How Venture Capitalists Bet: Evidence from Two Randomized Controlled Trials*

Ye Zhang †

October 31, 2020

Abstract

Understanding the importance of both human and non-human assets in firms' early-stage financing process is crucial to examining theories of the firm. However, it is difficult to empirically generate causal evidence due to data limitations and the lack of exogenous variations. This paper uses two randomized controlled trials with real venture capitalists mainly from the U.S. to identify characteristics of both the project and their teams that causally affect venture capitalist funding. Specifically, I also check their relative importance on investors' decisions. I find that multiple team characteristics (i.e. founder's educational background and previous entrepreneurial experiences) and project characteristics (i.e. traction, business model, location, comparative advantages, etc.) causally affect investors' contact and investment interests by influencing their evaluation of startups' potential financial returns, risk, and loyalty. Although project traction matters the most in my experimental setting, it is fundamentally the investors' belief in the startup's profitability that matters the most. I also find the traditional correspondence test method, to an extent, inappropriate in testing the significance of project characteristics in virtue of the different "signal-to-noise ratio" problem.

Key Words: Venture Capital, Entrepreneurship, Portfolio Selection Criteria, Field Experiments

JEL Classification: C93, D83, G24, G40, J71

*I would like to express my deepest appreciation to Jack Willis, Harrison Hong, and Wei Jiang for their guidance, support, and profound belief in the value of my work. I am also grateful to Donald Green, Mark Dean, Sandra Black, Junlong Feng, Shi Gu, Bingyi Jing, Michael Best, Bentley MacLeod, Alessandra Casella, and Corinne Low for their valuable comments. I thank the participants at the PhD Colloquiums at Columbia University and the investors who participated in these experiments. They provided extremely valuable feedback on how to improve the experimental methods in the future. Special thanks go to Corinne Low, Colin Sullivan, and Judd Kessler for sharing their IRR Qualtrics code package. This project was supported by PER funding from the Columbia University Economics Department, by the Columbia University Eugene Lung Entrepreneurship Center Fellowship, and by the Columbia CELSS Seed Grant. The project is registered at AEA RCT Registry (AEARCTR-0004982). All errors are my own. Please email the author for code examples.

†Columbia University Economics Department. Email: yz2865@columbia.edu.

1 Introduction

Compared with the rich empirical evidence of investment strategies in the secondary market, causal evidence of venture capitalists' startup selection strategies is limited. Kaplan, Sensoy and Strömberg (2009) provide a systematic description of how firm characteristics evolve from early stage to public firms, suggesting that early-stage investors should place more weight on the project ("the horse") than on the founding team ("the jockey"). However, Bernstein, Korteweg and Laws (2017) implement the first field experiment to study the selection process of early-stage investors, showing that investors weight a team's characteristics more strongly than the project's in the AngelList email setting. Understanding whether the team matters more than the project or vice versa for early stage firms' financing is important for understanding what resources constitute a firm in their early stages and what types of assets determine a firm's growth. As pointed out by Rajan (2012), studying how firms are financed in their very early stages, where the link between the nature of firms and the availability of financing is most clear, is important for examining several fundamental premises of existing theories of the firm.[1]

Examining these questions is empirically challenging due to data limitations and the lack of exogenous variations. Most existing data sources fail to observe startups' unique comparative advantages despite the progress in obtaining the nearly "ideal data" which cover all potential young firms (Ewens and Townsend (2020)).[2] This means that non-experimental studies conducted without considering any exogenous variations may suffer from severe omitted variable bias, making them difficult to find appropriate control and treatment groups in the entrepreneurial setting. Although previous field experiments (Bernstein et al. (2017), Gornall and Strebulaev (2020)) use the correspondence test experimental design to remedy these limitations, they usually generate evidence in the initial contact stage and have difficulties in testing underlying mechanisms. Moreover, as pointed out by Bernstein et al. (2017), the correspondence test results may suffer from the problem of different "signal-to-noise" ratios.[3] Hence, it may not be appropriate to use the correspondence test experiment to identify the importance of startup project characteristics and compare the relative importance of team (human assets) and project (nonhuman assets) characteristics.

In this paper, I have constructed a global VC investor database and implemented two randomized control trials in 2020 by recruiting real VC investors from mainly the U.S. to identify key characteristics of startups which causally affect venture capitalists' investment process, and to also test why these startup characteristics affect investors' decisions. Specifically, this paper joins the debate on the relative importance between the "jockey" (i.e. the founding team) and "horse" (i.e. the project) in the entrepreneurial financing literature. These tested startup characteristics include both the startup's team characteristics (e.g. previous entrepreneurial experiences, educational background, etc) and

[1] Theories of the firm usually classify assets into human assets and non-human assets. In the framework of Hart and Moore (Hart (1995)), firms are defined by nonhuman assets. However, Zingales (2000) and Rajan and Zingales (2001) argue that human assets have become more and more important for the current "new firm" compared with traditional firms.

[2] Several recent papers make this data progress. see Guzman and Kacperczyk (2019), Ewens and Townsend (2020), Hu and Ma (2020), Hebert (2020) and other papers using Census data like Cen (2020).

[3] Although they find that investors only respond to the founding team rather than the project, they cannot fully separate the following two mechanisms: (a) A good team plays a more important role than a good project for investors' investment decisions; (b) Project characteristics are noisier and harder to verify than team characteristics in the pitch email setting. Therefore, experienced investors will rationally ignore these nosier project characteristic information in their experiment setting even though a good project can be more important than a good team.

the project-related characteristics (e.g. traction, business model, targeted market, etc.). I test both investors' initial contact interests and investment interests. In addition to the average treatment effect of these characteristics, I also show how investors respond to different startup characteristics on the distribution of investors' interests.

I first start with a lab-in-field experiment (i.e. Experiment A), which mainly follows the latest incentivized resume rating (IRR) experimental paradigm created by Kessler, Low and Sullivan (2019). By collaborating with several incubators, I build a machine learning matching tool and invite real U.S. investors to evaluate 16 randomly generated startup profiles. Investors know the profiles are hypothetical, but they are willing to provide truthful evaluations so that the algorithm works better to help them find real matched investment opportunities. Besides this "matching incentive", randomly selected investors also receive a "monetary incentive" following Armona, Fuster and Zafar (2019). I provide a lottery to these participants, and the lottery winners will receive $500 and extra financial returns, which are calculated based on their startup evaluation results and future 12-month U.S. startup performance. Basically, the more accurately their evaluation results are, the more financial returns each lottery winner will receive.[4] Investors need to evaluate each startup's quality, loyalty (defined by the startup's willingness to raise funding from them rather than other investors), risk, and indicate their contact interests and investment interests. This customized experiment is designed to test investors' startup selection criteria in a setting where all of the startups' characteristics have similar "signal-to-noise" ratios. Moreover, researchers can test investors' investment interests besides their initial contact interests.

Results of Experiment A show the following three main findings. First, both startup team characteristics (including the founder's education and previous entrepreneurial experiences) and project characteristics (including traction, business model, company age, comparative advantages) causally affect early-stage investors' contact and investment interests through influencing investors' evaluations of each startup's future financial returns, risk and loyalty. This confirms the importance of human assets and provides new evidence of the importance of non-human assets in early stage financing of firms. Secondly, project traction plays the most important role in this experimental setting where the "signal-to-noise" ratios of different startup characteristics are similar. Hence, it challenges the conclusion that the founding team matters the most in the previous correspondence test experimental setting, and supports the story of Kaplan et al. (2009). Strictly speaking, the relative importance of human and nonhuman assets should depend on investors' investment philosophy and the startup characteristics pool available for investors to evaluate.[5] Thirdly, although influential characteristics are very consistent between investors' contact and investment interests, evidence suggests that taste-driven factors can only affect investors' contact decisions. However, these factors have very limited influence on investors' later-stage investment decisions. This implies the importance of differentiating between taste-driven and belief-driven mechanisms for studies that only observe the initial contact interests.

[4] Although the "monetary incentive" is noisier than the "matching incentive", it is friendly for researchers without many social connections, and it helps increase the experiment sample size.

[5] This experiment does not test the importance of startup founders' personal traits, such as passion, charm, reliability, or leadership. In practice, these founders' characteristics may play an important role in investors' decisions and make the team stand out. Also, some investors may care about other startup characteristics, such as the timing of the startup's product. Therefore, the available characteristics pool provided in Experiment A is smaller than the information pool available to investors in the real world.

Experiment A makes the following improvements compared with a standard correspondence test design that studies early-stage investors' startup selection criteria. First, the latest experimental design mitigates the concern of different "signal-to-noise" ratios of different startup characteristics by providing investors with the incentive to indicate their true investment preferences for the randomized, hypothetical startup profiles. Therefore, this design is more suitable to test the relative importance of different startup characteristics. Secondly, in addition to the initial contact interests, researchers can also observe investors' investment interests. Hence, it is possible to test how preferences evolve from the initial contact stage to later investment stage. Thirdly, this experimental design tests a much larger variety of startup characteristics and provides researchers with more flexibility to answer various economic questions. However, before we challenge the previous correspondence test method that tests the relative importance of startup team and project, it is important to show that the previous experimental design does not identify the importance of startup project characteristics even for our sample investors. [6]

I follow up with a new correspondence test experiment (i.e. Experiment B) to test whether the standard correspondence test experimental design can successfully identify the importance of startup projects for our sample investors. To mimic the experimental design used by Bernstein et al. (2017), I sent hypothetical pitch emails with randomized team characteristics (i.e. the founder's gender and race indicated by the names, and the founder's educational background), and startup project characteristics (i.e. comparative advantages, traction, etc.) displayed in both the email's subject line and contents in 2020.[7] To solve the "low-response-rate" problem, I utilize a new email behavior tracking technology and record investors' detailed information acquisition behaviors, which include each investor's email opening behaviors, time spent on pitch emails, and click rates of inserted startup's websites. These experimental designs and behavior measurements not only help replicate results of Bernstein et al. (2017), but they also generate enough experimental power to survive in a hash experimental environment when early-stage investors dramatically slow down their investment pace (Howell, Lerner, Nanda and Townsend (2020)) during the COVID-19 pandemic.

Results of Experiment B confirm that investors only respond to startup teams' characteristics rather than project characteristics in the pitch email setting, which indicates that this experimental design may not be suitable to test the importance of startup project characteristics or any nosier startup characteristics. Revealing excellent educational background of the founding team can increase the pitch email opening rates by roughly 1%. Similarly, sending pitch emails using female names increases the email opening rate by 1% compared with male names. Sending pitch emails using Asian last names generally increases the email opening rate by 0.7% compared with white names. However, investors do not respond to startup project characteristics, such as traction or comparative advantages. The importance of these project characteristics has been proved in Experiment A. Therefore, results suggest that the previous correspondence test design may not be informative of the importance of non-human assets during the entrepreneurial

[6] Bernstein et al. (2017) recruited more angel investors while most our sample investors are institutional investors. Therefore, the different results we find can also come from the sample variability. To rule out this potential mechanism, we need to replicate their experimental design on our sample investors to check whether the correspondence test design can successfully identify the importance of startup projects.

[7] When this research project began at the beginning of 2018, I started with two alternative experimental designs. Unfortunately, both of them failed due to different reasons, and the related discussions of alternative designs are provided in Section 4.1. I choose the current version after long time discussions about the experiment feasibility and risk with Columbia's IRB.

financing process due to its noisy setting.

The contribution of this paper is both empirical and methodological. Empirically, it contributes to the following strands of literature. First, it provides new experimental causal evidence of how venture capitalists select and evaluate early-stage firms based on new characteristics that had not been previously tested. Especially, this paper proves the causal importance of non-human assets of firms in their early-stage financing process. Hence, it contributes to the entrepreneurial finance literature. Second, it provides new insights on the relative importance of startup team and project characteristics for investors' decisions, hence it contributes to the theory of firm literature. Third, it examines how preferences driven by different mechanisms evolve from the initial contact stage to later stage decisions. As a result, it is helpful to evaluate results of previous studies which only observe evaluators' initial contact interests (i.e. correspondence test studies). Methodologically, Experiment A provides an experimental design that can introduce more accurate variations in startup project characteristics while also controlling startup project quality. This is important for establishing causality using RCT methods in the entrepreneurial financing setting.

This paper is organized as follows. Section 2 discusses the construction of the individual-level global VC investor database by merging multiple commercial databases with manually collected data. 3 presents the design of Experiment A and analyzes investors' evaluations of multiple startup team and project characteristics. Section 4 describes the design of Experiment B and analyzes investors' information acquisition behaviors. Section 5 concludes.

2 Data

I construct a cross-sectional individual-level global venture capitalists' database, which contains 17882 investors' most recently updated demographic information and contact information before 02/2020. This database contains only valid investors in the English-speaking areas whose email addresses are verified by the testing email used in the correspondence test. Considering that the experiments are implemented in English, I did not include investors from the Europe and most Asian areas. Therefore, strictly speaking, the database used in this paper is a subset of a more comprehensive constructed global venture capitalist database that also contains investors from the Europe and China.

This global database combines the following commercial databases including Pitchbook, ExactData, CB Insight, SDC New Issues Database VentureXpert, and ZDatabase.[8] For investors whose contact information is not available in these commercial databases, I supplement this database with contact information collected from RocketReach. All key variables used in the analysis, including gender, location and industry, are manually verified through multiple social platforms including LinkedIn, company websites, personal websites and online news if such information is not available on Pitchbook. Detailed database description and key variable construction process is provided in Appendix 2.

Despite of the granular information provided by this database, it is important to realize the following three limitations.

[8] Many of these used commercial databases are not free and require researchers to sign a data contract for academic purposes.

First, this database contains systematically more investors from the U.S. and more senior VCs due to the data availability online and data collection method used by the data companies.[9] Hence, it may not be representative of the true geographical distribution of all the venture capitalists in the world. Second, because of the high turnover rates within the VC industry, the contact information and the status of these investors need to be updated frequently before using it. Third, except for the key variables, such as gender, seniority and location, other demographic variables are only available for relatively famous investors whose biographies are more available online.

The Summary Statistics of the 17,882 investors' demographic information is provided in Table 1. Panel A reports the location distribution of these investors, showing that the U.S. investors account for 84.91% of my sample investors. The map of investors' global geographical distribution is provided in Figure 1 and the U.S. geographical distribution is provided in Figure 2. Panel B shows that most of investors are interested in the Information Technology industry. Other important preferred industries include Healthcare, Consumers and Energy. Panel C summarizes investors' background information. On average, female investors account for 24% of total investors. This is consistent with NVCA/Deloitte survey's results showing that women have accounted for 21% of investment professionals of the U.S. VC industry in 2018 due to recent progress of increasing diversity.[10] Senior investors, who are partners, president, C-level managers, or vice president and above, account for 84% due to the online information availability. Most of investors are institutional investors and angel investors only account for 11%. Therefore, it is important to note that our sample investors can be different from the angel investors recruited by Bernstein et al. (2017). 61% of investors attend graduate schools and more than 30% of them participated in top universities. This is consistent with Gompers and Wang (2017), which shows that VC investors are usually better educated.

3 Experiment A: Lab-In-Field Experiment

Experiment A is designed to identify venture capitalists' preferences towards a rich number of startup team and project characteristics in a clean experiment setting where information of startup characteristics is less noisy compared with previous correspondence test experiment designs. Specifically, this experiment also tests how investors' contact interests transform into the investment interests. The tested startup team characteristics include founders' educational backgrounds, previous entrepreneurial experiences, the number of founders, etc. The tested startup project characteristics include the company's business model, comparative advantages, location, company age, funding situation, and etc. These startup characteristics are not sensitive characteristics because any investment strategies based on these startup characteristics are acceptable in the U.S. society. The key challenge here is to provide more accurate startup information, especially project information, to investors so that researchers can compare the relative importance of team and project characteristics in a cleaner experimental environment. For the studies about more sensitive startup characteristics evaluation results (i.e. biases based on founders' gender, race, and age, or evaluations of startup's social

[9] Most of the commercial databases used here are provided by the U.S. data companies and collected by English speakers except ZDatabase, which is the most comprehensive and timely database covering VC and PE activities in China.

[10] see https://www2.deloitte.com/content/dam/Deloitte/us/Documents/audit/us-audit-egc-nvca-human-capital-suvey-2018.pdf Gompers, Mukharlyamov, Weisburst and Xuan (2014) also shows that women are under-representative among senior investment professionals in the VC industry.

responsibilities), please see Zhang (2020a) and Zhang (2020b).[11]

The main part of Experiment A uses the latest preference elicitation techniques, which is the IRR experimental paradigm. I invite real U.S. venture capitalists to try a "Nano-Search Financing Tool", which is a machine learning, algorithm-based matching tool that seeks for potential investment opportunities. In the profile evaluation part of this tool, investors need to evaluate multiple randomly generated startup profiles, which they know to be hypothetical, in order to be matched with high-quality real startups from the collaborative incubators. With the carefully designed evaluation questions and incentive structures, Experiment A has very strong internal validity, which means that the design provides stronger incentives to elicit investors' investment preferences and can directly test detailed belief-based mechanisms. Importantly, startup characteristics in this experiment setting have similar "signal-to-noise" ratios.

The experimental setting that develops data-driven methods to help investors evaluate potential deals is not special in the venture capital industry. A few incubators and VC funds have done extensive work on developing machine learning algorithms to help evaluate investments.[12] However, considering that several important startup characteristics, like the founder's passion and confidence, cannot be fully quantified by the data, these data-driven methods are usually designed to complement existing mainstream person-to-person multiple stage investment strategies rather than to fully substitute the existing due diligence methods.

This section is organized as follows. Section 3.1 introduces the experimental design and implementation details. Section 3.2 describes the analysis of the results of investors' evaluations of startup team and project characteristics. Section 3.3 discusses the limitations of this experiment.

3.1 Experimental Design

3.1.1 *Investor Characteristics and Recruitment Process*

Experiment A was implemented during 03/2020-09/2020 using only the online recruitment method.[13] I sent invitation emails together with the instruction posters to the 15000+ U.S. venture capitalists who also participated in Experiment B (see Appendix B Figure B5 and Figure B6 for the recruitment emails, Figure B7 and Figure B8 for the instruction posters). Both the recruitment emails and posters emphasized the matching purpose of this tool. However, investors were also notified of the research purposes, and they understand that their anonymized data would be used for studying investors' preferences for different startups characteristics as required by IRB. Therefore, this study has the ecological validity of a "natural field experiment" except that subjects understand that the academic research is also taking place in order to avoid deception.

[11] These startup characteristics are sensitive because investors have the motivation to distort their revealed preferences if no incentive is provided. For example, most people will claim that they would love to support female founders and ESG-related startups although their true thoughts may be different.

[12] For example, Techstars, Social+ Capital, and Citylight Capital, etc. Also, Open Scout, a startup working with Angel Capital Association (ACA), is designing platforms to connect founders with investors based on shared interests rather than shared network on their platforms.

[13] During the COVID-19 pandemic, Columbia's IRB paused all the field work which involved person-to-person activities due to the impact of COVID-19.

There are in total 69 real U.S. investors from 68 different funds participating in this project, which provides 1216 startup profile evaluation results.[14][15] The number of recruited experiment participants is comparable with Kessler et al. (2019) and one advantage of the IRR experimental design is that researchers can obtain a large enough sample size through recruiting a relatively small number of participants. This advantage is crucial for the experiment to survive in an environment where it is hard to recruit a large number of subjects.

Similar to the majority of experiments, Experiment A with roughly 0.5% response rate also has the sample selection bias during the recruitment process.[16] Based on the observable investor information, Table 2 reports the summary statistics of participants' backgrounds, showing that our sample investors are more likely to come from larger VC funds and be minority founders.[17] The average asset under management (AUM) of these VC funds is $547.46 million, which is larger than the average AUM $444.44 million in 2019 based on an NVCA survey. During our recruitment period, only larger funds still have money to look for new investment opportunities in the COVID-19 recession when most of VC funds have shifted to "survival mode". 42% of investors in the sample are from minority groups (i.e. Asian, Hispanic, African, and etc.), which is higher than the percentage of minority investors in the U.S.[18] However, our sample investors are representative in other dimensions. Recruited investors are mainly early-stage investors with preferences covering almost all major industries that VC focuses on. 86% of recruited investors are in the senior positions, and about 20% of investors are female. This is consistent with the situation described by the global investor's database.

Sample selection bias can also arise from the following unobservable reasons. First, participants are likely to be more pro-social and willing to help academic research studies. Second, our sample investors are likely to have preferences towards Ivy League College because the research project is supervised by Columbia University. Third, recruited investors are more likely to be interested in understanding how data-driven methods can help investment evaluations. Many investors also reject to participate because they do not believe that the algorithm can help startup portfolio selection process if it does not quantify the founder's personality and the chemistry during the meetings. These sample selection biases do not hurt the experiment internal validity, but imply that it is important to replicate this experiment in different settings in order to check the external validity.

[14] Recruiting real venture capitalists is crucial to understanding startup investing strategies because venture capital investment involves very specific skills. Carpenter, Connolly and Myers (2008) documents that lab experiment results provided by college students are very different from results provided by community members, which confirms the importance of implementing lab experiments in the field.

[15] At the beginning of the study, each investor evaluated 32 profiles and 6 investors finished the 32-profile version evaluation task. However, to recruit more investors, later participants only need to evaluate 16 profiles. One investor participated twice for different funds. Results are similar after removing the first 6 investors. As more investors participate in Experiment A, I will update the results again in the future.

[16] Future researchers can recruit investors by participating in different real events after the COVID-19 pandemic or collaborating with certain associations (i.e Angel Capital Association (ACA) or National Venture Capital Association (NVCA)) to increase the response rate.

[17] Recruited investors are likely to be the still active investors during the recession. Based on many investors' replied emails, investors usually refuse to participate in this research because they have shifted to "survival mode", which implies that they focus on helping their invested startups to survive rather than "purchases" new undervalued startups in 2020.

