I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17537

Are Women Blamed More for Giving Incorrect Financial Advice?

Martin Abel
Emma Bomfim
Izzy Cisneros
Jackson Coyle
Song Eraou
Martha Gebeyehu
Gerardo Hernandez
Julian Juantorena
Lizzy Kaplan
Danielle Marquez
Jack Mullen
Peyton Mulhern
Ayana Opong-Nyantekyi
Rin Osathanugrah
Joe Paul
Austin Philie
Luke Tingley
Jingyi Wang

DECEMBER 2024

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17537

Are Women Blamed More for Giving Incorrect Financial Advice?

Martin Abel
Bowdoin College, JPAL and IZA

Emma Bomfim, Izzy Cisneros, Jackson Coyle, Song Eraou, Martha Gebeyehu, Gerardo Hernandez, Julian Juantorena, Lizzy Kaplan, Danielle Marquez, Jack Mullen, Peyton Mulhern, Ayana Opong-Nyantekyi, Rin Osathanugrah, Joe Paul, Austin Philie, Luke Tingley, Jingyi Wang
Bowdoin College

DECEMBER 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Are Women Blamed More for Giving Incorrect Financial Advice?*

We conduct an incentivized experiment with a nationally representative sample to investigate gender discrimination among people receiving advice on risky investments. Participants learn about actual start-up firms they can invest in. Before deciding how much of their endowment to invest, they receive recommendations from either female or male professionals. We find that before outcomes are revealed, participants are equally likely to follow recommendations of female and male advisors. Likewise, we observe no gender discrimination following advice that proves correct. However, for advice that turns out to be incorrect, advisor gender significantly impacts the decisions made by male participants. They invest 47% less in the direction of this advice compared to situations where male advisors were incorrect. These differences are not explained by participants' stated views on gender roles and advisors' ability as well as the level of attention towards female advisors.

JEL Classification: J70, G11
Keywords: gender discrimination, investment decisions, financial advising

Corresponding author:
Martin Abel
Bowdoin College
255 Maine Street
Brunswick
Maine 04011
USA
E-mail: m.abel@bowdoin.edu

* This paper greatly benefited from comments by Willa Brown, Jeff Carpenter, Dan Stone, and participants at various conferences and seminars. The experiment was registered under registry number AEARCTR-0011303. IRB approval and funding was obtained from Bowdoin College. All errors and omissions are our own. Declarations of interest: none.

1 Introduction

Women remain underrepresented in leadership positions across various segments of the private sector. For example, despite comprising 46 percent of the financial sector's workforce, women hold only 15 percent of executive roles in finance, and fewer than 2 percent of CEO positions in financial institutions (Schwab et al., 2017; Sahay and Cihak, 2018). In addition, women make up only 23.6 percent of Certified Financial Planner professionals, despite evidence showing no gender differences in the performance or investment behavior of mutual fund managers (Atkinson et al., 2003; Ellingrud et al., 2021). Our study explores gender discrimination in responses to advice on risky investment projects as a potential factor contributing to the underrepresentation of women in finance. Specifically, we examine gender differences in the extent to which investors follow advice and whether female advisors receive more blame if their recommendations prove incorrect.

To study these questions, we recruit a nationally representative sample of 1,150 participants via Prolific. Each participant plays a total of four investment rounds. In each round, they first watch a short video about a real start-up company seeking investments and are then asked their beliefs about whether this firm will be successful. Next, an advisor gives participants a binary investment recommendation as well as a short justification for their advice. These advisors are actual people with extensive private sector experience who also watched the videos and had additional information about each firm. Participants are then asked for their updated belief in the company's success and decide how much of their endowment of 100 cents to invest in the company. The outcome is then revealed. If the company failed, the investment is lost. If it is successful, the investment is doubled. The advisor assigned to each participant switches after completing two of the four investment rounds.

Our experimental design features three levels of randomization. First, we randomize the gender of the assigned advisor. Participants are informed that the advisor's identity is not disclosed for privacy reasons. Instead, participants are presented with a picture of their advisor that *"closely matches their demographic profile (gender, age, race)"*. To control for factors other than gender, we employ computer-generated photos of siblings and randomly assign people to see either the male or female version of the picture. We further conduct an out-of-sample assessment of these photos and select sibling pairs so that key characteristics

such as perceptions of age, attractiveness, and how artificial the picture looks are balanced across advisor gender. Second, while every advisor gives one correct and one incorrect recommendation, we randomize the order in which they receive the correct or incorrect advice. This allows us to test the effect of (in)correct advice on behavior in the subsequent investment round. Third, the randomly assigned two projects under a given advisor either both fail or both succeed, which allows us to test if female advisors are assessed differently based on investment success, holding the correctness of advice constant.

We present three findings, following a registered pre-analysis plan. First, we find no evidence of gender discrimination when we only look at decisions *before* outcomes are revealed. These estimates are precisely estimated, allowing us to rule out advisor gender differences of less than 1.53 cents, equivalent to 0.03 standard deviations (sd).

Second, we find important differences across advisor gender depending on the outcome of the risky investment. If advice turns out to be *correct*, both female and male advisors get a similar boost in investors following their recommendations in the subsequent investment decision. When advice proves *incorrect*, investors are less likely to follow advisors' recommendations in the next round. However, key differences emerge in how women and men react to incorrect advice from female advisors. Female investors update their beliefs and invest amounts independent of advisor gender. Male participants change their investment amount 47% (5.5 cents) *less* in the direction recommended by a female advisor compared to a male advisor. This difference in how female and male participants react to incorrect recommendations from female advisors is both statistically significant and large in magnitude. Notably, following incorrect advice there is also a significant difference between how male participants invest and how they change their stated beliefs of investment success, which they update by almost 38% (2.5 pp) *further* in the direction of the advice by female advisors.

Third, we investigate underlying mechanisms for our results. Following our pre-analysis plan, we explore the role of investors' attention paid to advisors, their perceptions of advisors' ability and their views on gender roles and gender quotas. Attention discrimination, documented in other settings (e.g. Bartos et al. (2016)), cannot explain our findings. We find no differences in the attention paid to female advisors, measured by the time participants spent reading the recommendations, in the first investment round. After receiving incorrect advice, both female and male participants spend about 0.65 seconds (6.4%) *more* attention

to female compared to male advisors.

Participants do not evaluate female advisors more negatively than male advisors. Point estimates for the effect of female advisors on a standardized assessment index are positive but insignificant, and we can rule out that advisor gender differences are smaller than 0.035 sd. Moreover, in a post-experiment survey, participants rate gender discrimination as the most important reason for the gender gap among executives, while a lack of skills was rated as the least important. These results are consistent with recent evidence from Bohren et al. (2019) and Ayalew et al. (2021), showing that, at a certain level of seniority, women are perceived as having *higher* ability because people recognize their need to overcome discrimination to obtain these positions. However, one limitation of this test is that we only collect assessments after both rounds with a given advisor are complete and thus cannot isolate the effect of incorrect advice.

In addition to general ability levels, gender stereotypes about the ability to perform in specific roles may also be an important factor in explaining investor behavior, given that finance is a male-dominated field (Eagly and Karau, 2002).[1] To test this hypothesis, we ask participants whether they agree that *"Men and women have the same natural abilities and strengths."* and whether there are average differences in "*how well [women and men] are suited for certain types of work*". While we find substantial variation in people's responses and that male investors hold more gendered views about ability, we do not find that discrimination following incorrect advice differs across investors' gender views.

In contrast, we find that gender discrimination is limited to participants who are less supportive of gender quotas. Studies link resistance to gender quotas with holding sexist beliefs (Kane and Whipkey, 2009; Krook, 2006) and show that the introduction of gender quotas can contribute to gender discrimination (Leibbrandt et al., 2018). However, our data does not allow us to determine the exact mechanism behind these results. These findings nevertheless suggest that consciously held attitudes, distinct from beliefs about ability or stereotypes, play a role in explaining gender discrimination in our setting.

[1] An extensive literature studies the importance of gender role congruity in explaining discrimination. A meta analysis by Eagly et al. (1992) finds that devaluation of women is more prevalent in male-dominated roles and when evaluators are men. Recent studies (e.g. Bordalo et al. (2019); Coffman et al. (2021)) have documented the importance of stereotypes related to task-specific roles.

Our paper relates to a set of recent papers that study the role of gender in advice-giving.[2] Cooper and Kagel (2016) show that women are reluctant to contribute ideas in gender incongruent fields due to a lack of confidence. Studies also demonstrate that people are often less likely to follow the advice of women. Ayalew et al. (2021) find that people are less likely to follow the advice of female leaders on how to solve a task unless they learn that this person is highly qualified. A set of papers also explores the impact of advisor gender in investment settings (Klein and Shtudiner, 2016; Klein et al., 2021). Most closely related to our study, Klein et al. (2021) present study participants in Israel with a binary investment recommendation from either a female or male advisor. Their findings indicate that the gender of the advisor plays a crucial role, with participants being more inclined to follow recommendations of male advisors. In line with our results, this preference for male advisors is strongest among male investors.

We expand on these studies in several important ways: rather than studying binary advice from fictitious advisors, we utilize more detailed recommendations provided by actual professionals. And instead of using hypothetical investment projects, we utilize real (past) start-up firms seeking funding, which enables us to incentivize investment behavior on actual outcomes. Furthermore, participants learn about these firms and we collect their beliefs before receiving advice, allowing us to measure the effect of advisor gender on belief updating. We also intentionally collect both incentivized (investment-related) and unincentivized (belief-related) measures. That fact that discrimination is limited to investment outcomes underscores that relying on metrics that do not affect participants' payouts can potentially yield misleading evidence concerning discrimination. Last, we collect additional data on investors' attitudes and attention to test potential mechanisms proposed in the gender discrimination literature.[3]

We contribute to a second strand of literature on gender biases in the attribution of failure, which may present another reason why women shy away from advising roles or other

[2] A separate strand of literature studies the role of gender in advice-taking. While some studies find gender differences, e.g. in the context of competition entry (Brandts et al., 2015), a recent meta-analysis by Bailey et al. (2023) concludes that there are no systematic differences in advice-taking across gender.

[3] In addition, we study gender discrimination in the United States while Klein et al. (2021) conduct their study in Israel. This replication across contexts is important given that cultural differences may affect gender discrimination. For example, Abel (2022) finds evidence of gender discrimination against female managers in the United States, while Abel and Buchman (2023) do not find that workers discriminate in a replication of this study in India.

positions that include decision making with risky outcomes (Chakraborty and Serra, 2023). Sarsons (2019) finds that people lower their appraisal of female surgeons' ability more than their male counterparts after a bad patient outcome. Along similar lines, we find that (male) investors penalize women more for giving incorrect advice. Interestingly, Erkal et al. (2023) find that people are *less* likely to blame women for failed outcomes that involve prosocial tasks. These findings suggest that gender bias in the attribution of failure depends on how much the activities are associated with a certain gender.

We acknowledge some potential limitations of our design. First, while participants were able to double their overall compensation for participating in the experiment, the investment amounts in our game are small due to budget constraints. Although recent evidence suggests that behavior in experiments is not sensitive to the incentive amount (Pulford et al., 2018), this may limit the external validity of our results. Second, as with any study on discrimination, there are concerns about people hiding discriminatory behavior. While we cannot rule out this possibility, we find two pieces of evidence against reporting biases affecting our results: less than 10% of investors suspect the study is about gender and results are similar across a measure of participants' tendency to give socially desirable answers (Crowne and Marlowe, 1960). Third, we present evidence against the importance of attention discrimination and gender stereotypes in our setting. However, due to data limitations, we can only test a subset of potential mechanisms that may explain why male investors discriminate against female advisors following incorrect advice.

