on robust standard errors are reported in parentheses.

9. Robustness Checks for Horse-race Results

In this section, we present additional evidence for the horse race by considering alternative empirical specifications and subsample analyses.

9.1. Including Branch Fixed Effects

Table A8 reports the results for the horse race when we include branch fixed effects. As the table shows, virtually all the results are unchanged, and gambling preference and perceived information advantage remain the most powerful explanatory variables for trading. One caveat is that the significance of gambling preference drops from significant at the 1% level to 5%, and the only trading motive that is significant at the 1% level in this specification is perceived information advantage.

Dependent Variable: Average Monthly Turnover Ratio (%) (October 2018 – June 2019)			
Actual rank of 2017 performance	4.347*** (5.085)	Gambling preference, blockbusters	10.570** (2.428)
Over-placement, performance	12.718** (2.120)	Gambling preference, lotteries	-1.882 (-0.402)
Financial literacy, dummy	7.161* (1.687)	Sensation seeking, novelty	8.506* (1.651)
Over-placement, literacy	-5.285 (-1.183)	Sensation seeking, volatility	5.671 (1.207)
Miscalibration of uncertainty	-3.009 (-0.723)	Perceived information advantage	14.511*** (2.642)
Do not consider transaction cost	-5.980 (-1.506)	Dismissive of others' information	4.016 (1.059)
Underestimation of transaction cost	-2.111 (-0.513)	Social influence	-7.979 (-1.499)
Do not think bid-ask spread is a cost	-9.364** (-2.438)	Advisor influence	-12.094 (-1.632)
Extrapolation, up	0.306 (0.059)	Portfolio rebalancing needs	8.706 (1.558)
Extrapolation, down	-2.718 (-0.555)	Liquidity	-6.778 (-1.111)
Realization utility, winners	5.039 (1.246)	Risk aversion	-2.313 (-0.506)
Realization utility, losers	-2.224	Expected future one-year market return	0.514

	(-0.490)		(1.311)
Gender: male	22.262***	Branch fixed effects	YES
	(5.998)	N	4,648
Controls	YES	R2	0.177

t-statistics in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A8 Horse Race with Branch Fixed Effects

9.2. Bootstrapped Standard Errors

Table A9 reports the results for the horse race using bootstrapped standard errors. Significance is virtually unchanged.

Dependent Variable: Average Monthly Turnover Ratio (%) (October 2018 – June 2019)			
Actual performance in 2017	4.198*** (5.314)	Gambling preference, blockbusters	11.764*** (2.808)
Overplacement, performance	11.549** (2.018)	Gambling preference, lotteries	-1.159 (-0.265)
Financial literacy, dummy	7.065* (1.829)	Sensation seeking, novelty	6.598 (1.356)
Overplacement, literacy	-2.621 (-0.643)	Sensation seeking, volatility	3.632 (0.824)
Miscalibration of uncertainty	-2.989 (-0.827)	Perceived information advantage	15.660*** (2.973)
Do not consider trading costs	-3.989 (-1.168)	Dismissive of others' information	2.942 (0.822)
Underestimation of trading costs	-4.029 (-0.987)	Social influence	-7.839 (-1.496)
Do not think bid-ask spread is a cost	-9.456*** (-2.667)	Advisor influence	-12.089* (-1.896)
Extrapolation, up	-1.255 (-0.256)	Portfolio rebalancing needs	12.571** (2.285)
Extrapolation, down	-1.208 (-0.253)	Liquidity needs	-7.651 (-1.247)
Realization utility, winners	7.049* (1.824)	Risk Aversion	-2.943 (-0.743)
Realization utility, losers	-2.321 (-0.468)	Expected 1-year market return	0.709** (2.274)
Gender: male	21.488*** (6.236)	Controls	YES
		N	4,648
		R2	0.089

t-statistics in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A9 Horse Race with Bootstrapped Standard Errors

9.3. Subsample Analysis

9.3.1. Small investors

Table A10 reports the horse-race results for small investors, that is, investors with a maximum account balance of less than 100K RMB by the end of 2017. Again, gambling preference and perceived information advantage remain the most powerful explanatory variables for trading, and their coefficients increase in magnitude. This suggests that these factors are more powerful among small investors. Furthermore, over-placement based on performance becomes insignificant.

