Pre-registration Data Analysis Plan

Investor Characteristics and Founders' Collaboration Interest*

Junlong Feng†, Ye Zhang‡, Weijie Zhong§

February 28, 2022

*We would like to thank Stanford GSB for providing the research funding.

†Hong Kong University of Science and Technology, Department of Economics
‡Stockholm School of Economics Finance Department
§Stanford Graduate School of Business

1 Research Question

Attracting great startups is considered to be crucial to the success of a venture capital (VC) fund's performance. This paper investigates how different investor-level characteristics and fund-level organizational characteristics causally impact startup founders collaboration preferences to venture capitalists using the incentivized resume rating experimental method. Given the importance of different investor characteristics, we aim to answer the following three research questions.

a. Are U.S. startup founders biased against female and Asian venture capitalists? For gender discrimination questions, we care specifically about whether gender discrimination exists in male-dominated or high-tech industries, such as the IT industry.

b. Are impact VC funds aiming for ESG criteria more attractive to startup founders?

c.1 What other investor-level human capital characteristics (i.e., educational background, etc) and fund-level organizational capital characteristics causally impact VC's attractiveness to startup founders?

c.2 What are the relative importance of these characteristics?

Through eliciting startup founders' preferences for various venture capitalists, this project completes a field experimental system with Zhang (2020), which provides empirical insights for explaining several unique equilibrium phenomenon in the entrepreneurial finance literature. For example, why is the gender gap more persistent in the entrepreneurial community? Also, why is the VC fund performance also more persistent compared with other financial assets?

2 Experimental Design

2.1 Recruitment & Experimental Subjects

The main recruitment method used for this experiment is to collaborate with Qualtrics Panel Partners, which helps researchers to recruit real U.S. entrepreneurs and small business owners. However, researchers must provide monetary compensation (i.e. $47) to Qualtrics, who will pay each survey participant later if the experiment is successfully completed by them. Moreover, researchers cannot collect any identifiable information from each survey participant unless participants volunteer to contact the research team.

Target Sample Size: approximately 1000 experimental subjects. (However, if the main experimental results are already highly significant, we may also stop the recruitment process when the sample size is smaller than 1000.)

Inclusion/Exclusion rules: To obtain less noisy evaluation results from qualified participants, we further added two filter questions and some screeners to target founders satisfying the following criteria: 1) being a startup founder or business owner who plans to raise external funding from the venture capital industry, 2) understand the incentive structure provided in the experiment, 3) spending at least 60 seconds on evaluating the first investor profile and at least 15 seconds on evaluating the second investor profile, and 4) successfully answer the attention check question. If any of these criteria fails, Qualtrics will automatically terminate the survey process and inform the founders that they are no longer qualified for this experiment. Unqualified participants do not have a second chance to join this project. We will also consider removing those who spend very little time (defined as the bottom 1% or 5% of total evaluation time) on evaluating investor profiles as a robustness check.[1]

Given that the first few profile evaluations are always noisy, in the final analysis, besides using the full sample, we will also consider testing results after removing the first two or four profile evaluations.

2.2 Experimental Design

Incentives — To implement this startup-side Incentivized Resume Rating (i.e., IRR) experiment, we invite real U.S. startup founders to evaluate multiple randomly generated investor profiles.[2] Founders know that these profiles are hypothetical, but they are willing to provide truthful evaluations due to this "matching incentive" provided to all participants. Following Kessler et al. (2019), the "matching incentive" means that each founder will obtain an algorithm-generated investor recommendation list based on their revealed preferences.

Investor Profile Construction —The experimental design essentially follows the random-

[1] When a survey participant does not fill out parts of its demographic information (such as gender), his/her evaluation results do not enter the relevant heterogeneous effect analysis (for example, gender homophily effect analysis).

[2] IRR experimental method was first created by Kessler et al. (2019).

ization process of a factorial experimental design. To generate VC investors' hypothetical profiles, we randomize multiple investors' characteristics simultaneously and independently across profiles. Each characteristic is dynamically populated from a pool of options, and the matching tool combines these randomly selected characteristics together to create an investor profile. Specifically, participants were told to assume that all the hypothetical investors work in their area of operations and the investors' typical investment size matched their funding needs.

