# Gender, Race, and Entrepreneurship: A Randomized Field Experiment on Venture Capitalists and Angels*

**Will Gornall**

University of British Columbia Sauder School of Business

**Ilya A. Strebulaev**

Graduate School of Business, Stanford University

and

NBER

August 2019

**Abstract**

We study gender and race in high-impact entrepreneurship using a tightly controlled randomized field experiment. We sent out 80,000 pitch emails introducing promising but fictitious start-ups to 28,000 venture capitalists and angels. Each email was sent by a fictitious entrepreneur with a randomly selected gender (male or female) and race (Asian or White). Female entrepreneurs received a 9% higher rate of interested replies than male entrepreneurs pitching identical projects and Asian entrepreneurs received a 6% higher rate than their White counterparts. Our results suggest that investors do not discriminate against female or Asian entrepreneurs when evaluating unsolicited pitch emails.

*We thank Tania Babina, Michael Ewens, Yael Hochberg, Sabrina Howell, Kai Li, Song Ma, Richard Townsend, and participants at the London Business School Private Equity conference, National Bureau of Economic Research Summer Institute, Private Equity Research Consortium, Rome Junior Finance Conference, the Stanford Financing of Innovation Summit, and the University of British Columbia and Stanford Graduate School of Business research seminars for comments and suggestions on the project; Stella Chen, Michelle Gan, Marco Soohyeong Lee, Yingxiang Li, Justin Norden, Alexander Qu, Ashutosh Synghal, Eric Tolocka, Brian Triana, and Andy Wardle for excellent research assistance, including writing pitches; numerous VC industry practitioners for helping us develop pitches and providing helpful guidance; the Stanford GSB Venture Capital Initiative and the SSHRC for financial support; and most importantly, we thank all of the receivers of our pitch emails. This study is registered in the AEA RCT Registry, with the unique identifying number of AEARCTR-0003277. Both authors have consulted for general partners and/or limited partners investing in venture capital. Gornall: will.gornall@sauder.ubc.ca; Strebulaev: istrebulaev@stanford.edu.

# 1 Introduction

By almost any measure, women are starkly underrepresented in high-impact entrepreneurship. For example, in 2017, for every dollar invested in start-ups with female founding teams, there was 35 dollars invested in start-ups with male founding teams. As a whole, there were 16 times as many male-founded start-ups as female-founded start-ups.[1] By these and many other measures, high-impact entrepreneurship is one of the most male-dominated sectors of the economy: females are substantially less likely than males to found the innovation-driven companies that are critical to job creation, economic growth, and prosperity around the world. The prevailing wisdom is that biased investors play a significant role in that. This belief is bolstered by the gender composition of the investor body. Females make up just 9% of the thousands of U.S. venture capitalists (VCs) and only about 10% of new hires in the VC industry; and 79% of VC firms have never had a female senior investment professional.[2] Further buttressing that belief is a stream of sexual discrimination and harassment scandals coming from VC firms and the innovative companies they back.[3]

Studying gender discrimination in entrepreneurship is therefore of critical importance. Previous studies of gender bias in entrepreneurship have shown mixed results. Their biggest challenge has been controlling for characteristics that are unobservable to researchers but that are observable to decision-makers. Specifically, investors in start-ups have immeasurably more information than researchers about the companies and entrepreneurs they evaluate. This unobserved variation makes it difficult for researchers to establish causal relationships and measure discrimination.

In this paper, we address this challenge using a randomized field experiment that uses real investors and tightly controls the information they receive on the entrepreneurs and companies. Our empirical

[1]As reported by *Fortune*, based on data assembled by PitchBook. Retrieved November 23, 2018, from fortune.com/2018/01/31/female-founders-venture-capital-2017.

[2]See Gompers, Mukharlyamov, Weisburst, and Xuan (2014) and Gompers and Wang (2017), who augment VentureSource with hand-collected data. Similar numbers are widely reported by industry publications using PitchBook data. See, for example, techcrunch.com/2017/10/04/announcing-the-2017-update-to-the-crunchbase-women-in-venture-report, accessed November 23, 2018. Similar numbers are also reported for angel investors by Huang, Wu, Lee, Bao, Hudson, and Bolle (2017) and for other countries. See, e.g., british-business-bank.co.uk/wp-content/uploads/2019/02/British-Business-Bank-UK-Venture-Capital-and-Female-Founders-Report.pdf, accessed June 14, 2019.

[3]See, e.g., representative articles from 2017 in *The New York Times*, *The New Yorker*, and *The Atlantic* (all accessed November 24, 2018): nytimes.com/2017/06/30/technology/women-entrepreneurs-speak-out-sexual-harassment.html, newyorker.com/magazine/2017/11/20/the-tech-industrys-gender-discrimination-problem, theatlantic.com/magazine/archive/2017/04/why-is-silicon-valley-so-awful-to-women/517788.

design follows classical audit studies in labor economics.[4] We sent 80,000 email pitches from the founders of 50 fictitious start-ups to more than 28,000 VCs and angel investors. Each email was sent by a "founder" with a randomly assigned male or female name. We measure bias by comparing the response rate to the male and female names. Because the investors receiving the emails were unaware that they were taking part in a study, we get an accurate picture of whether they react differently to email messages from female entrepreneurs.

Our study benefits from the multi-stage investment process that is common in VC and angel investing. This multi-stage process is often thought of as a funnel, with introductions and a brief review of many start-ups at one end, in-person meetings and extended due diligence in the center, and investment in a select few start-ups at the other end. Very few start-ups make it through this process. Gompers, Gornall, Kaplan, and Strebulaev (2018) report that VCs invest in only 1% of the start-ups they consider. We concentrate on the very first step of this process, where entrepreneurs first connect with VCs and angels. This often takes the form of a pitch email, a short introductory communication that attempts to attract investor interest by presenting the main idea of a nascent company and some information about its founders. If, after having been pitched, investors are interested in the idea and the founding team, they will follow up to start a conversation about a potential investment. Past research shows that the perceived quality of the founders is crucial to sparking this interest.[5] Our emails are all at this first step, where entrepreneurs first reach out to potential investors.

This correspondence study design is frequently used by economists to test for discrimination in labor, housing, or product markets. Such studies typically use generic introductory messages or machine-generated resumes to assess interest. Our experimental intervention was more challenging to set up because actual start-ups work on innovative projects, thus we had to devise fictitious start-ups that might appeal to VCs and angels. Adding to the difficulty, these investors often specialize across industries, stages of investment, and geographies.

We composed 50 different email pitches covering the multitude of investment industries typical in the venture capital space. All of our pitches purported to offer a novel product or service that addressed

[4] Bertrand and Duflo (2017) provide an overview of this literature.

[5] Gompers et al. (2018) survey VCs and report that the founding team is the most important factor when deciding to invest. Huang et al. (2017) report similar results for angel investors. Bernstein, Korteweg, and Laws (2017) find that information on the founding team is the most important variable that triggers investor interest using an experiment on AngelList.

an existing problem, and did so without too closely resembling existing companies. These pitches were judged by professional investors as representative of the typical pitches coming from high-potential start-ups.

We sent out these pitches, randomly varying the entrepreneurial senders' demographics, between October 30 and November 16, 2018. Investors manually replied to 5,200 of our 80,000 emails, for a 6.53% response rate. We focus on the replies that clearly demonstrate investor interest to follow up with an entrepreneur by proposing a meeting or phone call, offering a referral to another investor, or, most commonly, requesting a pitch deck or other information. 3,400 of our pitch emails received interest replies, for a 4.3% interested reply rate.

Emails that were randomly assigned to be sent by a female entrepreneur had a 4.5% chance of getting an interested reply from investors, while emails assigned to a male entrepreneur had a 4.1% chance. Thus, female entrepreneurs received interested replies 0.4% more frequently than males in absolute terms and 9.5% more frequently in relative terms. This difference is statistically significant at the 1% level and can be attributed only to the randomized gender manipulation, as there are no other sources of variation. To the extent that gender perception affects investor decisions to follow up to an unsolicited email, our investor sample slightly favors female entrepreneurs; this preference persists for the wide variety of subsamples we examine.