Dependent Variable: Average Monthly Turnover Ratio (%) (October 2018 – June 2019)			
Actual rank of 2017 performance	3.486*** (4.238)	Gambling preference, blockbusters	11.266*** (2.695)
Over-placement, performance	7.020 (1.202)	Gambling preference, lotteries	-1.097 (-0.240)
Financial literacy, dummy	6.339 (1.564)	Sensation seeking, novelty	9.929** (1.974)
Over-placement, literacy	-3.591 (-0.828)	Sensation seeking, volatility	0.677 (0.149)
Miscalibration of uncertainty	-2.434 (-0.603)	Perceived information advantage	20.687*** (3.705)
Do not consider transaction costs	-4.714 (-1.224)	Dismissive of others' information	3.099 (0.819)
Underestimation of transaction costs	-3.770 (-0.946)	Social influence	-9.454** (-1.973)
Do not think bid-ask spread is a cost	-7.993** (-2.175)	Advisor influence	-10.549* (-1.666)
Extrapolation, up	-2.014 (-0.391)	Portfolio rebalancing needs	8.959 (1.598)
Extrapolation, down	0.055 (0.011)	Liquidity needs	-6.621 (-1.127)
Realization utility, winners	8.877** (2.253)	Risk aversion	-4.704 (-1.066)
Realization utility, losers	-3.351 (-0.752)	Expected future one-year market return	0.771** (1.981)
Gender: male	23.338*** (6.481)	Controls	YES
		N	4,241
		R ²	0.094

t-statistics in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A10 Horse Race for Small Investors

9.3.2. Large investors

Table A11 reports the horse-race results for large investors, that is, investors with a maximum account balance greater than 100K RMB by the end of 2017. Unlike in the sample of small investors, in this sample of large investors, the most significant explanatory variables are overconfidence based on performance and portfolio rebalancing needs.

Dependent Variable: Average Monthly Turnover Ratio (%) (October 2018 – June 2019)			
Actual rank of 2017 performance	10.499*** (3.017)	Gambling preference, blockbusters	23.300 (1.517)
Over-placement, performance	52.447** (2.589)	Gambling preference, lotteries	-6.862 (-0.372)
Financial literacy, dummy	26.847* (1.722)	Sensation seeking, novelty	-26.370 (-1.335)
Over-placement, literacy	11.502 (0.586)	Sensation seeking, volatility	32.999* (1.734)
Miscalibration of uncertainty	-9.376 (-0.563)	Perceived information advantage	-29.456* (-1.873)
Do not consider transaction cost	-5.314 (-0.370)	Dismissive of others' information	6.487 (0.495)
Underestimation of transaction cost	-2.900 (-0.219)	Social influence	17.155 (0.619)
Do not think bid-ask spread is a cost	-29.042* (-1.795)	Advisor influence	-12.954 (-0.413)
Extrapolation, up	0.645 (0.033)	Portfolio rebalancing needs	45.272** (2.044)
Extrapolation, down	-12.854 (-0.684)	Liquidity needs	-37.934 (-1.453)
Realization utility, winners	-11.929 (-0.749)	Risk aversion	7.232 (0.486)
Realization utility, losers	2.456 (0.134)	Expected future one-year market return	0.812 (0.567)
Gender, male	-4.092 (0.278)	Controls	YES
		N	407
		R2	0.165

t-statistics in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A11 Horse Race for Large Investors

9.3.3. More Invested Investors

We next split the main sample into two halves based on how much of an investor's net worth is invested in the stock market. For each investor, we first calculate the ratio of her entire portfolio

value at the Shenzhen Stock Exchange to her reported net worth. This, however, does not take into account her investment at the Shanghai Stock Exchange or mutual fund holdings. We then sort all investors into two groups of equal size, where investors with the ratio above (below) the median are considered more (less) invested in the stock market. Table A12 reports the horse race results for the more invested investors. A notable observation is that, in this subsample, portfolio rebalancing becomes a more important trading motive. This is consistent with these investors having to move money in and out of the stock market more frequently. Another observation is that, due to a smaller sample size, perceived information advantage is no longer significant, although the magnitude remains large. Gambling preference remains significant.

Dependent Variable: Average Monthly Turnover Ratio (%) (October 2018 – June 2019)			
Actual rank of 2017 performance	6.330*** (4.937)	Gambling preference, blockbusters	14.076** (2.264)
Over-placement, performance	16.820** (1.990)	Gambling preference, lotteries	-3.279 (-0.476)
Financial literacy, dummy	7.378 (1.207)	Sensation seeking, novelty	6.600 (0.853)
Over-placement, literacy	-5.092 (-0.777)	Sensation seeking, volatility	-2.396 (-0.343)
Miscalibration of uncertainty	-2.498 (-0.407)	Perceived information advantage	10.652 (1.319)
Do not consider transaction cost	2.893 (0.502)	Dismissive of others' information	3.024 (0.517)
Underestimation of transaction cost	-6.834 (-1.116)	Social influence	-1.485 (-0.178)
Do not think bid-ask spread is a cost	-12.834** (-2.277)	Advisor influence	-4.972 (-0.432)
Extrapolation, up	1.858 (0.234)	Portfolio rebalancing needs	30.106*** (3.228)
Extrapolation, down	-5.378 (-0.737)	Liquidity needs	-16.607* (-1.850)
Realization utility, winners	9.048 (1.499)	Risk aversion	-4.485 (-0.706)
Realization utility, losers	3.985 (0.559)	Expected future one-year market return	0.661 (1.191)
Gender, male	28.625*** (5.070)	Controls	YES
		N	2,321
		R2	0.075