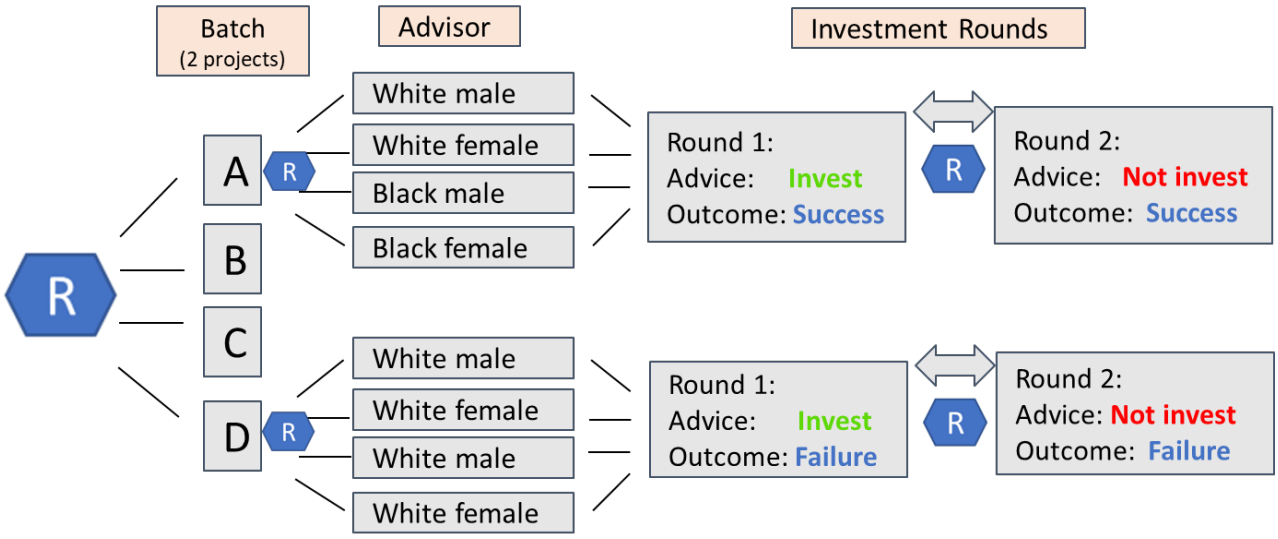
Our paper proceeds as follows: Section 2 describes our study design and empirical strategy. Section 3 reports our main results and Section 4 explores mechanisms. Section 5 concludes.

2 Study Design

This section describes the experimental design, summarized in Figure 1, and discusses the empirical strategy. Our design features three levels of randomization. First, participants are randomized to batches of two investment projects described in section 2.1. Second, they are randomly assigned to a real-world advisor who recommends to either invest or not invest (Section 2.2). Advisor identity is communicated via computer-generated images (Section

2.3). For each advisor, one of the recommendations is correct and one is incorrect. The order of advice correctness is randomized.

Figure 1: Randomization Design



2.1 Investment Project Selection

For the investment projects included in our study, we select eight start-up firms that were launched between 2016 and 2019. We identified potential firms through Y Combinator, a startup accelerator company, and Kickstarter, a crowdfunding platform. To mitigate concerns that investors systematically attribute domain-specific expertise to female or male advisors, we select firms across a range of industries including technology, health care, and food.[4] The success of the firm is determined based on the operating status of the company in 2023. Four companies were still active and thus considered successful and four were not active anymore and thus considered failed.

We create videos introducing each start-up firm. These videos are based on promotional footage provided by the firm, similar to what is used to raise investments on these sites,

[4]Our final selection include a health app, an instant coffee company, a camera accessory, an interactive boxing movement tracker, a smart oven, a device to guarantee internet anonymity, and a sleep monitoring gadget.

and include a basic description of the product and its (comparative) advantages. We edited videos to be 45 to 60 seconds long and to exclude details of the startup companies' founders. In exploratory work, we confirmed that people's perceived success probabilities for these firms range from 30 to 70%. This ambiguity is important as it mimics real-world situations in which people rely on the advice of outside experts for their financial decisions.

2.2 Advisor Recruitment

We recruit a sample of 60 people through Cloud Research to act as advisors. To boost their credibility as financial advisors, we restrict the sample to people holding at least a four-year college degree and having at least ten years of private sector experience. To create an information advantage for advisors, we provide them with details about their product collected from the firms' website (in addition to watching the video). Advisors are then asked to provide a binary investment recommendation (invest / not invest) as well as a justification for their advice.[5]

While providing justifications for investment advice is an important aspect of real-world advice giving, one crucial concern with including it in experimental studies is that it is impossible to hold the exact wording constant. This complicates isolating the role of advisor gender as female and male advisors may communicate differently (Manian and Sheth, 2021). Even subtle differences in the choice of wording may affect investors' decisions. We therefore follow Brandts and Rott (2021) and provide advisors with a list of three (pre-scripted) potential justifications for each recommendation. These are one paragraph long and refer to the product quality, competitors, or market demand (see Figure B5 for examples). Advisors still have the option to provide an open-ended response if their justification does not align with any of the pre-scripted options. For each of the eight projects included in our study, described in more detail below, we match advisors of varying gender that give both identical advice and justifications.

[5]We informed participants how we would use their advice in the research projects and that we reveal some of their demographic characteristics. To mimic real-world situations, advisors' payments did not depend on whether the advice was correct or not.

2.3 Advisor Images

One key challenge in studying discrimination is how to clearly signal the aspect of identity one is investigating (e.g. gender). Existing studies either list gender explicitly (e.g. Klein et al. (2021)) or signal gender through first names (e.g. Chakraborty and Serra (2023)). The former approach signals gender clearly but raises concerns about revealing the focus of the study while the latter strategy poses the challenge that people have associations with names other than gender (Abel and Burger, 2023). Other studies, such as Barron et al. (2022), use gendered icons which is unambiguous but by design lacks realism.[6]

We use a novel strategy to signal gender via computer-generated images of siblings. In addition to standardizing details like facial expression and direction of the gaze, these images also hold many of the facial features constant within a sibling pair. To further ensure that images are perceived similarly, we conduct an out-of-sample data collection with 70 people recruited via Cloud Research in which we let people rate the perceived age, attractiveness, intelligence, and whether the image is realistic. We select "sibling pairs" so that these perceptions are balanced across advisor gender. Our final selection of advisors, displayed in Figure 2, includes six sibling pairs, four white pairs and two pairs of people of color.

Figure 2: Advisor Pictures

Notes: Computer-generated images of advisors.

[6]Another clever strategy is to use the voice of participants to signal gender (Bordalo et al., 2019). However, this approach is only possible when study participants are in the same room.

2.4 Participant Recruitment and Sample Description

With the notable exception of Klein et al. (2021), who collect a nationally representative sample in Israel, most related studies use convenience samples that tend to be skewed towards younger and white participants. To address this concern, we recruit a sample of 1,150 participants through Prolific that is nationally representative with respect to key demographics such as gender, age, and race. We inform people that the study aims to "learn how people incorporate advice from others in their decision making" in the context of financial investments (see Figure B1 for details). Before consenting to participate, they further learn that the study is expected to take ten minutes and that they will earn a guaranteed payout of 1.50 USD plus an additional bonus, which depends on the outcomes of their investment decisions.

Table 1 describes the average characteristics of our sample. The average age is 45 years and the average income is around 45,000 USD. 13.3 percent of the respondents are Black, 6.1 percent are Asian, and 4.9 percent are Hispanic.[7] While representative of demographic characteristics, the sample is more liberal than the national population, with 54.4 percent identifying as liberal and 23 percent as conservative. In addition, participants have on average 14.9 years of education, which is above the national average. We also collect data on participants' attitudes. The average risk attitude is 4.2 on a 0 to 10 scale (with lower scores representing risk aversion). There is also substantial variation in views on gender quotas and race-based quotas with average support of 51 and 53 (on a 0 to 100 scale), respectively.

2.5 Investment Task

Investors first watch a short video about the firm, and we elicit their perceived "probability that this company is successful" (see Figure B4. On the next page, they see an image of the advisor and their recommendation. Figure 3 shows an example of advisor images displayed next to their recommendation. As part of the introduction we informed participants that advisors are actual people with at least a four-year college education and more than ten

[7] While the share of Black and Asian respondents is representative of the population, the share of Hispanics is smaller than in the population since the sample is representative of race and most Hispanics are categorized as white in the census.

Table 1: Summary Statistics and Randomization Balance

	(1) Male Advisor		(2) Female Advisor		(3) Total		T-test P-value
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)
Age	586	45.728 (0.654)	589	45.061 (0.646)	1175	45.394 (0.459)	0.468
Gender: Female	587	0.526 (0.021)	589	0.467 (0.021)	1176	0.497 (0.015)	0.041**
Race: Black	602	0.120 (0.013)	602	0.143 (0.014)	1204	0.131 (0.010)	0.232
Race: Asian	602	0.061 (0.010)	602	0.060 (0.010)	1204	0.061 (0.007)	0.904
Race: Hispanic	602	0.058 (0.010)	602	0.040 (0.008)	1204	0.049 (0.006)	0.142
Educ (Yrs)	585	14.843 (0.091)	589	14.919 (0.090)	1174	14.881 (0.064)	0.554
Income	585	45585 (1437)	582	44974 (1447)	1167	45281 (1019)	0.764
Risk seeking (0-10)	584	4.200 (0.104)	588	4.340 (0.105)	1172	4.270 (0.074)	0.345
Conservative	579	0.231 (0.018)	580	0.228 (0.017)	1159	0.230 (0.012)	0.876
Liberal	579	0.553 (0.021)	580	0.538 (0.021)	1159	0.545 (0.015)	0.615
Support Gender Quotas	583	51.2 (1.32)	583	50.2 (1.34)	1166	50.7 (0.942)	0.573
Support Race Quotas	581	53.9 (1.402)	586	51.4 (1.384)	1167	52.63 (0.985)	0.217
Social Desirability	585	3.01 (0.044)	588	3.00 (0.043)	1173	3.00 (0.031)	0.889
F-test of joint significance (p-value)							0.614

Notes: This table presents summary statistics of the full sample (3), and of subsamples of those assigned to a male (1) or female (2) advisors in the first investment round. The last column reports p-values from a test of equal means of characteristics across advisor gender.

years of experience (see Figure B3 for details). We collect data on the exact time people spend on this advice stage, which we use as our measure of attention.[8]

[8]The time data is collected through our Qualtrics survey. To increase the probability that people pay attention to the video and advice, we impose a minimum time requirement of 30 seconds for the video and 4 seconds for the advice before they can go to they can proceed to the next stage.

Figure 3: Advice and Investment Task

Notes: The graph shows an example of advisor images and recommendations as well as the post advice data collection of investment amounts and success beliefs. The participants will only see one advisor image per decision.

On the next page, we elicit a new measure of investors' (posterior) beliefs of the success probability using slider questions (Figure 3). The difference between the prior and posterior beliefs will serve as our measure of belief updating. On the same page, we also ask people how much of their endowment of 100 cents they choose to invest in the company (Figure 3). They can use any amount in increments of one and the slider starting position is at zero (for beliefs it was set at 50%). In contrast to beliefs, the investment amount was incentivized, which will be important for the interpretation of our results. On the next page, investors learn whether the project succeeded or failed and the consequences for payout given their investment decision (see Figure B6 for details). If the project succeeded, their investment amount is doubled. If it failed, they lose their investment. For each round, they receive a new endowment of 100 cents. The outcome of a randomly selected investment round determines their bonus payment. We explain this protocol before the first investment round and provide participants with specific examples of how investment decisions affect their payout (see Figure B2 for details).[9]

[9]One potential concern is that participants may conduct research during the experiment to find out status of the company. To test for this possibility, we compare how investment behavior varies depending on the

2.6 Randomization

As depicted in Figure 1, our experimental design features three aspects that we randomize. For the first two investment rounds, participants are randomized to one of four batches (A-D), each including two distinct investment projects. Within each batch, they are further randomized to one of four advisors, drawn from two advisor "sibling pairs" (see Figure 2).[10] These pairs are selected so that their recommendation and justification is identical for the two projects in that batch. For round three and four, participants are randomly assigned to two projects from a different batch and a new advisor. We further select advisors so that participants never see the same justification or advisor pair twice. Since advisor recommendations and justifications are matched for a given project, we can isolate the role of advisor gender. Specifically, we can do a within-project comparison of how participants' attention, beliefs and investments vary across assigned advisors using data from all four project rounds.