[18] Considering that the researcher is Asian female, it is not surprising to find that more minority founders are willing to participate in this research study.

3.1.2 *Survey Tool Structure*

If investors are interested to participate in this experiment, they need to open the link inserted in the recruitment email to start the Qualtrics survey online using their browsers. The survey tool contains the following two sections. After reading the consent form, investors will first enter the profile evaluation section (i.e IRR experiment section) where they need to evaluate 16 randomly generated startup profiles and answer standard background questions. In the second donation section (i.e. dictator experiment section), investors will decide how much of an unexpected $15 Amazon Gift Card they want to donate to randomly displayed startup teams. For this paper, our focus is to discuss results from the first section. For results from the second donation section, please see Zhang (2020a). Figure 3 provides the experiment flowchart that demonstrates the tool structure.

A. Consent Form and Instruction Page

Both consent forms and recruitment emails invited investors to "try a matching tool that help identify matched startups" and also noted that the anonymized data from investor responses would be used for studying investors' startup selection criteria, which was framed as secondary. Before the first profile evaluation section starts, I also provide an instruction page emphasizing that "the more accurately they reveal the preferences, the better outcomes the matching algorithm will generate (and the more financial returns that the lottery winner will obtain)" so that participants understand how the incentive works. Moreover, since most of VC investors only invest in startups of their interested industries and stages (called "the quality/disqualify test,[19] I ask all the participants to assume that the generated startups to be evaluated work in their interested industries and stages.[20]

B. Section 1 (Incentivized Resume Rating Experiment)

B.1 Profile Creation and Variation

[19]The first step of the investment process is to implement the "quality/disqualify" test before investors go through startup team composition and financial performance. The test, as a quick decision making, is based on many things such as the industry, stage, prior market knowledge and other factors, which tells investors whether the startup is worth looking at. For example, an investor who invests exclusively in the B2B SaaS sector does not want to evaluate a healthcare startup. It is important to consider how to pass the "quality/disqualify" test when designing the IRR experiment as documented in Kessler et al. (2019) when they fail to replicate the IRR experiment in the University of Pittsburg.

[20]Another potential way to pass the "quality/disqualify" test is to provide several survey questions asking the interested industry, stages and even revenue range before the evaluation section as what Kessler et al. (2019) did. For each different industry, researchers need to create different customized generated startup profiles which can capture the special characteristics in that industry by providing more details. I did not do this due the following two reason: First, the market changes very quickly in the entrepreneurial community. However, it usually takes a long time to prepare a field work which needs the approval from IRB. Therefore, it is hard to predict whether the startup information created in the design stage is still valid when I send out the invitation emails. Such situation happens during the COVID-19 period when multiple industries got hit hard within a short period. Second, from the research purpose, I need the insights from investors focusing on different areas and industries. This requires that the information provided should be general enough to accommodate as many participants with diverse backgrounds as possible.

Given the restrictions mentioned above, I only choose to provide the information that is usually publicly available on LinkedIn, Crunchbase or AngelList. Also, it provides the description of each hypothetical startup's comparative advantages. Some investors like Plug and Play Tech Centers sometimes go to these public platforms and look for relevant startups that fit their portfolios. The current design is to mimic this type of startup seeking behaviors and provides data-driven methods for the pre-selection decisions rather than fully substitutes the mainstream person-to-person deal flow process. Future researchers can think about more dedicated ways to pass the "quality/disqualify" test

Following the factorial experiment design, multiple startup characteristics are dynamically varied simultaneously and independently, enabling me to test investor preferences over multiple important startup characteristics suggested by the existing theories.[21] I first create a set of team characteristics (including founding team's gender, race, age, education, previous experience, etc.), project characteristics (including market traction, comparative advantages, location, ESG criteria, etc.) and existing financing situations. Then the back-end JavaScript code will randomly draw different characteristics and combine them together to create a hypothetical startup when each participant evaluates a new startup profile.[22]

In order to generate relatively reasonable startup profiles, I implement the following three designs. First, each hypothetical startup profile is constructed using different components those ranges are based on data of PitchBook Database. Second, the information provided follows the format similar to CrunchBase format and captures most of the online publicly available information of each startup.[23] I did not provide extra private information like equity sharing plan because such information is usually not disclosed to the public in the pre-selection stage and is usually determined after several rounds of negotiation between investors and startup founders. Third, I also introduce a short break after investors evaluate the first half part of the startup profiles (i.e. the first 8 profiles) by providing them with a progress screen and a startup ID for each profile to indicate the evaluation progress. Such break is usually designed for testing implicit bias. All the randomization of different startup components, including startup team and project characteristics, is provided in Table 3. The detailed startup characteristics construction process is provided in Appendix B

B.2 Evaluation Questions

The evaluation questions include three mechanism questions designed to directly test belief-based sub-mechanisms, and two decision questions designed to compare investors' initial contact interest and later-stage investment interest. Considering that most venture capitalists are well educated and sophisticated enough, I usually ask a probability or percentile ranking question rather than a Likert scale question,[24] which has two advantages. First, these questions are more objective than Likert Scale questions. Second, the wide range from 1 to 100 provide richer and more detailed evaluation results and additional statistical power. This question design allows researchers to implement infra-marginal

[21] Introducing a rich set of randomly generated startup characteristics is usually not feasible in the correspondence test because of the following two reasons. First, the unusual combinations of characteristics might raise investor's suspicion. Second, all the varied information inserted in the pitch email may not be salient enough. For example, it is a reasonable idea to randomize the traction or comparative advantages of a startup in the correspondence test. However, investors do not respond to such randomized information either because they feel this information is not verified and super noisy, or because it is hard to compare the information with the benchmark because different founders may have different writing style and some founders do not want to disclose too much information about their traction before they meet the investors. Therefore, as Bernstein et al. (2017) mentions in their paper, failing to find significant results related to the project traction does not imply that project does not play a role in the investment process.

[22] Sometimes the random combination may generate unusual cases like a startup with 50+ employees still do not generate profits (see Amazon's history). Such cases account for a small percentage of total generated cases. However, future researchers can think about how to mitigate this issue when a rich set of characteristics are randomly varied and combined at the same time. It is helpful to collect as many as such uncommon cases first to generate many filter criteria when writing the randomization code in order to capture the most common situations.

[23] Crunchbase is a commercial platform that provides public information of startups mainly in the US.

[24] Similarly, Brock and De Haas (2020) uses probability questions to replace Likert Scale questions when they recruit real Turkish bankers to evaluate different loan profiles in their lab-in-field experiment.

analysis and distributional analysis that explore how investor preferences change across the distribution of contact and investment interest. The screenshots of how these questions look like are provided in Appendix B Figure B3 and Figure B4.

Mechanism Questions

The three mechanism questions are designed to test the following three standard belief-based mechanisms which potentially explain why investors care about certain startup characteristics. First, some startup characteristics can be indicators of the startup's future financial return. To test this channel, investors need to evaluate the percentile rank of each startup profile compared with their previously invested startups, which is the quality evaluation question ($Q1$). Second, some startup characteristics may be suggestive of the startups' willingness to collaborate with certain investors rather than using other financial tools for their fundraising purposes, which is the "loyalty" evaluation question ($Q2$). Similar to the marriage market, entrepreneurial financing process is also a two-sided matching process. Therefore, this type of "loyalty" potentially also matters. To test this channel, investors need to evaluate the probability that the startup will accept their investment rather than other investors. Third, investors may use certain startup characteristics as indicators of the startup's risk (i.e. the second moment). Therefore, investors also evaluate the risk percentile rank of each startup profile compared with their invested startups,[25] which is the risk evaluation question (Q_5).

The risk evaluation question is added when I recruit investors using only the matching incentive for robustness test purposes. During the recruitment process, I received feedback to add such questions from several investors. Therefore, when recruiting the rest investors using only matching incentive, such risk evaluation question is added at the end of all the evaluation questions to minimize its impact on all the other questions while collecting information about such important channel.[26]

Q1. (Quality Evaluation, First Moment) Imagine that [Founder Name] team is guaranteed to accept your investment offer. Compared with firms you have previously invested in, which percentile do you feel this startup belongs to considering its quality?

0 (extremely low quality) —— 100(extremely high quality)

Q2. (Collaboration Likelihood Evaluation, Strategic Channel) Considering the potential network and negotiation power of Timothy Cheng's startup team, what's the probability that this startup team will accept your investment offer rather than that of another investor (Angel, VC, Loans, etc.)?

0 (guaranteed rejection) —— 100(guaranteed acceptance)

[25] For special characteristics like founder's gender, race and age, the first mechanism question (Q_1) tests one of the most common statistical discrimination mechanisms. The second mechanism question (Q_2) tests a typical confound channel in a two-sided matching market in the discrimination literature. The third mechanism question (Q_5) sheds light on whether the belief of expected variance affect investor's decision or not, which is discussed in detail in Neumark (2012) and Heckman (1998)

[26] Similar to evaluating variance when testing discrimination in the labor market, obtaining investors' evaluation of risks for different startups is difficult using traditional empirical methods. However, considering its importance, such mechanism is important to test if researchers need to fully understand investor's investment decisions. An alternative way to obtain such information is to implement a new field project (for example, send an extra survey) as what Bartoš, Bauer, Chytilová and Matějka (2016) did. However, since the alternative method cannot guarantee to collect information from the same group of investors, I decided to add such question after adjusting the pre-registration plan and making modification in the IRB proposal before implementing this change.

Q5. (Risk Evaluation, Second Moment) Compared with your previous invested startups, which percentile do you feel this startup belongs to considering its risk level (i.e. the level of uncertainty of achieving the expected finance returns)?

0 (No risk) —— 100(Highest risk)

Decision Questions

The two decision questions are designed to examine how the investors' preferences evolve from the initial contact interest to the investment interest. Traditional experimental method, like the correspondence test, generally observes the initial contact interest from candidate evaluators. However, it is still unknown that whether the contact interest can fully transform into the investment interest or affect any later-stage decisions. Therefore, I ask each experiment participate to indicate both their contact interest ($Q3$) and investment interest ($Q4$).[27]

Q3. (Contact Interest) If you consider both the team's attractiveness and their likelihood of collaboration, how likely would you be to ask for their contact information or pitch deck?

0 (will not ask) —— 100(will ask)

Q4. (Investment Interest) Considering both the team's attractiveness and their likelihood of collaboration, how much money would you invest in this startup compared to your average investment amount? Imagine that the startup asks for the amount of money that you can afford.
(For example, if your average amount of investment per deal is $1M and you would invest $0.5M to the team, drag the bar to 0.5.)

0 —— 1.0 (benchmark) —— >2.0

C. Section 2 (Background Questions)

At the end of the matching tool, I also collect participant's standard background information to check how representative my sample investors are and implement potential heterogeneous effect based on predetermined investors' characteristics. Such background information includes investors' preferred industries, stages, special investment philosophy (for example, only invest in social ventures and women-led startups) and standard demographic information, which include gender, race and educational background. It is important to ask the background questions after the evaluation section in order to avoid priming subjects to think about any particular characteristics that the research project aims to test. There is one extra section of a donation game in Experiment A. However, this paper does not discuss its design or results. For readers who are interested in this donation section, please see Zhang (2020a).

3.1.3 *Incentives*

As an incentivized preference elicitation technique, the key point of the IRR experiment design is that when subjects evaluate randomly generated hypothetical startup profiles, they understand that the more accurately they reveal their preferences, the

[27] The investment interest question asks the relative investment interest rather than the investment magnitude mainly because different investors have different range of targeted investment amount. In order to accommodate more investors, I try to make the question as standardized and generally applicable as possible.

more benefits they will obtain based on the incentive provided. Therefore, for all investors, I provide a "matching incentive" used Kessler et al. (2019). To increase the sample size, for a randomly selected subset of investors, I provide both the same "matching incentive" and a "monetary incentive" used by Armona et al. (2019). Considering the amount of time required for participating in this experiment, most of participants should value the incentive.[28] The details and justifications of both incentives are provided in the following two subsections.

A. Matching Incentive

For the randomly selected 4000 investors who receive the recruitment email (Version 1), I only provide a "matching incentive" which means that after each investor evaluates 16 hypothetical startup profiles, we will use a machine learning algorithm to identify matched startups from our collaborative incubators and such startups will contact them for a potential collaboration opportunity if they are also interested in the investor's investment philosophy. The matching algorithm will use their evaluation answers to identify their preferences for different startup characteristics similar to Kessler et al. (2019). Therefore, all the five evaluation questions are incentivized by providing this incentive and the description of the algorithm is provided in the consent form.

The matching incentive has the following three merits. First, it can be applied to any two-sided matching market, like entrepreneurial financing market and marriage market. Second, it can be used to incentivize all the evaluation questions compared with the monetary incentive. Third, if the designed matching algorithm can improve the matching efficiency, such incentive can bring real value to both sides of the matching market. Despite of the merits mentioned above, such incentive often requires researchers to have certain social resources and connections to implement.

B. Matching Incentive + Monetary Incentive

In order to increase the sample size, I provided both a "matching incentive" and a "monetary incentive" to randomly selected 14000 investors who received the recruitment email (Version B). Following Armona et al. (2019), the "monetary incentive" is essentially a lottery that 2 experiment participants will be randomly selected to receive $500 each plus an extra monetary return closely related to their evaluation of each startup's quality. Based on this monetary incentive, the more accurate their evaluations of each startup's quality are, the more financial return they will obtain as a lottery winner.[29] The evaluation results will be determined based on the Pitchbook data published in the next 12 months after the recruitment process is finished. I informed the two randomly chosen investors of the award separately by email at the end of July 2020 and also contacted the other participants by email separately that they are not chosen at the end of July 2020. The evaluation algorithm is provided in the consent form (Version 2).

[28] Some may be concern that there are potential two alternative motivations for investors to participate in this experiment. The first alternative incentive is to understand the algorithm and research method I am using for this matching tool. For such investors, the optimal decision is to read the consent form, evaluate a few startups and stop because the evaluation process is repetitive and time-consuming. The second alternative incentive is that some investors are very pro-social and willing to help the research on entrepreneurial activities. However, this survey tool takes at least 20-30 minutes to finish and some investors even replied to me that they would love to participate only if they are provided $5000 consulting fee. Therefore, none of these alternative motivations should be a serious concern.

[29] For example, Peter Smith participated in this survey study and was chosen as one of the two lucky draw winners. In his survey, he indicated that on average, he felt that male teams are of higher quality and more likely to generate higher financial returns. Then we would construct a portfolio containing more real startups with male teams. After one year, based on the financial performance of the portfolio on PitchBook Platform, this portfolio containing more startups with male teams would generate 10% return. Then Peter Smith would receive $500 + $500*10% = $550 as his finalized monetary compensation one year after he participated in the survey. $500*10%=$50 is the "extra monetary return". The historical return of VC industry is between -15% and +15%, which means that the range of expected monetary compensation is roughly between $425 and $575.

Such monetary incentive has the following merits and limitations. First, it mimics the real investment process that given certain amount of the principal, investors need to evaluate different startups accurately to generate maximum return. Second, compared to the matching incentive, such monetary incentive does not require too much social resources and is easy to implement. Third, this incentive can be applied to more general situations besides a two-sided matching market. However, the current version cannot incentivize all the evaluation questions. When designing this monetary incentive, only the evaluation of startup's quality (i.e. Q1) is incentivized to avoid distorting participants' evaluation on other questions.[30] The related incentive structure for each evaluation question is provided in the Table 4. Both incentives impose costs for making inefficient and inaccurate evaluations.

C. Justification

One concern of adding the "monetary incentive" to increase the sample size is that it will attract participants who do not value the matching incentive and generate noisy outcomes for Q2, Q3, Q4 and Q5. This implies that some insignificant startup characteristics can also be important in the real investment process when the sample size is larger. However, it does not affect the relative "signal-to-noise" ratio of each startup characteristic in Experiment A.

3.1.4 *"Signal-to-noise" Ratio and Project Characteristics*

Although RCT methods can solve the omitted variable bias caused by data limitations in the entrepreneurial financing research, another remaining key challenge is how to display credible startup project information in the experimental setting. In the real business world, startup projects' information, such as recent traction, is generally harder to verify compared with startup team's information, especially in the pitch email setting. Founders have the motivation to exaggerate their project advantages to raise funding. Therefore, the information of startup project is noisier than the information of startup team, especially before investors implement detailed investigations.

Experiment A's design can mitigate this "signal-to-noise" ratio problem and display accurate information of startup projects due to its special design. Investors know that all the startup profiles they evaluate are hypothetical profiles. Therefore, all the information of these hypothetical startups are accurate because it is generated by the computer. However, the incentive structure in Experiment A can make sure that they have the motivation to reveal their true preferences, especially for these non-sensitive startup characteristics. Therefore, compared with the correspondence test, Experiment A provides a cleaner experimental environment similar to the lab environment.

3.2 Results

3.2.1 *Both Team and Project Characteristics Matter in the Early Stage, Especially Traction.*

Table 5 reports regression results of how investors' evaluation results respond to multiple startup team characteristics and startup project characteristics. In column (1), the dependent variable is the quality evaluation (Q1), which indicates the percentile rank of each startup profile compared with an investor's previous invested startups in terms of its potential financial returns. In column (2), the dependent variable is the loyalty evaluation (Q2), which indicates how likely the investors think the startup

[30] If the collaboration likelihood (i.e. Q2) is added to the financial return algorithm, then all the participants will claim that best startups are willing to collaborate with them even if it is not true. Similarly, if the contact interest (i.e. Q3) and investment interest (i.e. Q4) are added to the financial return algorithm, participants have the motivation to distort their true evaluation in order to maximize the financial return because both the contact interest (i.e. Q3) and investment interest (i.e. Q4) can be affected by the collaboration likelihood (i.e. Q2).

team will accept his/her investment rather than other investors. In columns (3) and (4), the dependent variables are the contact interest, which describes the probability that the investor wants to contact this startup. In columns (5) and (6), the dependent variables are the relative investment interest ranging from 1 to 20, which describes the relative investment amount compared with the investor's general investment amount. The unit is one-tenth of the relative investment compared with investors' average investment amount. For example, if the investor's average invested deal is \$1M and Q4 is equal to 5, then it means the investor only wants to invest \$1M $\times$ 5 $\times$ 10% = \$500,000 in this startup. If Q4 is 20, then the investment amount is \$1M $\times$ 20 $\times$ 10% = \$2M. In column (7), the dependent variable is the risk evaluation, which describes the percentile rank of each startup profile compared with an investor's previous invested startups in terms of its risk level. "Serial Founder", "Ivy", "US Founder", "Has Positive Traction", "Is B2B" and "Domestic Market" are indicative variables that equal to one if the founder is a serial entrepreneur, alumni from Ivy College, lives in the U.S., and the project has positive traction, is a Business-to-business startup, focuses on the domestic market. These variables are equal to 0 if the startup does not have such characteristics. Number of founders is either 1 or 2; Number of Comparative Advantages and Company Age can be {1,2,3,4}; Company Age^2 is the square of the company age. All the regression results add investor fixed effect and use the robust standard errors reported in parentheses. I use Bonferroni Method in Table 5 and q-value in Table 6 to implement the multiple hypothesis testing. Results are robust to different multiple hypothesis testing methods.

In columns (1), (3) and (5) of Table 5 and Table 6, I find multiple startup characteristics and project characteristics that are causally important to investors' profitability evaluation results, contact interest and investment interest. Such important team characteristics include founder's educational background and previous entrepreneurial experiences. Important project characteristics include the startup's traction, location, comparative advantages and its business models.

It should be noted that coefficients of project traction are the largest among all the startup characteristics in Table 5, which indicates that the project matters the most.[31] Similarly, in Table 7, the standardized coefficients of all startup characteristics also show that project traction is the most influential characteristics among all the influential factors, which challenges the experiment results obtained by Bernstein et al. (2017). Actually, the coefficients of these project characteristics can also be interpreted as the lower bound of the real effects if someone is also concerned about the "signal-to-noise" ratios in the setting of Experiment A. Although the noise appears to be much less in the current experiment compared with the correspondence test, the matching outcomes still depend on the information collected from the real candidate startups, who have the motivation to exaggerate their startup project quality. Therefore, the importance of project can be even larger in this experimental setting.

3.2.2 *Essentially Investors' Beliefs of Profitability Matter the Most.*

Although Experiment A finds that project characteristics, especially traction, matters the most here, columns (4) and (6) of Table 5 and Table 7 show that it is essentially investors' beliefs of startup's future profitability that matters the most in their decision making process. After adding Q1 (i.e. quality evaluation results) and Q2 (i.e. loyalty evaluation results) into the regression, most of the startup characteristics are no longer significant for predicting investor's contact interest and investment interest. This implies that essentially, those important team and project characteristics mainly serve as the profitability indicators that our sample investors use for their startup selection decisions among all the characteristics that researcher provide.

It should not be surprising for future similar lab-in-field experiments to find that team characteristics are more important due to

[31] Considering that all the variables indicating startup team characteristics and the project traction are dummy variables randomized independently, it is safe to compare their coefficients directly to tell the relative importance of each startup characteristics.

the following two reasons. First, different investors have very different investment philosophy. Anecdotal evidence suggests that angel investors are more likely to bet on the "jockey" (i.e. team) rather than the "horse" while institutional investors are likely to bet on the "horse" rather than "jockey". Second, if the researcher includes certain extremely attractive team characteristics, like the founder successful experience of establishing a unicorn startup and has extensive connections in the industry, then such team characteristics can dominate other provided project characteristics. Also, this lab-in-field experiment does not include many important founders' characteristics, which are hard to quantify. For example, some early-stage investors care about founders' passion, confidence or charm. However, it is harder to include these personal traits of founders in this experiment. Therefore, the relative importance of team and project characteristics really depends on the characteristics pool available for investors to evaluate. However, no matter whether team matters more or project matters more, these characteristics fundamentally serve as the factors that affect investors' expectation of the startup's future financial returns (i.e. belief-based mechanisms) and also the preferences towards certain startups (i.e. belief-based mechanisms).