t-statistics in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A12 Horse Race for Highly Invested Investors

9.3.4. Less Invested investors

Table A13 reports the horse race results for the less invested investors. In this subsample, perceived information advantage is more important, while gambling preference remains significant. However, portfolio rebalancing becomes an insignificant factor in explaining trading.

Dependent Variable: Average Monthly Turnover Ratio (%) (October 2018 – June 2019)			
Actual rank of 2017 performance	2.038** (2.010)	Gambling preference, blockbusters	10.746** (2.145)
Over-placement, performance	2.517 (0.342)	Gambling preference, lotteries	-0.619 (-0.114)
Financial literacy, dummy	6.151 (1.231)	Sensation seeking, novelty	7.160 (1.251)
Over-placement, literacy	0.289 (0.054)	Sensation seeking, volatility	7.705 (1.412)
Miscalibration of uncertainty	-2.247 (-0.452)	Perceived information advantage	17.560*** (2.607)
Do not consider transaction cost	-12.097** (-2.576)	Dismissive of others' information	3.420 (0.770)
Underestimation of transaction cost	-1.534 (-0.329)	Social influence	-15.371*** (-2.925)
Do not think bid-ask spread is a cost	-7.940* (-1.816)	Advisor influence	-15.509** (-2.522)
Extrapolation, up	-6.465 (-1.095)	Portfolio rebalancing needs	-2.057 (-0.333)
Extrapolation, down	4.402 (0.758)	Liquidity needs	0.166 (0.023)
Realization utility, winners	4.430 (0.944)	Risk aversion	-0.416 (-0.077)
Realization utility, losers	-9.229* (-1.805)	Expected future one-year market return	0.837* (1.697)
Gender, male	14.740*** (3.449)	Controls	YES
		N	2,327
		R ²	0.106

t-statistics in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A13 Horse Race for Less Invested Investors

9.4. Other Robustness Checks

Finally, we consider four additional robustness checks. First, we consider an alternative measure of turnover, where, instead of value-weighting monthly turnover using monthly balance,

we equal-weight monthly turnover. Results are reported in Table A14. Second, we consider including investors with zero holdings at the Shenzhen Stock Exchange, and their turnovers are coded as zeros. Results are reported in Table A15. Third, we consider restricting to “active” investors, who were actively trading prior to the survey. Results are reported in Table A16. Fourth, we consider measuring turnover using transactions from January 2018 through September 2018, and results are presented in Table A17. In all four sets of analyses, gambling preference and perceived information advantage remain the most significant and powerful variables for trading.

Dependent Variable: Average Monthly Turnover Ratio (%), Equal-Weighted (October 2018 – June 2019)			
Actual rank of 2017 performance	2.928*** (4.125)	Gambling preference, blockbusters	10.715*** (2.994)
Over-placement, performance	9.420* (1.890)	Gambling preference, lotteries	-3.053 (-0.791)
Financial literacy, dummy	5.658* (1.650)	Sensation seeking, novelty	6.868 (1.589)
Over-placement, literacy	-3.750 (-1.043)	Sensation seeking, volatility	4.250 (1.085)
Miscalibration of uncertainty	-2.915 (-0.839)	Perceived information advantage	9.221** (2.013)
Do not consider transaction cost	-3.380 (-1.021)	Dismissive of others' information	2.805 (0.870)
Underestimation of transaction cost	-1.707 (-0.505)	Social influence	-5.761 (-1.325)
Do not think bid-ask spread is a cost	-8.431*** (-2.677)	Advisor influence	-7.642 (-1.367)
Extrapolation, up	-1.613 (-0.368)	Portfolio rebalancing needs	10.590** (2.212)
Extrapolation, down	-1.308 (-0.322)	Liquidity needs	-7.077 (-1.433)
Realization utility, winners	5.651* (1.668)	Risk aversion	-1.457 (-0.381)
Realization utility, losers	-2.930 (-0.773)	Expected future one-year market return	0.582* (1.775)
Gender, male	18.187*** (5.889)	Controls	YES
		N	4,648
		R ²	0.104

t-statistics in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A14 Using Equal-weighted Turnover as Dependent Variables