Each advisor in a given batch gives advice that is once incorrect and once correct. Batches further vary in whether the (in)correct advice recommends investing or not investing. We randomize the order of the correct and incorrect advice within a batch so that in round 2 and 4 half the sample received incorrect advice in the previous round. We can use data from these rounds to test the effect of previous (in)correct advice and whether the reaction varies across advisor gender.

The role of advisor gender may not only depend on whether or not the advice is correct, but also on investment outcomes. To test this hypothesis, we stratified the selection of batches so that each participant makes investments for two companies that failed and two that succeeded.[11] This allows us to test whether more or less blame is attributed to female advisors if an investment fails while holding the correctness of advice constant. For example, female advisors may receive harsher assessments when participants are frustrated because

time participants spend. We find no evidence that those who spend more than then median, above the 75th percentile or above the 90th percentile of time invest more (less) in successful (unsuccessful) projects compared to those below the median, which assuages this concern. In addition, we note that this would bias us *against* finding an effect of manager gender.

[10] These four advisors comprise of two "sibling pairs" of varying gender. Three batches have one Black and one white sibling pair and one batch has two white pairs.

[11] For example, those assigned to batch A first are then assigned to an advisor from batch D (and vice versa). Both projects in batch A succeeded while both projects in batch D failed.

they lost money.

2.7 Empirical Strategy

We can estimate the effect of advisor gender through the following specification:

$$y_{itp} = \beta_1 Fem_{it} + \gamma_i X_i + \lambda_p + \omega_t + \epsilon_{it} \tag{1}$$

y_{itp} measures outcome y of participant i for investment project p in round t. Our main outcomes, described in more detail below, measure participants' investment behavior, perceived probability of success, and attention. Fem_{it} represents a binary variable for whether the advisor in that round is female. X_i represents a vector of participant characteristics. We include a set of fixed effects for the randomly assigned project batch (ω_t) and the specific "sibling pair" of images we use for advisors (λ_p). The coefficient β_1 can thus be interpreted as the average effect of being assigned a female advisor, holding participant characteristics, investment project, advisor image characteristics (including race), and advice content constant. We cluster standard errors at the participant level.

To test if the effect of correct advice differs across advisor gender, we estimate:

$$y_{itp} = \beta_1 Fem_{it} + \beta_2 Fem_{it} x Corr_{t-1} + \beta_3 Corr_{t-1} + \gamma_i X_i + \lambda_p + \omega_t + \epsilon_{it} \tag{2}$$

$Corr_{t-1}$ is a binary variable measuring whether the advice in the previous round was correct. Coefficient β_3 thus captures the effect of correct advice for male advisors and β_2 measures the *additional* effect of correct advice for female advisors. The sample in these regressions is restricted to the second investment round with a given advisor.

To test whether the effect of advisor gender varies across participant characteristics, we estimate:

$$y_{itp} = \beta_1 Fem_{it} + \beta_2 Fem_{it} x S_i + \beta_3 S_i + \gamma_i X_i + \lambda_p + \omega_t + \epsilon_{it} \tag{3}$$

S_i refers to a group participant i belongs to. Coefficient β_2 measures whether the effect of advisor gender varies for this group. Our pre-analysis plan specifies two primary subgroups of interest: participant gender and whether participants believe that women and men have different abilities.[12] To test whether the interaction of advice correctness and advisor gender differs for subgroups, we estimate equation 2 separately in subgroups and test whether the difference in coefficients is significant across groups.

3 Results

This section presents results on the effect of advisor gender on investment behavior. Section 3.1 describes the effect of having a female advisor in the pooled data and separately by investment round. Next, Section 3.2 reports the effect of advice correctness and whether recommendations of female advisors are followed differently after incorrect advice. Section 3.3 explores how these patterns vary across participant gender.

3.1 Pooled Results

Table 2 presents the main results from specification 1 on the effect of advisor gender and the role of advice correctness. Columns 1 to 4 explore how much advice affects investment decisions. We multiply decisions by a factor of negative one for investment rounds in which the advisor recommended not to invest. With this transformation, coefficients measure effects on investment decisions in the *direction of the advice*.

We first pool decisions across all rounds and estimate the effect of having a female advisor. Col. 1 shows that while participants invest around 15.6 cents in the direction of the advice (see mean), advisor gender does not affect investment amounts. Estimates are relatively precise, and we can rule out with 95% confidence that the effect of a female advisor is greater than 1.53 cents. Next, we analyze investment amounts in round one and two separately. Differences across advisor gender are small (around 4%) and not statistically significant

[12] As specified in the pre-analysis plan, we combine responses to two questions about gender differences in ability into an index and divide participants by the median value.

Table 2: Gender Effects: Investment, Changing Opinions

	Invest (Cents)				Prob. Change (pp)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advisor Female	0.073	0.574	-0.604	-1.200	0.696	0.429	0.803	1.750*
	(0.820)	(1.265)	(1.264)	(1.876)	(0.484)	(0.640)	(0.692)	(0.943)
Fem Adv. x Prev. Correct				1.185				-1.883
				(2.536)				(1.366)
Previously Correct				10.629***				9.709***
				(2.099)				(1.411)
Observations	4551	2276	2275	2275	4546	2276	2270	2270
Mean (Male Adv.)	15.6	15.5	15.8	15.8	11.2	12.4	10	10
SD	49.4	48.6	50.2	50.2	16.1	16.4	15.8	15.8
R square	0.63	0.63	0.64	0.64	0.06	0.08	0.07	0.07
Round	Pooled	Round 1	Round 2	Round 2	Pooled	Round 1	Round 2	Round 2
Control Var	Y	Y	Y	Y	Y	Y	Y	Y
Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
P-value				0.99				0.89

Notes: The dependent variable is the investment amount in cents (Col 1-4) and the change in probability of success in percentage points (Col 5-8). These outcomes are rescaled such that they show changes in the direction of the advise. Rounds refer to whether outcomes were collected from the first (round 1) or second (round 2) interaction with a given advisor. Control mean measures outcomes for male advisors. We control for demographic variables, as well as product batch and advisor image type fixed effects. P-value measures whether the sum of female advisor and the interaction term is different from zero. Robust standard errors clustered at individual level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(Col. 2-3).[13]

We observe a similar pattern for changes in beliefs about projects' success (Col. 5-8). As with investment decisions, we transform the variable so that coefficients measure the change in beliefs in the direction of the advice. Pooling across advisor gender, participants update their beliefs about project success by 11.2 percentage points in the direction of the recommendations. Differences across advisor gender are again small and insignificant in the pooled sample (Col. 5). We can rule out that the effect of a female advisor is greater than 0.25 p.p.. Estimating effects separately by investment round, gender differences are small (4-8%) and not significant (Col. 6-7).

[13]There exists an extensive literature on the intersectionality of race and gender (see Cho et al. (2013); Collins and Bilge (2020) for reviews). This literature points towards a distinct set of potential mechanisms for discrimination. While testing these hypotheses goes beyond the scope of the present paper, we explore the effect of advisor race by estimating specification 1 separately for Black and white advisors. Results (not reported) show that the effect of advisor gender on investment behavior is very similar in statistical significance and magnitude, with differences in investment amounts of less than one cent.

3.2 The Role of Advice Correctness

This section explores the role of advice correctness from specification 2. As expected, participants are much more likely to change their investment (Col. 4) and beliefs (Col. 8) in the direction of the recommendation if the advisor was correct in the previous round. Surprisingly, we observe opposing effects for incorrect advice from female advisors. People are stating that they change their beliefs 1.75 p.p. (p-value=0.06) more in the direction of the recommendation than they do for incorrect male advisors (Col. 8). However, the coefficient for investment in the direction recommended by female advisors is negative and insignificant (Col. 4).[14] It is also notable that the sum of coefficients on the advisor gender term and interaction term is very close to zero, indicating that there is no gender difference in responses to correct advice. In fact, throughout our study, we do not find any form of discrimination following correct advice.[15]

3.3 The Role of Investor Gender

This section explores how the effects of advisor gender and advice correctness vary across female and male participants (which we specified as our key demographic characteristic of interest in the pre-analysis plan). Results from specification 3 are reported in Table 3.

[14]The top of Figure A1 reports these results in graphical format by estimating the effect of advisor gender separately across subgroups and advice correctness. The circle presents the effect of having a female advisor on following investment decisions (top panel) and belief updating (bottom panel). A point estimate of zero, marked by the dashed vertical line, indicates a zero effect of manager gender. While this follows the regression specified for our main analysis, we did not specify this strategy for subgroup analysis in our preanalysis plan.

[15]This finding is similar to Abel (2022) who finds that the gender of the manager matters for how workers respond to criticism but not praise.

Table 3: Effect of Advisor Gender by Participant Gender

	Invest (Cents)			Prob. Change (pp)		
	(1)	(2)	(3)	(4)	(5)	(6)
Advisor Female	1.68	2.91	0.34	0.37	0.84	-0.07
	(1.68)	(2.51)	(2.24)	(0.92)	(1.18)	(1.42)
Fem Adv. x Male Res.	-4.51*	-8.54**	-0.50	0.79	1.69	0.00
	(2.55)	(3.78)	(3.46)	(1.41)	(1.96)	(2.03)
Respondent Male	0.03	3.88*	-3.71	-2.91***	-2.85**	-2.70*
	(1.74)	(2.34)	(2.56)	(0.94)	(1.21)	(1.40)
Observations	2273	1132	1141	2268	1130	1138
Mean (Male Adv.)	15.6	11.1	20.5	11.2	6.48	13.6
SD	49.4	49.3	50.7	16.1	14.2	16.5
R square	0.64	0.63	0.65	0.08	0.02	0.07
Control Var	Y	Y	Y	Y	Y	Y
Round	R. 2	R. 2	R. 2	R. 2	R. 2	R. 2
Previous Correct	Pooled	Incor.	Correct	Pooled	Incor.	Correct
P-value: Male Partic.	0.139	0.048	0.949	0.270	0.101	0.961

Notes: The data is limited to the second round of interacting with a given advisor as we focus on the effect of incorrect and correct advise. The dependent variable is the investment amount in cents (Col 1-3) and the change in probability of success in percentage points (Col 4-6). These outcomes are rescaled such that they show changes in the direction of the advise. We control for demographic variables, as well as product and advisor image type fixed effects. P-values measure whether the sum of the female advisor and interaction term with male participants is significant. Robust standard errors clustered at individual level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In the pooled sample, the coefficient on female advisor, which measures the behavior of female investors, is positive but insignificant (Col. 1, Table 3). The coefficient on the male dummy indicates that there is also no overall difference in investments across participant gender. The interaction term between participant and investor gender is -4.51 (p-value=0.077), suggesting that male investors tend to follow recommendations less than female investors. However, the p-value of 0.139, presented in the bottom row of the table, shows that, in aggregate, male investors are not less likely to follow female than male advice.