Table 5 also indicates other mechanisms affecting the contact decisions and investment decisions. First, investors' contact decisions can be affected by both investors' belief of startup's profitability (accounts for 90%) and also the startup's willingness to collaborate (accounts for 10%) based on the coefficients before Q1 and Q2 in column (4). This is consistent with Sørensen (2007), which shows that the entrepreneurial process is a two-sided matching process, and sorting is almost twice as important as direct influence from investors for the difference in IPO rates. Second, most of the influential characteristics (except the location) are very consistent in both the initial contact stage and the investment stage, which means previous literature's evidence in the initial contact stage can be very informative to the later decisions if no frictions exist in the communication process.

Although Experiment A documents that investors' beliefs of the startup's future profitability or financial return matters the most, such beliefs or judgement of investors can be inaccurate and even wrong. Hu and Ma (2020) analyze the pitch video data and implement a lab experiment on students from Yale Business School. They find that investors do not seem to correctly form beliefs about startup quality based on founders' delivery features, and such biases can be explained by a taste-based channel (18 percent) and inaccurate beliefs (82 percent). Therefore, besides improving the startups' profitability, it is helpful for founders to obtain good persuasion skills during their fundraising process.

3.2.3 *Taste-driven Contact Interest May Not Transform into Money.*

Columns (4) and (6) of Table 5 and Table 7 also provide suggestive evidence that our recruited investors have taste-driven preferences towards startups located in the U.S. and graduating from Ivy League Colleges although such taste-driven contact interest does not transform into investment interest. After controlling Q1 and Q2 in the regression, whether locating in the U.S. and graduating from Ivy League Colleges are still highly significant. Considering these two factors are not predictors of risk as shown in column (7), this suggests that investors still prefer contacting startups team with excellent educational background after controlling the startup's expected profitability, which makes sense given that they participate in research project supervised by Columbia University. Our investors mainly come from the U.S.; thus, it is also not surprising that they prefer contacting U.S. based startups, which has been documented by the finance literature studying home bias. However, what is interesting is that these two factors are no longer significant in the regression of investment interest after controlling Q1 and Q2, which indicates that such taste-driven preferences do not enter the investment decision. Different from contact interest, investment decisions are more rational for professional investors. Therefore, it is important for studies that only observe initial contact interest to separate whether the contact interest is driven by belief or taste in order to understand its implication in the later

round decisions. Also, it is important to realize that results in Section 3.2.3 is suggestive evidence. To establish more rigorous causal evidence, it is helpful for future researchers to design separate RCTs (for example, dictator games) that directly tests taste-driven preferences.

3.2.4 *Effects across the Distribution of Contact and Investment Interest*

The previous regression specifications only provide the average effect of startup team and project characteristics on investors' contact and investment interest. However, as shown pointed out by Neumark (2012) and shown in Zhang (2020a), such preferences can differ in magnitude and even direction across the distribution of investors' indicated interest. Understanding the distribution effect is helpful in predicting how generalized these experiment results are in different fundraising methods and different market conditions. For example, when the economy is booming or lots of hot money is flowing into the VC industry, investors' preferences can be shifted to the relatively left part of the distribution. However, when the economy is experiencing recession and venture capitalists has to increase their investment bars, their preferences can be shifted to the right tail of the distribution.

Figure 5 shows that investors' preferences towards certain important team characteristics (for example, educational background and entrepreneurial experiences) and project characteristics (for example, traction and business models) are causally important along the whole distribution of investors' contact interest. Figure 6 shows that investors' preferences towards these characteristics are causally important for most parts of the distribution of investors' investment interest. However, for the right tail of the investment interest, such preferences are no longer salient. This happens potentially because more information, which is usually revealed during the communication stages, needs to be provided for investors to make their investment decisions. Generally, startups with these attractive team and project characteristics will enjoy the advantages in most of the market conditions and fundraising settings. Specifically, having a positive traction plays an important role in the fundraising process. This also implies that for future empirical research to establish causality in the entrepreneurial financing setting, it is crucial to take these identified startup characteristics into account for their model specifications.

3.3 Discussion

Experiment A, which follows the latest RCT paradigm, has the following four merits when studying the investors' preferences to important startup characteristics. First, it provides a cleaner experimental environment to reveal investors' preferences where the signal-to-noise ratio is more similar to investors. Second, it is extremely powerful to directly test belief-driven mechanisms and investment interest by carefully designing the evaluation questions and incentive structures. Third, the design is more ethical as it provides real benefits to experiment participants. Lastly, the experiment design is very helpful in generating a large enough sample size when it is hard to recruit experiment participants. In the entrepreneurial finance setting, senior VC investors are very hard to recruit, especially during the recession when most people focus on surviving the economic difficulties. By allowing each individual to provide multiple evaluations, the experiment can have enough sample size from a limited number of participants.

However, when implementing such lab-in-field experiments, it is important to realize the following limitations. First, similar to any experiments requiring voluntary participation, this lab-in-field experiment design also has potential sample selection bias during the recruitment process. It does not hurt the internal validity of the experiment. However, it is important to

checking the external validity by running a complementary experiment. Future researchers can also replicate this experiment in different settings by recruiting different investor groups. Moreover, any recruitment process which increases the response rate (i.e. collaborating with prestigious institutions, recruiting investors face to face, etc.) can help mitigate such sample bias. Second, the incentive structure used in the experiment requires more social resources, which may not be user-friendly to junior researchers without many social connections. Therefore, any innovation on providing more cheaper incentive structures is important to lowering the experiment cost. Lastly, this experiment design does not generate real economic outcomes as a preference elicitation technique, which makes it hard to implement the welfare analysis. Any attempts to obtain real economic outcomes or develop quasi-experiment design are important to help better understand the entrepreneurial finance process.

The lab-in-field experiment design in this paper can also be improved from the following dimensions based on investors' valuable feedback. First, the survey tool can be shorter if it is possible to recruit larger number of investors. Second, it is helpful to ask for some simple investment criteria (i.e. revenue range, industry, etc.) before the evaluation section to make each profile more customized to different investors. For example, series B investors can evaluate more mid-stage companies. Such design can improve the user experience of the survey tool although it costs researchers more time and efforts. Third, more relevant information can be provided to investors, which includes founder's experience in the industry, whether the previous startups succeeded or not, more background of existing investors, monthly burn rates, and others. Also, any effort to improve the realism of each startup profile is also helpful. However, researchers need to be aware that too much information provided may dilute investors' attention, which makes it harder to test your major interested variables. Fourth, considering that most investment decisions are made on a relative basis at a specific moment, future researchers can ask each investor to compare multiple startups at the same time rather than evaluate one by one. This "Netflix style" evaluation method is a more realistic way to capture such relative investment strategy. A better format to visualize the startup information and questions at the same time is also helpful. Lastly, researchers can ask more questions about the investor types, for example, whether they are financial investors or strategic investors.

Before we use the experiment results from the latest RCT method to challenge the validity of correspondence test in testing startup project characteristics, it is important to check whether the results from Bernstein et al. (2017) are replicable for our sample investors. Based on the distributional effect shown in Section 3.2.4, well designed RCTs with good internal validity should be able to discover that project characteristics are causally important for investors' decisions for our sample investors. Only when we find that the previous correspondence test design fails to reveal the importance of startup project characteristics for the same investors can we conclude that different results of the latest RCT method and the previous RCT method come from method limitations rather than sampling variability Therefore, I follow up with Experiment B, which essentially mimics the experiment design of Bernstein et al. (2017).

4 Experiment B: Correspondence Test

In Experiment B, I mainly test whether investors only respond to startup team's characteristics rather than project characteristics for institutional investors following the correspondence test design. I sent generated hypothetical pitch emails with randomized startup team and project characteristics in both the email subject line and email content, and traced detailed investors' information acquisition behaviors to these pitch emails. The experiment design follows a factorial experiment design,

which orthogonally randomizes the startup founder's gender, race, education background and startup project's comparative advantages. Recruited investors come from VC industry of mainly the U.S. and other English-speaking areas. [32]

Sending out cold call pitch emails to investors for fund raising purposes is more popular recently following the trend of removing the barriers to funding. For example, deck sender,[33] an online platform helping entrepreneurs to send pitch decks to the right investors for free, is designed to democratize access to funding and has sent out 90,000+ decks by 2020/06. compared to the mainstream fundraising methods from warm networks and in-person interactions, sending cold call emails does not require startup founders to have close connections with practitioners in the VC industry, hence lowers the entry cost.[34] To some extent, it helps to increase the diversity of the entrepreneurial community. [35] However, considering the potential risk of idea stealing and the lower response rate,[36] I would recommend that young startups try the mainstream fund-raising methods first before joining this probability game.[37]

Section 4 is organized as follows. Section 4.1 introduces the experiment design, including the email sending process and the email behavior tracing techniques. Section 4.2 describes the sample selected, including the summary statistics of both the real investors and fictitious startup founders. Section 4.3 discusses the analysis results of investors' information acquisition behaviors. Section 4.4 discusses the limitations of this correspondence test.

4.1 Experiment Design

Manipulating Startup Team Characteristics. — In order to introduce meaningful variation of startup team characteristics, I randomized founder's educational background in both the email subject line and the email content. For the control group, I do not mention the founders' education background at all. For the treatment group, the email subject line and the email content indicate that the startup team members come from prestigious universities in the US. Therefore, this experiment design identifies whether startup founders' excellent education background attract investors more compared to average startups who do not mention these when sending cold call pitch emails to investors.[38] I also randomized the founders' gender and race in this experiment setting. However, the discussions about how founders' gender and race affect investors' interest are provided in

[32] Include UK, Canada, Australia, Singapore, Hong Kong, Israel, India etc. Considering the global trend in the entrepreneurial activities and VC investment activities, I also recruit investors from other English-speaking areas. However, I do not include investors from China, Korea and Japan due to the language used in this experiment and also do not recruit European Union investors because of the EU General Data Protection Regulation (GDPR).

[33] https://decksender.com/

[34] Gompers, Gornall, Kaplan and Strebulaev (2020) show that unsolicited approaches by founders account for 12% of early-stage VCs' deals, and the majority of deals (62%) still come from professional networks and referrals.

[35] One concern of investing through personal network is that minority founders may face more financing difficulties due to the lack of connections.

[36] Gornall and Strebulaev (2020) shows that the cold email response rate of angle investors and VC investors are 6% in December 2018 (economics boom). In this experiment which was implemented between 2020/03-2020/09 (economics recession), the cold email response rate is 1.5%. This phenomenon is consistent with the Howell et al. (2020), which documents that early-stage investors are significantly more responsive to business cycles than later-stage investors.

[37] Thanks for the advice from a managing director participating in the experiment, who inform us of the risk related to sending out cold call pitch emails. Some investors who seem to be interested in the cold call pitch emails are likely to be just fishing around to get undeveloped but decent ideas worth stealing. Therefore, startup teams should be aware of such risks before sending out large scale of cold emails. However, connecting investors within your own network from Alumni, events or friends after careful due diligence sometimes work well. See the discussions on Quora: https://www.quora.com/How-do-I-pitch-a-startup-idea-by-email

[38] Prestigious universities used in this experiment include Ivy Colleges, MIT and Stanford. In the first-round experiment implemented between 2020/03 and 2020/04, I also included Northwestern University, Caltech, John Hopkins University, Juilliard School and other top schools in the field related to the startup. For example, if the startup is related to music, I will mention that the founding team members come from Columbia University and Juilliard School. However, investors did not appreciate such "mixed Ivy education background" as strongly as the "pure Ivy education background". Therefore, in the second-round experiment implemented between 2020/08-2020/09, I only used Ivy Colleges, MIT and Stanford as the education background in the treatment group.

Zhang (2020a).

Manipulating Startup Project Characteristics. — In order to introduce meaningful variation of startup project characteristics, I also randomized startup projects' impressive advantages in both the email subject line and the email content. The control group does mention any specific comparative advantages of the startup and the treatment group mentions the comparative advantages like "22% MOM Growth Rate" or "Patent Registered".[39] The experiment design essentially mimics Bernstein et al. (2017), which only mentions the attractive startup characteristics in the treatment group and provide no information to the control group.

Manipulating Access to Information. —The randomization of startup's characteristics (i.e. founder's educational background and project's advantages) is implemented in the following two stages. For the first stage before the investor opens the pitch email, she will see the randomly generated email subject lines indicating whether the startup has a well-educated founding team and a project with impressive advantage. For the second stage after the investor opens the pitch email, she will decide how much attention to spend on reading this pitch email. In the email content, if the email subject line mentions the Ivy education background or project advantages, there are extra sentences inserted to emphasize this information again in the email content while keeping all the rest of email contents the same. After reading the email content, the investor can decide whether to reply or forward the email to other related investors.

To make sure that the i.i.d. assumption holds for the experiment, the randomization is implemented in the following steps. First, to increase the response rate, I match investors with pitched startup ideas based on her industry/vehicle preferences so that healthcare related pitch emails are sent to investors who are interested in healthcare industry.[40] Second, considering the potential spillover effect within each VC fund, investors receiving the same pitch email ideas come from different VC funds.[41] Each startup pitch email is sent to roughly 1000 investors all working in different funds. Among these 1000 investors, they are randomly divided into 16 groups because based on the factorial experiment design, founder's gender, race, education and project advantages should be randomized independently.[42] Hence, we have $2 \times 2 \times 2 \times 2 = 16$ groups. Third, it usually takes more than 2 weeks for us to send two sequential pitch emails to the same investor to avoid unnecessary attention and keep the i.i.d. assumption along time. Each investor receives 3 to 5 pitch emails between 2020/03-2020/09.[43]

Pitch Email Design and Website Construction—The pitch emails covering 67 startup ideas written for this experiment follow the template and structure provided by Gornall and Strebulaev (2020) and good pitch email template examples posted on Quora. The startup ideas are provided by my research team members,[44] who are usually young startup founders or members of startup-related clubs at Columbia and other Ivy Colleges who are interested in this research project. We use Wix, a commercial

[39] Abbreviated form for "month over month" growth in finance.

[40] For investors recorded in Pitchbook Database, I use the recorded industry preference for the matching purpose. For investors from other databases, I manually collected their industry preferences from information on their company websites, LinkedIn and CBInsight. If the manually collected industry information is not accurate, this will increase the noise of the experiment results and reduce the email response rate. However, it does not affect investors' email opening behaviors.

[41] For some VC funds, they usually have a weekly meeting to discuss promising investment opportunities before replying to cold call pitch emails. If investors receiving the same startup idea come from the same fund, their responses are likely to be correlated. However, this situation will not affect the email opening behaviors and email reading time when they just receive pitch emails.

[42] This randomization that the number of treatment group observations is equal to the control group size is mainly to increase the experiment power.

[43] Gornall and Strebulaev (2020) waited at least five days to send a sequential email, which raises the attention of some investors who draw attention to these cold emails on twitter in the middle of the experiment. Their experiment was finished between 2018/11-2018/12. To avoid such situation, I slowed down the pace of sending cold emails and extend the experiment implementation period.

[44] I only choose the valid startup ideas with relatively good coverage of key industries after discussion with practitioners.

website builder, for making the related startup websites which are in the under-construction stage. I do not create any LinkedIn accounts for these hypothetical startup founders because LinkedIn Community does not allow to create suspicious accounts even for research purposes. However, the believability concern should not affect the email opening rate and the email reading time although it may affect the response rate and the replied contents. The pitch email example is provided by Figure 7 and the website example is provided by Figure 8.

Emailing Process—I mainly implement the following two steps to solve the technical difficulties of sending large amount of cold call emails to investors' email inboxes and pass the existing spam filters.[45] First, before sending large-scale pitch emails in 2020/03, I sent out a testing email (see Figure 9 in Appendix C) which introduces the public information about COVID-19 in 2020/02. The testing email is to identify which email addresses are invalid and check the opening rate of cold emails irrelevant to investment opportunities.[46] The opening rate of the testing email after 2 weeks is 2.8% while the average opening rate of the investment related pitch email in this experiment is 11.8%. This indicates that investors only open the emails that they are interested in based on the email subject line and senders.

Second, I used the Mailgun's Managed Service,[47] a third-party commercial email API delivery service provider, for sending a large amount of emails. compared to the traditional method of using multiple web hosts to combat spam policies, Mailgun is designed for developers and businesses, with extremely powerful functionality of tracing the status of each email sent and achieving high delivery rate through its sending infrastructure. It also provides developers with complete freedom to customize the email sender names, setting the back-end database structure and developing new email tracing functionalities with a user-friendly price compared to G-suite,[48] which is an email provider from Google. Before automatically sending pitch emails, I used GlockApp, a spam filter testing service provider, to test and improve my pitch email templates.

Following the two-step email sending procedures mentioned above, the response rate is very stable along the whole recruitment process. Gornall and Strebulaev (2020) used traditional methods of sending out a large amount of cold call pitch emails and the email response rate declined from 9.0% for the first 4,000 emails to 5.3% for the last 4,000 emails. Such situation did not happen in this experiment. Moreover, the email sending procedures in this experiment allow monitoring multiple investors' information acquisition behaviors without hurting too much of the email delivery rate.

Email Behavior Measurements—I traced the following email behavior measurements, including both the new behavior measurements used for this paper and the behavior measurements used in previous correspondence test. These measurements include the email opening status and the corresponding time stamp, the email staying time measured in milliseconds, the

[45] Different email providers usually use different spam filtering algorithms. However, there are some common patterns for detecting spam emails. First, if there are many invalid email addresses sent out from the same domain at an extremely high frequency (for example, 10 emails sent out per second), then the emails sent are more likely to be labeled as spams. To avoid this, it is helpful for researchers to send a safe testing email identifying the invalid email addresses and remove them in the formal recruitment process. Second, if the email contains unverified website links or common words used in spam emails like "Dear", these emails are likely to fail the spam filter. Hence, it is important to use spam filter testing service to double check the email contents. However, none of these spam filtering algorithms are correlated with email senders' gender and race.

[46] Invalid email addresses are those that no longer exist or no longer being frequently checked by investors based on the bounced back email notifications. The investor database was constructed between 2018/04-2019/12. Therefore, more than 20% of the collected email addresses are no longer valid due to the high turnover rate.

[47] https://www.mailgun.com/ Mailgun has more than 150,000+ customers in 2020. It was founded in 2010 and was a part of Y Combinator Winter 2011 cohort.

[48] If researchers have abundant research funding, they can also create multiple G-suite accounts to combat spam policies. G-suite is a "company-version" Gmail and is user-friendly to people without strong coding skills. The only drawback is its relatively expensive price, costing $6 per account per month starting in 2020.

sentiment of the replied emails analyzed through LIWC,[49] the click rate of the inserted startup websites, the response rate and whether the response is a positive response or a negative response.[50] Despite of these rich behavior measurements, only email opening rate and email staying time generate enough power to analyze investors' responses. All the other traditionally used behavior measurements do not survive in the recession period when the "low-response-rate" problem is more severe than before. The detailed mechanisms of recording different email behaviors and whether such behavior measurements are used in the previous literature are described in the Appendix C Table C4. The flow chart of the first correspondence test is provided in Figure 11.

Alternative Experiment Design without Deception—This experiment follows the mainstream practice of the correspondence test, which usually involves deception. There are two alternative experiment designs without deception implemented by me in 2018 but both of them failed due to different reasons. For the first experiment design, I collaborate with real startup teams with excellent educational background for this experiment. Such design brings legal risks involved with working with real businesses. Therefore, it requires researchers to be extremely careful to design all the consent documents and consent procedures. For the second experiment design, I organized a startup Pitch Night in 2018/10 by inviting multiple VC investors to evaluate eight real startup teams pitching their ideas.[51] In the formal invitation email of this event, I introduce multiple exogenous variation of each real startup team's characteristics and trace investors' email behaviors. This design failed because investor's response rate and the inserted startup website click rate were extremely low. However, these failures provide crucial insights for the current version design of the correspondence test.

4.2 Sample Selection and Data

This correspondence test experiment was implemented twice for testing the external validity purposes. The first round was implemented between 2020/03-2020/04 during the outbreak of COVID-19 around the world. Considering the unusualness of this period, I implemented a later-round correspondence test in 2020/10 (*Results to be added*) during the economy re-opening when people have calmed down from the COVID-19 shock. Running the correspondence test in multiple times helps test the external validity since the investors' investment preferences can be affected by different social events as shown in Zhang (2020a).

Investors recruited for this experiment are mainly early stage venture capitalists in the U.S. and other English-speaking areas in the world as documented in Section 2. Table 8 provides the industry distribution of the created hypothetical startups. There are 67 startup ideas created with more than 200 names used in order to make sure that all experiment results are not driven by any special names or startup ideas. These ideas cover the majority of mainstream industries that venture capitalists are interested in, which include Information Technology, Healthcare, Consumers, Energies, and etc. Note that with this correspondence test design, I cannot observe the later stage decisions like investment interest for each startup.

4.3 Results

Investors Respond to the Team Characteristics Rather Than Project Characteristics.

[49] LIWC (Linguistic Inquiry and Word Count) is a text analysis program used for sentimental analysis.

[50] Positive response indicates either a direct invitation to a call or interest in the pitch deck.

[51] One startup successfully received investment from an event participant. Based on the feedback from practitioners, participating in these person-in-person pitch events are more likely for startups to build connections with investors and receive funding compared to sending cold call pitch emails.

Table 9 Panel A summarizes investors' major information acquisition behaviors in the first-round correspondence test. On average, the pitch email opening rates are 12.03% and each investor spent roughly 24 seconds on reading a cold call pitch email in 2020/03-2020/04. However, both the startup website click rates and the email response rates are very low,[52] which indicates that early-stage investors are sensitive to business cycle as documented by Howell et al. (2020). Therefore, traditional measurements used in the correspondence test, like the email response rate, do not generate enough experiment power during the COVID-19 Pandemic. All of our experiment results rely on the new behavior measurements created in this paper.