Dependent Variable: Average Monthly Turnover Ratio (%) (October 2018 – June 2019)			
Actual rank of 2017 performance	3.102*** (3.777)	Gambling preference, blockbusters	12.040*** (3.180)
Over-placement, performance	12.285** (1.978)	Gambling preference, lotteries	-4.415 (-1.073)
Financial literacy, dummy	1.566 (1.466)	Sensation seeking, novelty	6.264 (1.381)
Over-placement, literacy	-6.528* (-1.690)	Sensation seeking, volatility	5.351 (1.282)
Miscalibration of uncertainty	-3.109 (-0.845)	Perceived information advantage	14.163*** (2.801)
Do not consider transaction cost	-3.042 (-0.900)	Dismissive of others' information	-5.870 (-1.220)
Underestimation of transaction cost	-3.029 (-0.848)	Social influence	-10.383** (-2.327)
Do not think bid-ask spread is a cost	-8.722*** (-2.601)	Advisor influence	-7.451 (-1.264)
Extrapolation, up	-3.997 (-0.859)	Portfolio rebalancing needs	10.300** (2.009)
Extrapolation, down	0.911 (0.208)	Liquidity needs	-3.872 (-0.732)
Realization utility, winners	6.896* (1.931)	Risk aversion	-4.405 (-1.106)
Realization utility, losers	-1.569 (-0.385)	Expected future one-year market return	0.421 (1.194)
Gender, male	20.732*** (6.302)	Controls	YES
		R2	0.101
		N	4,937

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A15 Horse Race When Including Accounts with Zero Holdings

Dependent Variable: Average Monthly Turnover Ratio (%) (October 2018 – June 2019)			
Actual rank of 2017 performance	4.989*** (5.726)	Gambling preference, blockbusters	12.148*** (2.842)
Over-placement, performance	11.929* (1.937)	Gambling preference, lotteries	-3.846 (-0.831)
Financial literacy, dummy	8.265** (1.979)	Sensation seeking, novelty	6.538 (1.279)
Over-placement, literacy	-3.450 (-0.779)	Sensation seeking, volatility	5.366 (1.155)
Miscalibration of uncertainty	-2.884 (-0.697)	Perceived information advantage	15.954*** (2.889)
Do not consider transaction cost	-4.334 (-1.107)	Dismissive of others' information	1.897 (0.487)
Underestimation of transaction cost	-3.570 (-0.878)	Social influence	-8.870* (-1.733)
Do not think bid-ask spread is a cost	-10.395*** (-2.765)	Advisor influence	-11.379* (-1.744)
Extrapolation, up	-1.378 (-0.261)	Portfolio rebalancing needs	13.537*** (2.332)
Extrapolation, down	0.306 (0.062)	Liquidity needs	-5.889 (-0.966)
Realization utility, winners	7.171* (1.763)	Risk aversion	-3.349 (-0.742)
Realization utility, losers	-4.156 (-0.908)	Expected future one-year market return	0.740* (1.877)
Gender, male	23.440*** (6.306)	Controls	YES
		R ²	0.066
		N	4,398

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A16 Horse Race Among Active Investors

Dependent Variable: Average Monthly Turnover Ratio (%) (January 2018 – September 2018)			
Actual rank of 2017 performance	6.476*** (8.726)	Gambling preference, blockbusters	10.812*** (2.837)
Over-placement, performance	14.416*** (2.722)	Gambling preference, lotteries	-2.568 (-0.624)
Financial literacy, dummy	4.260 (1.201)	Sensation seeking, novelty	-1.018 (-0.237)
Over-placement, literacy	-4.472 (-1.190)	Sensation seeking, volatility	6.988* (1.717)
Miscalibration of uncertainty	-6.911* (-1.845)	Perceived information advantage	11.644*** (2.476)
Do not consider transaction cost	-1.993 (-0.578)	Dismissive of others' information	5.925* (1.744)
Underestimation of transaction cost	-4.171 (-1.146)	Social influence	-13.520*** (-3.328)
Do not think bid-ask spread is a cost	-10.033*** (-3.059)	Advisor influence	-8.331 (-1.563)
Extrapolation, up	-3.037 (-0.715)	Portfolio rebalancing needs	8.235* (1.654)
Extrapolation, down	5.305 (1.304)	Liquidity needs	-5.508 (-1.006)
Realization utility, winners	4.057 (1.143)	Risk aversion	-2.366 (-0.605)
Realization utility, losers	-2.898 (-0.709)	Expected future one-year market return	0.587* (1.713)
Gender, male	19.623*** (5.991)	Controls	YES
		R ²	0.079
		N	4,648