These aggregate results mask important differences in behavior across investor gender for incorrect advice. Col. 2 of Table 3 shows that, following incorrect advice, male participants invest about 5.5 cents less in the direction recommended by female advisors compared to

male advisors. The estimate for female investors is positive but not significant. Both the difference across investor gender of 8.5 cents and the overall effect of advisor gender for male participants (5.6 cents) are significant at the 5 percent level. They are also sizable in magnitude: men invest almost 50% less in the direction recommended by female compared to male advisors after incorrect advice. By contrast, Col. 3 shows that the effect of advisor gender after positive advice is close to zero and does not vary across participant gender. The top panel of Figure A1 presents these results graphically. Appendix Section A.3 provides several robustness tests for these results. It suggests that our main findings are not driven by participants' tendency to give socially desirable answers, their suspicion that the research is related to discrimination, and choice of specific investment project domains.[16]

These patterns look different when we estimate the role of advisor gender on how people update their (stated) beliefs. In aggregate, the effect of advisor gender is close to zero and insignificant (Col. 4). Following incorrect advice, the effect of a female advisor is close to zero for female investors (Col. 5). Male investors update their beliefs 2.5 p.p. more in the direction of what female advisors recommend (Col. 5). While this magnitude is large (almost 40% of mean value), it is estimated imprecisely and not statistically significant (p-value: 0.101). However, a test of equal coefficients across regressions shows that the difference in investment behavior and stated beliefs is significant at the 5% level for male investors. These results provide evidence against an increase in uncertainty following incorrect advice as an explanation for gender discrimination.[17] As before, we do not find that advisor gender matters after participants receive correct advice (Col. 6) (Figure A1, bottom panel).

Differences in the effect of advisor gender across our main two outcomes can have important consequences. Male investors say that they are, if at all, *more* forgiving of incorrect advice from female advisors as reflected by the larger influence of their recommendation on their stated beliefs of success. However, in their costly investment decisions they penalize female

[16] We also show evidence that people who spend more time on our experiment did not make better investment decisions, which assuages concerns that participants spend time researching the firms.

[17] A related study by Brandts and Rott (2021) finds that in the context of advice about competition entry, women are less likely to give advice that maximizes advisee's expected earnings in settings with uncertainty. One explanation for this behavior is that women and men have different attitudes towards risk-taking as documented e.g. by Friedl et al. (2020). To test for the role of risk attitudes, we divide the sample by the median level of risk aversion. Results (not reported) show that gender discrimination following incorrect advice is similar between participants with high and low levels of risk aversion, providing evidence against the role of uncertainty. However, this specification was not part of the preanalysis plan.

advisors who were previously incorrect more heavily, even for the modest amounts used in our experiment. This pattern is concerning as it suggests that any evidence on the role of advisor gender that relies on stated beliefs or other forms of cheap talk is likely to be unreliable in detecting forms of gender discrimination. Examples of this type of data may include increasingly common customer satisfaction surveys and workplace feedback forms. Sections 4 explores underlying reasons for this discrepancy in outcomes.

4 Mechanisms

This section will explore underlying reasons why (male) investors are less likely to follow recommendations of female advisors after they were incorrect. Following our registered pre-analysis plan, we will focus on the role of participants' level of attention (4.1), attitude towards gender roles (4.2), and differential attribution of skills (4.3). While this choice is based on important potential mechanisms proposed in the literature, it is essential to note that there are other explanations we are not testing due to the lack of available data.

4.1 Attention

People may allocate less attention to certain groups due to the stereotypes they hold (Bartos et al., 2016; Carli, 2001). In the context of (financial) advice, investors may pay less attention to female advisors, which limits their ability to influence behavior. Inattention can be both the reason why investors are less influenced by female advisors and a contributor to future gender discrimination. For example, investors may respond more negatively to incorrect recommendations from female advisors if they did not carefully listen to the underlying rationale of and potential caveats for their recommendation. Our study provides a unique opportunity to test this hypothesis as we collect data, unannounced to investors, on the exact time they spend reading and thinking about the advice.[18]

[18] As per our analysis plan, we winsorize this outcome to account for outliers.

Table 4: Effects on Attention

	Attention (sec.)					
	(1)	(2)	(3)	(4)	(5)	(6)
Advisor Female	-0.139	-0.404	0.655*	0.015	0.532	-0.594
	(0.290)	(0.253)	(0.341)	(0.373)	(0.522)	(0.525)
Fem Adv. x Prev. Correct			-2.104***			
			(0.501)			
Previously Correct			1.556***			
			(0.591)			
Respondent Male				0.062	-1.068**	1.050*
				(0.395)	(0.490)	(0.609)
Fem Adv. x Male Res.				-0.849*	0.230	-1.740**
				(0.515)	(0.667)	(0.759)
Observations	2276	2276	2276	2274	1132	1142
Mean (Male Adv.)	12.7	10.2	10.2	11.3	8.19	12.2
SD	7.29	6.59	6.59	7.13	4.7	8.1
R square	0.11	0.15	0.16	0.15	0.08	0.20
Round	Round 1	Round 2	Round 2	Round 2	Round 2	Round 2
Control Var	Y	Y	Y	Y	Y	Y
Previous Correct		Pooled	Pooled	Pooled	Incor.	Correct
P-value			0.00	0.02	0.07	0.00

Notes: Attention is measured by the time (in seconds) participants spend reading the advice. Rounds refer to whether outcomes were collected from the first (round 1) or second (round 2) interaction with a given advisor. Control mean measures outcomes for male advisors. We control for demographic variables, advisor image type and product batch fixed effects. P-values report if the effect of female advisors if significant for correct advice (Col. 3) and whether the effect of female advisors is significant for male investors (Col. 4-6). Robust standard errors clustered at individual level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We find little evidence that inattention is an explanation for gender discrimination discussed in the previous section. In each round, the time spent reading recommendations is almost identical across advisor gender (Table 4, Col. 1 and 2). Looking at the role of advice correctness, investors pay more attention if the previous recommendation was correct (Col. 3). This boost in attention is driven by male investors paying more attention to male advisors (as indicated by the negative interaction term in Col. 6). By contrast, male investors spend almost 0.8 seconds (p-value=0.07) more reading recommendations from female than male advisors (Col. 5).

In summary, for female investors we observe consistency across our outcomes: attention, belief updating, and investing and there are no difference across advisor gender. By contrast, for male investors we observe inconsistent results across our outcomes, especially after receiving incorrect advice. They spend more time on the advice of female advisors and update their beliefs more in the direction of their recommendations.[19] However, they invest more money in the direction recommended by male advisors. Our results suggest that attention discrimination is not a mechanism for this behavior.

4.2 Views on Gender Roles and Gender Quotas

Our pre-analysis plan specifies analyzing results along two dimensions: participants' views on gender roles and support for gender quotas.

Views on gender roles: While the gender composition of the private sector has been changing over the last decades, it is possible that views of gender roles in the workplace affect behavior, especially in an industry like finance that has historically been dominated by men. We measure participants' attitude regarding gender roles by asking whether they agree that there are innate differences in ability across gender and whether men and women are better suited for certain jobs. We categorize people as holding traditional gender views depending on whether their responses are above or below the median.

Overall, we find substantial variation in gender attitude across demographic characteristics, with respondents who are older, male, without a college degree, and identify as conservative more likely to agree with these statements (Table A4). However, Table A2 shows that views on gender roles are not correlated with how participants react to having received either correct or incorrect advice. One potential explanation for this pattern we explore next is that while these views may explain whether people expect to see female financial advisors, they do not think they are less qualified *conditional* on holding these positions. The representation of women in private sector leadership roles prompted several countries to adopt gender quotas. Studies have raised concerns that these policies might result in women in leadership being seen as less competent (Neschen and Hügelschäfer, 2021) and experiencing

[19]We note that inattention can also reflect more trust in the advice given. However, this positive interpretation would imply that investors are *more* likely to base their investment decision on the recommendation.

backlash (Leibbrandt et al., 2018). We measure participants' support for gender quotas on a scale of 0 to 100, categorizing those with scores above the median value of 51 as supporters. Column 2 of Table 5 indicates that discrimination against female advisors following incorrect recommendations is limited to participants who oppose quotas. For correct advice (Col. 3) and stated beliefs (Col. 4-6), we do not find significant differences across advisor gender.

Table 5: Effect of Advisor Gender by Support Gender Quota

	Invest (Cents)			Prob. Change (pp)		
	(1)	(2)	(3)	(4)	(5)	(6)
Advisor Female	-2.16	-5.66**	1.07	1.21	0.90	1.48
	(1.77)	(2.64)	(2.41)	(1.02)	(1.41)	(1.44)
Support Gender Quota	-1.54	-5.45**	2.52	0.36	-0.51	1.41
	(1.75)	(2.38)	(2.53)	(0.94)	(1.24)	(1.38)
Fem Adv. x Support Quota	3.39	8.89**	-1.80	-0.45	1.62	-2.58
	(2.54)	(3.76)	(3.44)	(1.39)	(1.92)	(1.98)
Observations	2257	1124	1133	2252	1122	1130
Mean (Male Adv.)	15.6	11.1	20.5	11.2	6.48	13.6
SD	49.4	49.3	50.7	16.1	14.2	16.5
R square	0.64	0.63	0.65	0.08	0.02	0.07
Control Var	Y	Y	Y	Y	Y	Y
Round	R. 2	R. 2	R. 2	R. 2	R. 2	R. 2
Previous Correct	Pooled	Incor.	Correct	Pooled	Incor.	Correct
P-value: Male Partic.	0.498	0.229	0.768	0.418	0.049	0.416

Notes: The data is limited to the second round of interacting with a given advisor as we focus on the effect of incorrect and correct advise. The dependent variable is the investment amount in cents (Col 1-3) and the change in probability of success in percentage points (Col 4-6). These outcomes are rescaled such that they show changes in the direction of the advise. We control for demographic variables, batch fixed effects and product fixed effects. 'Support Gender Quota' is a binary variable measuring if respondents' support for gender quotas is above the median. P-values measure whether the sum of the female advisor and interaction term with gender quota support is significant. P-values measure whether the sum of the female advisor and interaction term with support for quotas is significant. Robust standard errors clustered at individual level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A key challenge in interpreting this result is the difficulty in isolating factors that explain opposition to gender quotas. While some argue that resistance to gender quotas stems from sexist beliefs (Kane and Whipkey, 2009; Krook, 2006), there may be additional factors correlated with support for quotas driving discriminatory behavior. For example, we find that support for gender quotas is substantially lower among male participants (Table A4, Col. 5). It is also highly correlated with participants' tendency to hold traditional gender views (correlation coefficient=0.4). To account for these correlations, Table A5 jointly controls for the three interaction terms of interest. This specification was not part of the pre-analysis

plan. Our findings indicate that differences in investment behavior following incorrect advice, based on both participant gender and their support for quotas, remain significant and similar in magnitude. While these results suggest that attitudes play a distinct role in explaining our findings, the exact underlying mechanisms remain unclear.

4.3 Advisor Assessment

This section presents results from advisor assessments conducted after participants completed two investment rounds with a given advisor. As per our analysis plan, we focus on two outcomes. Firstly, a standardized index of investor assessments comprising their perceived experience, professionalism and ability to judge and communicate risk (see Figure B7). And secondly we use a binary measure whether participants want to keep the same advisor in a future investment round they may be invited to.[20]

We find no evidence for gender discrimination in advisor assessments. Coefficients for assessment (Col. 1) and choice to keep advisors (Col. 3) are positive but not significant (Table 6). We see similar patterns of positive but statistically insignificant estimates across the categories comprising the index (Table A3).

While every advisor is once correct and once incorrect, participants are randomly assigned to groups of projects that either succeeded or failed for a given advisor. We find that the success of projects has large effects on how advisors are evaluated. Those assigned to successful projects receive 0.57 sd more positive assessments and participants are 30.6 p.p. more likely to want to interact with them in the future (Table 6, Col. 2, 4). However, the effect of project success does not vary across advisor gender as indicated by the insignificant interaction term reported in Col. 2 and 4.

These results are consistent with recent evidence showing that women are perceived as more competent conditional on having been promoted (Bohren et al., 2019). However, we want to highlight two caveats. First, all advisors are correct 50% of the time and we may

[20]Specifically, we inform participants that we conduct a lottery and that winners receive the chance to play additional investment rounds for which they can earn money. Similar to Chakraborty and Serra (2023), we ask if, conditional on winning, they want to have the same advisor or want to be matched with a randomly chosen new advisor. See Figure B8.