Insert Breaks — To test the implicit gender/racial bias, we randomize the number of breaks inserted. The control group receives 0 breaks in total. The treatment group receives 1 break (i.e, after 10th). These short breaks are designed to indicate the current progress and encourage survey participants to finish this experiment. Based on the psychology literature, a short break will significantly increase subjects' cognitive load and make people fatigued. Therefore, inserting breaks is crucial to test any implicit bias or preferences.

Procedure

Step 1: Consent Form

Real startup founders who are willing to participate in this survey experiment will first read a consent form at the beginning of the survey tool and decide whether they want to participate in this study. After they click the button "Yes, I consent", they will enter the survey. If they click the button "No, I do not consent", they will go to the end of the survey directly. The research team does not collect any identifiable information about these subjects.

Step 2: Evaluation Section

Real startup founders who are willing to participate in this survey experiment will start evaluating 20 investor profiles which they know to be hypothetical and randomly generated in order to be matched with real recommended investors. Subjects need to rate how much they would like to collaborate with different types of investors and how likely these investors will collaborate with them rather than other startup teams.

This experimental method preserves incentives while avoiding the deception necessary in traditional audit studies. It can help examine hypothesis 1,2,3. We will clearly inform the startup founders that all profiles they will evaluate are hypothetical before they fill out the survey. However, the more truthfully they reveal their preferences, the more benefits they can obtain from this study. The benefit is mainly a better matched investor recommendation list, containing 10 most matched real investors' public information based on our designed ML algorithm.

To test the implicit bias, we also randomly insert some short breaks in this evaluation session following the standard IRR experimental design. Basically, entrepreneurs will see some randomly inserted pages that inform them about their study progress. For example, after they evaluate the first 10 profiles, there is a page mentioning that "You have rated 10 profiles of the 20 profiles. Keep up with the good work!"

Step 3: Background Questions

After the evaluation section, experimental subjects need to answer some standard background questions asking their demographic information and their startup background information. Such information can help us evaluate how representative the experiment subjects are.

Step 4: Payment Game

To generate real economic outcomes and test the existence of taste-driven gender bias in the IRR experiment, we add a classical payment game after the evaluation section. Subjects will be randomly assigned to the control group, treatment group 1 and treatment group 2.

For the control group, we tell subjects that we will provide a lottery opportunity to them and randomly pick 2 participants as the lottery winners. The lottery winners have the following two options.

Option 1: receive $500

Option 2: receive ($500-price) and a full investor recommendation list containing 200 most matched venture capitalists' information. The price is randomly drawn from [$20, $80].

We want to know which option the subject will choose.

For the Treatment Group 1, everything is the same as the control group except that to promote gender equality, we would prefer to recommend female investors conditional on the same matching quality based on their indicated beliefs.

For the Treatment Group 2, everything is the same as the control group except that to promote the social responsibility campaign in the entrepreneurial community, we would prefer to recommend impact investors conditional on the same matching quality based on their indicated beliefs.

This is a real lottery opportunity and there is no deception in this experiment. If entrepreneurs have taste-driven bias against female investors or impact investors, we expect less people to choose option 2 in Treatment Group 1 and Treatment Group 2. In order to recover the demand curve for different types of recommendation lists (i.e., a normal recommendation list, a list with more female investors, or a list with more impact investors), price has to be randomized. This can also help us to test whether taste-driven preferences can vary based on stakes involved in the experiment.

Step 5: Social Preference Elicitation Questions (Donation Game)

Following recent papers on sustainable finance (Riedl and Smeets (2017)), we also add two questions to elicit subjects' social preferences. Basically, we would like to provide another independent lottery opportunity to all participants in this section. We will randomly choose another 2 lottery winners and each will receive $1000. If subjects win the lottery, one of their following donation decisions will be randomly chosen to determine their finalized lottery payment. Therefore, it is important to reveal their truthful donation preference.