The original impetus for our research is the stark underrepresentation of women in high-impact entrepreneurship and venture capital firms and the important question of how much of this underrepresentation is due to investor bias. Our surprising results lead us to largely rule out discrimination against females by investors at the initial contact stage of the investment process, insofar as this bias can be discerned from the reply rates to unsolicited email pitches. We emphasize that our findings only apply to email pitches—short emails introducing a start-up—and are silent on discrimination in the subsequent investment process. Thus, our results should not be interpreted as showing a lack of discrimination against females overall or at other steps of the investor decision-making process. We discuss this and other limitations of our experiment in Subsection 2.6.

The difference in interested reply rates between genders is strongly statistically significant. The extent of its economic significance depends in large part on what transpires at later stages of the investor decision-making process. In light of the substantial gender imbalance in real-world investments, one way to interpret our results is that a bias against female entrepreneurs materializes after the initial introductions, perhaps during in-person meetings. For example, Brooks, Huang, Kearney, and Murray (2014) show that experimental subjects react more favorably to a male voice dubbing a video of an entrepreneurial pitch than a female voice. Studying the evidence and causes of such bias in a randomized controlled setting should be the objective of further research.

Given the uniqueness of our empirical design among entrepreneurship studies, we also decided to test for any differences in reply rates due to race, which has been the main variable of interest in audit studies in labor economics. After long and thoughtful consideration, we discarded the idea of assigning African-American names to our fictitious entrepreneurs. First, there has been a concern about the ability to disentangle race and socioeconomic perception by using distinctly African-American names (Fryer and Levitt (2004)). Second, and most importantly, African-American entrepreneurs are woefully underrepresented in high-impact entrepreneurship (less than 1%, see Gompers and Wang (2017)). This poses a substantial challenge for our setup, because even names that are disproportionately likely to be African-American in the general population (such as "Martin Jackson") may not be perceived as African-American by investors in this context.[6] Studying discrimination against African Americans and other underrepresented minorities in high-impact entrepreneurship remains an important problem.

For our entrepreneurial email senders, we opted to use distinctively East-Asian (i.e., Asian) and Western-European (i.e., White) last names, both paired with common American first names that are distinctively female or male. Asians are overrepresented in the U.S. venture capital space, relative to their fraction of the U.S. population. Gompers and Wang (2017) report that 11% of U.S. entrepreneurs backed by venture capital firms are Asian, as are 11% of U.S. VCs, compared to 5% of the U.S. labor force. Despite this, past audit studies uniformly find strong discrimination against Asians (see Section 4 for more details). Correspondence studies that look at discrimination against both Asian and Black applicants often find that both groups face similar level of discrimination (e.g., Wood, Hales, Purdon, Sejersen, and Hayllar (2009); Booth, Leigh, and Varganova (2012); and Milkman, Akinola, and Chugh

[6]For example, Jackson is the most common majority-Black last name in the U.S. (53% according to 2000 U.S. Census Data) and was used by Bertrand and Mullainathan (2004) as one of their five names to suggest African American identity. Crunchbase, a platform actively used by entrepreneurs to seek very early stage investors, lists 34 Jacksons in the U.S. who have raised at least one round of start-up financing and have a profile picture. When we asked 100 users of Amazon Mechanical Turk to identify the race of these 34 people, only one was identified as Black.

(2012)). This bias extends to Asians who do not appear to be foreigners. For example, Oreopoulos (2011) finds that Asians with English first names, Canadian university degrees, and several years of Canadian work experience get 21% fewer callbacks than similar White Canadian job applicants.

While prior correspondence studies find significant discrimination against Asians, we find a mild preference for Asian entrepreneurs. Emails that were randomly assigned to be sent by an Asian entrepreneur had a 4.4% chance of getting an interested reply from investors, while those from a White entrepreneur had a 4.1% chance. This translates into Asian entrepreneurs receiving 6.1% more interested replies in relative terms, suggesting that investors have a marginal preference for Asian entrepreneurs in the initial contact stage of the investment process.

Our gender and race results are surprising in light of existing studies in other fields, widely held beliefs, and expectations. Our gender results could be driven by either statistical discrimination (being female is taken as a signal of quality) or taste-based discrimination (investors seek to support female entrepreneurs due to either personal values or external pressure). Our racial results appear more likely to be driven by statistical discrimination, as most taste-based stories work against Asians. Romantic or sexual interest by male investors does not appear to be the explanation for the preference for female entrepreneurs, as that preference is driven by requests for pitch decks, not requests for meetings; out-of-state investors show a similar preference for female entrepreneurs; and none of the replies received appear to suggest this motive. Homophily (e.g., female investors show preference to female entrepreneurs) could be part of the story but does not provide the entire explanation, as male investors show a preference for female entrepreneurs and White investors show a preference for Asian entrepreneurs. This calls for more rigorous research to understand decisions made by investors backing high-impact entrepreneurial companies at all stages of their investment process and the impact of those decisions on real outcomes.

Our study also addresses investor response to cold pitches. Our entrepreneurial email senders had an overall average 6.5% chance of a reply and a 4.3% chance of an interested reply. Many of the rejections were the result of poor matching between the start-up and the investor (such as a mismatch on geography or industry). In fact, of our 50 start-up pitches, the top decile got a response rate of 14% from investors overall and an 18% response rate from VCs. We also believe many of our emails did not pass through spam filters, without which the response rate likely would have been non-trivially higher.

These rates are substantially higher than the anecdotal evidence on cold calls would suggest. Taking this into account, an entrepreneur with a promising start-up and a well-structured and thoughtful pitch might get a 10% or higher rate of interested replies from investors. We take this as good news for budding entrepreneurs who lack strong networks and need financial backers and advisers.

Our study contributes to several strands of the literature. Social scientists have long studied and debated the evidence for and impact of discrimination by gender, race, and other characteristics, such as age and religion. In an important empirical study, Bertrand, Goldin, and Katz (2010) study the careers of male and female graduates of a prestigious business school. They show that their earnings diverge soon after graduation and provide suggestive evidence for three factors explaining the findings: differences in training prior to graduation, career interruptions, and weekly hours worked. Goldin and Rouse (2000) find discrimination against female musicians auditioning for orchestras by comparing the results of non-blind and blind auditions. Bagues and Esteve-Volart (2010) examine how the success of females on the Spanish judiciary entry exam was impacted by the gender composition of the evaluation committee. A causal study is made possible by the fact that people are randomly assigned to a committee. They find that females are less likely to succeed when their committee assignment has more females, while the opposite is true for male candidates.

Field experiments play an important part in furthering our understanding of discrimination (Bertrand and Duflo (2017) provide a literature review as well as detailed methodological discussion about challenges researchers encounter; Riach and Rich (2002) describe a number of earlier studies). One particularly powerful tool is audit or correspondence studies, such as ours, which are randomized controlled studies of discrimination by real decision-makers who are unaware they are being studied. In an early study, Neumark, Bank, and Van Nort (1996) study gender discrimination in restaurant hiring using fictitious male and female applicants. They find that females were discriminated against when they applied for server positions at expensive restaurants. Bertrand and Mullainathan (2004) find that resumes assigned distinctively White names receive substantially higher callback rates than those assigned distinctively African American names. Agan and Starr (2017) find that online job applications from people with Black-sounding names have a lower response rate after the introduction of policies which prohibited employers from asking about the criminal records of job applicants. Edelman, Luca, and Svirsky (2017) find that Airbnb guests with distinctively African American names are 16% less likely to be accepted by landlords than guests with White names. Milkman et al. (2012) sent emails from fictitious prospective doctoral students to college faculty requesting a meeting and find that females are less likely to get a response from private colleges and higher paying disciplines. They also find a substantially lower response rate for Asian senders, especially Asian females.

Our study also builds upon the entrepreneurship and venture capital literature. Recent studies in entrepreneurship explore the empirical evidence for gender imbalance. Gompers and Wang (2017) provide time series evidence of a substantial underrepresentation of females in high-impact entrepreneurship and VC. Robb and Robinson (2014) show that female entrepreneurs raise less outside debt and equity. Herbert (2018), using empirical data from France, shows that female entrepreneurs are at the aggregate level less likely to raise external equity including venture capital financing. Ewens and Townsend (2018) show that on AngelList, a matching platform for angel investors and start-ups, female-led start-ups experience more difficulty garnering interest and raising capital from male investors compared to male-led start-ups. Raina (2017) shows that female-founded start-ups are less successful and perform worse than male-founded start-ups, although the effect seems to attenuate when there is at least one female partner on the VC's investment team. Hellmann, Mostipan, and Vulkan (2019) find that female teams on an equity crowdfunding site seek lower valuations and face lower investment flows. Lee and Huang (2018) study judge evaluations at start-up competitions and find that female-led ventures were generally evaluated more negatively than male-led ventures. Alesina, Lotti, and Mistrulli (2013) find that female entrepreneurs seeking bank loans pay more for credit than male entrepreneurs.