Table 6: Gender Effects: Advisor Assessment, Keep Advisor

	Assessment (std.)		Keep Advisor	
	(1)	(2)	(3)	(4)
1=Advisor Female	0.043	0.037	0.024	0.009
	(0.040)	(0.059)	(0.020)	(0.026)
Fem Adv. x Proj success		0.000		0.024
		(0.084)		(0.039)
1=Projects succeed		0.567***		0.306***
		(0.073)		(0.033)
Observations	2276	2276	2276	2276
Mean (Male Adv.)	-.0218	-.0218	-.61	-.61
SD	.995	.995	.488	.488
R square	0.03	0.07	0.03	0.09
Control Var	Y	Y	Y	Y
Fixed Effects	Y	Y	Y	Y
P-value		0.50		0.27

Notes: The data set consists of two advisor assessments per participant. The dependent variable is a standardized assessment index (Col 1-2) comprising of level of experience, ability to assess and communicate risk, and professionalism. The dependent variable in Col. 3-4 is a binary variable measuring whether a participant wants to keep the advisor. We control for demographic characteristics and batch fixed effects. Robust standard errors clustered at individual level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

observe more discrimination in advisor assessments if we had more variation in how correct recommendations are. Second, advisor assessments are by design not incentivized and may thus be prone to participants giving answers that are socially desirable. While we cannot rule out this possibility it is reassuring that results do not vary across participants' social desirability score.

5 Conclusion

We conduct an investment experiment that utilizes actual risky business projects and advice from actual professionals. We find that advisor gender does not affect investment behavior before outcomes are revealed or how investors react if advice proves to be correct. For incorrect advice, male investors are less likely to follow subsequent recommendations from female advisors. Implications for how to address discrimination depend on the underlying

reason for people's behavior. We conduct three tests of mechanisms that may explain gender discrimination: people may not pay attention, investors may think female advisors are less qualified, or women holding these positions may be incongruent with their gender views. We do not find that any of these explanations help explain our results, suggesting that they operate through channels we did not investigate. While the policy implications are therefore unclear, our findings highlight that relying on self-reported assessments is not sufficient to detect discrimination.

Overall, the pattern that male investors blame female advisors more for incorrect advice can have important consequences, especially given that men are not only over-represented in the finance sector but also tend to dominate saving and investment decisions of households (Meier et al., 1999; Kim et al., 2017). Receiving more blame for incorrect advice on risky projects may help explain the lack of women in financial advising positions. In a related influential study, Chakraborty and Serra (2023) show that women are more hesitant to pursue leadership positions because they are concerned that their decisions will lead to backlash from unhappy employees. Being blamed may thus be one of the reasons why women are less interested in pursuing these careers.

References

Abel, M. (2022), 'Do workers discriminate against female bosses?', Journal of Human Resources pp. 1120–11318R3.

Abel, M. and Buchman, D. (2023), 'The effect of manager gender and performance feedback: Experimental evidence from india', Economic Development and Cultural Change .

Abel, M. and Burger, R. (2023), 'Name-based race discrimination: The role of heuristics', mimeo .

Atkinson, S. M., Baird, S. B. and Frye, M. B. (2003), 'Do female mutual fund managers manage differently?', Journal of Financial Research **26**(1), 1–18.

Ayalew, S., Manian, S. and Sheth, K. (2021), 'Discrimination from below: Experimental evidence from ethiopia', Journal of Development Economics **151**, 102653.

Bailey, P. E., Leon, T., Ebner, N. C., Moustafa, A. A. and Weidemann, G. (2023), 'A meta-analysis of the weight of advice in decision-making', Current Psychology **42**(28), 24516–24541.

Barron, K., Ditlmann, R., Gehrig, S. and Schweighofer-Kodritsch, S. (2022), 'Explicit and implicit belief-based gender discrimination: A hiring experiment'.

Bartos, V., Bauer, M., Chytilová, J. and Matějka, F. (2016), 'Attention discrimination: Theory and field experiments with monitoring information acquisition', American Economic Review **106**(6), 1437–1475.

Bohren, J. A., Imas, A. and Rosenberg, M. (2019), 'The dynamics of discrimination: Theory and evidence', American Economic Review **109**(10), 3395–3436.

Bordalo, P., Coffman, K., Gennaioli, N. and Shleifer, A. (2019), 'Beliefs about gender', American Economic Review **109**(3), 739–773.

Brandts, J., Groenert, V. and Rott, C. (2015), 'The impact of advice on women's and men's selection into competition', Management Science **61**(5), 1018–1035.

Brandts, J. and Rott, C. (2021), 'Advice from women and men and selection into competition', Journal of Economic Psychology **82**, 102333.

Carli, L. L. (2001), 'Gender and social influence', Journal of Social Issues **57**(4), 725–741.

Chakraborty, P. and Serra, D. (2023), 'Gender and leadership in organisations: The threat of backlash', The Economic Journal .

Cho, S., Crenshaw, K. W. and McCall, L. (2013), 'Toward a field of intersectionality studies: Theory, applications, and praxis', Signs: Journal of women in culture and society **38**(4), 785–810.

Coffman, K. B., Exley, C. L. and Niederle, M. (2021), 'The role of beliefs in driving gender discrimination', Management Science **67**(6), 3551–3569.

Collins, P. H. and Bilge, S. (2020), Intersectionality, John Wiley & Sons.

Cooper, D. J. and Kagel, J. H. (2016), 'A failure to communicate: an experimental investigation of the effects of advice on strategic play', European Economic Review **82**, 24–45.

Crowne, D. P. and Marlowe, D. (1960), 'A new scale of social desirability independent of psychopathology.', Journal of Consulting Psychology **24**(4), 349.

Eagly, A. H. and Karau, S. J. (2002), 'Role congruity theory of prejudice toward female leaders.', Psychological review **109**(3), 573.

Eagly, A. H., Makhijani, M. G. and Klonsky, B. G. (1992), 'Gender and the evaluation of leaders: A meta-analysis.', Psychological bulletin **111**(1), 3.

Ellingrud, K., Krivkovich, A., Nadeau, M. and Zucker, J. (2021), 'Closing the gender and race gaps in north american financial services', McKinsey & Company. Retrieved February **23**, 2022.

Erkal, N., Gangadharan, L. and Koh, B. H. (2023), 'Do women receive less blame than men? attribution of outcomes in a prosocial setting', Journal of Economic Behavior & Organization **210**, 441–452.

Friedl, A., Pondorfer, A. and Schmidt, U. (2020), 'Gender differences in social risk taking', Journal of Economic Psychology **77**, 102182.

Kane, E. W. and Whipkey, K. J. (2009), 'Predictors of public support for gender-related affirmative action: Interests, gender attitudes, and stratification beliefs', Public Opinion Quarterly **73**(2), 233–254.

Kim, J., Gutter, M. S. and Spangler, T. (2017), 'Review of family financial decision making: Suggestions for future research and implications for financial education', Journal of Financial Counseling and Planning **28**(2), 253–267.

Klein, G. and Shtudiner, Z. (2016), 'Trust in others: does it affect investment decisions?', Quality & Quantity **50**, 1949–1967.

Klein, G., Shtudiner, Z. and Zwilling, M. (2021), 'Uncovering gender bias in attitudes towards financial advisors', Journal of Economic Behavior Organization **189**, 257–273.

Krook, M. L. (2006), 'Gender quotas, norms, and politics', Politics & Gender **2**(1), 110–118.

Leibbrandt, A., Wang, L. C. and Foo, C. (2018), 'Gender quotas, competitions, and peer review: Experimental evidence on the backlash against women', Management Science **64**(8), 3501–3516.

Manian, S. and Sheth, K. (2021), 'Follow my lead: Assertive cheap talk and the gender gap', Management Science **67**(11), 6880–6896.

Meier, K., Kirchler, E. and Hubert, A.-C. (1999), 'Savings and investment decisions within private households: Spouses' dominance in decisions on various forms of investment', Journal of Economic Psychology **20**(5), 499–519.

Neschen, A. and Hügelschäfer, S. (2021), 'Gender bias in performance evaluations: The impact of gender quotas', Journal of Economic Psychology **85**, 102383.

Pulford, B. D., Colman, A. M. and Loomes, G. (2018), 'Incentive magnitude effects in experimental games: Bigger is not necessarily better', Games **9**(1), 4.

Reynolds, W. M. (1982), 'Development of reliable and valid short forms of the marlowe-crowne social desirability scale', Journal of Clinical Psychology **38**(1), 119–125.

Sahay, M. R. and Cihak, M. M. (2018), Women in finance: A case for closing gaps, International Monetary Fund.

Sarsons, H. (2019), 'Interpreting signals in the labor market: Evidence from medical referrals'.

Schwab, K., Samans, R., Zahidi, S., Leopold, T. A., Ratcheva, V., Hausmann, R. and Tyson, L. D. (2017), The global gender gap report 2017, World Economic Forum.

A Appendix

A.1 Tables

Table A1: Effect of Advisor Gender by Participant Gender (No Control Variables)

	Invest (Cents)			Prob. Change (pp)			Attention (sec.)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1=Advisor Female	1.83	3.04	0.32	0.06	0.80	-0.81	0.10	0.83	-0.74
	(1.63)	(2.44)	(2.15)	(0.91)	(1.18)	(1.41)	(0.37)	(0.51)	(0.54)
1=Male Respondent	0.07	3.90*	-4.18*	-2.89***	-2.58**	-3.32**	-0.11	-1.22**	0.94
	(1.70)	(2.29)	(2.50)	(0.93)	(1.22)	(1.40)	(0.40)	(0.48)	(0.62)
Fem Adv. x Male	-4.57*	-8.11**	-0.31	0.67	1.15	0.65	-0.87*	-0.03	-1.53**
	(2.51)	(3.71)	(3.39)	(1.39)	(1.93)	(2.00)	(0.51)	(0.66)	(0.76)
Observations	2339	1168	1171	2334	1166	1168	2340	1168	1172
Mean (Male Adv.)	14.6	10.6	18.6	9.67	5.22	12.2	11.3	8.19	12.2
SD	51.2	49.4	52.9	15.6	13.9	16.3	7.13	4.7	8.1
R square	0.64	0.63	0.65	0.08	0.02	0.08	0.11	0.04	0.17
Control Var	N	N	N	N	N	N	N	N	N
Previous Correct	Pooled	Incorr.	Correct	Pooled	Incorr.	Correct	Pooled	Incorr.	Correct
P-value	0.15	0.07	0.99	0.48	0.20	0.91	0.03	0.06	0.00

Notes: The dependent variable measures how much participants change their investment amount (Col. 1-3) and beliefs of success (Col. 4-6) in the direction recommended by the advisor, and how many seconds they spend reading advice (Col. 7-9). Robust standard errors clustered at individual level are in parentheses. Regressions do control for demographic characteristics. Regressions include round, project, and picture (twin) fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2: Effect of Advisor Gender by Gender Views

	Invest (Cents)			Prob. Change (pp)		
	(1)	(2)	(3)	(4)	(5)	(6)
Advisor Female	-0.07	-0.53	0.64	1.28	2.00*	0.59
	(1.70)	(2.50)	(2.33)	(0.88)	(1.09)	(1.38)
Traditional Gender Views	-0.69	-0.73	-0.26	0.07	0.39	-0.27
	(1.73)	(2.37)	(2.54)	(0.94)	(1.25)	(1.40)
Fem Adv. x Tradit. Views	-1.03	-1.31	-1.07	-0.66	-0.54	-0.81
	(2.55)	(3.76)	(3.49)	(1.40)	(1.95)	(2.02)
Observations	2257	1124	1133	2252	1122	1130
Mean (Male Adv.)	15.6	11.1	20.5	11.2	6.48	13.6
SD	49.4	49.3	50.7	16.1	14.2	16.5
R square	0.64	0.63	0.65	0.08	0.02	0.07
Control Var	Y	Y	Y	Y	Y	Y
Round	R. 2	R. 2	R. 2	R. 2	R. 2	R. 2
Previous Correct	Pooled	Incor.	Correct	Pooled	Incor.	Correct
P-value: Male Partic.	0.564	0.519	0.865	0.575	0.369	0.883