Table 9 Panel B reports regression results of global investors' email opening behaviors of randomized pitch emails in Experiment B. The dependent variable is a dummy variable, which is one when an investor opens the pitch email, and zero otherwise. Female Founder = 1 is an indicator variable that equals one if the first name of the email sender is a female name, and zero otherwise. Similarly, Asian Founder =1 is an indicator variable that equals one if the last name of the email sender is an East Asian name, and zero otherwise. Ivy = 1 is an indicator variable for Ivy education background. Project Advantage = 1 is an indicator variable which is one when the email subject line includes the corresponding comparative advantages. March Chinese Virus = 1 is an indicator variable which is one when the email was sent between 2020/03/18-2020/03/24 when President Trump uses the phrase "Chinese Virus". US Investor =1 and Female Investor = 1 are indicator variables for being a U.S. investor and being a female investor. Column (1) (2) (3) (5) uses all the observations collected in the first-round correspondence test. In Column (4), results are reported for the sub-sample where the startup team's graduated university are from purely the Ivy colleges, Stanford and MIT. "Pure_Ivy" indicates cases like "Team from Columbia University" while 'Mixed_Ivy" indicates cases like "Team from Columbia University and Juilliard Music School". For some startups in the music or medical industry, I combined an Ivy college with a good university in that specific area for the treatment group. All the regressions include start-up fixed effects to control for any idiosyncratic characteristics of each start-up pitch email, like the business models, etc. Hence, I am comparing investors' email opening rates within the same start-up's pitch email and all the results are similar after including investor fixed effects. Following Bernstein et al. (2017), the standard errors are clustered at the investor level to account for the correlated opening decisions across different pitch emails received by the same investor.

Results of Table 9 Panel B show that investors' email opening behaviors respond to startup team characteristics, but not respond to startup project characteristics. Column (1) shows that a pitch email sent by a female first name raises the opening rate by 1%, which is statistically different from zero. Such difference is between pitch emails sent by female first names and pitch emails sent by male first names. Considering that the base opening rate is 12.03% for all the pitch emails, this represents an 8% increase of opening rates. Column (2) shows that a pitch email sent by an Asian last name raises the opening rate by 0.7% after President Trump stops using the phrase "Chinese Virus", which is statistically significant at 10% and represents a 6% increase of opening rates compared to using a white last name. Importantly, Column (3) shows that mentioning good education background in the email subject line increases the opening rate by 0.7%. Such effect increases to 1.2% if I focus on the sub-sample that only the pure Ivy education background is mentioned (i.e. "Team from Columbia University" rather than "Team from Columbia University and Juilliard Music School "). This represents a 10% increase of email opening rates compared to not mentioning anything in the email subject. However, Columns (4) and (5) show that mentioning the project advantages in the email subject line does not significantly increase the email opening rate. Results provided by Table 9 confirms the surprising results found in Bernstein et al. (2017) that investors only respond to the startup team characteristics rather the project characteristics in the

[52] Gornall and Strebulaev (2020) documents that the cold call pitch email response rate is roughly 6% in 2018 when the economy goes well.

email setting even for the institutional investors.

4.4 Discussion

4.4.1 Different Signal-to-noise Ratios

Based on results of Experiment A, it is clear that project characteristics, such as traction or comparative advantages, should play an important role in our sample investors across the distribution of investors' contact interests. However, in Experiment B, the importance of project characteristics disappears. This seemingly contradictory results can be explained by the different "signal-to-noise" ratios of team characteristics and project characteristics in the email setting.

Essentially, the email experimental setting is very noisy. compared to team characteristics, startup's project information is much harder to verify and founders generally have the motivation to exaggerate their project quality or traction information. Therefore, investors, especially experienced investors, may rationally ignore such noisy information about the project characteristics. However, it is very easy to verify founders' gender, race or educational backgrounds online by using the university directory. Therefore, the accuracy of team's information is generally more credible. Due to this reason, correspondence test experimental design may not appropriate to test the importance of project characteristics, and also the relative importance of startup team and project characteristics.

4.4.2 Alternative Interpretations

One alternative interpretation of the different results from Experiment A and B is the sample variability. Although subjects of Experiment A are a subset of participants of Experiment B, it is theoretically possible that only subjects in Experiment A care about project characteristics in their indicated decisions while majority of other investors only care about the team characteristics in their investment decisions. However, this alternative story is less plausible because I did not find any strong evidence to show that participants of Experiment A had any special reasons to behave extremely differently from the rest of the investors. Specially, most of them are institutional venture capitalists, and preferences towards different startup project characteristics are not sensitive preferences.

4.4.3 Limitations of Experiment B

It should also be noted that standard correspondence test has the following limitations. First, the setting is noisy, and all the email behavior measurements are not perfect. The noisy setting also limits the number of research questions researcher can test because many introduced variations are not salient enough to generate significant results. Second, this correspondence test, as a preference elicitation technique, only observes initial contact interests rather than later stage investment interests. Third, I do not observe real economic outcomes, which makes it hard to implement welfare analysis and to transform various email behaviors into real money analysis. Lastly, this correspondence test design involves deception similar to most of traditional correspondence test designs despite the efforts to attempt two alternative designs as discussed in Section 4.1. However, the limitations mentioned above can be mitigated by the latest RCT method used by Experiment A.

Also, this correspondence test experiment in the cold call pitch email setting only mimics the experimental design of Bernstein et al. (2017) rather than fully replicates their experiment by collaborating with AngelList. On the AngelList platform, investors

understand that if a team does not disclose their educational background or project advantages, such as traction, it means that the founders' educational background or the project traction does not exceed the threshold. However, in the cold call pitch email setting, if a team does not disclose their educational background or project characteristics in the email subject lines or email contents, their background can still exceed investors' threshold. Therefore, our experimental setting is even more nosier than the experimental setting of Bernstein et al. (2017) because investors' unobservable internal threshold is more opaque and subjective in our setting. However, the logic of this experimental design is essentially the same as the logic of Bernstein et al. (2017). We both assume that if the founders' educational background or projects' advantages are not disclosed, investors consider these startups to have lower quality than startups that disclose such information. Also, the different "signal-to-noise" ratio problem applies to both Experiment B and the correspondence test of Bernstein et al. (2017).

5 Conclusion

In this paper, I study the startup selection criteria of early-stage investors and identify multiple startup team and project characteristics that causally affect investors' contact interests and investment interests. Although this research question is one of the core questions in entrepreneurial financing literature and one that is frequently debated among practitioners and academics, we have a very limited understanding of the topic due to the lack of exogenous variations and granular informative data.

This paper implements two RCTs by recruiting real venture capitalists mainly from the U.S. The purpose is to understand investors' preferences towards different startup characteristics and why they care about these characteristics. Experiment A invites U.S. investors to evaluate 16 hypothetical startup profiles in order to match investors with appropriate, real startups from real incubators. I find that both startup team characteristics (including human capital assets such as educational background, previous entrepreneurial experiences) and startup project characteristics (including their business models, traction, locations, etc.) causally affect investors' contact interests and investment interests on the aggregate level. This is because these identified characteristics are viewed as indicators of the expected financial returns of each startup. Investors' investment decisions are mainly influenced by the judgment of each startup's quality while the contact decision is influenced by the both the evaluation of the startup's quality and the startup team's collaboration willingness. Moreover, if the influential startup characteristics are belief-driven in the initial contact stage, suggestive evidence shows that they will also affect investors' later round investment interests. However, any types of taste-driven preferences in the contact stage may disappear in the investment stage.

Experiment B mimics the experimental design of Bernstein et al. (2017) and sends randomly varied pitch emails to global VC investors. By utilizing a new email behavior tracking technology, I observe global investors' information acquisition behaviors for different types of pitch emails. I find that investors only causally respond to the startup team characteristics rather than the project characteristics in the email setting, which confirms the results of Bernstein et al. (2017). However, this also indicates that the previously used correspondence test experimental design may not be suitable to test the importance of project characteristics due to its noisy setting and startup project characteristics' lower "signal-to-noise" ratios compared with team characteristics.

Overall, results in the paper present causal evidence for the importance of nonhuman assets for the financing success of early-stage firms. It also identifies new human assets, such as entrepreneurial experiences, that casually affect investors' interests. Specifically, the fact that project traction plays the most important role in Experiment A supports the story of Kaplan et al.

(2009) even in the very early stage financing process of firms, and it challenges the validity of the previous correspondence test method in testing the relative importance of startup team and project characteristics. The relative importance of human assets and nonhuman assets should depend on the investors' background and also the startup characteristics pool available for investors to evaluate. Future research can replicate these experiments in different settings to test the external validity. Also, any effort to study the communication process between investors and founders using real economic outcomes is vital to understanding the full picture of the investment process and strategies of the VC industry.

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.

Barber, Brad M, Adair Morse, and Ayako Yasuda, "Impact investing," *Journal of Financial Economics*, 2020.

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.

Bernstein, Shai, Arthur Korteweg, and Kevin Laws, "Attracting Early-Stage Investors: Evidence from a Randomized Field Experiment," *The Journal of Finance*, 2017, *72* (2), 509–538. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12470.

Bertrand, Marianne and Sendhil Mullainathan, "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination," *American economic review*, 2004, *94* (4), 991–1013.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of political Economy*, 1992, *100* (5), 992–1026.

Brock, Michelle and Ralph De Haas, "Discriminatory Lending: Evidence from Bankers in the Lab," 2020.

Carpenter, Jeffrey, Cristina Connolly, and Caitlin Knowles Myers, "Altruistic behavior in a representative dictator experiment," *Experimental Economics*, 2008, *11* (3), 282–298.

Cen, Xiao, "Household Wealth and Career Choices: Evidence from Natural Disasters," *Working Paper*, 2020.

Ewens, Michael and Richard R. Townsend, "Are early stage investors biased against women?," *Journal of Financial Economics*, March 2020, *135* (3), 653–677.

Gompers, Paul A and Sophie Q Wang, "Diversity in innovation," Technical Report, National Bureau of Economic Research 2017.

——, **Vladimir Mukharlyamov, Emily Weisburst, and Yuhai Xuan**, "Gender effects in venture capital," *Available at SSRN 2445497*, 2014.

Gompers, Paul A., Will Gornall, Steven N. Kaplan, and Ilya A. Strebulaev, "How do venture capitalists make decisions?," *Journal of Financial Economics*, January 2020, *135* (1), 169–190.

Gornall, Will and Ilya A. Strebulaev, "Gender, Race, and Entrepreneurship: A Randomized Field Experiment on Venture Capitalists and Angels," SSRN Scholarly Paper ID 3301982, Social Science Research Network, Rochester, NY March 2020.

Guzman, Jorge and Aleksandra Olenka Kacperczyk, "Gender gap in entrepreneurship," *Research Policy*, 2019, *48* (7), 1666–1680.

Hart, Oliver, *Firms, contracts, and financial structure*, Clarendon press, 1995.

Hebert, Camille, "Gender Stereotypes and Entrepreneur Financing," SSRN Scholarly Paper ID 3318245, Social Science Research Network, Rochester, NY March 2020.

Heckman, James J, "Detecting discrimination," *Journal of economic perspectives*, 1998, *12* (2), 101–116.

Hong, Harrison and Inessa Liskovich, "Crime, punishment and the halo effect of corporate social responsibility,"

Technical Report, National Bureau of Economic Research 2015.

Howell, Sabrina T., Josh Lerner, Ramana Nanda, and Richard Townsend, "Financial Distancing: How Venture Capital Follows the Economy Down and Curtails Innovation," SSRN Scholarly Paper ID 3594239, Social Science Research Network, Rochester, NY May 2020.

Hu, Allen and Song Ma, "Human Interactions and Financial Investment: A Video-Based Approach," *Available at SSRN*, 2020.

Jr, Roland G Fryer and Steven D Levitt, "The causes and consequences of distinctively black names," *The Quarterly Journal of Economics*, 2004, *119* (3), 767–805.

Kaplan, Steven N, Berk A Sensoy, and Per Strömberg, "Should investors bet on the jockey or the horse? Evidence from the evolution of firms from early business plans to public companies," *The Journal of Finance*, 2009, *64* (1), 75–115.

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.

Neumark, David, "Detecting Discrimination in Audit and Correspondence Studies," *The Journal of Human Resources*, 2012, *47* (4), 1128–1157. Publisher: [University of Wisconsin Press, Board of Regents of the University of Wisconsin System].

Rajan, Raghuram G, "Presidential address: The corporation in finance," *The Journal of Finance*, 2012, *67* (4), 1173–1217.

—— **and Luigi Zingales**, "The firm as a dedicated hierarchy: A theory of the origins and growth of firms," *The Quarterly Journal of Economics*, 2001, *116* (3), 805–851.

Sørensen, Morten, "How smart is smart money? A two-sided matching model of venture capital," *The Journal of Finance*, 2007, *62* (6), 2725–2762.

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

—— , "ESG and Venture Capital Investment: Experimental Evidence," 2020.

Zingales, Luigi, "In search of new foundations," *The journal of Finance*, 2000, *55* (4), 1623–1653.

Tables

Table 1: Summary Statistics for Investors

Panel A: Investor Location Distribution

Country	N	Percentage	Female Percentage
US	15,184	84.91%	23.57%
Canada	647	3.62%	29.68%
Israel	456	2.55%	29.39%
UK	93	0.52%	22.58%
India	514	2.87%	18.87 %
Singapore & Hong Kong	454	2.54%	21.59%
Australia & New Zealand	228	1.28%	25.44%
Others	306	1.71%	21.57%
Total	17882	100%	

Panel B: Investor Industry Distribution

Industry	N	Percentage
Information Technology	13,628	76.21%
Healthcare	6,056	33.87%
Consumers	6,256	34.98%
Energy	4,234	23.68%
Life Sciences	3,347	18.72%
Finance	3,023	16.91%
Media & Entertainment	2,533	14.17%
Agriculture & Food	2,072	11.59%
Transportation	1,743	9.75%
Education	1,359	7.60%
Clean Technology	1,201	6.72%
Others	3,271	18.29%

Panel C: Investor Characteristics

	N	Mean
Female Investor=1	17,882	0.24
Senior Investor=1	17,882	0.84
Angel Investor=1	17,882	0.11
Top University=1	13,785	0.31
Graduate School=1	9,232	0.61
Not-for-profit Fund=1	13,156	0.02

Notes. This table reports descriptive statistics for the active venture capitalists (defined as those whose email addresses are verified by the testing email) who received the cold call pitch emails in the correspondence test and the recruitment emails in the lab-in-field experiment. Panel A reports the geographical distribution of the sample investors. "Others" includes South Africa, Cayman Islands, Malaysia, and etc. Panel B reports the industries that these investors have stated that they are interested in investing. An investor can indicate multiple preferred industries. "Others" includes special industries like packaging technologies. 3.8% of the investors' industry preferences cannot be found online and I have assumed that they are interested in all of the industries when sending out pitch emails. Panel C reports the investors' demographic information and investment philosophy. 'Female = 1' is an indicator variable that equals one if the investor is female, and zero otherwise. 'Senior = 1' is an indicator variable that equals one if the investor is senior (defined as C-level positions, principals, vice president, partners, etc.), and zero otherwise. 'Angel = 1' is an indicator variable that equals one if the investor is an angel investor or belongs to angel group, and zero otherwise. If an investor is both an an angel investor and also an institutional investor, I treat her as an angel investor. 'Not-for-profit Fund = 1' is an indicator variable that equals one if the investor works in a not-for-profit impact fund based on the "primary investor type" in Pitchbook. 'Top University=1' and 'Graduate School =1' are indicator variables that equal one if the investor attended a top university (i.e. Ivy League Colleges, MIT, Duke, Caltech, Amherst, Northwestern, Stanford, UC Berkeley, University of Chicago and Williams College) or attended graduate school.

Table 2: Experiment A Summary Statistics of Investors

Panel A: Investor Stated Interest Across Sectors

Sector (Repeatable)	N	Fraction (%)
Information technology	39	55.7%
Consumers	10	14.3%
Healthcare	17	24.3%
Clean technology	3	4.3%
Business-to-business	7	10.0%
Finance	11	15.7%
Media	4	5.8%
Energy	5	7.1%
Education	3	4.3%
Life sciences	2	2.9%
Transportation & Logistics	4	5.7%
Others	6	8.6%
Industry Agnostic	6	8.6%

Panel B: Investor Stated Interest Across Stages

Stage (Repeatable)	N	Fraction (%)
Seed Stage	47	67.1%
Series A	45	64.3%
Series B	17	24.3%
Series C or later stages	5	7.1%

Panel C: Investor Stated Demographic Information

	N	Mean	S.D
Female Investor	69	0.20	0.40
Minority Investor	64	0.42	0.50
Senior Investor	69	0.86	0.37

Panel D: Investor Stated Investment Philosophy

	N	Mean	S.D
Cold Email Acceptance	69	0.74	0.44
Prefer ESG	69	0.17	0.03
Direct Investment	69	0.94	0.24

Continued

Panel E: Available Fund's Financial Performance

				Percentile		
	N	Mean	S.D	10	50	90
Total Active Portfolio	54	41.40	44.51	10	24	102
Total Exits	46	32.74	48.39	1	9	110
Fund Age	52	11.75	8.95	3	8.5	25
AUM (Unit: $1 Million)	33	547.46	1029.10	30	111.7	1700
Dry Power (Unit: $1 Million)	33	163.86	307.04	6.43	44.35	313.59

*Notes.*This table reports descriptive statistics for the investors who participated in the lab-in-field experiment. In total, 69 different investors from 68 institutions, mostly venture funds, provided evaluations of 1216 randomly generated startup profiles. Panel A reports the sector distribution of investors. Each investor can indicate their interest in multiple industries. "Others" includes HR tech, Property tech, infrastructure, etc. "Industry Agnostic" means the investor does not have strong preferences based on sector. Panel B reports the stage distribution of investors, and each investor can invest in multiple stages. "Seed Stage" includes pre-seed, angel investment, and late-seed stages. "Series C or later stages" includes growth capital, series C, D, etc. Panel C reports the demographic information of the recruited investors. "Female" is an indicator variable which equals to one if the investor is female, and zero otherwise. "Minority" is an indicator variable which equals to one if the investor is Asian, Hispanic, or African Americans, and zero otherwise. Investors who prefer not to disclose their gender or race are not included in these variables. "Senior" is equal to one if the investor is in a C-level position, or is a director, partner, or vice president. It is zero if their position is as an analyst (intern) or associate. "Cold Email Acceptance" is an indicator variable which equals to one if the investor feels that sending cold call emails is acceptable as long as they are well-written, and zero if the investor feels that it depends. "Prefer ESG" is an indicator variable which equals to one if the investor prefers ESG related startups, and zero otherwise. "Direct Investment" is an indicator variable which equals to one if the investor can directly make the investment, and zero if their investment is through limited partners or other channels. Panel E provides the financial information of the 68 funds that these investors work for. However, we can only recover parts of their financial information from Pitchbook.

Table 3: Experiment A Design, Randomization of Profile Components

Profile Component	Randomization Description	Analysis Variable
Startup Team Characteristics		
First and Last Names	Drawn from list of the same names given selected race and gender as used in Experiment 1 (See names in Tables B1)	White Female[a] (25%) Asian Female (25%) White Male (25%) Asian Male (25%)
Number of Founders	The team can have 1 founder or 2 co-founders	Single Founder (8/16)
Age	Founders' age is indicated by the graduation year Young VS Old=50% VS 50% Young: uniformly distributed (2005-2019) Old: uniformly distributed (1980-2005)	Age
Educational Background	Drawn from top school list and common school list (See school list Table B2)	Top School (8/16)
Entrepreneurial Experiences	The team can have serial founder(s) or only first-time founder(s)	Serial Founder (8/16)
Startup Project Characteristics		
Company Age	Founding dates are randomly drawn from the following four years {2016, 2017, 2018, 2019}	Company Age
Comparative Advantages	Randomly drawn from a comparative advantage list (See Tables B3), the number of drawn advantages is between 1 to 4	1 Advantages (4/16) 2 Advantages (4/16) 3 Advantages (4/16) 4 Advantages (4/16)
Traction	Half randomly selected profiles generate no revenue Half randomly selected profiles generate positive revenue. Previous monthly return: uniform distribution [5K, 80K]; Growth rate: uniform distribution [5%, 60%]	Positive traction (8/16)
Company Category	Randomly assigned as either B2B or B2C	B2B (8/16)
Number of Employees	Randomly assigned with one of four categories	0-10 (8/16) 10-20 (8/16) 20-50 (8/16) 50+ (8/16)
Target Market	Randomly assigned as either domestic market or international market	Domestic (8/16)
Mission	Randomly assigned with one of three categories "For profit", "For profit, consider IPO within 5 years", "Besides financial gains, also cares ESG"	For profit (8/16) For profit, IPO Plan (4/16) For profit, ESG (4/16)
Location	Randomly assigned as either U.S. or Outside the U.S.	U.S. (70%)
Previous Funding Situation		
Number of Existing Investors	Randomly assigned as one of the four categories with equal probability {0,1,2,3+}	Number of investors

[a]The randomization distribution is to increase the experimental power. Considering that our collaborative incubators have more Asian and female founders than the normal gender and race distribution, I increased the ratio of female and Asian founders in this experiment to mimic the distribution of the collaborative incubators, which provides the pool of potential matched startups. Although some investors feel that providing more information would be helpful, no one complains that the distribution of founding team gender and race is unrealistic.
[b]If there are two co-founders in the same founding team, all the founders' background information is similar to each other. For example, if the first founder's age belongs to the young founder category, then the second founder's age also belongs to the same age category.

Notes. This table provides the randomization of each startup profile's components and the corresponding analysis variables. Profile components are listed in the order that they appear on the hypothetical startup profiles. Weights of characteristics are shown as fractions when they are fixed across subjects (e.g., each subject saw exactly 8/16 resumes with all female team members) and percentages when they represent a draw from a probability distribution (e.g., for startups with positive revenue records, the revenue follows a uniform distribution between [5K - 80 K]). Variables in the right-hand column are randomized to test how investors respond to these analysis variables.