Experimental settings in entrepreneurship are challenging to set up, and thus only a few studies report experimental evidence. Using an AngelList platform, Bernstein et al. (2017) conduct a randomized field experiment on angel investor preferences by varying the description of the start-up. They find that angel investors respond strongly to information about the founding team, but not to firm traction or existing lead investors. Brooks et al. (2014) use data from three U.S. pitch competitions to show that investors give lower ratings to the companies pitched by female entrepreneurs. In experiments to test the mechanism, subjects (who were not investors) watched two start-up video pitches in the field of veterinary technology: one with a male narrator and the other with a female narrator. 68% of subjects chose pitches dubbed by the male voice compared to 32% choosing a female voice, which suggests a bias against females at the in-person pitching stage. Tinkler, Whittington, Ku, and Davies (2015) asked business school students to evaluate start-up business plans that were randomly assigned to come from male or female entrepreneurs, with varying backgrounds and social ties. They find that technical female entrepreneurs are not disadvantaged as compared to their technical male counterparts, but that non-technical female entrepreneurs are evaluated less favorably than non-technical male entrepreneurs.

The rest of the paper proceeds as follows. In Section 2, we discuss our experimental setting. We present the baseline experimental results in Section 3. In Section 4, we compare our results to those from other correspondence studies. We discuss and test mechanisms for discrimination in Section 5. Concluding remarks are given in Section 6.

## 2 Experimental Methodology

### 2.1 Entrepreneur Email Senders

Our experiment relied on the investors who received our pitch emails being able to infer the gender and race of the email sender. We accomplished this by creating 200 different sender names. Each first name was a common and ethnically neutral American first name that was either masculine or feminine, to signal gender.[7] Each last name was either White-sounding or Asian-sounding, to signal race.

For first names, we started with the 1,000 most frequently used first names from the Social Security Administration dataset of male and female baby names in 1995.[8] To ensure that our selection of names does not suffer from even a slight degree of gender ambiguity, we performed a number of additional checks. First, we removed ambiguous names, which we 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 20 times. This removed names such as Taylor and Alexis. Next, we removed names that were not

[7] In entrepreneurship research, women who are entrepreneurs are usually referred to as female entrepreneurs. For this reason, we refer to man and woman genders as male and female throughout this study.

[8] The data are available at ssa.gov/cgi-bin/popularnames.cgi, accessed on September 10, 2018.

unambiguously of one gender in the United Kingdom, such as Tyler or Lauren, using the list published by Jörg Michael.[9] Prior audit studies using names as the experimental variation, such as Bertrand and Mullainathan (2004), struggle with the issue that distinctively ethnic first names convey social status and background in addition to ethnicity. To address this, we used the same list to remove names that were as popular in Spain, Portugal, or Israel as in the United Kingdom. This step removed names that might be perceived as Hispanic or Jewish, such as Alexandra or Maria.

For last names, we started with the most common 1,000 last names in the 2010 U.S. Census.[10] We considered last names White-sounding if at least 85% of the people with that last name were White and less than 3% were Hispanic. Similarly, we considered names Asian-sounding if at least 85% of the people with that last name were Asian. Based on those criteria, we took the 50 most common White-sounding last names and the 26 most common Asian-sounding last names.

This procedure produced a list of first and last names that were either female or male and either White or Asian with a very low degree of ambiguity according to actual demographic data. To confirm that the perception of gender and race elicited by these names was in line with demographic data, we recruited 100 U.S.-based Amazon Mechanical Turk users to assess the gender and race of our chosen names. The first names were correctly classified as male or female 98.9% of the time. For the last names, we excluded any last names that were not correctly classified more than 90% of the time. This resulted in the exclusion of 3 Asian names; in particular, South Asian names (e.g., Shah and Patel), as well as 15 predominantly-White names that respondents perceived as potentially African American (e.g., Powers and Carlson). After these removals, we had 23 Asian last names and 35 White last names that were correctly classified an average of 95.1% of the time. We duplicated names as needed to create a list of 50 Asian and 50 White last names.

We randomly paired the 50 male and 50 female first names and randomly paired Asian last names and White last names. Each pair of first names was randomly matched with a pair of last names, to give a quartet of four names representing the four possible conditions. For example, Adam was paired with Jennifer and Jensen was paired with Liu, which were then paired together giving us Adam Jensen (WM), Adam Liu (AM), Jennifer Jensen (WF), and Jennifer Liu (AF).

[9] Retrieved from ftp.heise.de/pub/ct/listings/0717-182.zip on September 9, 2018.

[10] The data are available at census.gov/topics/population/genealogy/data/2000_surnames.html, accessed Sept 9, 2018.

Each of these quartets was then matched to a particular pitch. Online Appendix A provides the list of names we employed. As we used common first and last names, there are many real people who have the same names as our fictitious entrepreneurs. To prevent our pitches being associated with real people, we searched LinkedIn and online university directories to ensure that there were no real people who had the same name and matched the key details in the pitch. If a conflict was found, we re-sorted name pairs to remove the conflicts while preserving the quartet form. For example, if we assigned the Adam-Jennifer-Jensen-Liu quartet to a pitch that was purported to come from a graduate student at a certain university, we ensured that there were no students at that university at the time of the experiment with any of those four names. If there were, we changed the pitch to a pitch that did not produce conflicts.

### 2.2 Pitch Emails

Our experimental design involved sending out email pitches to real VCs and angels. The challenge we faced was to create pitches that these investors would find appealing and would perceive as coming from actual start-ups. We ended up creating 50 pitches that covered all of the major industry classes, from energy to healthcare to information technology.

The purpose of our experiment was to test whether investors treated entrepreneurs differently based on their gender or race. Because we relied on investors inferring gender and race from the entrepreneur's name, we made it as salient as possible. All of the pitch emails identified the sender in the "From" field, in the sender's email address, in the body of the email, and at the close of the email. Figure 1 provides a sample pitch email. This structure is not uncommon for pitches by early-stage start-ups.

We needed investors to perceive these pitch emails as originating from real companies. To make our start-ups more believable, all of them were presented as very early-stage companies ("pre-seed," using industry jargon). That would explain the lack of any additional information on the start-up if an email recipient was to search for an online or offline presence. It is well known that very early stage start-ups are often in stealth mode and minimal information is available about them. We also had all of the founders purport to be graduate students at U.S. academic institutions, as practitioners told us that they expect less information would be available about students.

**Figure 1: Example Pitch Email.** This figure shows one of the 50 email pitches we sent to VCs and angels. We use John Smith in place of the name of a real recipient. The actual emailed pitch had the sender claim to be from a specific university, the name of the university is changed to <University Name> in the figure.

**To:** John Smith <John.Smith@topventurefund.com>
**From:** Samantha Huynh <shuynh@gemducate.com>
**Subject:** Investment in Gemducate Inc.

Hi John,

My name is Samantha Huynh, and I am currently studying computer science at <University Name>. I recently started a company, Gemducate, that is developing a technology and platform to revolutionize language learning.

Gemducate is a data-science driven education software that optimizes and accelerates language learning by analyzing users' daily communication patterns. The software acquires data from several digital communication methods such as online message, e-mail, note, and online post. Then it analyzes the data to get information such as high-frequency words, preferred sentence structure, favorite communication topics, etc. Gemducate then uses these insights to customize the learning experience, giving users the most efficient and fun way of learning new languages.

We are getting ready to raise our seed round to accelerate adoption and bring Gemstone to more users. If you are interested, I would be happy to share our pitch deck with you.

Thank you for your time.

Samantha Huynh
Gemducate

Our pitches were structured based on real pitches received by VCs and angels. The drafts of all the pitches were written by students at the Stanford Graduate School of Business, Stanford Medical School, and the University of British Columbia. Their task was to write email pitches describing innovative products or services that resolve an important problem in a way not too closely related to an existing product or service. All of our start-up companies were fictitious, and we verified through online searches that the company names we used were not used by similar companies.