Notes: The data is limited to the second round of interacting with a given advisor as we focus on the effect of incorrect and correct advise. The dependent variable is the investment amount in cents (Col 1-3) and the change in probability of success in percentage points (Col 4-6). These outcomes are rescaled such that they show changes in the direction of the advise. We control for demographic variables, batch fixed effects and product fixed effects. 'Traditional Gender Views' is a binary variable measuring if respondents' views are above the median value. P-values measure whether the sum of the female advisor and interaction term with gender quota support is significant. P-values measure whether the sum of the female advisor and interaction term with male participants is significant. Robust standard errors clustered at individual level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3: Gender Effects: Detailed Advisor Assessment

	Experience		Professionalism		Judging Risk		Commun. Risk	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advisor Female	0.051	0.079	0.033	0.029	0.004	-0.008	0.049	0.053
	(0.034)	(0.051)	(0.034)	(0.053)	(0.037)	(0.055)	(0.038)	(0.057)
Fem Adv. x Proj success		-0.057		0.009		0.023		-0.008
		(0.070)		(0.074)		(0.079)		(0.080)
1=Projects succeed		0.068		0.179*		-0.017		0.068
		(0.094)		(0.099)		(0.107)		(0.107)
Observations	2276	2276	2276	2276	2276	2276	2276	2276
Mean (Male Adv.)	2.26	2.26	2.46	2.46	2.08	2.08	2.2	2.2
SD	.852	.852	.872	.872	.954	.954	.964	.964
R square	0.08	0.08	0.07	0.07	0.11	0.11	0.10	0.10
Control Var	Y	Y	Y	Y	Y	Y	Y	Y
Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
P-value		0.63		0.44		0.77		0.40

Notes: The data set consists of two advisor assessments per participant. The dependent variables are the assesment of advisors with regard to experience (Col 1-2), professionalism (Col 3-4), judgment of risk (Col 5-6), and communication of risk (Col 7-8). We control for demographic characteristics as well as product batch and advisor image type fixed effects. Robust standard errors clustered at individual level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Correlates: Gender Role Attitudes and Support for Gender Quotas

	Non-Traditional Gender Views				Support Gender Quotas			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.24***				10.06***			
	(0.06)				(1.87)			
Age (Yrs)		0.00**				-0.06		
		(0.00)				(0.06)		
Educ (Yrs)			0.04***				0.08	
			(0.01)				(0.43)	
Conservative				-0.49***				-16.40***
				(0.08)				(2.64)
Liberal				0.56***				15.70***
				(0.07)				(2.16)
Observations	1174	1173	1173	1159	1166	1165	1165	1151
Mean (Male Adv.)	0	0	0	0	51	51	51	51
SD	1	1	1	1	32.2	32.2	32.2	32.2
R square	0.03	0.00	0.01	0.19	0.03	0.00	0.00	0.17

Notes: The dependent variable in Col 1-4 is a standardized index measuring non-traditional gender attitudes. The dependent variable in Col 5-8 measures support for gender quotas on a scale from 0 to 100. Robust standard errors clustered at individual level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

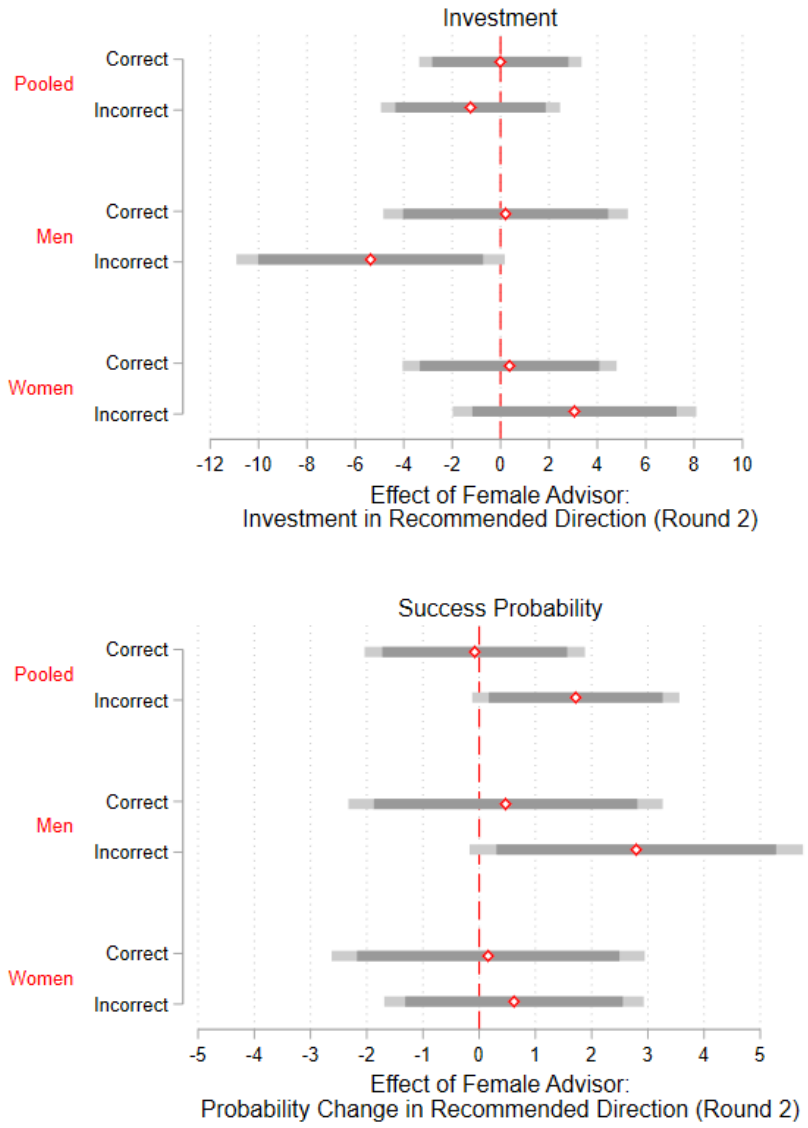
Table A5: Effect of Advisor Gender: Combined Interaction Terms

	Invest (Cents)			Prob. Change (pp)		
	(1)	(2)	(3)	(4)	(5)	(6)
Advisor Female	-0.23	-2.95	2.26	1.31	-0.01	2.58
	(2.62)	(3.80)	(3.60)	(1.40)	(1.97)	(1.99)
Respondent Male	0.40	3.79	-3.24	-2.67***	-2.81**	-2.26
	(1.83)	(2.54)	(2.62)	(0.90)	(1.19)	(1.40)
Fem Adv. x Respondent Male	-4.00	-6.95*	-0.54	0.78	2.06	-0.31
	(2.59)	(3.79)	(3.51)	(1.36)	(1.93)	(1.98)
Support Gender Quota	-1.74	-5.84**	2.61	0.37	-0.54	1.42
	(1.87)	(2.64)	(2.64)	(0.92)	(1.24)	(1.37)
Fem Adv. x Support Quota	3.08	8.59**	-2.28	-0.60	1.78	-3.08
	(2.65)	(3.86)	(3.64)	(1.39)	(1.95)	(1.95)
Traditional Gender Views	-1.44	-2.64	0.35	0.22	0.29	0.05
	(1.87)	(2.62)	(2.63)	(0.92)	(1.23)	(1.42)
Fem Adv. x Tradit. Views	0.27	1.45	-1.51	-0.91	-0.31	-1.62
	(2.65)	(3.85)	(3.65)	(1.39)	(1.95)	(2.00)
Observations	2257	1124	1133	2252	1122	1130
Mean (Male Adv.)	15.6	11.1	20.5	11.2	6.48	13.6
SD	49.4	49.3	50.7	16.1	14.2	16.5
R square	0.64	0.63	0.65	0.08	0.02	0.07
Control Var	Y	Y	Y	Y	Y	Y
Round	R. 2	R. 2	R. 2	R. 2	R. 2	R. 2
Previous Correct	Pooled	Incor.	Correct	Pooled	Incor.	Correct

Notes: The data is limited to the second round of interacting with a given advisor. The dependent variable is the investment amount in cents (Col 1-3) and the change in probability of success in percentage points (Col 4-6). These outcomes are rescaled such that they show changes in the direction of the advise. We control for demographic variables, batch fixed effects and product fixed effects. 'Traditional Gender Views' is a binary variable measuring if respondents' views are above the median value. Support Gender Quotas is measured on a 0-100 scale with higher values indicating more support. Robust standard errors clustered at individual level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.2 Figures

Figure A1: Advisor Gender and Correctness of Advice



Notes: The graph shows the results from specification 1 for investment amounts (top) and change in perceived success probability (bottom) estimated separately for incorrect and correct advise (in the previous round) and across participant gender. The circle presents the point estimate of having a female compared to male advisor on following recommendations. The bars present 90% and 95% confidence intervals.

A.3 Robustness

This section presents several robustness tests for our main results. Since participants are compensated based on the success of the investment opportunities, one may be concerned that they spend time researching the firms online. To address this concern, we test how investment behavior varies across the time participants spent completing the experiment. We find that those who spent more than the median or 90th percentile of time on the study invest slightly less money in successful firms. While this test is not conclusive and was not part of the pre-analysis plan, it suggests that they did not spend time researching the firms. We further note that this should bias us against finding an effect of advisors.

One concern with studies on gender discrimination is that participants may hide their true attitudes and behaviors. We therefore deliberately did not reveal the research question at the beginning of the study. Nevertheless, when asked at the end of the experiment, almost 12% of participants suspected that the research was related to testing some form of bias.[21] Figure B10 replicates our main analysis separately for those who suspected the research was related to discrimination and those who did not. While estimates are relatively imprecise given the smaller sample size, we find that those suspecting that the study is related to bias are, if at all, less likely to follow advice of female advisors who were previously incorrect.

A related robustness test investigates whether results vary depending on participants' tendency to give socially desirable responses to surveys. Specifically, we administer a version of the Marlowe-Crowne social desirability scale in our survey (Crowne and Marlowe, 1960).[22] Figure B9 shows that results do not vary between participants who score high or low on this scale, suggesting that results are not affected by social desirability bias.

Another question related to the external validity of our study is how specific results are to our choice of investment projects. We address this question by estimating our results separately for each investment projects. Results (available on request) show that we observe

[21] We ask respondents after the experiment but before asking questions related to race (e.g. affirmative): "In a few words, what do you think is the main question that this research is trying to answer?" We categorize people depending on whether their answer includes words like race, bias, discrimination, Black.

[22] Given time constraints, we administered a shortened module with five of the original 33 questions included in Crowne and Marlowe (1960). These include statements such as "*I'm always willing to admit it when I make a mistake.*" and "*I am always courteous, even to people who are disagreeable.* Reynolds (1982) shows that shorter versions of the module show high degrees of internal consistency."

similar patterns across each project with investments of male investors decreasing between five and ten cents more following incorrect advice when paired with female advisors. While we cannot conclusively test how domain-specific our findings are, these results suggests that they are consistent within the domains included in our design. Last, results (available on request) show that the discrimination among male investors following incorrect recommendations of female advisors does not differ depending on whether participants lost money above or below the median investment amount.

B Online Appendix

Figure B1: Consent Form

Informed Consent Form

You have been asked to complete this survey as part of a research project led by Martin Abel, a faculty member at Bowdoin College.

The research project is designed to learn how people incorporate advice from others in their decision making. Specifically, participants receive a monetary endowment and watch short videos about start-up firms they can invest in. They also receive advice about whether to invest from another person. Participants can earn money if the firm succeeds and lose money if it fails. All examples are drawn from actual firms and people.