Table 4: Experiment A Design, Incentives Design for Different Evaluation Questions

Evaluation Questions	Matching Incentive (Version 2)	Monetary Incentive	Matching and Monetary Incentive (Version 1)
Q1 (quality evaluation)	Yes	Yes	Yes
Q2 (collaboration likelihood)	Yes	No	Yes
Q3 (contact interest)	Yes	No	Yes
Q4 (investment interest)	Yes	No	Yes
Q5 (risk evaluation)	Yes	N/A	Yes

Notes. This tables describes how different types of incentives affect each evaluation question. Column 1 shows that the matching incentive, which identifies the matched startups using the matching algorithm, works for all five of the evaluation questions. I sent Version 2 recruitment emails, instruction posters, and consent forms to investors who only receive this matching incentive. Column 2 shows that the monetary incentive, which provides a lottery opportunity, only incentivizes Q1 (the evaluation of the startup quality evaluations) because the financial returns for the lottery winners only depends on the belief of the startup's financial return. Column 3 shows that combining the matching and the monetary incentive together can also incentivize all five questions. I sent Version 1 recruitment emails, instruction posters, and consent forms to investors who received both incentives.

Table 5: Evaluation Results (Team VS Project)

Dependent Variable	Q1 Quality (1)	Q2 Collaboration (2)	Q3 Contact (3)	Q3 Contact (4)	Q4 Investment (5)	Q4 Investment (6)	Q5 Risk (7)
Serial Founder	5.23***	-0.81	5.64***	1.26	0.76***	0.13	-0.65
	(1.08)	(0.88)	(1.28)	(0.91)	(0.19)	(0.15)	(3.05)
Ivy	5.36***	-1.06	7.44***	3.01***	0.87***	0.20	-6.44**
	(1.10)	(0.87)	(1.31)	(0.93)	(0.20)	(0.15)	(3.26)
Number of Founders	1.56	-1.21	1.17	-0.11	0.21	0.04	-5.32*
	(1.07)	(0.88)	(1.29)	(0.91)	(0.20)	(0.15)	(3.06)
US Founder	0.95	0.02	4.23***	3.69***	0.08	0.03	-0.91
	(1.18)	(0.91)	(1.39)	(1.00)	(0.21)	(0.16)	(3.48)
# Comparative Adv	3.10***	-0.22	2.76***	0.34	0.55***	0.15**	0.91
	(0.54)	(0.43)	(0.64)	(0.43)	(0.10)	(0.07)	(1.48)
Has Positive Traction	12.70***	1.75**	13.35***	1.91*	1.81***	0.28*	-9.51***
	(1.07)	(0.86)	(1.28)	(0.99)	(0.20)	(0.16)	(3.15)
Number of Employees [0-10]	0.67	2.37**	-1.73	-2.57**	-0.19	-0.29	-1.18
	(1.43)	(1.16)	(1.69)	(1.18)	(0.26)	(0.20)	(3.94)
Number of Employees [10-20]	-1.08	0.94	-3.26	-2.08	-0.46	-0.33	
	(1.64)	(1.35)	(1.99)	(1.39)	(0.30)	(0.23)	
Number of Employees [20-50]	-0.47	-0.02	-1.21	-0.72	-0.16	-0.12	-1.28
	(1.45)	(1.17)	(1.71)	(1.17)	(0.27)	(0.19)	(3.59)
Company Age	-4.59*	-5.99***	-7.39**	-2.19	-1.26**	-0.54	-3.41
	(2.72)	(2.19)	(3.19)	(2.26)	(0.49)	(0.37)	(7.74)
Company Age^2	0.75	1.12**	1.27**	0.42	0.23**	0.10	0.77
	(0.54)	(0.44)	(0.64)	(0.45)	(0.10)	(0.07)	(1.52)
Is B2B	3.90***	3.73***	6.10***	1.47	0.81***	0.32**	-4.91
	(1.07)	(0.86)	(1.28)	(0.89)	(0.20)	(0.15)	(3.01)
Domestic Market	-0.10	-0.60	0.09	0.57	0.08	0.13	-3.32
	(1.08)	(0.86)	(1.28)	(0.90)	(0.20)	(0.14)	(3.19)
Q1				0.88***		0.12***	
				(0.03)		(0.01)	
Q2				0.18***		0.01	
				(0.03)		(0.01)	
Constant	49.75***	78.20***	66.20***	-4.19	5.62***	-0.33	67.01***
	(6.56)	(6.02)	(4.93)	(7.50)	(1.43)	(0.63)	(11.66)
Investor FE	Yes	Yes	Yes	Yes	Yes	Yes	
Observations	1,216	1,184	1,216	1,184	1,176	1,154	176
R-squared	0.44	0.55	0.56	0.80	0.44	0.70	0.34

Notes. This table reports the regression results of how the evaluation results respond to other startup team characteristics and startup project characteristics. In column (1), the dependent variable is the evaluation results of Q1 (quality evaluation). For Q1, the unit is 1 percentile compared with the investor's previous investment portfolio. 0 is the worst and 100 is the best startup that the investor has invested in before. In column (2), the dependent variable is the evaluation results of Q2 (collaboration interest). For Q2 , the unit is 1% of the possibility of collaboration with the investor. 0 means no possibility of collaboration and 100 is guaranteed to collaborate. In columns (3)-(4), the dependent variable is the evaluation results of Q3 (contact interest). For Q3 , the unit is 1% of the possibility to contact the startup team. 0 means that the investor definitely does not want to contact the team and 100 means the investor is guaranteed to contact them. In columns (5)-(6), the dependent variable is the evaluation results of Q4 (investment interest). For Q4, 1 unit means 1 hundredth (0.1) of the average investment. "Serial Founder", "Ivy", "US Founder", "Has Positive Traction", "Is B2B" and "Domestic Market" are all indicative variables that equal to one if the founder is respectively a serial entrepreneur, graduated from an Ivy League College, lives in the U.S., the project has positive traction, is a Business-to-Business startup, and focuses on the domestic market. These variables are equal to 0 if the startup does not match these characteristics. Number of founders is either 1 or 2; Number of Comparative Advantages and Company Age can be {1,2,3,4}; Company Age^2 is the square of the company's age. Q1 is the evaluation results of the startup's quality. Q2 is the evaluation results of the likelihood of collaboration. All the regression results add investor fixed effect and use the robust standard errors reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1 indicate statistical significance at 1%, 5%, and 10% levels, respectively.

Table 6: Evaluation Results (Team VS Project) q-value

Dependent Variable	Q1 Quality (1)	Q2 Collaboration (2)	Q3 Contact (3)	Q3 Contact (4)	Q4 Investment (5)	Q4 Investment (6)	Q5 Risk (7)
Serial Founder	5.23***	-0.81	5.64***	1.26	0.76***	0.13	-0.65
	(1.08)	(0.88)	(1.28)	(0.91)	(0.19)	(0.15)	(3.05)
Ivy	5.36***	-1.06	7.44***	3.01****	0.87***	0.2	-6.44
	(1.10)	(0.87)	(1.31)	(0.93)	(0.20)	(0.15)	(3.26)
Number of Founders	1.56	-1.21	1.17	-0.11	0.21	0.04	-5.32
	(1.07)	(0.88)	(1.29)	(0.91)	(0.20)	(0.15)	(3.06)
US Founder	0.95	0.02	4.23****	3.69***	0.08	0.03	-0.91
	(1.18)	(0.91)	(1.39)	(1.00)	(0.21)	(0.16)	(3.48)
# Comparative Adv	3.1***	-0.22	2.76***	0.34	0.55***	0.15	0.91
	(0.54)	(0.43)	(0.64)	(0.43)	(0.10)	(0.07)	(1.48)
Has Positive Traction	12.7***	1.75*	13.35***	1.91	1.81***	0.28	-9.51**
	(1.07)	(0.86)	(1.28)	(0.99)	(0.20)	(0.16)	(3.15)
Number of Employees [0-10]	0.67	2.37*	-1.73	-2.57*	-0.19	-0.29	-1.18
	(1.43)	(1.16)	(1.69)	(1.18)	(0.26)	(0.20)	(3.94)
Number of Employees [10-20]	-1.08	0.94	-3.26	-2.08	-0.46	-0.33	0
	(1.64)	(1.35)	(1.99)	(1.39)	(0.30)	(0.23)	(0.00)
Number of Employees [20-50]	-0.47	-0.02	-1.21	-0.72	-0.16	-0.12	-1.28
	(1.45)	(1.17)	(1.71)	(1.17)	(0.27)	(0.19)	(3.59)
Company Age	-4.59	-5.99***	-7.39**	-2.19	-1.26**	-0.54	-3.41
	(2.72)	(2.19)	(3.19)	(2.26)	(0.49)	(0.37)	(7.74)
Company Age^2	0.75	1.12**	1.27*	0.42	0.23**	0.1	0.77
	(0.54)	(0.44)	(0.64)	(0.45)	(0.10)	(0.07)	(1.52)
Is B2B	3.90***	3.73***	6.1***	1.47	0.81***	0.32	-4.91
	(1.07)	(0.86)	(1.28)	(0.89)	(0.20)	(0.15)	(3.01)
Domestic Market	-0.10	-0.60	0.09	0.57	0.08	0.13	-3.32
	(1.08)	(0.86)	(1.28)	(0.90)	(0.20)	(0.14)	(3.19)
Q1				0.88***		0.12***	
				(0.03)		(0.01)	
Q2				0.18***		0.01	
				(0.03)		(0.01)	
Constant	49.75***	78.2***	66.2***	-4.19	5.62***	-0.33	67.01***
	(6.56)	(6.02)	(4.93)	(7.50)	(1.43)	(0.63)	(11.66)
Investor FE	Yes	Yes	Yes	Yes	Yes	Yes	
Observations	1,216	1,184	1,216	1,184	1,176	1,154	176
R-squared	0.44	0.55	0.56	0.80	0.44	0.70	0.34

Notes. This table reports regression results of how the evaluation results respond to other startup team characteristics and startup project characteristics. It's the same as Table 5 except that I report the q-value adjusted by the Bonferroni method and Simes method (red *) to implement the multiple hypothesis testing. Since the Simes method is less conservative than the Bonferroni method, I use * to indicate the significance level of the q-value generated by the Simes method whenever the significance level of the Simes method q-value is smaller than that of the Bonferroni method q-value. Standard errors are in parentheses. *** q-value<0.01, ** q-value<0.05, * q-value<0.1 indicate statistical significance at 1%, 5%, and 10% levels, respectively.

Table 7: Standardized Coefficients of Evaluation Results (Team VS Project)

Dependent Variable	Q1 Quality (1)	Q2 Collaboration (2)	Q3 Contact (3)	Q3 Contact (4)	Q4 Investment (5)	Q4 Investment (6)	Q5 Risk (7)
Serial Founder	0.109***	-0.019	0.087***	0.019	0.089***	0.015	-0.014
	(0.022)	(0.021)	(0.020)	(0.014)	(0.023)	(0.017)	(0.067)
Ivy	0.111***	-0.025	0.114***	0.046****	0.101***	0.023	-0.142
	(0.023)	(0.021)	(0.020)	(0.014)	(0.023)	(0.017)	(0.069)
Number of Founders	0.033	-0.029	0.018	-0.002	0.024	0.005	-0.118
	(0.023)	(0.021)	(0.020)	(0.014)	(0.023)	(0.017)	(0.067)
Located in US	0.018	0.000	0.061****	0.053***	0.009	0.003	-0.019
	(0.022)	(0.020)	(0.020)	(0.014)	(0.023)	(0.017)	(0.068)
# Comparative Adv	0.131***	-0.010	0.087***	0.011	0.132***	0.036	0.041
	(0.022)	(0.020)	(0.020)	(0.014)	(0.023)	(0.017)	(0.067)
Has Positive Traction	0.265***	0.041*	0.207***	0.030	0.211***	0.033	-0.211**
	(0.022)	(0.021)	(0.020)	(0.015)	(0.023)	(0.018)	(0.068)
Number of Employees [0-10]	0.012	0.048*	-0.023	-0.034*	-0.020	-0.030	-0.023
	(0.026)	(0.024)	(0.023)	(0.016)	(0.027)	(0.020)	(0.071)
Number of Employees [10-20]	-0.018	0.018	-0.040	-0.026	-0.043	-0.031	0.000
	(0.027)	(0.025)	(0.024)	(0.017)	(0.027)	(0.020)	(0.000)
Number of Employees [20-50]	-0.009	0.000	-0.016	-0.010	-0.016	-0.012	-0.025
	(0.026)	(-0.317)	(0.023)	(0.016)	(0.027)	(0.020)	(0.072)
Company Age	-0.214	-0.317**	-0.256**	-0.076	-0.330**	-0.142	-0.170
	(0.127)	(0.116)	(0.112)	(0.078)	(0.129)	(0.097)	(0.387)
Company Age^2	0.177	0.301**	0.224*	0.074	0.300**	0.128	0.195
	(0.127)	(0.116)	(0.112)	(0.078)	(0.129)	(0.097)	(0.386)
Is B2B	0.081***	0.088***	0.095***	0.023	0.095***	0.037	-0.109
	(0.022)	(0.020)	(0.020)	(0.014)	(0.023)	(0.017)	(0.067)
Domestic Market	-0.002	-0.014	0.001	0.009	0.009	0.015	-0.074
	(0.022)	(0.020)	(0.020)	(0.014)	(0.023)	(0.017)	(0.068)
Q1				0.639***		0.659***	
				(0.018)		(0.023)	
Q2				0.121***		0.040	
				(0.020)		(0.025)	
Investor FE	Yes	Yes	Yes	Yes	Yes	Yes	
Observations	1,216	1,184	1,216	1,184	1,176	1,154	176
R-squared	0.44	0.55	0.56	0.80	0.44	0.70	0.34

Notes. $Y_{ij}^{(k)} = X_{ij}\beta_i^{(k)} + \alpha_i + \epsilon_{ij}^{(k)}$ Investor i evaluates the kth question of the jth profile. This table reports the q-value (multiple hypothesis testing), which is adjusted by Bonferroni method or Simes method (blue *, use more information). Standardization applies to all the independent variables except for the indicator variables used for the fixed effect. In column (1)-(7), the dependent variable is the evaluation results of Q1 (quality evaluation), Q2 (collaboration interest), Q3 (contact interest), Q4 (investment interest) and Q5 (risk evaluation). "Serial Founder", "Ivy", "US Founder", "Has Positive Traction", "Is B2B" and "Domestic Market" are indicative variables that equal to one if the founder is respectively a serial entrepreneur, graduated from an Ivy League College, lives in the U.S., the project has positive traction, is a Business-to-Business startup, and focuses on the domestic market. These variables are equal to 0 if the startup does not have any such characteristics. Number of founders is either 1 or 2; Number of Comparative Advantages and Company Age can be {1,2,3,4}; Company Age^2 is the square of the company age. Q1 is the evaluation results of startup quality. Q2 is the evaluation results of the collaboration likelihood. All the regression results add investor fixed effect and use the robust standard errors reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1 indicate statistical significance at 1%, 5%, and 10% levels, respectively.

Table 8: Summary Statistics for Hypothetical Startups

Panel A: 1st round

	N	Industry Covered
B2B	13	Media, Music, Fashion, Advertisement, Real Estate, Construction, SAAS, Education, Logistics, Energy, Agriculture
B2C	12	Media, Fashion, Sports, Food, SAAS, Traveling, Pets, Chemical products, Education
Healthcare	8	Healthcare
Total	33	

Panel B: 2nd round

B2B	13	Entertainment, Media, Packaging, Advertisement, Finance, Management, Education SAAS
B2C	14	Entertainment, Media, Energy, SAAS, Sports, Chemical Product, Food
Healthcare	7	Healthcare
Total	34	

Panel C: Total

B2B	26	Media, Music, Fashion, Advertisement, Real Estate, Construction, SAAS, Education, Logistics, Energy, Agriculture, Entertainment, Packaging, Finance, Management
B2C	26	Media, Fashion, Sports, Food, SAAS, Traveling, Pets, Chemical products, Education, Entertainment, Energy,
Healthcare	15	Healthcare
Total	67	

Notes. This table reports descriptive statistics for the 67 startups used in the first-round and second-round correspondence test. All the startups are classified into B2B (Business to Business), B2C (Business to Consumer) and Healthcare following the classification category of Gornall and Strebulaev (2020). I also provide more granular industry information about the created startups in the table. Panel A reports the startup category distribution of the first-round correspondence test, which was implemented between 03/2020 and 04/2020 during the outbreak of COVID-19. During the "Chinese Virus" period between 03/18/2020 to 03/23/2020, the three pitch emails sent out include an AI logistics startup (B2B), a healthcare startup and a startup developing a financial management platform targeting U.S. schools (B2B). Panel B reports the startup category distribution of the second-round correspondence test, which was implemented between 07/2020 and 08/2020 when the economy began to reopen. Panel C reports the startup category distribution of the 67 startups used in the two rounds of correspondence test. If a startup belongs to both B2B and B2C, I have labeled it as "B2B". In the first round experiment, there were 2 startups belonging to both B2B and B2C. In the second round experiment, there were 3 startups belonging to both B2B and B2C.

Table 9: Investor Response to Randomized Emails

Panel A: Response Summary Statistics

	Observations	Mean	Median	S.D.	min	max
Open Rate	3,720	12.03%	0	0.33	0	1
Staying Time (Unit: s)	3,381	24.10	10.33	26.73	0.01	86.63
Click Rate	519	1.68%	0	0.13	0	1
Replied Emails	472	1.53%				

Panel B: Email Opening Behaviors

	Dependent Variable: 1(*Opened*)				
	(1)	(2)	(3)	(4)	(5)
	Full	Full	Full	"Pure Ivy"	Full
Female Founder=1	0.010***				0.010***
	(0.004)				(0.004)
Asian Founder=1		0.007*			0.006
		(0.004)			(0.004)
Ivy=1			0.007*	0.012**	0.007*
			(0.004)	(0.005)	(0.004)
Project Advantage=1					0.001
					(0.004)
Asian Founder=1 × March Chinese Virus=1		-0.009			
		(0.010)			
March Chinese Virus=1		-0.040**			
		(0.020)			
US Investor=1	-0.016***	-.016***	-0.016***	-0.023***	-0.016***
	(0.006)	(0.006)	(0.006)	(0.008)	(0.006)
Female Investor=1	-0.019***	-0.020***	-0.019***	-0.017***	-0.019***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)
Constant	0.193***	0.194***	0.194***	0.108***	0.186***
	(0.019)	(0.019)	(0.019)	(0.017)	(0.019)
Startup FE	Yes	Yes	Yes	Yes	Yes
Observations	30,909	30,909	30,909	16,578	30,909
Adjusted R-squared	0.005	0.005	0.005	0.006	0.005

Notes. This table summarizes investors' email responses to the first-round correspondence test and reports the regression results of global investors' email-opening behavior of randomized pitch emails in Experiment B. Panel A summarizes important investors' information acquisition behaviors in the pitch email setting. Panel B reports regression results of how startup characteristics affect investors' email opening behavior. In Panel B, the dependent variable is a dummy variable, which is one when an investor opens the pitch email, and zero otherwise. "Female Founder = 1" is an indicator variable that equals one if the first name of the email sender is a female name, and zero otherwise. Similarly, "Asian Founder =1" is an indicator variable that equals one if the last name of the email sender is an East Asian name, and zero otherwise. "Ivy = 1" is an indicator variable for Ivy League educational background. "Project Advantage = 1" is an indicator variable which is one when the email subject line includes the corresponding comparative advantages. "March Chinese Virus = 1" is an indicator variable which is one when the email was sent between 2020/03/18-2020/03/24 when President Trump uses the phrase "Chinese Virus". "US Investor =1" and "Female Investor = 1" are indicator variables that are one for U.S. investors and female investors. Columns (1) (2) (3) and (5) use all the observations collected in the first-round correspondence test. In Column (4), results are reported for the sub-sample where the startup team's graduated university is from purely an Ivy League college, Stanford or MIT. "Pure_Ivy" indicates cases like "Team from Columbia University" while 'Mixed_Ivy" indicates cases like "Team from Columbia University and Juilliard Music School". For some startups in the music or medical industry, I combined an Ivy League college with a good university in that specific area for the treatment group. R^2 is the adjusted R^2 for all OLS regressions. Standard errors in parentheses are clustered at the investor level. *** p<0.01, ** p<0.05, * p<0.1 indicate statistical significance at 1%, 5%, and 10% levels, respectively.