A very low baseline response rate (unconditional on gender and race) would attenuate our statistical power. This was a key concern because experienced investors cautioned us that cold emails receive a low response rate. Thus, we attempted to make the pitches as appealing as possible. All of the proposed pitches were reviewed by practitioners for feedback and iteratively improved. In addition, we had our senders come from prestigious U.S. academic institutions, as practitioners and evidence from Bernstein et al. (2017) suggested this would improve our baseline response rate.

To assess the quality of our pitches, we had them rated by VCs, active angel investors, and more than 100 graduate students at the Stanford Graduate School of Business. We eliminated pitches that were poorly rated, which resulted in 50 well-reviewed pitches. The practitioners confirmed that our pitches were qualitatively similar to the pitches they receive every day and belonged to the subset of thoughtful, well-composed, and well-written pitches.

We also made simple websites for all of the companies.[11] A sample company website is shown in Online Appendix B. We did not create LinkedIn profiles or social media accounts for our entrepreneurs due to the biases potentially induced by pictures. We used some of these websites as sources for our emails, with email addresses of the form adamjensen@domain.com, adamj@domain.com, or ajensen@domain.com. Other emails were sent from the domain of an academic institution, with emails of the form adamljensen1@university.edu.

We ensured our pitches did not overweight areas traditionally associated with one gender. Research (Petit, 2007; Booth and Leigh, 2010; Carlsson, 2011; Herbert, 2018; Lee and Huang, 2018) suggests that females face larger hurdles in areas that are often perceived as masculine. That effect could bias our results if our start-ups were systematically in areas perceived as masculine or feminine. To alleviate this concern, we measured the perceived femininity of our pitches (without names or university affiliation) and of the descriptions of a random sample of 50 VC-backed start-ups from Dow Jones VentureSource. We asked Amazon Mechanical Turk users in the U.S. to rate the founder, employees, and customers of these companies on a scale from 1 = definitely male to 5 = definitely female. Our pitches were rated with average scores of 2.55, 2.80, and 2.48 for founders, employees, and customers, respectively, while the VentureSource pitches had nearly identical average ratings at 2.58, 2.78, and 2.53. This suggests our pitches are not systematically perceived as more feminine or more masculine than the typical start-up.

[11]Some initial emails included company logo images that tracked email opens. Because the emails with tracking images were disproportionately blocked by spam filters, we stopped sending them and 97.5% of our sent emails are text-only emails that do not track email opens.

### 2.3 Email Recipients

We sent our pitches to a diverse group of VCs and angels. Our list of investors to target was assembled from several databases of investors. We restricted our list to investors interested in U.S. companies: VCs at U.S.-based firms and angels who had a U.S. location or made a U.S. investment. To be included in the list, VCs had to be current members of an institutional or a corporate VC firm. For angel investors, we removed all the individuals for whom we could not find at least one recorded investment. If an investor was featured in both VC and angel lists, we assigned that investor a VC identity.

We gathered email addresses for as many investors as possible. If the databases we used included investor contact emails, we used those. Otherwise, we hired a lead generation service to retrieve the email addresses of the target investors. We discarded investors for whom we could not locate an email address. Because we combined databases, we ended up with multiple records for some investors. We dealt with that by discarding all but one of the investors with the same name. For example, if we initially had two distinct Peter Smith entries and one Pete Smith entry, we removed all but one of those names.

Table 1 presents information on our investor sample. Our email list contains 9,577 investors identified as VCs and 22,467 investors identified as angels. We had to exclude some of these either due to bounced email messages or errors (see Section 2.4 for details). Our final sample contained 28,433 investors (8,561 VCs and 19,872 angels) and is used for the analyses.

We categorized investors by gender and race. We classified gender using genderize.io, which classifies the gender of names based on their frequency of use. We supplemented this by inferring race from photos using Face++ for the 70% of investors for whom we had images. Two research assistants coded cases where there was disagreement between the two programs or genderize.io had less than 80% confidence in the classification.[12] Race was more difficult to code given that racial categories are inherently imprecise. In a similar procedure, we used NamePrism to determine investor race from their name, Face++ to determine their race from their photo, and had research assistants code unclear cases and cases of disagreement between the name and face codings.[13] A small number of

[12]For example, the name "Sam" has a 76% probability of being male according to Nameprism.

[13]We thank Steven Skiena and Junting Ye for giving us access to the NamePrism API, which is based on Ye, Han,

**Table 1: Investor Sample.** This table reports the number of investors in our sample. Mailing list consists of the initial set of investors that we collected across databases. Sample is our final sample of investors. Gender is classified using genderize.io API, Face++, and manual coding. Race is classified using NamePrism, Face++, and manual coding. Emails received is the number of investors who received the given number of emails.

| | Angels | | VCs | | Total | |
|---|---|---|---|---|---|---|
| | # | % of list | # | % of list | # | % of list |
| Mailing list | 22,467 | 100.0% | 9,577 | 100.0% | 32,044 | 100.0% |
| Not deliverable | 2,481 | 11.0% | 955 | 10.0% | 3,436 | 10.7% |
| Emailing error | 97 | 0.4% | 59 | 0.6% | 156 | 0.5% |
| Wrong person | 17 | 0.1% | 2 | 0.0% | 19 | 0.1% |
| Sample | 19,872 | 88.4% | 8,561 | 89.4% | 28,433 | 88.7% |
| | # | % of sample | # | % of sample | # | % of sample |
| Gender | | | | | | |
| Female | 1,876 | 9.4% | 1,082 | 12.6% | 2,958 | 10.4% |
| Male | 17,992 | 90.5% | 7,478 | 87.3% | 25,470 | 89.6% |
| Unclassified | 4 | 0.0% | 1 | 0.0% | 5 | 0.0% |
| Race | | | | | | |
| Asian | 3,157 | 15.9% | 1,368 | 16.0% | 4,525 | 15.9% |
| White | 15,481 | 77.9% | 6,797 | 79.4% | 22,278 | 78.4% |
| Other | 1,231 | 6.2% | 394 | 4.6% | 1,625 | 5.7% |
| Unclassified | 3 | 0.0% | 2 | 0.0% | 5 | 0.0% |
| Emails received | | | | | | |
| 1 | 1,228 | 6.2% | 1,197 | 14.0% | 2,425 | 8.5% |
| 2 | 4,669 | 23.5% | 2,532 | 29.6% | 7,201 | 25.3% |
| 3 | 9,069 | 45.6% | 3,309 | 38.7% | 12,378 | 43.5% |
| 4 | 4,906 | 24.7% | 1,523 | 17.8% | 6,429 | 22.6% |

investors were not classified due to ambiguous names and limited online presences. Consistent with the underlying population of investors, our sample was 90% male. Our sample was 78% White, 16% Asian, and 6% other races—71% of our investors were White men.

The line between angels and VCs can be blurry as angel investing can be taken to mean either investing in very early stage start-ups or investing one's own money. Adding to the confusion, many angels are current or former VCs. We call people VCs if they are either associated with an investment fund or referred to as current VCs in one of our databases. That classification made our sample 30% VCs and 70% angels. In addition to the separation of investors by their VC or angel identity, we

Hu, Coskun, Liu, Qin, and Skiena (2017).

stratified angels by activity using the number of investments they had made. Our average angel made 4.1 recorded investments but the distribution was skewed, so we defined experienced angels as those with at least three recorded investments (39% of angels) and inexperienced angels as those with fewer than three (61% of angels). Note that the number of recorded investments is a noisy lower bound for investments actually made and some angels had 10 unrecorded investments for every recorded investment.

We matched investors to pitches that were similar to their past investments and stated interests. First, we mapped the industry codes in each database we used to the 226 VentureSource industries. We did that by finding companies that were in multiple databases and comparing the industry codes assigned to them. For example, if companies classified as Advertising in one database were overwhelmingly classified as Advertising/Marketing in VentureSource, we mapped Advertising to Advertising/Marketing. Second, we used that mapping to link all companies and investor industry preferences to the VentureSource industries. Investors were then assigned a propensity score to invest in each industry based on their past investments and their preferences. Third, we mapped all of our pitches to the VentureSource industries. Fourth, we estimated the interest each investor would have in each pitch using the product of the pitch-industry mappings and the investor-industry mappings. Fifth, we partially normalized these ratings so that pitches in narrow industries would not be dramatically underrepresented. Sixth, we used the resulting ranking to find the ten best pitches for each investor. Seventh, we removed all pitches that were a better fit for someone else at the same VC fund or angel group. Finally, we took the four best resulting pitches for each investor. While we believe this procedure resulted in the best practical match given the data, this match is still imperfect because of changing investor interests and subjective industry classifications.