The first question is if they win the lottery, what percentage of the $1000 would they like to donate to an NGO that supports gender equality.

The second question is if they win the lottery, what percentage of the $1000 would they like to donate to an NGO that aims for generating positive environmental, social and governance (ESG) impact on the entrepreneurial community.

This lottery opportunity is also real, and the research team will donate the money for subjects to the corresponding NGOs. There is no deception in this part.

3 Statistical Model Specification

3.1 Part A: Profile Evaluation Section

3.1.1 Primary Outcomes

We have five co-primary outcome variables, which are the five designed evaluation questions:

Mechanism Questions.

Q1 (First Moment: Quality Evaluation)

What's the probability that you feel [investor name] can help your company generate higher financial returns based on [his/her] quality? (Think only about your perception of [his/her] quality and attractiveness when gauging your interest level in the investor– imagine that [he/she] is guaranteed to finance your startup.)?

Probability of Collaboration

(Not Interested)0-10%-20%-30%-40%...-80%-90%-100% (Want to collaborate for sure)

Q2 (Strategic Mechanism)

What's the probability that you think [investor name] would show interest (e.g. offer a meeting or further discussion) in providing funding for your startup? (Think only about whether you feel he would finance you or not when gauging how likely [he/she] would be to finance your startup, imagining that [he/she] has many startups to choose from.)

Probability of Collaboration

(Will not show interest) 0-10%-20%-30%-40%-...70%-80%-90%-100% (Show interest for sure)

Q5 (Second Moment: Informativeness & Variance)

Imagine that you have access to a professional online profile or resume of the investor. To what extent do you think the profile is informative for evaluating [Investor Name] as a prospective collaborator?[3]

Informativeness

(Not informative) 0-10%-20%-30%-40%...70%-80%-90%-100% (Provide all the information)

Decision Questions

[3] This evaluation question comes from the complementary survey used in Bartoš et al. (2016).

The two decision questions are designed to capture the most important two dimensions in startup's fundraising: likelihood of contact (i.e., external margin) and the proposed funding plan (i.e., internal margin). Moreover, it can re-examine how a founder's preferences evolve from initial contact interest to a later stage fundraising plan (see Zhang (2020)). The proposed fundraising plan question asks the relative funding magnitude rather than the absolute funding magnitude mainly because different startups have different ranges of targeted fundraising amounts. In order to accommodate more founders, we try to make the question as standardized and generally applicable as possible.

Q3 (Contact Likelihood)

How likely would you be to contact [investor name] (e.g. send an email, build networks and relationships) for a meeting to discuss your startup financing, considering both [his/her] potential interest in your startup and your collaboration interest with [him/her]? (Remember that you have limited energy and the algorithm will generate top 10 recommended investors to you based on your preference.)

Probability of Contact

(Will not contact) 0-10%-20%-30%-40%-50%-60%-70%-80%-90%-100% (Contact for sure)

Q4 (Funding plan)

How much money are you comfortable with asking for from [investor name] compared to your original funding plan, considering both [his/her] potential interest in your startup and your collaboration interest with him?

(For example, if you feel it is safe to ask for 80% of your original planned funding needed from [investor name], you can move the bar to 0.8.)[4]

(Percentage) 0-0.1-0.2-0.3——1——1.8-1.9-≥ 2 (Contact for sure)

3.1.2 Secondary Outcomes

We also record the time spent on evaluating each investor profile in order to test founders' implicit bias based on gender, race or investors' ESG-related characteristics. Therefore, the evaluation time (measured in milliseconds) also serves as the outcome variable.

[4] Q4 helps to evaluate the cost of collaborating with certain groups of investors, especially those very attractive ones.

3.2 Statistical Models

3.2.1 Balanced Table

Variables used in the Balanced Table include all the randomized investor characteristics and the reported demographic information of each survey participant. Specifically, we want to include the following experiment participant demographic characteristics:
Gender, race, industry, stage, location (state), political attitudes, entrepreneurial experience, funding experience,[5] nature of the enterprise (for profit vs hybrid combinations vs non-profit), firm size, etc.