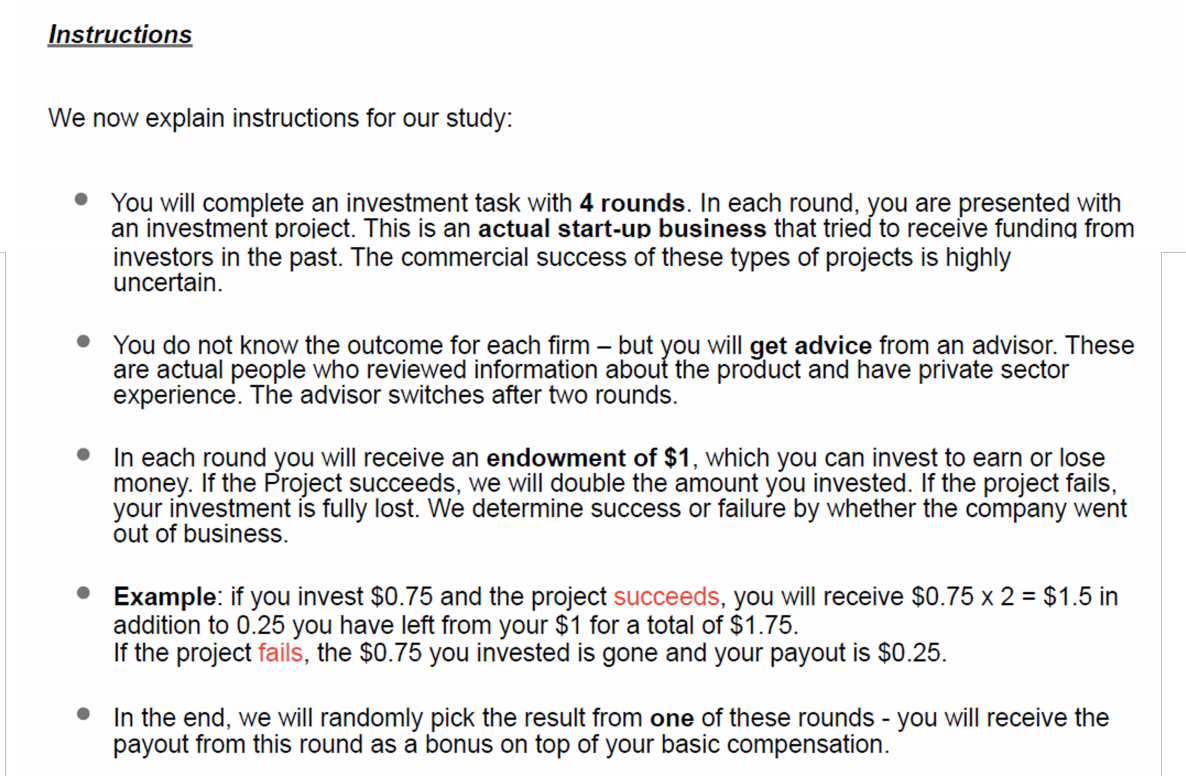
The expected time for this study is **10 minutes**. You will earn a **guaranteed payout of $1.50** and can earn an **additional bonus of up to $2.00** depending on the investment decisions and outcomes.

Your participation is entirely voluntary, and you may refuse to complete any part or all of this survey without any effect on the guaranteed payout. This survey is designed to be confidential, meaning that we will protect subjects' identifying information from access by anyone.

By completing and submitting the survey, you affirm that you are at least 18 years old and that you give your consent for Martin Abel to use your answers in his research. If you have any questions about this research before or after you complete the survey, please contact Martin Abel at m.abel@bowdoin.edu.

If you have any concerns or questions about your rights as a participant in this research, please contact the Chair of the Bowdoin Institutional Review Board at irb@bowdoin.edu.

Figure B2: Introduction



Instructions

We now explain instructions for our study:

- You will complete an investment task with **4 rounds**. In each round, you are presented with an investment project. This is an **actual start-up business** that tried to receive funding from investors in the past. The commercial success of these types of projects is highly uncertain.
- You do not know the outcome for each firm – but you will **get advice** from an advisor. These are actual people who reviewed information about the product and have private sector experience. The advisor switches after two rounds.
- In each round you will receive an **endowment of $1**, which you can invest to earn or lose money. If the Project succeeds, we will double the amount you invested. If the project fails, your investment is fully lost. We determine success or failure by whether the company went out of business.
- **Example**: if you invest $0.75 and the project succeeds, you will receive $0.75 x 2 = $1.5 in addition to 0.25 you have left from your $1 for a total of $1.75.
  If the project fails, the $0.75 you invested is gone and your payout is $0.25.
- In the end, we will randomly pick the result from **one** of these rounds - you will receive the payout from this round as a bonus on top of your basic compensation.