## 2.4 Emailing Process

Between August 22 and October 19, 2018, we conducted a series of small pilot studies using an additional pitch. We excluded those emails from the main study. Our results are robust to excluding all investors emailed as part of the pilot.

The main phase of the experiment took place between October 30 and November 16, 2018. Each investor received between one and four different pitch emails. To maximize statistical power, we assigned gender race using recipient-level random permutations, so that investors received no more than one email from each of the four sender demographic categories. Crucial to our experiment, the randomization was independent of all other sources of variation. We sent emails on weekdays during the day and waited for an average of five days between sending two sequential pitches to the same investor.

Delivering a large number of emails is difficult due to the aggressive spam policies of most email providers. To combat that, we sent the emails using a variety of web hosts, with a variable degree of deliverability and thus reply rates (the assignment of web hosts was independent of gender and race variation). Despite this, reply rates declined from 9.0% for the first 4,000 emails to 5.3% for the last 4,000 emails. Comparing the reply rate of our first emails to the reply rate of the overall sample suggests approximately one-quarter of our emails were blocked by spam filters.[14] Importantly, we have every reason to be certain that these filters were completely independent of gender and race.[15]

If there was an issue with an investor's email address, we excluded any emails sent to them from our sample and attempted not to send them further emails. In total, 3,611 or 11.3% of investors were excluded. By far the most common reason for this was automated replies suggesting our emails would not be read (3,436 investors or 10.7%, see Section 2.5). A small number of investors (156 or 0.5%) were excluded due to an error during the sending process. We also excluded nineteen email addresses (0.1%) that were associated with a different person. There was no relationship between any of these issues and the sender's gender or race and our results are robust to including all of these problem emails in the sample. In total, of the 86,453 emails sent, 79,677 (92.2%) ended up in our sample.[16]

[14]The email campaign monitoring tool GlockApps estimated that all of our early emails went to recipients' inboxes, while just one-quarter of our later emails went to recipients' inboxes, with one quarter going to spam folders and one half discarded by the recipients' email servers. We confirmed these email blocking results using manual tests of our emails with Gmail, Outlook, and Yahoo, some of the most popular U.S. email providers.

[15]We checked this in several ways. First, we created test accounts at Gmail, Outlook, and Yahoo and tested the deliverability of the messages. We did not identify any gender or race pattern in deliverability. Second, we tested our emails using commercial spam filters before and after the test and found no patterns. Third, we used GlockApps to test deliverability and found no patterns. Fourth, we consulted the literature and computer science faculty and found no evidence consistent with gender- or race-based filters. All of these support the conclusion that there are no gender- or race-specific spam filters. Finally, we performed robustness checks on our data. We confirmed that our gender results hold when looking at the first pitch email sent to each investor (early emails were less subject to spam filters) and when looking at interested replies as a fraction of total replies (non-interested replies indicate an email was read, refer to Section 2.5 for details).

[16]Including these excluded emails does not change our results.

In other cases, we did not send investors further emails but kept them and their replies in the sample. If recipients expressed displeasure about being emailed (38 investors or 0.1% of the sample), we did not email them again but kept them in the sample. If an email was met with a reply suggesting the recipient was not actively investing (367 or 1.3%), we similarly stopped emailing them.[17] A small fraction of our investors questioned the company's legitimacy (33 or 0.1%). Far more frequently, investors asked how a recipient got their email or knew them (469 or 1.7%). Out of an abundance of caution, we excluded both types of questioning investors (490 or 1.7% of the sample) from receiving future emails, even though most of these inquiries were probably benign as 63% of them were interested replies and are associated with a request for a pitch deck or a call.

### 2.5 Response Collection and Processing

No subsequent communication was sent to investors, independent of the nature of responses that were received. Instead, all the responses were recorded and matched back to our original pitch emails.[18] We recorded all the responses received between October 30 and November 28, 2018. We gave recipients more than two weeks to reply, despite the vast majority of replies (72%) coming within one day.

The responses were then coded. Table 2 reports the statistics. The analysis of the content of email responses was inevitably partially subjective. Therefore, every reply was classified independently by two coders. In all the cases where the coders disagreed on the coding, one of the authors reconciled the responses. To ensure the absence of any gender and race biases in this process, the entrepreneur email sender names were removed from the reply before coding and gender-specific words and phrases were obscured. Coders were asked to flag any replies where they could guess from the context the gender or race of the entrepreneur sender.[19] For those replies, the entrepreneur sender's identity was removed and the replies were coded again by a different group of research assistants.

In a number of cases, the email respondents either copied other people (such as investors' business partners) who then responded or investors forwarded the pitch email to others who responded. We aggregated all the emails up to the original pitch email sent and therefore processed all the information

[17]Our results are robust to excluding these recipients from the sample.

[18]We matched all manual replies and 97.4% of automated replies. The unmatched automated replies were generally delivery delay or delivery failure messages that provided too little information to allow for matching.

[19]For example, many investors expressed an interest in supporting female entrepreneurs.

**Table 2: Email Replies.** This table reports the categorization of the manual replies we received to our pitch emails. The columns under "Emails with at least one reply" report the emails we sent that received at least one reply that fell into each of the listed categories. The columns under "Investors with at least one reply" report the investors emailed who yielded at least one reply in each category. Investors may have received multiple emails, emails sent may have provided multiple replies, and the replies themselves may fall into multiple categories. For example, if Samantha Huynh emailed John Smith who replied that he was not interested due to the stage, but then forwarded the email to Jane Doe, the email sent by Samantha Huynh to John Smith would be classified as both "Not interested - Too early stage" and "Interested - Referred to others" in the "Emails with at least one reply" columns and John Smith would also be classified that way in the "Investors with at least one reply" columns. The category of "Not emailed again" includes all of the categories marked $^\Diamond$.

| | Emails with at least one reply | | Investors with at least one reply | |
|---|---|---|---|---|
| | # | % of sample | # | % of sample |
| Interested | 3,407 | 4.3% | 2,777 | 9.8% |
| Face-to-face meeting | 152 | 0.2% | 140 | 0.5% |
| Phone or video call | 397 | 0.5% | 360 | 1.3% |
| Request for deck | 2,095 | 2.6% | 1,770 | 6.2% |
| Follow-up questions | 476 | 0.6% | 441 | 1.6% |
| Referred to others | 650 | 0.8% | 577 | 2.0% |
| Other interest | 204 | 0.3% | 190 | 0.7% |
| Questioning (and interested)$^\Diamond$ | 330 | 0.4% | 307 | 1.1% |
| Not interested | 1,966 | 2.5% | 1,599 | 5.6% |
| No, too early stage | 305 | 0.4% | 263 | 0.9% |
| No, wrong industry/geography | 702 | 0.9% | 597 | 2.1% |
| No, not actively investing$^\Diamond$ | 408 | 0.5% | 367 | 1.3% |
| No, unsubscribe$^\Diamond$ | 38 | 0.0% | 38 | 0.1% |
| No, other reason | 526 | 0.7% | 463 | 1.6% |
| Unclassified$^\Diamond$ | 136 | 0.2% | 125 | 0.4% |
| Questioning (and not interested)$^\Diamond$ | 225 | 0.3% | 196 | 0.7% |
| of which alleged spam or fraud$^\Diamond$ | 34 | 0.0% | 33 | 0.1% |
| Not emailed again | 971 | 1.2% | 863 | 3.0% |
| Total replies | 5,188 | 6.5% | 3,980 | 14.0% |
| Total sample | 79,677 | 100.0% | 28,433 | 100.0% |

contained in multiples emails of this kind as if they were one email response.

### 2.5.1 Automatic Replies

We ignored all automatic email responses. These generally fell into the following categories: email is not deliverable because the email address no longer exists; the person no longer works at the company; email was waitlisted or delivery delayed; email produced non-personal automatic responses, such as confirmation of submitting a help ticket; responses were out of the office messages. We exclude from our sample any email message we sent that triggered an automated reply suggesting our email would not be read by the target investor. For example, we exclude undeliverable emails but do not exclude emails that triggered out of office replies.