Figures

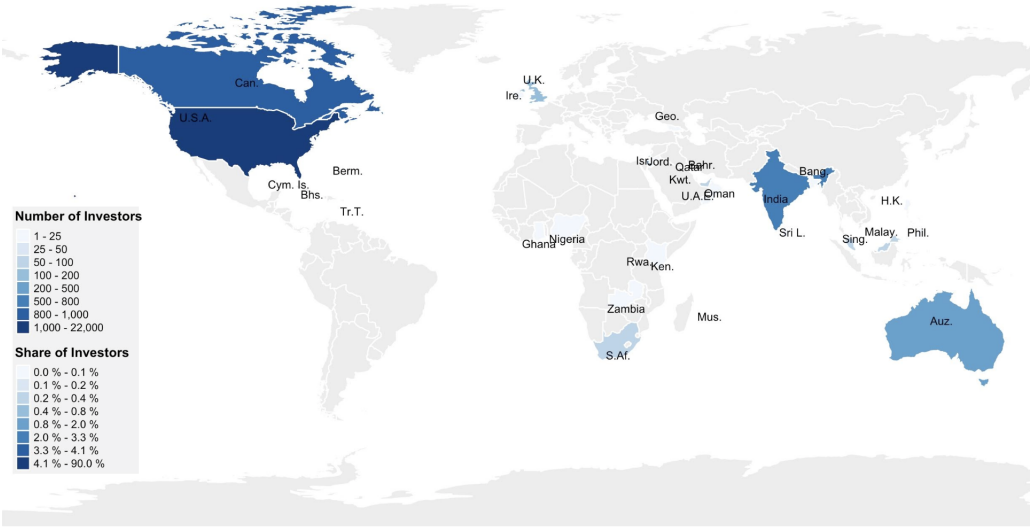

Figure 1: Geographical Distribution of Global Investors

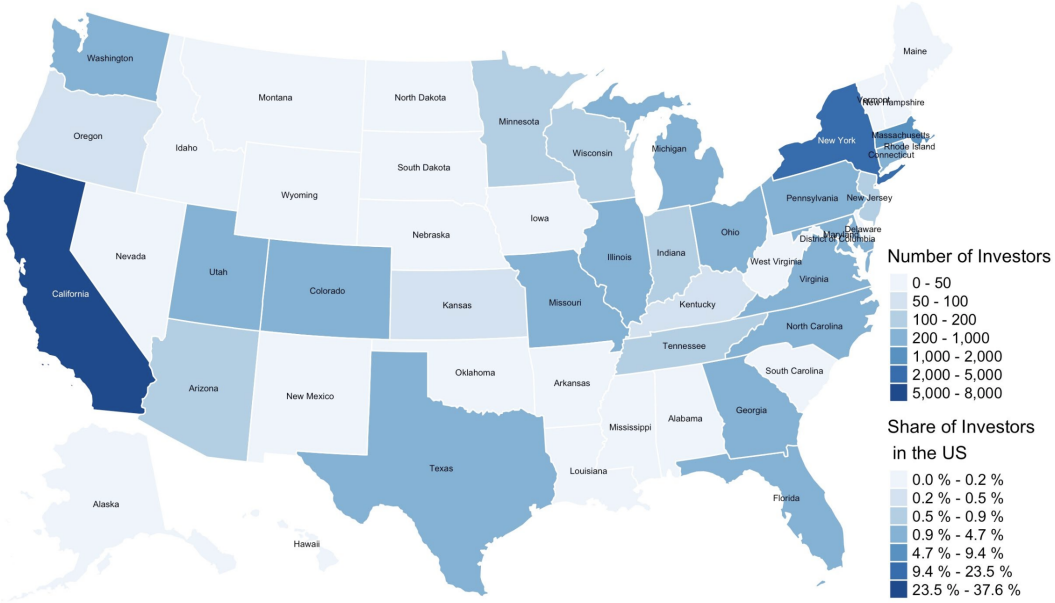

Figure 2: Geographical Distribution of US Investors

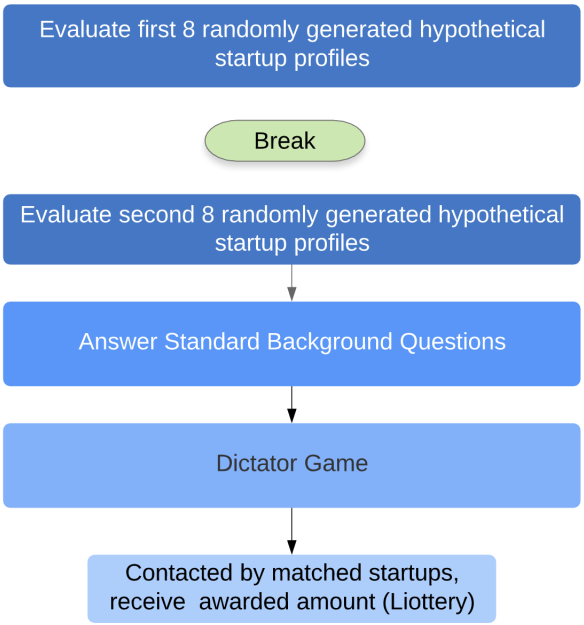

Figure 3: Lab-in-field Experiment Flow Chart

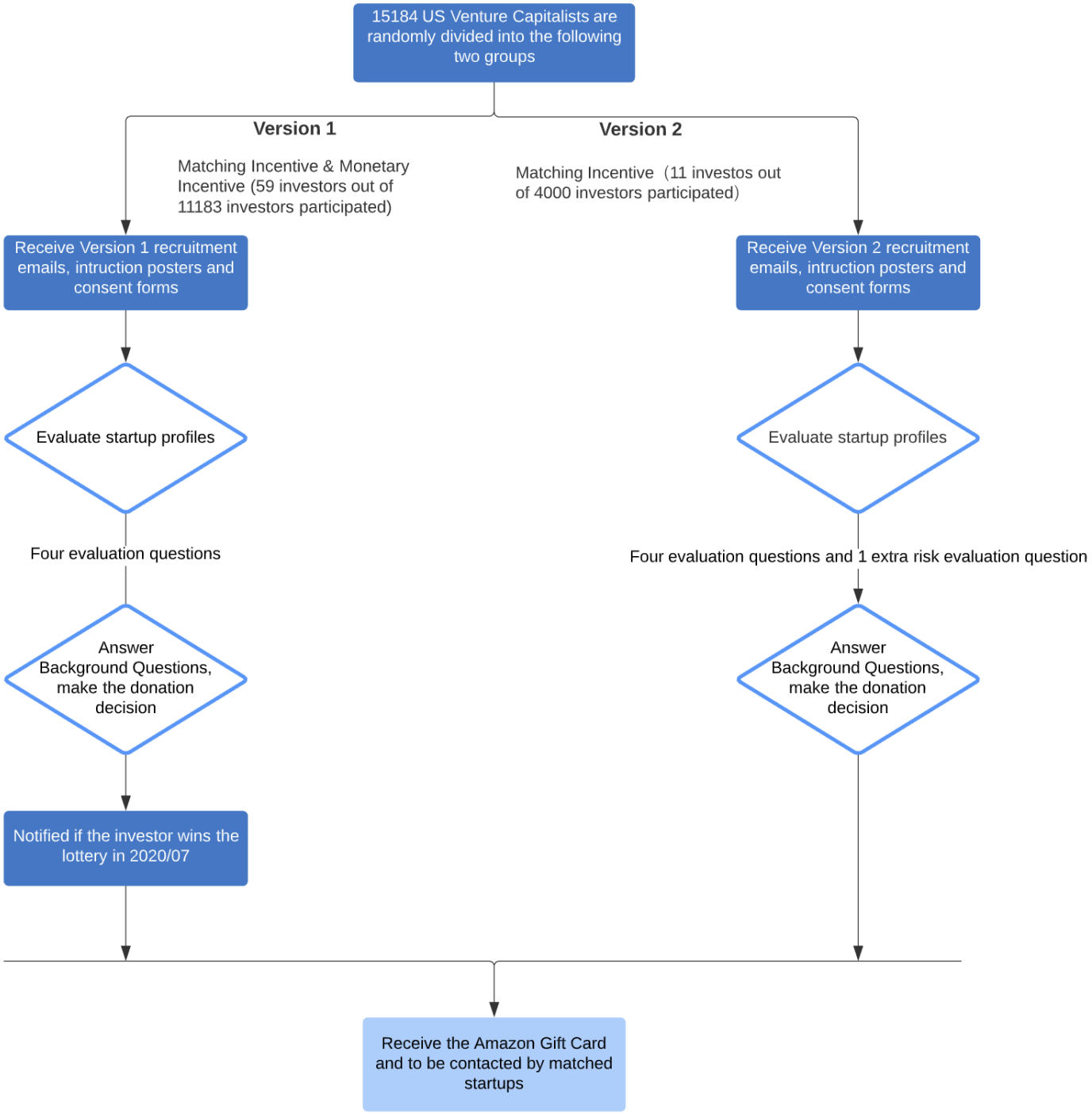

Figure 4: Lab-in-field Experiment Incentive Structure

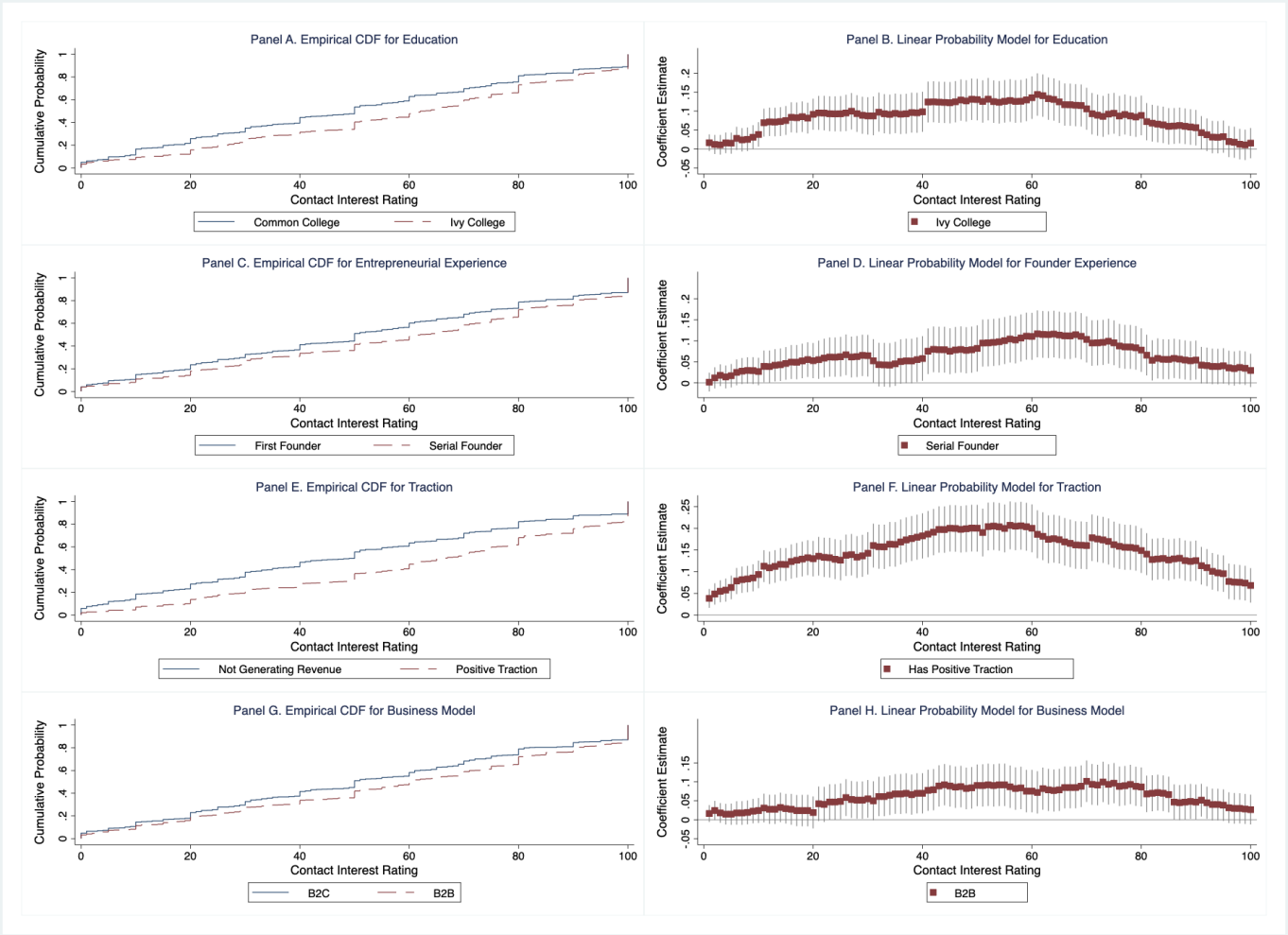

Figure 5: Distributional Effect across Contact Interest

Notes: This figure demonstrates the effect of a startup's team and project characteristics across the contact interest distribution using the total profiles evaluated in Experiment A. Panel A provides the empirical CDF for founders' educational background of investors' contact interest rating (i.e. $Pr(\text{Contact Interest} > x|$ Graduate from Ivy League College) and $Pr(\text{Contact Interest} > x|$ Graduate from Common College)). Panel B provides the OLS coefficient estimates (i.e. $Pr(\text{Contact Interest} > x|$ Graduate from Ivy League College) $-$ $Pr(\text{Contact Interest} > x|$ Graduate from Common College)) and the corresponding 95% confidence level. Similarly, Panels C E, and G provide the empirical CDF for the founder's entrepreneurial experiences, the project's traction and for the business model. Panels D, F and H provide the OLS coefficient estimates for the founder's entrepreneurial experiences, the project's traction and for the business model.

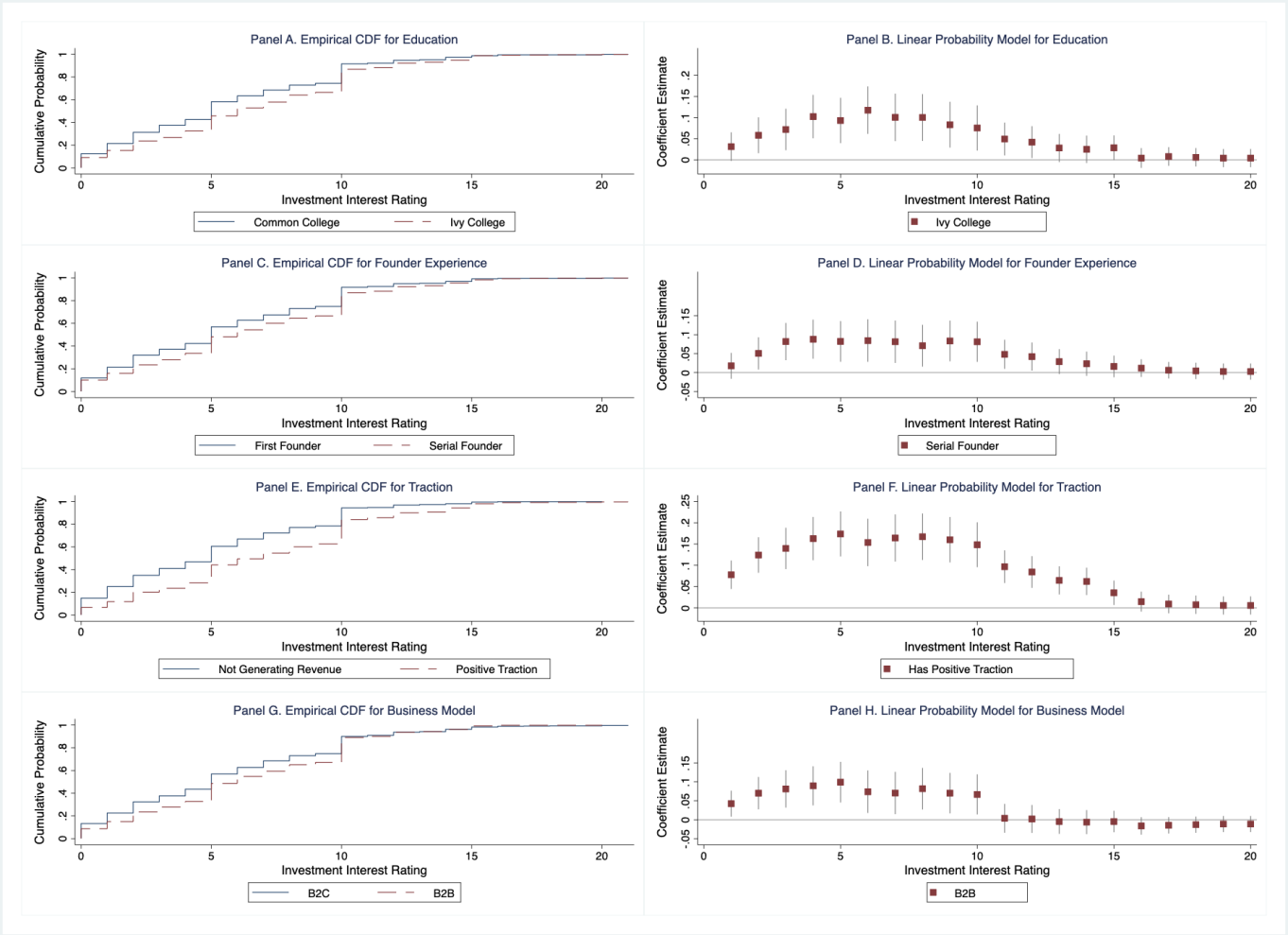

Figure 6: Distributional Effect across Investment Interest

Notes: This figure demonstrates the effect of startup team and project characteristics across the investment interest distribution using the total profiles evaluated in Experiment A. Panel A provides the empirical CDF for founder's educational background of investors' investment interest rating (i.e. $Pr(\text{Investment Interest} > x|$ Graduate from Ivy League College) and $Pr(\text{Investment Interest} > x|$ Graduate from Common College)). Panel B provides the OLS coefficient estimates (i.e. $Pr(\text{Investment Interest} > x|$ Graduate from Ivy League College) $-$ $Pr(\text{Investment Interest} > x|$ Graduate from Common College)) and the corresponding 95% confidence level. Similarly, Panels C E, and G provid the empirical CDF for founder's entrepreneurial experiences, project's traction and business model. Panels D, F and H provide the OLS coefficient estimates for the founder's entrepreneurial experiences, the project's traction and for the business model.

Email Subject Line: Invest in StudioFinder {University1} {Advantage1}

Hello {Name},

My name is {FounderName}. I am the co-founder of StudioFinder, a music studio search app in New York City. Our team found your information in the VCPro Database, and we feel you might be interested in our startup.

StudioFinder is simply a music studio Airbnb that matches a studio holder and an artist. Our platform helps individuals who have a studio set in their house to make profits while they are not using it. StudioFinder provides new artists with affordable studio settings. The music studio rental corporations in New York cannot compete with our commission, which is the lowest (1.5 %), because we just match individuals. We have been collaborating with a few studios for a year and have had positive outcomes from new artists and video makers.

{Advantage2}{University2}

We are getting ready to raise funding to accelerate software adoption and bring StudioFinder to more users. If you are interested, we would love to share our pitch deck with you. Any feedback is also highly appreciated.

Thank you for your time. We look forward to your reply!

Sincerely,

{FounderName}

StudioFinder

Figure 7: Example of the Pitch Email

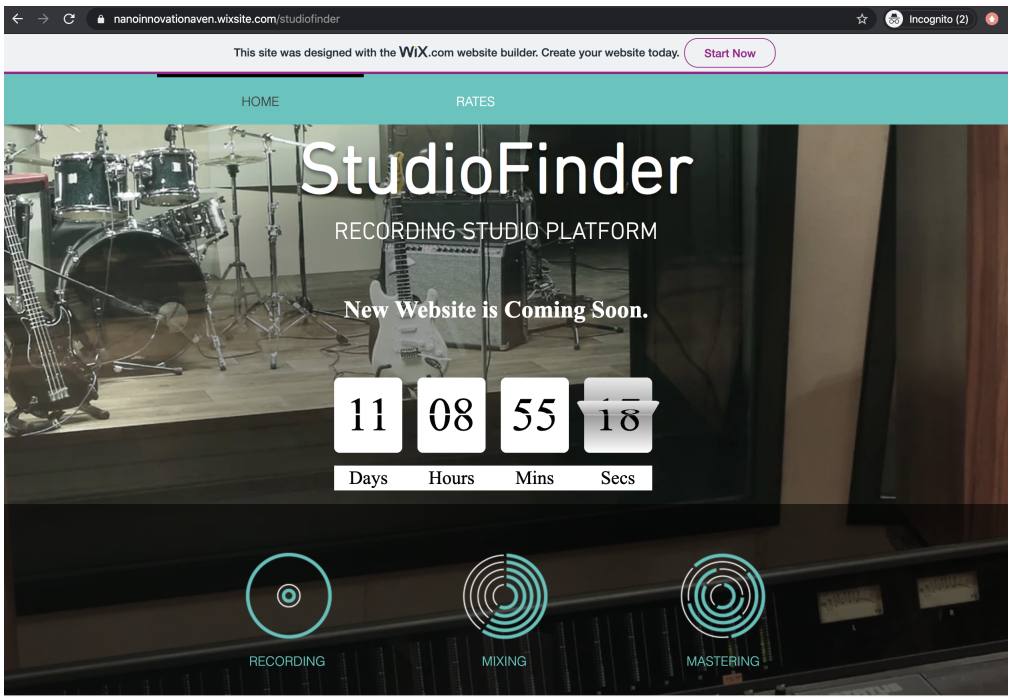

Figure 8: Example of the Startup Website

Subject Line: Monitoring the 2019 Novel Coronavirus (2019-nCoV)
Dear [First Name]:

We are actively monitoring the 2019 Novel Coronavirus (2019-nCoV) and want to share with you important information about the virus's symptoms and current recommendations. The Centers for Disease Control and Prevention (CDC) is working with the World Health Organization as this outbreak, originating in December 2019 in Wuhan City, Hubei Province, China, continues to expand.

Currently, there are few known cases in the U.S. and other countries. However, we want to provide some additional information as this situation evolves. This virus belongs to a family of viruses called "coronavirus." There are other viruses in the coronavirus family that can cause illness in both humans and animals. These viruses can cause either mild illness like a cold or can make people very sick with pneumonia. This particular coronavirus has not been seen previously in humans. There is no vaccine available for this or other coronaviruses.

How is it transmitted?
Since this virus is very new, health authorities continue to carefully watch how it spreads. It is spread from animals to humans and also appears to be spread from person to person. Incubation is likely 5-7 days, but may be up to 14 days.

What are the symptoms?
Fever, cough, and shortness of breath are the most common symptoms. If you have any of these symptoms and have been traveling or in contact with someone that has been traveling in the Asia-Pacific region, please seek medical attention (see below).

Recommendations:

- Please review the CDC Travel Health Notice. The CDC recommends that travelers avoid all nonessential travel to Wuhan, China.
- If you have traveled recently, especially to the Asia-Pacific region, and are experiencing the above symptoms please seek medical attention immediately:
- Wash hands often with soap and water for at least 20 seconds. Use an alcohol-based hand sanitizer, if soap and water is not available.
- Expect additional time at airports and transportation hubs throughout Asia and in major US cities for health screening to prevent spread.