3.2.2 Average Treatment Effect

Startup Founder $i \in \{1, 2, ..., I\}$ evaluates the $j^{th} \in \{1, 2, ..., J\}$ randomly generated investor profile. We consider the following two regressions:

$$Y_{ij}^{(k)} = X_{ij}\beta^{(k)} + \alpha_i + \epsilon_{ij}^{(k)}, \tag{1}$$

$$Y_{ij}^{(k)} = X_{ij}\beta_i^{(k)} + \alpha_i + \epsilon_{ij}^{(k)} \tag{2}$$

where $Y_{ij}^{(k)}$ means startup founder i evaluated the k^{th} question for the j^{th} generated profile, and $k \in \{1, 2, 3, 4, 5\}$ denotes the question number as each founder needs to provide the answers to Q1 (belief of investors' quality/value added), Q2 (belief of investors' investment interest), Q3 (founders' contact likelihood), Q4 (amount of funding to be raised) and Q5 (uncertainty or informativeness). Covariates X_{ij} contains the tested variable of interest, mainly including investors' gender, race, and the ESG characteristic. For example, to test whether founders have any gender bias, $X_{ij} = 1$ if the gender of the investor in the j^{th} generated profile evaluated by founder i is female, and $X_{ij} = 0$ if the gender is male. Founder fixed effects are captured by α_i. For the unobservable ϵ, given our experimental design, $\epsilon_{ij}^{(k)} \perp \epsilon_{ij'}^{(k)}$ if $j \neq j'$ while $\epsilon_{ij}^{(k)} \not\perp \epsilon_{ij}^{(k')}$ if $k \neq k'$. Regression (1) pools the founders together and estimates the overall effect. Regression (2) estimates the founder-specific average effect.

[5] Defined as how many investors have invested in their firms.

3.3 Mechanisms

3.3.1 Implicit Bias

Recall that we insert breaks in this experiment. After evaluating some investor profiles (the exact number is determined by whether the participant is in the randomly assigned treatment group or not; see Section 2.2 for details), participants will see a screen displaying the current progress and encouraging them to finish the experiment. Based on psychology literature, inserting breaks might reduce or increase subjects' cognitive burdens and help to test the implicit gender/race/ESG bias.

To test this implicit bias, we will compare the control group with 0 inserted break and treatment group with 1 inserted break. Our hypothesis is that the existence of implicit bias will make the treatment group founders more biased against minority investors (i.e., female, Asian, impact investors). We will also compare the evaluation results of the first half evaluations with those of the second half evaluations following Kessler et al. (2019) across the treatment and the control groups. We will also compare the evaluation results between the pre-break profiles and the post-break profiles for the treatment and the control group respectively. We include 1 or 2 profiles immediately before and after each break and pool them together for each group. If we find that subjects give lower ratings to minority groups in the second half profiles compared with the first half profiles, and or in the second half profiles within a 10-profile block compared to the first half profiles within a 10-profile block, and/or in the post-break profiles compared with the pre-break profiles, that's generally interpreted as the evidence of implicit bias.

3.3.2 Heterogeneous Effect

We estimate various heterogeneous effects.

a. Homophily based on gender, race and ESG characteristics. Aim: Testing whether subjects with similar background is more friendly to investors with the same background.

Note: Specifically, it is interesting to check whether startups aiming for pure financial goals are more attracted by impact funds or not. If not, what's the explanations for the underlying mechanisms.

b. Industry-based heterogeneous effect. Aim: Testing whether founders focusing on tech sectors or male-dominated sectors have more bias against female investors.

c. Stage-based heterogeneous effect (early stage vs later stage). It tests whether founders with different experiences and stages have different preferences. This helps to check whether experiences reduce/strengthen the potential bias.

d. Experience-based heterogeneous effect (novice vs experienced founders).

e. Political attitudes based heterogeneous effect.

f. Educational background based heterogeneous effect. Testing whether subjects with better educational background have less discrimination and favor impact funds.

e. Startup size-based heterogeneous effect.