### 2.5.2 Interested Replies

We examined whether each manual email response suggested that the recipient was interested in learning more about the company as an investor. If they were, the response was coded as *Interested Reply.* These emails generally fell into the following categories:

**Meeting or call request.** Asking to set up a meeting or more commonly a telephone call.

- Example 1: "When is the best time to reach you by phone?"[20]
- Example 2: "Happy to meet you and happy to have a chat! Do any of the times below work for you for a quick call? Show me times I'm free at [calendar link provided]."
- Example 3: "Let me know over the course of the next few weeks a date/time that might be convenient and we can grab a coffee."

**Request for more information.** Asking for a pitch deck or asking for specific information of concern to early-stage investors, such as the customer traction, valuation of the company, amount of money to be raised, etc.

- Example 1: "Thanks for reaching out. Please send me your pitch deck. I've got questions but

[20] In all the examples in this subsection, we retained the original text of the emails, and only replaced identifying information with [...].

I'd like to see the deck first before providing a list of questions back. Please ensure that your deck includes: 1) What type of valuation your seeking; 2) Terms; 3) What you're going to use the funds for; 4) What are you looking for in an investor; 5) When are you looking to close the round; 6) Any co-investors who have already soft closed."

- Example 2: "Thanks for reaching out to me. Can you send me an executive summary when you have a chance."
- Example 3: "From your brief email, it's hard for me gauge what our interest level might be. Do you have a more detailed pitch deck we can review?"
- Example 4: "Hi [name] thanks for reaching out! How does [sender company name] differ from [company 1 name], [company 2 name] (b2b), [company 3 name] (b2b + b2c), et. al.?"

**Referral.** A referral or forward to another person or fund (excluding a secretary or an assistant).

- Example 1: "This sounds interesting, I'm adding in a colleague ([name]) who is more familiar with the space. Perhaps you two try to find a time to talk or share some of your information over email?"
- Example 2: "Thanks for thinking of us, my principal [name] will be in touch to learn more shortly."

### 2.5.3 Not Interested Replies

Some investors sent manual replies that did not indicate interest in investing in the company. These replies generally fell into the following categories:

**Too early stage or inconsistent with the investor's mandate.**

- Example 1: "We invest only in companies with monthly revenues of at least $50k. All the best with your venture!"
- Example 2: "Thanks, but I'm not interested. We focus on later stage investments, usually above $5m."
- Example 3: "Sounds very interesting the idea and the business. Unfortunately, our mandate is to invest only in Revenue Generating Companies. Please, keep us informed and it will be great

to evaluate an opportunity when the time is right."

**Different industry, geography, or focus.**

- Example 1: "Thanks for contacting us, but our group only considers investment in companies based in the Mid West. Good luck with your venture."
- Example 2: "Unfortunately this would not be a fit for us because we focus our investments on the east coast."
- Example 3: "Congratulations on your success. We are a FinTech focused venture fund so not a fit for us. We wish you well."
- Example 4: "Thanks. It's not really medtech enough for me."

**Overall lack of fit.**

- Example 1: "Thanks, this is not for me. Success with your project."
- Example 2: "Not interested right now, but your email was well put together and it sounds like a cool concept so I certainly wish you the best of luck! "

**Recipient is not currently an active investor.**

- Example 1: "Thanks for getting in touch but we are not looking for investments right now."
- Example 2: "Thanks, but I am raising my next fund right now and not making investments."
- Example 3: "I am fully invested at this point. I am advising companies."

### 2.5.4 Questioning and Unclassified Replies

We also classified all manual responses into whether they contained questions on how the investor's contact information was found, why the investor was chosen, or what the investor's connection to the company was. These responses and any other response that indicated any form of suspicion were classified as questioning. They generally fell into the following categories:

**Questioning and interested**. A question on how investor's email was located and why the investor was chosen that came as a part of a positive response, often providing a different email to continue the correspondence.

- Example 1: "Hi, yes pls send me the pitch. Happy to review. Curious to understand too how you got my name as I cannot recall we have met?"
- Example 2: "If you can tell me how and where you found me and got my email, then you can send me your deck."
- Example 3: "Interested in finding out a bit more. How did you hear of our firm?"
- Example 4: "Thanks for the note. Always happy to listen - tomorrow 9am? Do you have any background info you can send? And out of curiosity why contact me / how did you get my contact info?"
- Example 5: "Either I'm really poor using Internet search or you are really good at hiding your company info, I couldn't seem to find much on the web ;-) Sure I would love to learn more about what you're working on and see how I may assist, see below and attached to share a bit about what I do... May I also ask where you got my contact info. Looking forward to hearing more from you. Thanks!"
- Example 5: "Sounds neat! How did you find this e-mail? Send your deck to [email address] and I will follow up."

**Questioning and not interested.** A response that asked questions about why no information about the entrepreneur was found on the Internet, such as the absence of a LinkedIn profile.

- Example 1: "Hi, just curious, how did you get my email?"
- Example 2: "Is there a reason you chose to send this to me?"
- Example 3: "I'm curious what you think my investment strategy is exactly..."

In some infrequent cases, the nature of the response content was vague and difficult to interpret. These emails were left unclassified (136 or 0.2%).

**Table 3: Characteristics of Investors Who Were and Were Not Interested.** This table reports statistics on the investors that responded with interest to our unsolicited pitch emails (the Interested investors column), the investors that were emailed (Emailed investors), and the overall population of investors (All investors). We matched VCs to Pitchbook and angels to AngelList and present averages based on those databases. The All investors column includes all investors in those databases that meet our screening criteria (as described in Subsection 2.3), whether or not they were emailed. The first Difference column gives the difference between the average Interested investor and the average Emailed investor. The second Difference column gives the difference between the average Emailed investor and the average investor in Pitchbook. Tests of the null of no difference between groups are reported with * denoting significance at the 10% level, ** denoting significance at the 5% level, and *** denoting significance at the 1% level. Firm exit rates (IPO, M&A, failure, and private) are based on deals made 2009-2013. Connections refers to the number of connections an investor has on AngelList. All VC statistics except board seats are based on the firm they are associated with due to limited person-level data. We averaged firm-level measures across firms for VCs that were associated with multiple firms. All data are Winsorized at the 99.5% level.

| | Interested investors | Emailed investors | Difference | All investors | Difference |
|---|---|---|---|---|---|
| Angels | | | | | |
| connections | 118.1 | 99.8 | 18.3*** | 81.7 | 18.0*** |
| investments | 5.3 | 4.2 | 1.1*** | 3.7 | 0.5*** |
| exits | 0.50 | 0.45 | 0.05* | 0.39 | 0.06*** |
| exited % | 8.3 | 8.9 | -0.6 | 7.8 | 1.1*** |
| VCs | | | | | |
| board seats | 2.50 | 2.35 | 0.15*** | 2.40 | -0.05** |
| year firm founded | 2009 | 2007 | 2 | 2006 | 1 |
| # professionals at firm | 10.0 | 11.1 | -1.1 | 13.8 | -2.7 |
| # investments by firm | 124.7 | 124.1 | 0.6 | 175.7 | -51.6*** |
| # active investments by firm | 44.6 | 39.2 | 5.4*** | 50.9 | -11.7*** |
| round # of firm's average deal | 3.4 | 3.6 | -0.2** | 3.7 | -0.1** |
| size of firm's average deal ($m) | 18.4 | 20.9 | -2.5 | 22.9 | -2.0 |
| size of firm's average fund ($m) | 211.9 | 190.7 | 21.2 | 225.5 | -34.8 |
| vintage of firm's most recent fund | 2015 | 2014 | 1*** | 2014 | 0 |
| firm # open funds | 0.6 | 0.6 | 0.1 | 0.6 | 0.0 |
| firm # closed funds | 3.3 | 3.6 | -0.4 | 4.5 | -0.8*** |
| % of firm deals that IPO | 4.6 | 5.4 | -0.9 | 6.2 | -0.7* |
| % of firm deals that M&A | 35.9 | 35.7 | 0.3 | 35.9 | -0.2 |
| % of firm deals that fail | 9.9 | 11.5 | -1.5 | 11.3 | 0.2 |
| % of firm deals that are private | 49.6 | 47.4 | 2.1 | 46.6 | 0.8 |

### 2.5.5 Characteristics of Investors Who Were and Were Not Interested

We check whether respondents were systematically different from nonrespodents by matching 86% of our sample to either Pitchbook or Angellist. Table 3 compares the investors who replied with interest to the investors emailed. The investors who replied with interest are of slightly better quality than the average investor we emailed. Interested angels have more investments (5.3 vs 4.2), more exits (0.50 vs 0.45), and more Angellist connections (118 vs 100). Interested VCs have more board seats (2.50 vs 2.35) and work for firms with more active investments (45 vs 39), a more recently raised fund (2015 vs 2014), and earlier stage investments. Senior VCs using associates to screen deal flow may explain the apparent lack of adverse selection against our unsolicited pitches by VCs. Although the vast majority of replies came from the person emailed, the top quartile of VCs (as measured by board seats) was less likely to response personally (79%) than the bottom quartile (88%), despite both groups having the same total interested reply rate.