With care for our community,

Confidentiality Disclaimer: This e-mail message and any attachments are private communication and may contain confidential, privileged information meant solely for the intended recipient. If you are not the intended recipient, you are hereby notified that any use, dissemination, distribution or copying of this communication is strictly prohibited. Please notify the sender immediately by replying to this message, then delete the e-mail and any attachments from your system. Thank you.

Figure 9: Example of the Testing email

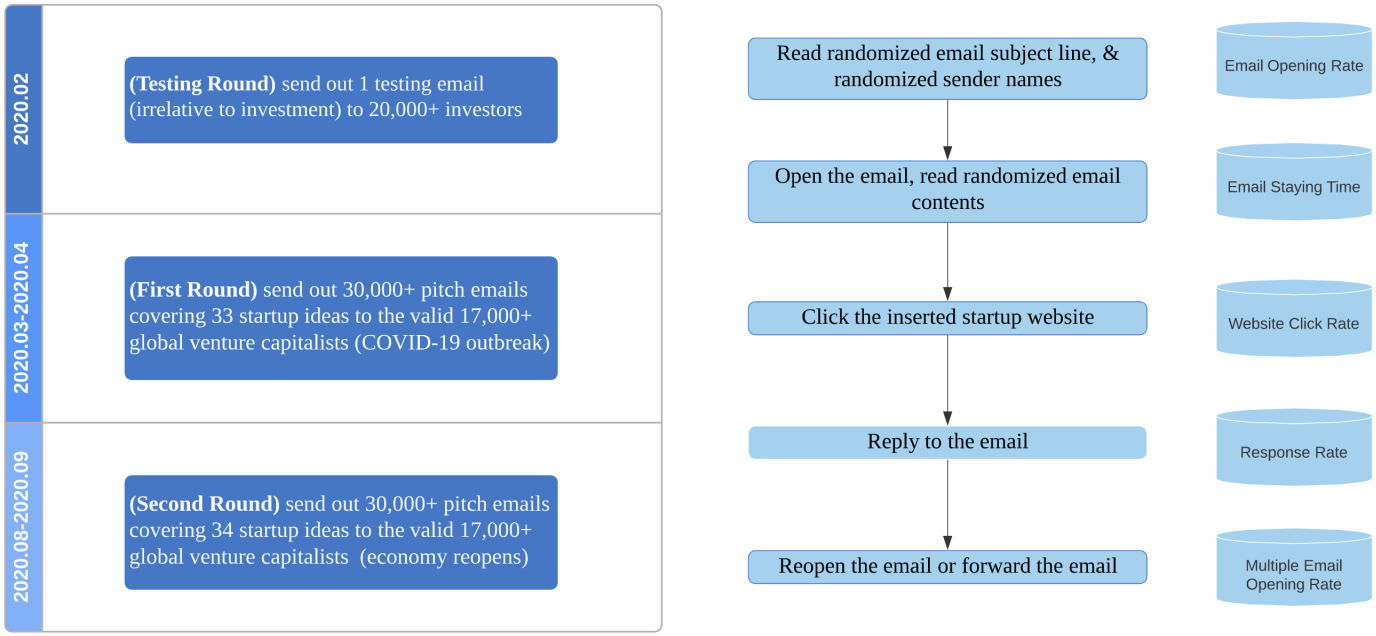

Figure 10: Correspondence Test Flow Chart

Notes: This figure describes the experiment timeline, experiment design and the traced investors' email behaviors.

Appendix

A Data Construction Process

A.1 Data Sources

In order to construct individual-level global venture capitalist database, containing the demographic information and contact information, I use the following commercial datasets and manually collected data.

A.1.1 PitchBook

PitchBook database contains extremely comprehensive information about venture capital and angel investors' demographic information and contact methods in the world, especially in the U.S. I purchased their individual-level data from 2017-2020 and selected the following types of investors from the PitchBook: Angel Group, Angel individual investor, corporate venture capital, family office and venture capital.

A.1.2 ExactData

I also purchased a database of VC practitioners in the U.S. from a professional data company "ExactData. Inc", which collects the information from online websites and various VC industry events or gatherings. My research team verified and cleaned the database during the summer 2018 and spring 2019, deleted those who have left the industry and corrected other invalid information. Moreover, we manually went through each firm contained in the database and added the contact information of new VC practitioners who was not contained in the original database through the following channels: personal websites, firm websites, LinkedIn, Zoominfo and RocketReach.

A.1.3 SDC New Issue Database & RocketReach

RocketReach is one of the largest platforms and data source providing contact information of company employees.[53] Given the company name list, it is feasible to extract the employees' contact information. Therefore, I implemented the following steps to further add investors' contact information:

Step 1: add new companies

I added many new venture capital funds to our previous database by checking the 2018 national venture capital association (NVCA) member list and Thompson Reuters SDC Platinum VentureXpert Database.

Step 2: collect investors' information

Based on the fund list, I searched for all the employees working in the corresponding funds and companies using RocketReach API. I only kept the investment related positions, like VC investor, analysis, associate, VP, MD, etc. RocketReach provided me with both the contact information (e.g. email and telephone) and also the demographic information (e.g. Facebook, twitter, LinkedIn, Position etc.). For investors not contained in PitchBook and ExactData, individual-level investor's demographic data were extracted manually from personal website, Facebook, firm websites, LinkedIn, Zoominfo and other social platforms.

[53] Using RocketReach to collect contact information of employees is a very efficient data collection method. Given company name list, researchers can extend the company level data to individual-level data by using RocketReach. Potentially this data collection method can be implemented in a board range of research in labor economics and corporate finance field.

A.1.4 ZDatabase

ZDatabase is provided by Zero2IPO Research Center and is currently one of the most comprehensive, accurate and timely databases covering VC and PE industry in China.[54] It covers rich information of active Chinese investment institutions and their management team starting from 1992. All the data were collected through regular surveys, daily phone call and verification through many other available channels. The database is updated daily to provide the accurate, promptly and authoritative data source. Considering that the research was implemented in English, I only included investors from the area of Hong Kong and excluded investors from the Mainland.

A.2 Key variables

A.2.1 Gender

PitchBook and ExactData contain each investor's gender information. For other investors not contained in these datasets, my research team manually verified their gender by searching online social platforms and company websites. For investors whose gender information is ambiguous, I excluded them from the recruitment list.

A.2.2 Location

PitchBook and ExactData contain each investor's location information. For other investors not contained in these datasets, my research team manually collected their location information on LinkedIn or company websites.

A.2.3 Industry

PitchBook contains each investor and their fund's detailed industry preference. For other investors not contained in PitchBook, my research team manually collected their individual-level preference from LinkedIn and other social platforms. If the individual-level industry preference is not available, I will use the fund's industry preference instead. If no preference information is found online or from CBInsight or PitchBook, I will assume such investor does not have any specific investment preference. Such assumption will bring extra noise and lower the email response rate in the correspondence test.

A.2.4 ESG

PitchBook contains each fund's investment philosophy and their types. In the heterogeneous analysis based on fund's ESG criteria, I treat those not-for-profit VC funds as impact funds and for-profit VC funds as common funds. Such classification method potentially underestimates the fraction of ESG-related VC funds. An alternative way is to classify VC funds through selecting ESG-representative key words in their company descriptions as what Barber, Morse and Yasuda (2020) did. However, the key word selection is very subjective and highly depends on context. Based on this more aggressive method, the ESG-related funds can account for roughly 7% of the total observations. However, basic heterogeneous effect analysis based on these two classification methods is similar.

[54] ZDatabase description: http://www.p5w.net/fund/smjj/201209/P020120905327816063973.pdf

B Lab-in-field Experiment

B.1 Startup Profile Construction Process

B.1.1 Startup Team Characteristics (Human Capital Assets)

Manipulating Gender and Race. — To indicate the gender and race of the startup founder, I randomly assign each hypothetical startup team member a first name highly indicative of gender (male or female) and a last name highly indicative of race (Asian or white).[55] In the same startup team, all the members are assigned the names of same gender and race to make such information more salient. Also, I emphasized the information of gender and race in both Q1 and Q2 by mentioning the founder's name again and use indicative words like "she/her/his/him/he". The list of full names used in the tool is provided in Table B1.[56] Similar to other components, the combination of first names and last names are dynamically implemented by Qualtrics.

Manipulating Age and Education. — The age of the startup founder is indicated by the graduation year of their colleges or graduate schools rather than listed directly.[57] If a team has two co-founders, their age falls into the same range, which belongs to with the elder group or the younger group. The randomization details are provided in Table 3. I also randomize the education background based on attending prestigious universities and common universities, whose school list is provided in Appendix B Table B2. All the universities selected have alumni who are real successful startup founders based on the biography information recorded in the Pitchbook Database.

Other related characteristics. —Besides the gender, race, age and education background, I also randomize the following startup team characteristics, which are usually available on public platforms like LinkedIn, AngelList or CrunchBase. Such characteristics include the number of startup founders (1 or 2) and the founder's previous entrepreneurial experience. In order to accommodate investors from different industries, I use the wording "serial entrepreneur" to indicate the founding team's previous experiences.

B.1.2 Startup Project Characteristics (Non-human Capital Assets)

Comparative Advantages. —To indicate the quality of the startup project, I randomly generated a subset of common comparative advantages in the startups and use the number of these advantages to suggest the quality. However, considering that different comparative advantages are valued by investors from different industries,[58] I also ask investors which of the

[55] Following the similar concern of the first correspondence test, I only add Asian entrepreneurs in the experiment because other minority groups account for a small percentage of the total entrepreneurial community. Considering our collaborative incubators and startups have relatively more Asian founders and female founders, the ratio of female and male startup founders is 50% and 50% to maximize the experiment power. Similar ratio is used for Asian founders and white founders.

[56] Names were selected uniformly and without replacement within the chosen column of the table. I use the variation induced by these names for the analysis variables Female, Asian; Female, White; Male, Asian; Male, White. I did not list the gender information explicitly as what the Crunchbase platform does (For example, add one more bullet point: Gender: Male) due to the experiment observer effect.

[57] It is suspicious to list the age directly in the startup profile because none of the public startup platforms did so before. Considering that age discrimination is a sensitive preference question, I use the graduation year as a proxy to age at the cost of accuracy in order to achieve more realism.

[58] For example, investors in the tech industry may care more about the registered intellectual properties in order to create the entry barriers while investors in the fashion industry may care more about the celebrity endorsement rather than any tech related advantages.

comparative advantages they would care among the 10 comparative advantages used at the end of the tool and used the number of such cared comparative advantages to confirm the results. The comparative advantage list is provided in Table B3.

Traction. —Traction is also an important indicator of the startup's financial situation and is measured by the previous monthly revenue and the annual revenue growth rate. Considering that we target early-stage investors, half of the startup profiles do not generate positive revenue yet and the other half have generated positive revenue. The range of the previous monthly revenue and the annual revenue growth rate comes from Pitchbook, which are biased towards more mature companies.[59]

Mission (ESG) —How ESG criteria affects investors' decision is an important institutional question that draws more and more attention from both practitioners and researchers recently. In order to randomize the company's ESG criteria, I introduced a random carriable called "Mission", which indicates whether such startups are purely profit driven (i.e. the control group, most commonly observed startups), profit driven with an IPO plan within 5 years (i.e. treatment 1 group) or also care about its environmental and social impact (i.e. treatment 1 group , social ventures). The description of the ESG related mission is extracted from real social ventures.

Other related characteristics —Besides the project characteristics mentioned above, I also added the following characteristics usually available on CrunchBase to enrich the startup profiles: startup founded date, company category (B2B or B2C),[60] number of employees, targeted market and location. Since the investors recruited in this experiment are the U.S. investors, I only create two categories in terms of location, which includes the U.S. and outside the U.S., in order to test any potential home bias channels.

B.1.3 Previous Fund-raising Situation

Number of existing investors — Some investors may reply on previous investors' behaviors to make their decisions rather than reply on their own private information, especially when the previous investors are successful. Such herding behavior is documented in the IPO setting where subsequent investors ignore their private information and imitate earlier investors (Bikhchandani, Hirshleifer and Welch (1992)) and explained by informational cascades (Bikhchandani et al. (1992)). In order to test this behavior in the primary market, I also randomize the number of existing investors to indicate other investors' decisions similar to . Existing investors' information is also available on multiple platforms like CrunchBase, Pitchbook or CB Insights. However, one limitation of such randomization is that I did not provide further background information of existing investors' financial background and reputation. Future researchers can provide more background information in order to better test such theoretical hypothesis.

[59] The growth rate of some early stage startups can be 100% to 200% while most of the startups recorded in Pitchbook are from 20% to 80%.

[60] Business to business or business to customers. Such category may affect investor's expectation since it is closely related to the startup's underlying business models. See the discussion of Tomasz Tungus' twitter, who is an investor at Redpoint.

Table B1: Full Names Populating Profile Tool

Asian Female	**White Female**	**Asian Male**	**White Male**
Cynthia Huynh	Amber Morris	Evan Liu	Patrick Kelly
Jennifer Tang	Erica Carpenter	Alan Wu	Stephen Bennett
Amanda Cheung	Anna Hoffman	Bryan Liang	Steven Martin
Christina Chang	Amanada Gray	William Chung	Jeremy White
Linda Chung	Tiffany Roberts	Nicholas Wang	Jason Adams
Brittany Yi	Lisa Taylor	Charles Luu	Donald Schultz
Megan Ho	Karen Carroll	Zachary Ho	Jack Wright
Emily Xu	Danielle Collins	Marcus Yoon	Victor Becker
Jacqueline Lin	Megan Bennett	George Thao	Michael Hughes
Kayla Wang	Brenda Cox	Vincent Huynh	Keith Meyer
Cassandra Kwon	Kathleen Phillips	Luke Yang	Anthony Roberts
Julie Chan	Amber Sullivan	Justin Dinh	Justin Cooper
Monica Luong	Madeline Walsh	Matt Hoang	Benjamin Hill
Amber Hoang	Abigail Kelly	Jacob Xu	Mark Myers
Sara Truong	Alicia Cook	Donald Choi	Phillip Baker
Katrina Tsai	Amanda Jensen	Dennis Lin	Vincent Peterson
Abigail Zhao	Angela Larson	Victor Kwon	Dennis Reed
Vanessa Choi	Hayley Thompson	Jason Pham	Frank Phillips
Patricia Li	Christine Campbell	Eric Duong	Shane Taylor
Lisa Zhou	Caroline Parker	Stephen Hsu	William Welch
Caroline Lu	Kristy Baker	Kevin Jiang	Bryan Ward
Melissa Hwang	Tina Reed	Jeffrey Chen	Ian Russell
Mary Pham	Sara Burke	Erik Luong	Brian Wilson
Amy Hu	Victoria Snyder	Philip Zhao	Seth Schwartz
Jenna Nguyen	Molly Weaver	Jeremy Yu	Jared Walsh
Margaret Liang	Melissa Stone	Seth Truong	Zachary Parker
Danielle Liu	Melanie Wilson	Ian Zhou	John Carpenter
Megan Dinh	Rachael Ward	Matthew Chang	Jeffery Cook
Melanie Yang	Elizabeth Miller	Scott Lu	Nathan Nelson
Amanda Thao	Mary Hill	Sean Hwang	Matthew Rogers
Sarah Yu	Amy Moore	Patrick Hu	George Barker
Nichole Liu	Vanessa Smith	Mark Chan	Sean Beck
Christine Cho	Teresa Anderson	Jack Zhu	David Hall
Victoria Xiong	Catherine Schultz	Timothy Cheng	Andrew Miller
Teresa Wong	Heather Martin	Benjamin Nguyen	Peter Keller
Kara Yoon	Kathryn Myers	Steven Tang	Luke Jensen

Continued

Asian Female	**White Female**	**Asian Male**	**White Male**
Kathleen Cheng	Katie Meyer	Travis Wong	Kevin Hansen
Angela Wu	Valerie Price	David Zheng	Dustin Sullivan
Catherine Zheng	Melinda Evans	Paul Ngo	Philip Morris
Hayley Huang	Sandra Wright	Anthony Yi	Evan Moore
Karen Ngo	Christina Russell	Shane Huang	Paul Burke
Elizabeth Duong	Kayla Allen	Robert Zhang	Matt Price
Laura Luu	Jacqueline Schmidt	Kenneth Tsai	Marcus Collins
Rebecca Hsu	Jennifer Welch	Richard Xiong	Richard Thompson
Melinda Zhang	Michelle Nelson	Brian Cho	Thomas Snyder
Katherine Le	Sarah Fisher	Joel Le	Christopher Larson
Tara Jiang	Brittany Rogers	Michael Li	Travis Gray
Alicia Zhu	Grace Keller	Trevor Cheung	Charles Hoffman
Molly Huynh	Julie Beck	Adam Liu	Joel Stone
Samantha Tang	Monica Cooper	Peter Wu	Joseph Allen

Notes. This table provides the name lists of hypothetical startup founders used in the survey tool. 50 names were selected to be highly indicative of each combination of race and gender. Considering the White and Asian startup founders account for most of the highly innovative startups, we only have four combinations listed above: Asian Female, White Female, Asian Male, White Male. A name drawn from these lists was displayed at the beginning part of the startup profiles and in the questions used to evaluate the resumes. First and last names were linked every time as they appeared and the combinations of first and last names are randomly generated. Considering that Asian and White Americans have very similar naming patterns as documented by Fryer Jr and Levitt (2004) I choose their first names from the same name pool. After I generated a list of potential full name candidates, we further checked these names to make sure that there are no names owned by famous startup founders or CEOs.

Table B2: **Education Background (School List)**

School Category	Universities	Percentage
(Top School) Example	Brown University	50%
	Columbia University	
	Cornell University	
	Dartmouth College	
	Harvard University	
	Princeton University	
	University of Pennsylvania	
	Yale University	
	California Institute of Technology	
	MIT	
	Northwestern University	
	Stanford University	
	University of Chicago	
(Common School) Example	Thomas Jefferson University(153)	50%
	University of Arkansas(153)	
	Hofstra University(162)	
	University of Mississippi (162)	
	Virginia Commonwealth University (162)	
	Adelphi University (166)	
	University of Maryland-Baltimore County(166)	
	University of Rhode Island(166)	
	St.John's University (179)	
	University of Detroit Mercy (179)	
	University of Idaho (179)	
	Biola University (185)	
	Chatham University (185)	
	Bellarmine University (197)	
	Bethel University (197)	
	Loyola University New Orleans (197)	
	Robert Morris University (202)	
	Regis University (202)	
	Widener University(202)	
	Laurentian University (Canada)	
	Auburn University (104)	
	Rochester Institute of Technology (104)	
	University of Tulsa (121)	
	DePaul University (125)	

Note: This table provides the school list used to generate the education background of each hypothetical startup founder. The percentage of top school and common school is 50% VS 50% to increase the power. Also for highly innovative startups, their founders are more likely to graduate from prestigious universities. Top schools refer to the Ivy League schools (Brown University, Columbia University, Cornell University, Dartmouth College, Harvard University, Princeton University, University of Pennsylvania, and Yale University) as well as other top U.S. schools (Amherst College, California Institute of Technology, Duke University, MIT, Northwestern University, Stanford University, University of California, Berkeley, University of Chicago, and Williams College). Considering the incubators we collaborate with have more connections with Columbia University and Stanford University, we give more weight on these universities. Common Schools are those who are ranked lower than the 150th based on the US News 2020 ranking results and I add a Canadian common school since one of the incubators is from Canada.

Table B3: **Company Comparative Advantage**

Advantage Category	Description
(Product)	trade secrets/patents registered
	celebrity endorsement
	exclusive partnerships
	accumulated many pilot consumers
	adoption of the latest technology
	pricing advantage
	great product design
	1st mover
(Cost)	lower cost
	economies of scale
Total	100%

Notes: I use the number of the corresponding comparative advantages as a measure of the quality of the startup project. For each startup profile, the subset of comparative advantages is randomly drawn from the 10 advantages listed above.

Startup Team Evaluation Section

Instructions:

All 16 startup teams are hypothetical and randomly generated. However, we will help you find real high-quality startup teams, which have connections with our collaborative incubators, based on your choices and ratings in this survey. The matched startup teams will contact you after 1 month.

We will use all evaluation answers to recommend highly matched startup teams from our collaborative incubators. All data will be kept strictly confidential and analyzed at the aggregate level after removing identifiable information.

Note:

1. **Assume that all the hypothetical startups work in the industry (or industries) and stage(s) of your interest and that all startup teams have adequate knowledge of the industry.**
2. **The more carefully and truthfully you evaluate each startup profile, the more benefits you can get.**

Back Next

Figure B1: Instruction Page (Version 2)

Nano-Search Financing Tool

Finding Startups, Your Way

Startup 1

Founding Team

Founder	Samantha Tang (graduated from Bellarmine University in 2004)
Previous Experience	Yes, the team has at least one serial entrepreneur.
Founded date	2018

Project Description

Competitive advantage	Accumulated many pilot consumers, 1st mover, Great product design
Traction	Previous Monthly Revenue: $9K, Annual Revenue Growth Rate: 42%

Additional Information

Company Category	B2C
Number of Employees	10-20
Target Market	Domestic Market
Mission	For profit
Location	U.S.
Number of Existing Investors	3 or more

*Assume that all the hypothetical startups work in the industry (or industries) and stage(s) of your interest.

Figure B2: Randomly Generated Startup Profile

1. Imagine that Jeffrey Chen and David Zheng's team is guaranteed to accept your investment offer. Compared with firms you have previously invested in, which percentile do you feel this startup belongs to considering its quality?

Extremely Low Quality — Extremely High Quality

0 10 20 30 40 50 60 70 80 90 100

Probability of Generating Higher Return (**Drag the bar**)

2. Considering the potential network and negotiation power of Jeffrey Chen and David Zheng's startup team, what's the probability that this startup team will accept your investment offer rather than that of another investor (Angel, VC, Loans, etc)?

Guaranteed Rejection — Guaranteed Acceptance

0 10 20 30 40 50 60 70 80 90 100

Probability of Accepting Your Offer (**Drag the bar**)

3. If you consider both the team's attractiveness and their likelihood of collaboration, how likely would you be to ask for their contact information or pitch deck?