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A17 Horse Race Using Pre-Survey Turnover

10. Gambling Preferences and Characteristics of Stocks Purchased

In the main text of the paper, we show that investors with a survey-based gambling preference tend to trade more and perform worse than others. In this section, we examine the characteristics of stocks purchased by different investors. Table A18 shows that investors with a survey-based gambling preference tend to buy stocks that are smaller, have a larger market beta, and have larger counts of daily up-limit hits, and higher past volatility and past returns. These stocks also perform worse subsequently, confirming that investors with a gambling preference trade in the wrong direction and their trading is excessive.

	(October 2018 – June 2019)						
	Past 30-day # of Up-limit Hits	Past 30-day Return Volatility (%)	Past 30-day Return (%)	Size (Billion RMB)	Beta	B/M	Future 30-day Return (%)
1. Strongly disagree	0.6	3.25	9.71	43.73	0.93	0.62	-0.03
2. Disagree	0.75	3.39	11.58	35.21	0.96	0.62	-0.87
3. Neutral	0.83	3.49	11.94	26.92	0.99	0.61	-1.53
4. Agree	0.89	3.56	12.45	26.29	1.00	0.61	-1.36
5. Strongly agree	0.92	3.55	12.74	26.65	1.02	0.62	-1.77
5 – 1	0.32***	0.30***	3.03**	-17.08**	0.09***	0.00	-1.74**

Table A18 Trading Characteristics for Investors Sorted on Gambling Preference, Blockbusters

11. A “Nudge” Experiment

Among all 500 brokerage branches we distributed the survey to, we randomly selected 250 branches to include an additional “nudge.” The “nudge” asked the respondent to read a short article that highlighted the negative consequences of excessive trading. As shown in Figure A3, the article contained a detailed calculation of how much investors lose from frequent trading along with a quote from Warren Buffett advising investors to buy and hold. Instead of presenting trading costs as a fraction of total transaction value, we made it more salient by presenting the annualized fee rate for a frequent trader. We also included a “validation” question after the article by asking the respondent to calculate the total trading costs of a given level of turnover. Answers to this question help identify those who have actually read the article and therefore were treated.



Figure A3 The “Nudge” to Reduce Trading Due to Transaction Costs

We study the effect of this “nudge” in a difference-in-difference framework; Table A19 reports the results. Column (1) shows that the interaction term is small and insignificant, suggesting that the treatment and control groups exhibit similar turnover rates one month after the survey. We repeat this exercise in columns (2) and (3) by expanding the window to three months and six months before and after the survey, and the interaction term remains insignificant. Overall, these results suggest that the nudge had no effect on reducing trading. One might argue that the “nudge” was not sufficiently strong and the treated group may not have read the article carefully. However,

we identify an investor as treated only if they were in the treated group and answered the “validation” question correctly.

Turnover Around the Survey (%)			
	1-month window	3-month window	6-month window
	(1)	(2)	(3)
<i>After*Treated</i>	0.672 (0.119)	-5.971 (-0.944)	-4.417 (-0.675)
<i>Treated</i>	-0.219 (-0.053)	4.153 (0.911)	0.583 (0.130)
<i>After</i>	-2.858 (-0.956)	-1.012 (-0.305)	16.144*** (4.612)
Controls	YES	YES	YES
<i>R</i> ²	0.056	0.058	0.056
N	6,628	6,628	6,628

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A19: Comparing Turnover Before and After the Survey for the Control and Treatment Groups

Note: Before distributing the survey, we randomly assigned 500 targeted branches of brokerage firms into treated and control groups. Investors in the two groups received questionnaires that were otherwise identical except for one difference: the questionnaire for the treated group included a “nudge” that highlighted the negative consequences of excessive trading. In this table, we study the effect of the “nudge” on investors’ trading frequencies using difference-in-difference tests. The dependent variables from Columns (1) to (3) are investors’ average monthly turnover rates in the one month, three months, and six months before and after the survey. The dummy, *Treated*, equals one if an investor is in the treated group and correctly answered the follow-up question designed to test whether the respondent understands the content of the message. The dummy, *Treated*, equals zero if an investor is in the control group. The dummy, *After*, equals one for the periods after the survey month and zero for the periods before or in the survey month (September 2018). Control variables include age, gender, wealth, income, trading experience, account size, and education. T-statistics are based on robust standard errors and are reported in parentheses. See the Table A2 for more details about variable definitions.