We limited the number of emails sent to any one VC firm. That meant a VC at a large firm was less likely to be emailed than a VC at a small firm, despite the large firm receiving more overall emails than the small firm. Hence, the average VC we emailed was of slightly lower quality than the average VC in Pitchbook.

## 2.6 Experiment Limitations

This is the first randomized field experiment exploring the gender and racial preferences of investors. As such, this study possesses a number of significant strengths relative to the extant literature on the topic. At the same time, it also has a number of weaknesses and limitations that may limit its external validity. We discuss these concerns now.

**Stage of investment decision-making process.** The experiment concentrates on the first stage of a multi-staged investment decision-making process. This process progresses from initial introductions to extended in-person meetings to formal due diligence and culminates with investments in a select few start-ups. Investors may be subject to different biases at later stages of this decision-making process. For example, investors may exhibit discrimination at in-person meetings but not at the email stage. As with most correspondence studies, our experiment gives us an accurate picture of bias in initial interest but does not allow us to identify discrimination later in the decision-making process or to measure final decisions, which are what matter the most.

At the same time, even though a response to an unsolicited pitch email may not lead to an investment, there are other benefits for entrepreneurs, especially for those who are outside of investor networks. Even half an hour on the phone with an experienced investor can generate valuable feedback and advice. This is further corroborated by extensive literature suggesting that start-up investors add value beyond simply providing financing (Bernstein, Giroud, and Townsend, 2016; Howell, 2017).

**Unsolicited emails versus warm introductions.** Although there are many anecdotal stories of unsolicited pitches succeeding,[21] most consummated deals come via referrals from their professional networks (known in the industry as warm introductions).[22] We cannot test for discrimination in these other channels. However, because women and other underrepresented populations may have fewer advantageous connections that can enable warm introductions, cold approaches may be especially important for them.

**Nature of pitches.** Although we strived to make our pitches similar to the high-quality pitches early-stage investors receive, the pitches we wrote may not be fully representative of real pitches. As we are unable to access a large database of real-world pitches, we resorted to interviewing industry practitioners, reviewing many actual pitches, and asking experienced practitioners to read and comment on our pitches. Our best estimate is that our pitches are better written and structured than most email pitches received by the typical early-stage investor; however, we are unable to assess this conjecture empirically.

Perhaps the most important difference between our pitches and typical pitches is that all of our fictitious entrepreneurs are graduate students in highly ranked colleges. As discussed above, we chose this identity for our email senders to increase the baseline response rate. This group is also vastly overrepresented in high-impact entrepreneurship (Gompers and Wang, 2017). Our experimental design, however, does not directly extend to entrepreneurs from other educational backgrounds. Under

[21] See, for example, quora.com/Does-cold-emailing-VC-firms-ever-work-Are-there-any-tactics-to-help-increase-the-od ds or medium.com/tech-product-and-life/how-i-almost-closed-a-deal-with-mark-cuban-with-only-an-e-mail-d7eda956 a7a7, accessed June 1, 2019.

[22] Gompers et al. (2018) show that unsolicited approaches by founders account for only 10% of venture capitalists' deal flow and anecdotal evidence suggests the same pattern holds for angels (e.g. paulgraham.com/angelinvesting.html, accessed June 26, 2019).

many models of statistical discrimination, we would expect people with external certification, such as a prestigious education, to experience less discrimination (Agan and Starr, 2017). Thus, our results may understate the overall differences in reply rates.

Beyond that, the innovative projects in our pitch emails may be different in nature than the typical projects pitched. Because of the nature of our study, the fictitious companies in our pitches are at an earlier stage and less developed than the companies typically pitched to VCs and angels.

**Asian entrepreneurs and immigration status.** The race of the email senders in our experiment is indicated exclusively by the last name, resulting in our fictitious Asian entrepreneurs having names like Adam Liu and Jennifer Huang. Our target investors could have perceived such names either as Asian Americans or Asian immigrants. Our study does not allow us to differentiate between these two categories. Oreopoulos (2011) provides important empirical evidence in this regard. The author conducts a randomized resume experiment to study discrimination against immigrants in Canada, identified as those with foreign-sounding names. In some cases, the author used Chinese-sounding last names with English-sounding first names, as in our study. His results for English first names/Chinese last names are similar to his results for Chinese first and last names (he finds discrimination in both categories). The author concludes that "...many second generation Chinese adopt and use an English sounding name to make pronunciation easier for non-Chinese and to signal North American assimilation. Interestingly, this adoption does not improve one's chances for a callback."

**Believability.** Neither our start-ups nor the entrepreneurs who are pitching them actually exist. Thus, they have no online or offline presence (although we created simple uninformative websites for our start-ups). Several respondents questioned the lack of a web presence and LinkedIn profiles. It is likely that the fictitious nature of companies and founders reduced the overall response rate, but Females or Asians received the same rate of questioning replies as their counterparts. Females did receive significantly more replies alleging fraud or spam (0.06% versus 0.03%), suggestive of women receiving more suspicion (Egan, Matvos, and Seru, 2019). This may have changed the interested reply rate female entrepreneurs received; however, the effect is driven by less than a dozen messages. On the second to last day of the study, there were two Twitter threads by VCs questioning suspicious emails.[23] Restricting our analysis to replies received before that day does not change our main results,

[23]One thread is accessible at https://twitter.com/dkoretz/status/1063239065196085248 as of July 31, 2019. The

as the vast majority of data was collected by that point. Our results are also robust to considering only the first message each investor received.

**Gender and race are suggested rather than stated.** Correspondence studies often struggle to convey race to the study participants. First, it can be difficult to ensure that race and gender are noticed. We attempt to address this by repeating the sender's name (our cue for race and gender) several times: the body of the email; the email signature; the email address; and the email sender's name. Second, it can be difficult to choose names that convey gender and race without additional information. For example, Bertrand and Mullainathan (2004) use distinctively Black first names to convey race and Darolia, Koedel, Martorell, Wilson, and Perez-Arce (2015) argue that distinctively Black first names are markers of not just race but also socioeconomic status. This is less of an issue for our study because we use common ethnically neutral first names that reliably convey gender and ethnically associated last names that reliably convey race.

**Timing.** Our results pertain to the late-2018 period during which the experiment was run. There has been a long-term trend toward greater support of female representation, and that trend may have accelerated in past years. For example, the General Social Survey polls Americans on a number of questions, including whether it is "much better for everyone involved if the man is the achiever outside the home and the woman takes care of the home and family." In the 14 studies run between 1988 and 2010, 35%–40% of respondents agreed that it was better for women to take care of the home, but in 2018 that rate dropped to 25%, a larger drop in eight years than the preceding 32.[24] It is possible that our results would have been different if we had conducted this experiment even a few years earlier.

other thread was deleted shortly after it was posted.

[24]The data are available at gssdataexplorer.norc.org/trends/Gender&Marriage?measure=fefam, accessed on June 28, 2019.

# 3 Results

## 3.1 Is There a Gender Gap in Investor Interest?

Surprisingly, female entrepreneurs received a higher rate of interested replies than male entrepreneurs. Emails that were randomly assigned to be sent by a female entrepreneur had a 4.47% chance of receiving an interested reply, while emails assigned to a male entrepreneur had a 4.08% chance. This represents an absolute difference of 0.39%, or a relative difference of 9.5%, in favor of female entrepreneurs and is statistically significant at the 1% level. Importantly, this difference can be attributed only to gender manipulation through random name substitutions, as all the other sources of variation were excluded. This suggests female entrepreneurs are favored by investors, at least at this initial stage.