Will Not Ask — Will Ask

0 10 20 30 40 50 60 70 80 90 100

Probability of Asking for More Information (**Drag the bar**)

Figure B3: Evaluation Questions (Part 1)

4. Considering both the team's attractiveness and their likelihood of collaboration, how much money would you invest in this startup compared to your average investment amount? Imagine that the startup asks for the amount of money that you can afford.

(For example, if your average amount of investment per deal is $1M and you would invest $0.5M to the team, drag the bar to 0.5.)

Benchmark Investment

0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 >2.0

Relative Preferred Investment Amount (**Drag the bar**)

5. Compared with your previous invested startups, which percentile do you feel this startup belongs to considering its risk level (i.e. the level of uncertainty of achieving the expected finance returns)?

No risk Highest risk

0 10 20 30 40 50 60 70 80 90 100

Risk Level (**Drag the bar**)

Back Next

Figure B4: Evaluation Questions (Part 2)

Dear [Investor Name],

Our research team learned about your startup investment experience from Pitchbook and would like to invite you to participate in a research project conducted by the Columbia University Economics Department. Given your expertise in the startup investment, your insight would be indispensable to our research, which we hope would shed light on the entrepreneurial financing process in the U.S. and help the recovery of entrepreneurial activities from recession.

The research project is supervised by Prof. Jack Willis and led by a Columbia Economics Ph.D. student, Ye (Iris) Zhang, who is collaborating with Hash Outliers and the En Lab. The purpose of the project is to understand the current entrepreneurial financing process (for example, investors' preferences for future collaborative startups) and remove the frictions typically found in the fund-raising process using the matching algorithms we have developed. We have developed a matching tool (the "Nano-Search Financing Tool") that can match investors with the best fit startup teams.

Using the tool takes about 20 minutes and involves evaluating 16 hypothetical startup profiles in your invested industry. After evaluating these profiles, the tool uses a newly developed machine-learning algorithm to identify startups who could be a good fit for your investment portfolios from our collaborative incubators. The matched startup teams will try to contact you after 1 month.

Besides the potential investment and collaboration opportunities, we will offer a lucky draw opportunity to thank you for your support of this research project. At the end of July 2020, we will randomly pick 2 survey participants and inform them of the lucky draw results. These 2 participants will be paid in July 2021 according to the startup quality evaluation results they made in the financing tool (that is, the $500 and the extra return based on their quality evaluation results). Details are described on the instruction page and consent form in the matching tool.

To access the tool, please click the link; we have also attached the instruction poster for its use.

Our research team will also use a completely anonymized version of your data to research broader trends in what investors value when investing in startups. We will be glad to share these insights with you when the research is complete.

If you have any questions or would like more detailed information about how the tool will enhance your portfolio construction process, please contact the tool developer and project investigator, Ye (Iris) Zhang (yz2865@columbia.edu).

Thank you very much and have a nice day!

Sincerely,
Ye

--
Ye Zhang
Ph.D. Candidate
Economics Department, Columbia University
Email: yz2865@columbia.edu

Figure B5: Recruitment Email (Version 1)

Notes. Version 1 provides both matching incentive and monetary incentive to randomly selected 11183 U.S. venture capitalists.

Dear [Investor Name],

Our research team learned about your startup investment experience from Pitchbook and would like to invite you to participate in a research project conducted by the Columbia University Economics Department. Given your expertise in the startup investment, your insight would be indispensable to our research, which we hope would shed light on the entrepreneurial financing process in the U.S. and help the recovery of entrepreneurial activities from recession.

The research project is supervised by Prof. Jack Willis and led by a Columbia Economics Ph.D. student, Ye (Iris) Zhang, who is collaborating with Hash Outliers and the En Lab. The purpose of the project is to understand the current entrepreneurial financing process (for example, investors' preferences for future collaborative startups) and remove the frictions typically found in the fund-raising process using the matching algorithms we have developed. We have developed a matching tool (the "Nano-Search Financing Tool") that can match investors with the best fit startup teams.

Using the tool takes about 20 minutes and involves evaluating 16 hypothetical startup profiles in your invested industry. After evaluating these profiles, the tool uses a newly developed machine-learning algorithm to identify startups who could be a good fit for your investment portfolios from our collaborative incubators. The matched startup teams will try to contact you after 1 month.

To access the tool, please click the link; we have also attached the instruction poster for its use.

Our research team will also use a completely anonymized version of your data to research broader trends in what investors value when investing in startups. We will be glad to share these insights with you when the research is complete.

If you have any questions or would like more detailed information about how the tool will enhance your portfolio construction process, please contact the tool developer and project investigator, Ye (Iris) Zhang (yz2865@columbia.edu).

Thank you very much and have a nice day!

Sincerely,
Ye

--
Ye Zhang
Ph.D. Candidate
Economics Department, Columbia University
Email: yz2865@columbia.edu

Figure B6: Recruitment Email (Version 2)

Notes. Version 2 provides only matching incentive to randomly selected 4000 U.S. venture capitalists.

Figure B7: Recruitment Poster (Version 1)

Notes. Version 1 provides both matching incentive and monetary incentive to randomly selected 11183 U.S. venture capitalists.

Figure B8: Recruitment Poster (Version 2)

Notes. Version 2 provides only matching incentive to randomly selected 4000 U.S. venture capitalists.

C Correspondence Test

C.1 Name Generation Process

I generate a list of names that are highly indicative of race (Asian or white) and gender (male or female), combining the approach of Fryer Jr and Levitt (2004) and Gornall and Strebulaev (2020). I used Social Security Administration (SSA) dataset,[61] birth records for selecting first names highly indicating gender and used 2010 U.S. Census data for generating last names highly indicative of race.[62] The full lists of names are given in Appendix C.1 Table C1.[63] The following describes the detailed steps of generating these names.

First Names:

Step1: I started with the first names from the Social Security Administration (SSA) dataset of male and female baby names in the U.S. Common names are chosen to mitigate the concern that a distinctively ethnic first name can convey other information besides gender. For example, such confounding information can be social status and economic background of the person (Bertrand and Mullainathan (2004)). Considering that the naming pattern of Asian and white is very similar (Fryer Jr and Levitt (2004)), I did not select indicative first names within an ethnic group.

Step2: To avoid gender ambiguity, I did the following additional checks. First, I removed ambiguous names, which is defined as names that were in both the top 1,000 male and top 1,000 female lists with a difference in the frequency of less than 200,000 times.[64] Then I pick the most frequent 100 names for each gender after deleting the gender ambiguous names for further checks.[65]

Second, to remove names that might be perceived as Hispanic or Jewish, we manually checked each potential candidate name and its origin, keep all the popular Christian names and remove names whose origin is mainly from Jewish (countries like Spain, Portugal, or Israel).[66] I further removed names that were strongly indicative of religion (such as Moshe).

Last Names:

I followed exactly the method of Gornall and Strebulaev (2020) by starting with the most common 1,000 last names in the 2010 U.S. Census data. The white-sounding last names were the 50 most common last names that were more than 85% white

[61] The SSA dataset is available at https://www.ssa.gov/OACT/babynames/limits.html, accessed on July 27, 2019.

[62] 2010 Census surnames product:https://www.census.gov/topics/population/genealogy/data/2010_surnames.html accessed on July 27, 2019

[63] Birth Statistical Master File: https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm, accessed on July 27, 2019

[64] From the histogram of the frequency, we see the majority (74%) of the difference is lower than this number; to be conservative, I choose 200,000 to avoid gender ambiguity.

[65] **An alternative method** is to construct an index for each name of how distinctively the name was associated with a particular race and gender following Fryer Jr and Levitt (2004). Female name index (FNI) is constructed in the whole sample from SSA data and defined as follows.

$$FNI_{name,t} = \frac{Pr(name|Female,t)}{Pr(name|Female,t) + Pr(name|Male,t)} * 100$$

For selecting female names, I set the cutoff as 99 and keep all the names whose FNI is greater than 99. Among these names, I choose the most frequently used 100 names for female; For selecting male names, I set the cutoff as 3 and keep all the names whose FNI is less than 3. Among these names, I choose the most frequently used 100 names for male. I choose asymmetric cutoffs for female and male FNI due to the fact that the number of male names in the U.S. are much less than the number of female names. This method balances the name popularity and also the gender unambiguity.

[66] For Gornall and Strebulaev (2019), they used the name list published by Jorg Michael and removed names that were gender ambiguous in the United Kingdom and as popular in Spain, Portugal, or Israel as in the United Kingdom. We do not feel popular names in these countries are necessary religious and considering the size of our potential names, manually check is feasible here.

and less than 3% Hispanic. The Asian-sounding last names were all 26 last names on the most common list that were more than 85% Asian. I deleted the surnames which do not show up in venture capital investors' names. For each selected last name, I searched the key word "last name venture capital investor " or "last name angel investor " on Google and LinkedIn. If there is no investor with this such last name shown up, I choose to delete this surname from the name list. I also remove certain very religious last names. This removed some last names like "Kaur, Vang".[67]

Additional Check:

I also hired 107 Amazon Mechanical Turk users in the U.S. to confirm that the perception of gender and race elicited by these names was in line with demographic data. For both first names and last names, I excluded any names that were not correctly classified more than 90% of the time. If the number of first names and last names left are less than 50 each, I would duplicate the process to add names on the waiting list.

After generating names indicative of gender and race each, I randomly paired first names and last names to generate a list of full names assuming that last names do not convey enough information of gender. I selected 50 names for each race-gender combination for randomization. Names of hypothetical female investors are shown in Table C1; names of hypothetical male investors are shown in Table C2.

To prevent the generated investor names being associated with famous investor names, I searched LinkedIn and our combined investor database (Pitchbook database and other hand-collected data) to ensure that there were no real investors who had the same name and matched the key details in the profile. If a conflict was found, I deleted the full name and add a new name from the waiting list.

Gender and Race is randomized independently. The corresponding names used for each hypothetical startup for both rounds of correspondence test are provided in Table C3.

[67] An alternative method is to construct white name index (WNI) and Asian name index (ANI) following Fryer Jr and Levitt (2004), which is defined as follows.

$$
\begin{aligned}
WNI_{surname,t} &= \frac{Pr(surname|White,t)}{Pr(surname|White,t)+Pr(surname|Non-White,t)} * 100 \\
Pr(surname|White) &= \frac{Pr(surname,white)}{Pr(white)} = \frac{Pr(white|surname)\times Pr(surname)}{Pr(white)} \\
ANI_{surname,t} &= \frac{Pr(surname|Asian,t)}{Pr(surname|Asian,t)+Pr(surname|Non-Asian,t)} * 100
\end{aligned}
$$

I implemented similar checks to first names and require that the last name make up at least 0.1% of that race's population , to ensure that last names were sufficiently common.

Table C1: First Names Populating Profile Tool

Panel A: Female						
Jennifer	Elizabeth	Lisa	Laura	Megan	Emily	Erica
Natalie	Jacqueline	Victoria	Melanie	Tina	Kayla	Kristy
Melinda	Linda	Theresa	Kara	Amanda	Sarah	Amy
Angela	Christina	Rebecca	Tiffany	Mary	Brittany	Samantha
Katherine	Alicia	Monica	Kathryn	Patricia	Anna	Catherine
Veronica	Kathleen	Sandra	Cassandra	Valerie	Amber	Teresa
Allison	Amber	Katrina	Jenna	Megan	Jessica	Melissa
Nicole	Sara	Julie	Christine	Tara	Katie	
(Extra)						
Abigail	Danielle	Michelle	Rachael	Brenda	Margaret	Amanada
Hayley	Madeline	Molly	Vanessa	Rachael	Grace	Heather
Cynthia	Caroline	Karen				

Panel B: Male						
Robert	Brian	Kevin	Steven	Thomas	Adam	Patrick
Bryan	Keith	Donald	Peter	Jared	Phillip	Jeffery
Victor	Seth	Alan	Matt	David	Jason	John
William	Andrew	Justin	Anthony	Jonathan	Timothy	Nicholas
Jeremy	Richard	Jeffrey	Benjamin	Paul	Stephen	Nathan
Jacob	Gregory	Travis	Kenneth	Samuel	Edward	Derek
Ronald	Joel	Frank	Dennis	Erik	Philip	Christopher
James	Mark	Scott	Dustin	Zachary	Marcus	Gary
(Extra)						
Vincent	Jack	Luke	Michael	Evan	Joseph	Eric
Shane	Sean	Matthew	Ian	George	Trevor	Charles

Notes. All listed first names which are indicative of gender is used for both the correspondence test experiment and also the lab-in-field experiment. For the correspondence test, these names will be used to create fictitious startup founder's names. For the lab-in-field experiment, these names will serve as the hypothetical names of startup founders. It covers the popular first names of people who are between 24 years old and 45 years old. To make sure all the names are only indicative of gender, I hired 107 Amazon Mechanical Turk to classify potential names into different genders and provide their feedback on whether these names remind them of other information besides gender (e.g. economic background, race, immigration status, etc). For the all the selected names listed above, more than 98% of Amazon Mechanical Turk correctly classified the names into the corresponding gender. I also delete the names which are indicative of other information. For example, "Chelsea" was deleted because some M-turks feel it is associated with upper-class; "Luis", "Carlos" or "Antonio" were deleted because they are perceived more likely to be Hispanic. I also added the first names and last names used in Gornall and Strebulaev (2020) in the "extra" part.

Table C2: **Last Names Populating Profile Tool**

Panel A: Asian				
Yu	Zhao	Zhang	Jiang	Hwang
Huynh	Luong	Cheung	Hsu	Liang
Li	Hu	Xu	Zhu	Huang
Yang	Kwon	Choi	Nguyen	Pham
Hoang	Luu	Liu	Lu	Chen
Lin	Chang	Chung	Zheng	Xiong
Zhou	Ngo	Truong	Wu	Duong
Cho	Cheng	Yi	Dinh	Tang
Wong	Chan	Ho	Thao	Tsai
Le	Yoon	Wang		
Panel B: White				
Nelson	Russell	Roberts	Rogers	Adams
Cooper	Wright	Cox	Kelly	Phillips
Bennett	Bailey	Collins	Thompson	Stewart
Parker	Evans	Allen	Martin	Anderson
Clark	Campbell	Morris	Reed	Wilson
White	Taylor	Sullivan	Myers	Peterson
Murphy	Fisher	Cook	Hughes	Price
Gray	Moore	Hill	Baker	Hall
Smith	Miller	Ward		
(Extra)				
Hansen	Welch	Hoffman	Meyer	Schmidt
Burke	Beck	Walsh	Carpenter	Schultz
Jensen	Keller	Snyder	Stone	Cohen
Barker	Becker	Schwartz	Larson	Weaver
Carroll				

Note: The table contains selected last names indicating ethnicity identity for hypothetical startup founders. I first create a list of candidate last names combining the results from Method I and the last name list from Kessler et al. (2019). To make sure all the names are only indicative of race and perceived correctly by people, I further hired 107 Amazon Mechanical Turk to classify potential names into different races and provide their feeds on whether these names remind them of other information besides race (e.g. economic background, immigration status ,etc). For the all the selected last names listed above, more than 95% of Amazon Mechanical Turk correctly classified the Asian last names into the corresponding race and more than 92% of Amazon Mechanical Turk correctly classified the white last names . I delete all the ambiguous last names, for example, "Shah" was deleted because many M-turks feel it can also be the Middle-eastern names; "Patel" was deleted because they feel it is an Indian names and may not be perceived as typical Asian names; "Long" was deleted because it can serve as both White and Asian names. I also delete last names that are related to religion or very rare in venture capital industry, like "Kaur" and "Vang". I also added the first names and last names used in Gornall and Strebulaev (2020) in the "extra" part.

Table C3: Startup and Entrepreneur Names Used

Panel A: the 1st round

Startup Names	White Female	Asian Female	White Male	Asian Male
VoiceFocus	Kathleen Jensen	Kathleen Yi	Joseph Adams	Kevin Truong
Light Run	Lisa Thompson	Stephanie Lu	Vincent Snyder	Jeffrey Luong
Instrument Tell	Molly Weaver	Jennifer Dinh	Sean Miller	Justin Huang
Sign Reader	Megan Schwartz	Valerie Yu	Evan Meyer	Shane Chan
Bross	Catherine Welch	Rachael Pham	Eric Burke	Ryan Le
Chicky	Rachael Smith	Vanessa Zhu	Robert Reed	Trevor Thao
LoopuDeck	Mary Meyer	Melissa Liu	George Price	Vincent Xu
EasySample	Melissa Larson	Catherine Yang	Matthew Russell	Ian Zheng
YouTubys	Grace Clark	Christine Tang	Justin Hansen	Bryan Hu
OSS	Veronica Russell	Emily Thao	Shane Snyder	Luke Zhao
CPRX	Danielle Cook	Margaret Dinh	Scott Parker	Eric Pham
All-in	Julie Barker	Karen Wong	Marcus Becker	Derek Yoon
SkatED	Kathryn Beck	Abigail Chang	Andrew Moore	George Cheng
GeniusPlot	Christina Parker	Katie Kwon	David Sullivan	Marcus Wang
EasyTry-On	Katherine Snyder	Angela Ho	Richard Cook	Mark Chung
Krysco	Valerie Baker	Amanda Jiang	Patrick Ward	Kevin Hoang
Lens Bioimage Technology	Emily Bennett	Erica Zhou	Adam Hoffman	Peter Cheung
Medprint	Jacqueline Hughes	Patricia Yoon	Ian Cooper	Brian Dinh
BM International	Vanessa Phillips	Mary Luu	Edward Keller	Jack Luu
Vet Technology	Michelle Gray	Natalie Hwang	Jeremy Carroll	Michael Wu
Freight Future	Amanda Meyer	Danielle Cheng	Christopher Cohen	Edward Lin
AfroLab	Madeline Hill	Nicole Xu	Steven Collins	Stephen Liu
SmartTeacher	Jessica Evans	Melanie Ngo	William Welch	Jason Chung
CleanPlanet	Christine Fisher	Megan Liang	Jeffrey Barker	Nicholas Lu
FancyTravel	Melanie Schultz	Rebecca Zhao	Ryan Schwartz	Sean Xiong
MeSafeMicro	Cynthia Keller	Allison Duong	Samuel Kelly	Samuel Ngo
Talently	Caroline Stone	Heather Zhang	Jack Moore	Richard Thao
AgriSoft	Rebecca Miller	Katherine Truong	Gregory Morris	Jonathan Duong
EduPar	Erica White	Caroline Chung	Derek Jensen	Jeremy Jiang
Milkless	Hayley Becker	Christina Hsu	Luke Thompson	William Hwang
Durabuddy	Brenda Bailey	Madeline Tsai	Brian Reed	James Le
Constructech	Samantha Peterson	Samantha Le	Michael Myers	Patrick Nguyen
SolarWat	Patricia Stewart	Brenda Hoang	Thomas Beck	Christopher Huynh

Note: 33 startups were created for the first round experiment, which was implemented between 2020/03-2020/04. All the startup founders' names are randomly generated using the commonly used first names and last names in the U.S. To prevent the fictitious startup founders being associated with real people, I searched LinkedIn, Google and available university directories to make sure that no real students from the corresponding universities are using the same names. If a conflict was discovered, I will replace the conflicted names with other randomly generated names to avoid such situation. Information of startups used in the later round correspondence test will be updated in the next version of draft.

Table C4: **Trace Investors' Email Behaviors**

This table provides detailed mechanisms of recording different email behaviors, the merits and limitation of each traced behavior measurements, and the previous correspondence tests in the literature that used similar participants' behaviors. To realize these functions, I used the Mailgun platform, which is a professionally designed platform for large email campaign activities founded in 2010.

Email Behaviors	Behavior Tracing Mechanisms	Merits	Limitations	Literature
1. Email Opening Rate (time stamp)	Write each pitch email using HTML with a unique ID and insert an one-pixel invisible transparent picture into the email. If the picture is downloaded from the server, I assume the investor opened the pitch email when the picture was downloaded	Increase the experiment power (high opening rate); only affected by the email subject line rather than the email contents	Noisy measurements (Some remote servers prevent users from downloading a picture while others automatically download a picture for their users. However, such server property is unrelated to the experiment treatment.)	
2. Email Reading Time (time stamp)	Write each pitch email using HTML with a unique ID and insert a large invisible transparent picture (i.e. 500 MB) into the email. Set the speed of downloading the picture from our server to 10KB/s. If only 200KB is download from the server, then the email staying time is 20s.	A continuous variable which measures the attention; Increase the experiment power;	Noisy measurements (Researchers cannot observe directly whether inventors are reading the email or is simply leaving the email open while having lunch.)	
3. Multiple Email Opening Rate	If the one-pixel transparent picture inserted in the pitch email is downloaded multiple times as recorded in the server, then I assume the email is opened multiple times. This happens if the same investor opens the email multiple times or the email is forwarded to others who open it later.	Increase the experiment power; a stronger indicator of investors' interest	Noisy measurement. Researchers cannot differentiate whether the email is opened multiple times by the same investor or the email is forwarded to others.	
4. Sentimental Analysis of Replied Emails	Use LIWC to analyze the sentiment of the content of each replied email. I used the following website which automatically generate analyzed results: http://liwc.wpengine.com/	Relatively objective measurement of the investors' attitudes towards each pitch emails	Low response rate during the recession, hence low experiment power	Hong and Liskovich (2015)
5. Website Click Rate	The Mailgun platform developed this function and researchers can use it directly. Click here for mechanism explanations provided by Mailgun.	Can be used when investors do not reply to the email	Low website click rate in the entrepreneurial financing setting	Bartoš et al. (2016); Bernstein et al. (2017)
6. Email Response Rate & Replied Contents	Collected directly from the inbox and spam box	Commonly used callback measurements	Low response rate; The replied contents may not represent true interest if investors try to be politically correct.	Gornall and Strebulaev (2020), etc.