Figure B3: Introduction Advisor

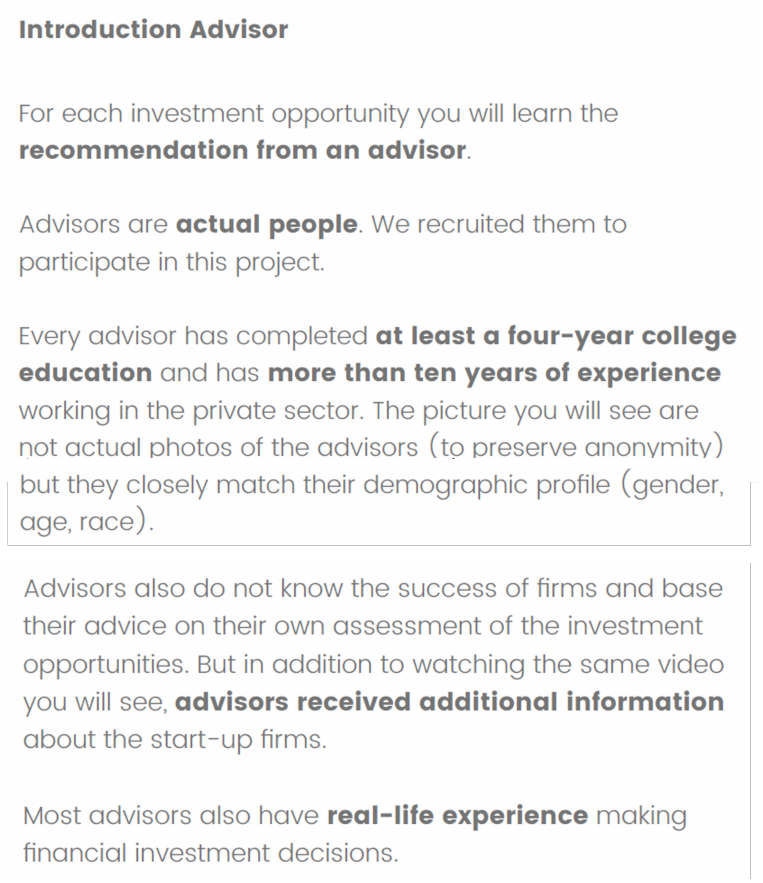


Figure B4: Video and Prior Belief Elicitation

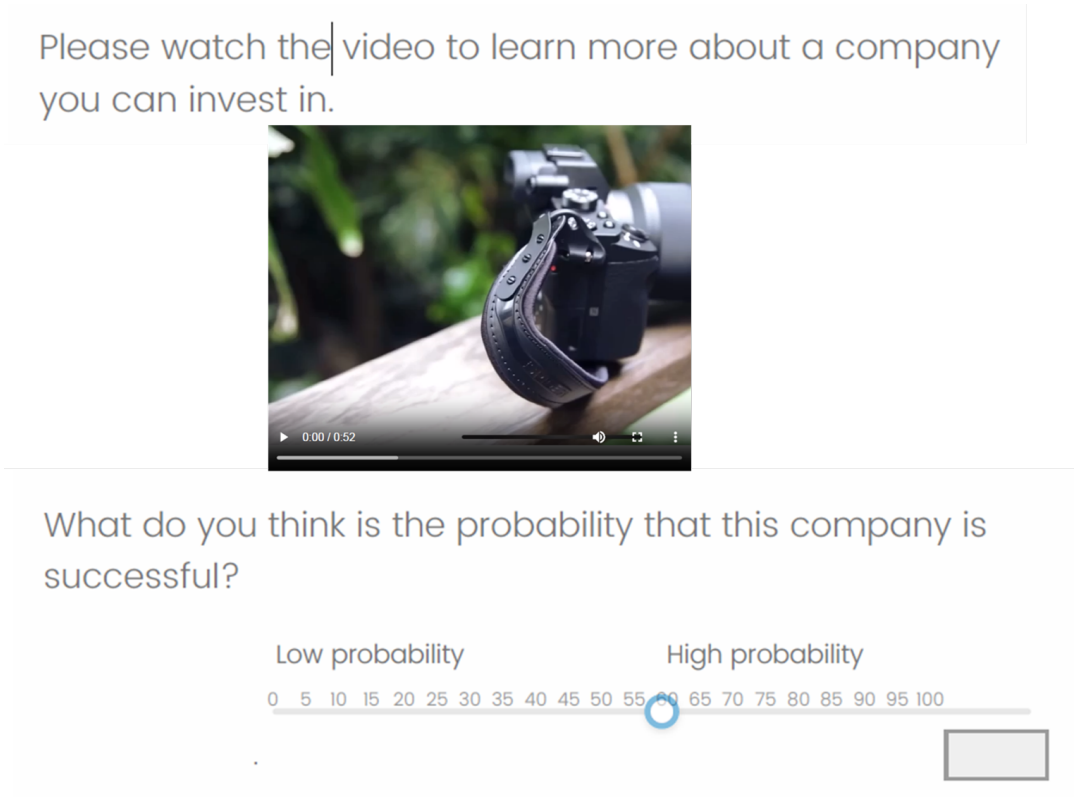


Figure B5: Advise Examples

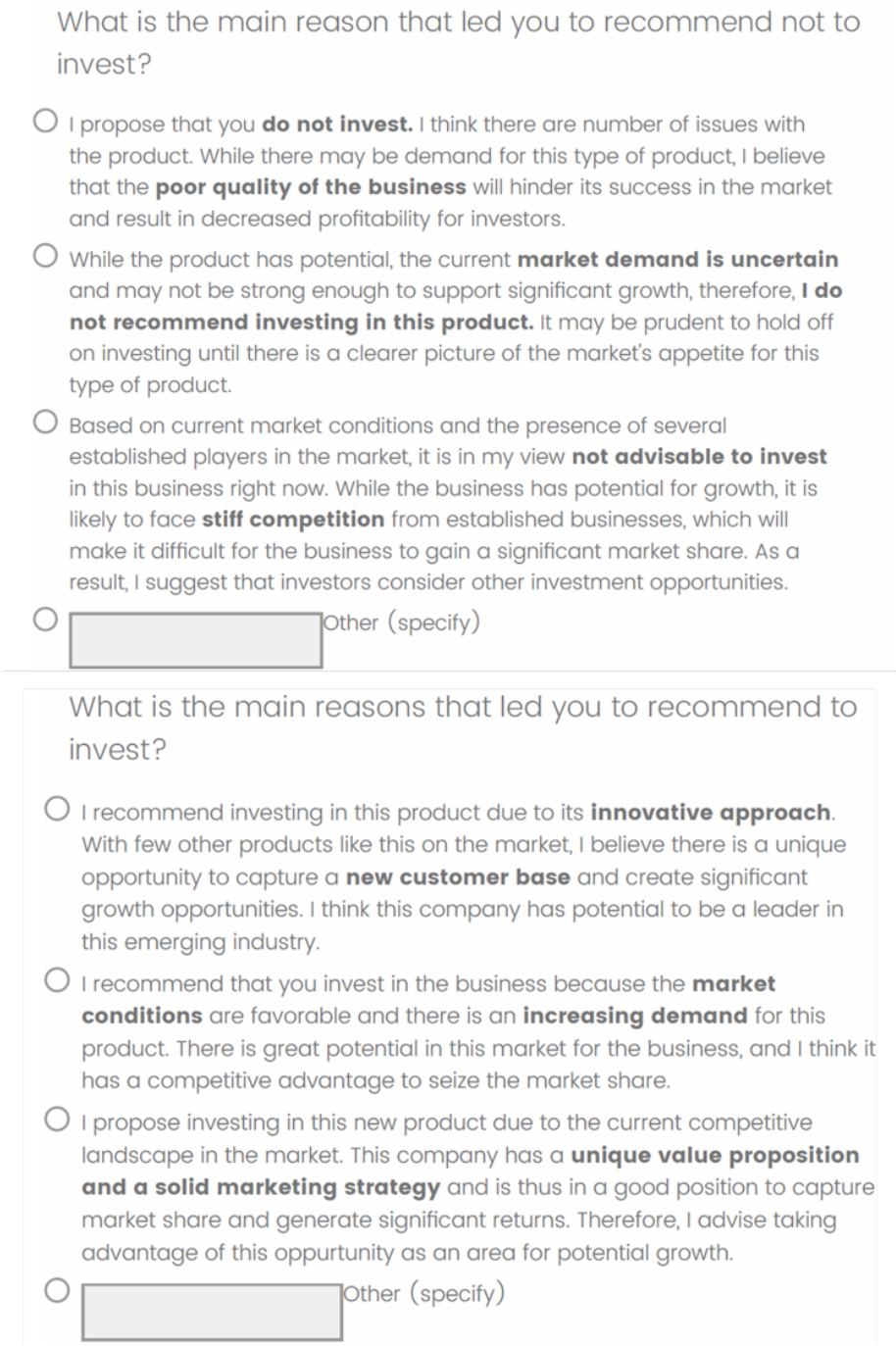
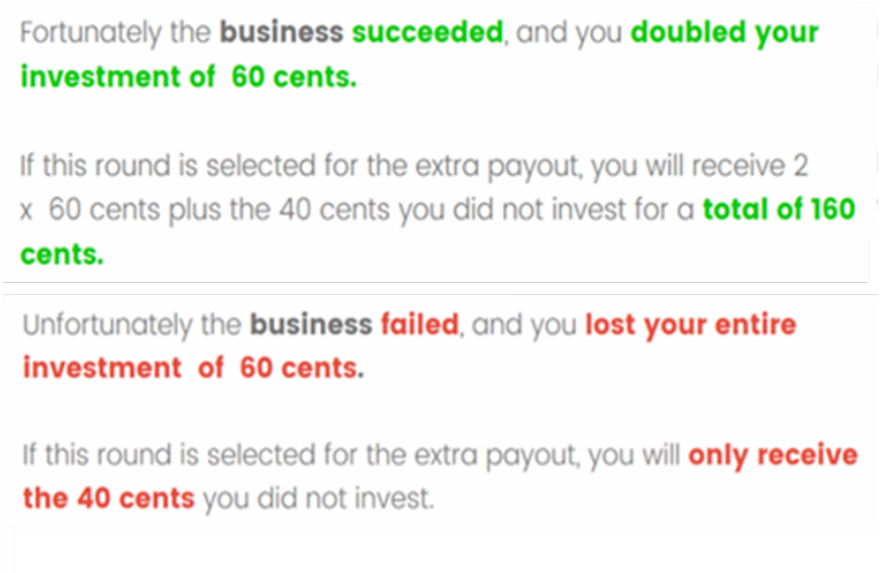


Figure B6: Communication of Investment Outcome



Notes: The graph shows how success (top) and failure (bottom) if investment outcomes are communicated=.

Figure B7: Advisor Assessment

Figure B8: Advisor Selection

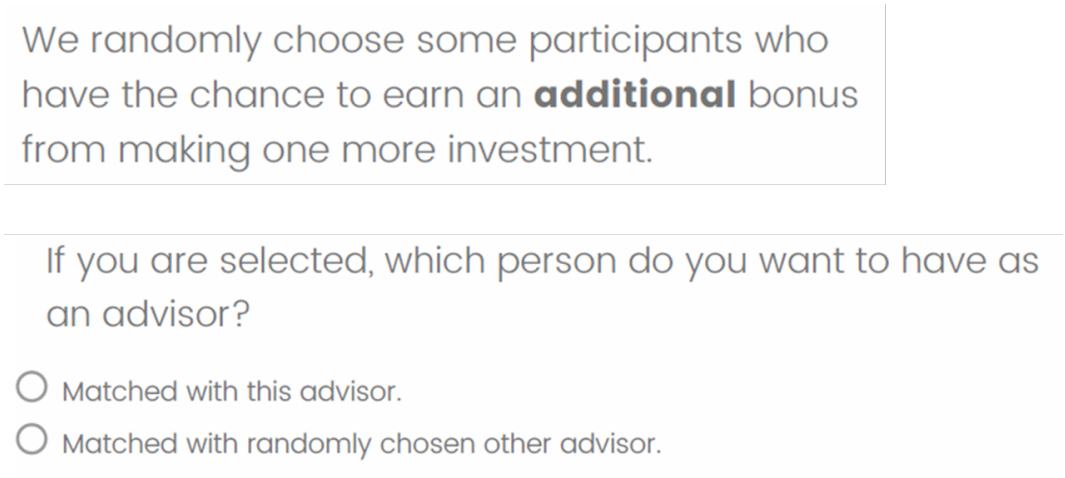
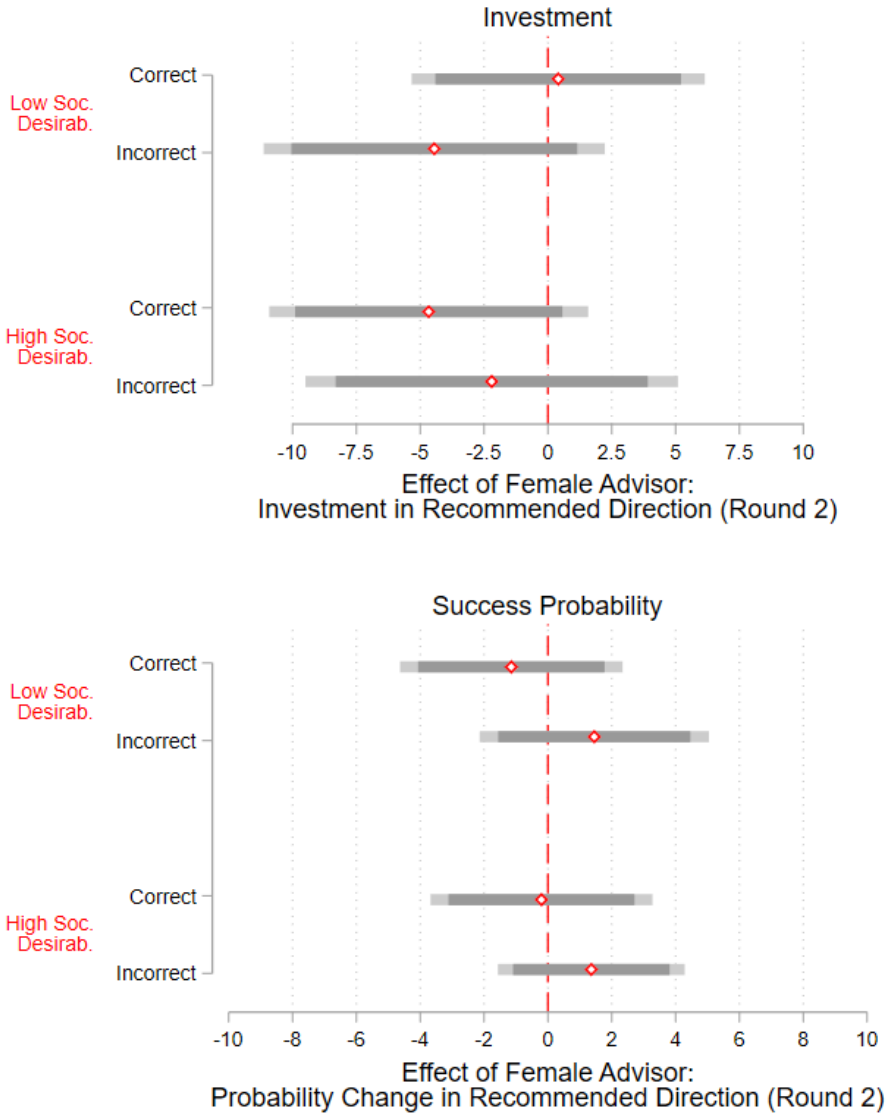
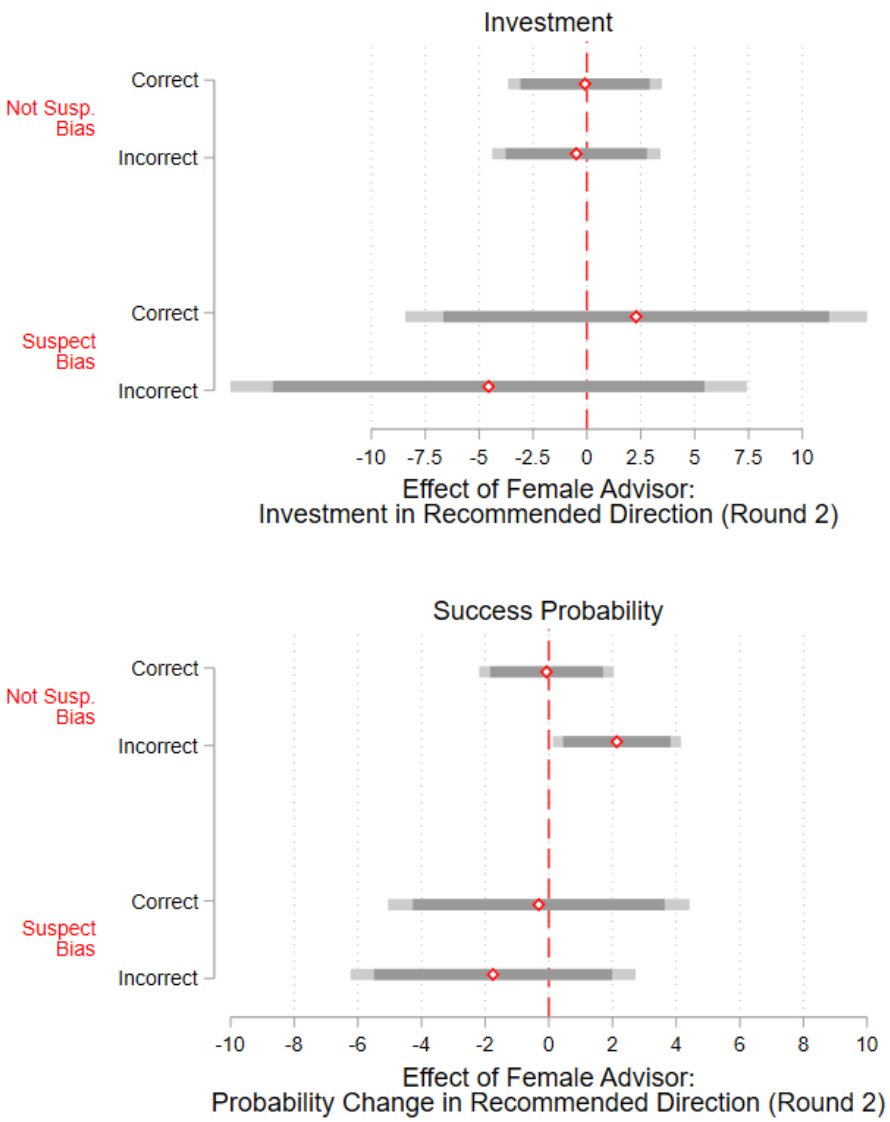


Figure B9: Robustness: Social Desirability Bias



Notes: We categorize people according to their social desirability score based on (Crowne and Marlowe, 1960). Participants are grouped based on the tercile of responses.

Figure B10: Robustness: Suspect Research Purpose



Notes: To categorize people we ask about what they suspect the research project tries to answer. We code people as suspecting bias if they they suspect that the research is related to study some form of bias.