Pre-registration Data Analysis Plan

Investor Characteristics and Founders' Collaboration Interest*

Junlong Feng†, Ye Zhang‡, Weijie Zhong§

September 20, 2021

*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?

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 provide 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 each survey participant according to Qualtrics business model. Moreover, researchers cannot collect any identifiable information from each survey participant unless participants volunteer to contact the research team. As described in detail later, we also provide both "matching incentives" and "monetary incentives" to reveal investors' true preferences

Target Sample Size: 1000 experimental subjects (However, if the main experimental results are already highly significant, we may 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 one screener to target founders satisfying the following three criteria: 1) being a startup founder or business owner who plans to raise external funding from the venture capital industry, 2) valuing the data-driven investor recommendation services and agreeing to provide truthful evaluation answers, and 3) spending at least 60 seconds on evaluating the first investor profile. If any of the three 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 to remove 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, we will provide the first 4 profiles for participants to practice. In the final analysis, besides using the full sample, we will also test results after removing the first 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 the "monetary incentive" and "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 if they contact the research team for this list. Following Armona et al. (2019), the "monetary incentive" means that the more correctly/accurately the founders' judgements are, the more monetary award they can receive in addition to the $47.

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. All the relevant investor characteristics are listed in Table 1. 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 will insert 10 breaks in this experiment. After evaluating 2 profiles, there is a short break indicating the current progress and encourages survey participants to finish this experiment. Based on the psychology literature, a short break will significantly increase the mental workload and make people fatigued. Therefore, inserting breaks is crucial to test any implicit bias or preferences.

3 Statistical Model Specification

3.1 Variable Construction & Definition

Detailed definitions of randomized investor characteristics are provided in Table 1. Survey participants' background information includes the following: gender, race, industry (Tech VS. non-Tech), stage, funding/entrepreneurial experiences, ESG goals,

3.1.1 Primary Outcomes

We have four co-primary outcome variables, which are the four 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 résumé 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

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

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

(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-$\geq$ 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.

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 (average annual budget), financial return's expectations

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

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

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. The following regression with founder-heterogeneous coefficients test group-level preferences.

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

where $Y_{ij}^{(k)}$ means startup founder i evaluated the k^{th} question for the j^{th} generated profile, and $k \in \{1, 2, 3, 4\}$ 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) and Q4 (amount of funding to be raised). 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. 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 $k \neq k'$

When testing the relevant importance of multiple individual-level human capital characteristics and the fund-level organizational capital characteristics, we will include all the variables in Table 1 in X_{ij} and implement the multiple hypothesis testing method (q-value method).[6]

3.3 Mechanisms

3.3.1 Implicit Bias

Recall that we insert 10 small breaks in this experiment. After evaluating each two investor profiles, participants will see a screen displaying the current progress and encouraging them to finish the experiment. Based on psychology literature, inserting breaks will significantly increase subjects' cognitive burdens and help to test the implicit gender/race/ESG bias.

To test this implicit bias, we will compare the evaluation results of the first half evaluations with those of the second half evaluations following Kessler et al. (2019). If we find that subjects give

[6] Ewens and Rhodes-Kropf (2015) uses AKM model to identify the relative importance of human capital and organizational capital in the venture capital industry. Building on their paper, we will investigate the importance of detailed characteristics through the sorting channel.

lower ratings to minority groups in the second half profiles compared with the first half 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 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. We may also use some machine learning techniques (with default parameters) to test heterogeneous effect, such as Causal Random Forest or BART.

f. Political attitudes based heterogeneous effect.

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

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

3.3.3 Interaction Effect

Race, gender or 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, 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 and the second half of the profiles evaluated to test these results.

3.3.4 Distributional analysis (especially the second half)

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.

3.3.5 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

Armona, Luis, Andreas Fuster, and Basit Zafar, "Home price expectations and behaviour: Evidence from a randomized information experiment," *The Review of Economic Studies*, 2019, *86* (4), 1371–1410.

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.

Ewens, Michael and Matthew Rhodes-Kropf, "Is a VC Partnership Greater than the Sum of its Partners?," *The Journal of Finance*, 2015, *70* (3), 1081–1113.

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.

Tables

Table 1: Variable Construction & Definition

Variable	Definition and Construction	Example
A. Individual-level Characteristics		
Gender*	(Female, Male) First name to indicate the investor's gender	Emily
Race*	(Asian, White) Last name to indicate the investor's race	Lin
Human Capital		
Investment Experience*	Total number of investments	23
	Years of experience as a venture capitalist (log)	log10
Past Business Experience*	whether have prior experience in industries such as accounting, consulting, legal and other (specify)	Experience Description.
Specialization*	Whether have science and engineer background	Dummy variable
Education Background*	Major	Law, Science/Humanities
	Degree	MBA
	University Ranking	Columbia
Investor Activism	Whether a venture capital firm is involved with recruiting the management team, assembling the company's board of directors, providing assistance with obtaining additional financing, playing a constructive role in building up the internal organization of the firm, or interacting with a portfolio company on a monthly or weekly basis	Dummy variable
B. Fund-level Characteristics		
Organizational Capital		
Investment Experience*	a. Total number of investments (Difference between the logarithm of the number of investments made by the firm prior to year t and the average made by all venture capitalists)	log()-log()
	b. Total number of deals (number of previous investments made by the firm)	70
	c. Total number of IPO (the number of past investments going public divided by the total number of past investments)	0.17
Previous Fund Performance*	Previous fund return[7]	100/103.5
Sequence Number	The sequence (First, Second, Third) of the fund raised by the particular general partner (log)	log2
Reputation and Affiliation*	Industry reputation rank (a ranking of 7 being first best) of VC reputation among offers received, rated by entrepreneur	4

Continued

Variable	Definition and Construction	Example
Fund Size*	a. Dollar value	$64 million
	b. The logarithm of the amount of capital a fund has under management	log(172)
	c. Whether a VC fund is the first fund raised by a particular VC management form or not	Dummy variable
	d. Parent firm experience (the log aggregate amount the parent firm has invested between the parent's creation and the fund's creation)	ln(112)
Portfolio Companies	a. Valuation(B/M and P/E ratios)	B/M=0.514, P/E=16.4
	Stage number of development of the company at the time of the investment (Early-stage, Late-stage)	Dummy variable
Network/Connection*	a. Firm connection (the number of other firms a firm has invested in)	6
	b. The number of ties a VC has(indegree, degree, outdegree, betweenness, eigenvector)	8
Specialization (industry, stage)*	The fraction of all previous investments that the VC organization made in a particular industry	0.74
Organizational Structure	Whether the VC firm is a private independent VC firm	Dummy variable
Active Monitoring*	a. Active ownership, governance and operational engineering (the last pre-PE-ownership year) b. Intensity of engagement of PE houses c. Employment of value-creation initiatives d. Complementing top management with external support e. Leverage and powerful incentives	6 years
Risk Preferences	Option value (anticipation of learning from future investment)	6.2%
Location*	Distance between the venture capitalist and the firm in which he invests	60 miles
Investment Philosophy*	Whether the VC fund cares about the ESG criteria	dummy

Notes: This table shows the construction process and definition of different variables used in this experiment.
The public market equivalent compares an investment in a private equity fund to an investment in the S&P500, calculated by discounting the actual cash outflows and cash inflows that the fund received with the returns on the S&P 500 over the same time period and forming the ratio of the discounted cash inflows over the discounted outflows.