These heterogeneous effects help us to create a portray of survey participants and check which groups prefer certain investor characteristics.

To estimate these effects, we build the following framework. Let $\beta_{im}^{(k)}$, estimated in regression (2), be the individual average treatment effect of investor characteristic m on startup founder i. Let W_i be the vector of founder i's characteristics that we describe earlier. Let $g^{(k)}(W_i) \equiv \mathbb{E}(\beta_{im}^{k}|W_i)$ be the conditional expectation of such effect. Then we have

$$\beta_{im}^{(k)} = g^{(k)}(W_i) + \eta_i^{(k)} \tag{3}$$

We can estimate and conduct inference about the function $g^{(k)}$ and its partial derivatives by parametric and nonparametric methods. For nonparametric estimation, we will adopt methods such as causal forests (Athey et al., 2019) which allow moderately high dimensional W_i. For implementation, we need to plug in the estimates $\hat{\beta}_{im}^{(k)}$ from equation (2).

Similarly, we can also recover the conditional quantile function $h^{(k)}(\tau; W_i) \equiv q_{\beta_{im}^{(k)}|W_i}(\tau)$ by parametric quantile regression or by applying, for instance, the generalized causal forest methods (Athey et al., 2019) for nonparametric estimation.

3.3.3 Payment Game

We use the payment game to estimate how preference changes with stakes. Recall that in the payment game, each participant faces a randomly generated price $p_i \in [\$20, \$80]$ for a recommendation list. We will divide the price into K equally spaced bins on the price interval. We will calculate the proportion of participants of different groups who select lottery option 2 in the same bin. Groups are defined as, for instance, how much they favor the minority group in profile evaluation. We then estimate the the proportion (difference) changes with price. For the number of bins K, we will set $K = 50, 60, 70$ for robustness. We can also adopt data-driven methods like cross-validation to determine K.

3.3.4 Donation Game

The answers to the two questions can be used to estimate interaction effects. See 3.3.5 below.

3.3.5 Interaction Effect

Race, gender, the ESG characteristics could also subconsciously affect how employers view other profile/resume components. We test for negative interactions between gender, race, ESG and other desirable candidate characteristics (such as the educational background, fund size, IRR, board composition etc). Specifically, the interaction effect of funds' ESG characteristics and gender/race is also an interesting question to investigate. We will use both the full sample, the second half of the profiles evaluated, the post-break profiles evaluated in the treatment group and the control group (as defined in Section 2.2 about inserting breaks) to test these results.

3.3.6 Distributional analysis (especially the second half and the post-break profiles)

Following Kessler et al. (2019), we will use Q3 (contact likelihood) to test the distributional analysis and check whether subjects evaluate female/Asian/ESG investors differently between attractive investors and unattractive investors. Specifically, given that the second half of the profile evaluations often give more accurate results based on implicit preferences/bias, we will also check the distributional analysis based on the second half profile evaluations and the post-break profiles in treatment and control groups.

3.3.7 Decision Based Heterogeneous Effect for Sensitive Characteristics

We will test the decision-based heterogeneous effect according to Zhang (2020) using methods like leave-one-out estimators. Mechanism based questions include Q1 and Q2, decision based questions include Q3 and Q4.

References

Athey, Susan, Julie Tibshirani, and Stefan Wager, "Generalized random forests," *The Annals of Statistics*, 2019, *47* (2), 1148–1178.

Bartoš, Vojtěch, Michal Bauer, Julie Chytilová, and Filip Matějka, "Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition," *American Economic Review*, June 2016, *106* (6), 1437–1475.

Kessler, Judd B., Corinne Low, and Colin D. Sullivan, "Incentivized Resume Rating: Eliciting Employer Preferences without Deception," *American Economic Review*, November 2019, *109* (11), 3713–3744.

Zhang, Ye, "Discrimination in the Venture Capital Industry: Evidence from Two Randomized Controlled Trials," 2020.