**Figure 2: Illustration of Gender Bias in Different Samples.** This figure presents the ratio of the interested reply rate of fictitious female entrepreneurs to that of fictitious male entrepreneurs. Ratios are reported for the overall sample and various subsamples. Experienced angels are angels with three or more listed investments; inexperienced angels are angels with fewer than three. B2B denotes business-to-business start-ups and B2C denotes business-to-consumer start-ups. Feminine pitches are those a group of 50 Mechanical Turk users said were more likely to be for a company founded by a female than a male; Masculine pitches are those that were more likely to be founded by a male. Point estimates are presented as squares; for example, 1.1 would describe females getting 10% more interested replies than males. The vertical line denotes zero bias. 90% confidence intervals are presented as horizontal lines.

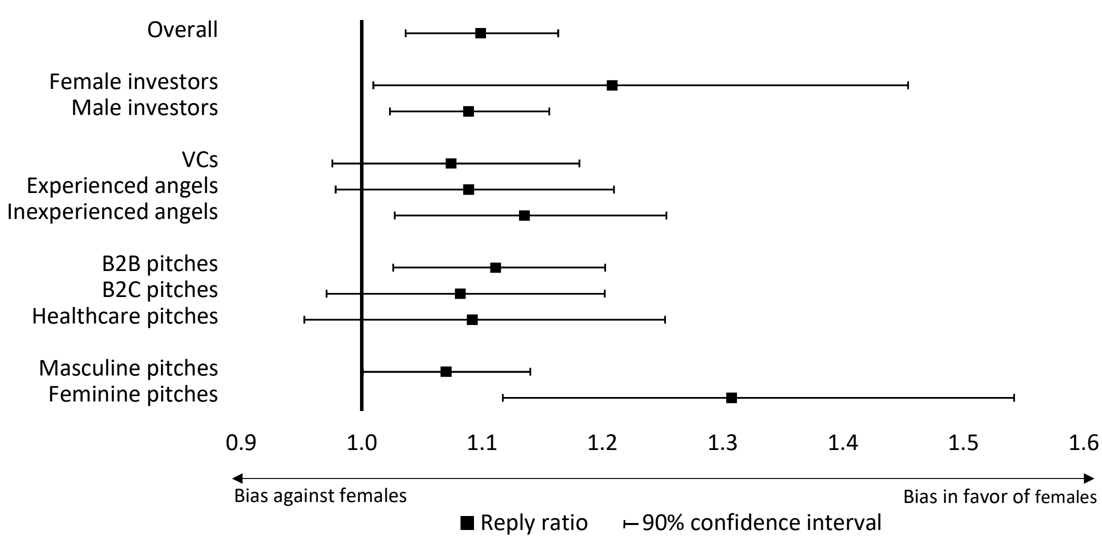


**Table 4: Gender Bias in Different Samples.** This table reports the rate of interested replies to pitches randomly assigned to be emailed from fictitious male and female entrepreneurs. The rows correspond to the subsample of investors receiving the pitches. The Emails sent column lists the number of emails we sent. The Interest column reports the fraction of our emails that generated interested replies. The Ratio column reports the ratio of the interested reply rate females received to the interested reply rate males received, with rates above one indicating a bias in favor of females. The Difference in interest column reports the difference between the interested reply rate for females and the interested reply rate for males. Experienced angels are angel investors we were able to link to three or more investments; Inexperienced angels are angels who were linked to fewer than three investments. B2B denotes business-to-business and B2C denotes business-to-consumer. Feminine pitches are those a group of 50 Mechanical Turk users said were more likely to be for a company founded by a female than a male; Masculine pitches are those that were more likely to be founded by a male. Standard errors are given in parentheses. Tests of the null hypothesis of no discrimination are reported with * denoting significance at the 10% level, ** denoting significance at the 5% level, and *** denoting significance at the 1% level.

| | Male entrepreneur | | Female entrepreneur | | | |
|---|---|---|---|---|---|---|
| | Emails sent | Interest % | Emails sent | Interest % | Ratio | Difference in interest % |
| Overall | 39,823 | 4.08 | 39,854 | 4.47 | 1.09 | 0.39*** |
| | | (0.10) | | (0.10) | | (0.14) |
| Female investors | 4,075 | 3.78 | 4,155 | 4.55 | 1.20 | 0.77* |
| | | (0.30) | | (0.32) | | (0.44) |
| Male investors | 35,740 | 4.11 | 35,692 | 4.46 | 1.08 | 0.35** |
| | | (0.11) | | (0.11) | | (0.15) |
| VCs | 11,127 | 5.46 | 11,153 | 5.85 | 1.07 | 0.38 |
| | | (0.22) | | (0.22) | | (0.31) |
| Angels | 28,696 | 3.55 | 28,701 | 3.93 | 1.11 | 0.39** |
| | | (0.11) | | (0.11) | | (0.16) |
| Experienced angels | 11,158 | 4.34 | 11,160 | 4.70 | 1.08 | 0.37 |
| | | (0.19) | | (0.20) | | (0.28) |
| Inexperienced angels | 17,538 | 3.04 | 17,541 | 3.44 | 1.13 | 0.40** |
| | | (0.13) | | (0.14) | | (0.19) |
| B2B pitches | 21,991 | 3.88 | 22,127 | 4.29 | 1.11 | 0.41** |
| | | (0.13) | | (0.14) | | (0.19) |
| B2C pitches | 12,351 | 3.87 | 12,277 | 4.17 | 1.08 | 0.30 |
| | | (0.17) | | (0.18) | | (0.25) |
| Healthcare pitches | 5,481 | 5.38 | 5,450 | 5.85 | 1.09 | 0.47 |
| | | (0.30) | | (0.32) | | (0.44) |
| Feminine pitches | 6,700 | 2.85 | 6,650 | 3.71 | 1.30 | 0.86*** |
| | | (0.20) | | (0.23) | | (0.31) |
| Masculine pitches | 30,856 | 4.20 | 30,897 | 4.47 | 1.07 | 0.28* |
| | | (0.11) | | (0.12) | | (0.16) |

Table 4 and Figure 2 show that the preference for female entrepreneurs persists across various subsamples. Both male and female investors are more likely to show interest in female entrepreneurs. Female investors reply with interest to female entrepreneurs 20% more frequently than to male entrepreneurs, while male investors reply to female entrepreneurs 8% more frequently. Female investors show a 11% higher preference for female entrepreneurs than male investors do, suggestive of homophily; however, this difference is not significant due to the low number of female investors in our sample.

Both VCs and angels are more likely to respond to female entrepreneurs, although the difference is statistically significant only for angels (70% of our sample). Some angels in our sample are less experienced, as gauged by the number of investments they are recorded to have made. These less experienced angels have a statistically significant 13% higher interested reply rate to our female entrepreneurs, while the more experienced angels have an 8% higher reply rate and VCs have a 7% higher rate.

We break down our full sample by the industry of the company described in our email pitches: business-to-consumer (B2C), business-to-business (B2B), and healthcare. (This industry breakdown is widely used by practitioners.) All three industry subsamples exhibit a higher response rate to the unsolicited pitch emails sent by females, although the difference is statistically significant only for the B2B industry.

Finally, we tested theories of gender congruence by looking at pitches perceived to be more feminine or masculine. We measure the perceived femininity of pitches by asking 50 Mechanical Turk workers whether the gender of the company founder is 1 definitely male, 2 most likely male, 3 equally likely to be male or female, 4 most likely female, or 5 definitely female. We call pitches with an average score above 3 feminine and pitches with an average score below 3 masculine. Although only a small number of pitches were perceived to be written by female founders, those pitches had a significantly larger pro-female bias than other pitches, with females getting 30% more responses than males. This is consistent with past work on gender congruence (e.g., Herbert (2018)). Note that even the pitches perceived to be written by a male founder had a significant pro-female bias of 7%.

Although theories testable using pitch-level variation are extremely interesting, they are difficult for us to test because our pitches were designed to maximize external validity and response rates. For example, there is limited power in our gender congruence test because we did not intentionally create distinctively feminine or distinctively masculine pitches. We chose not to do that and instead to maximize external validity by creating pitches that matched the normal pitches received by investors. Similarly, it would be interesting to test the interaction of gender and pitch informativeness by randomly removing information from pitches. We chose not to do that to avoid compromising the quality of our pitches.

Like most correspondence studies, we measure only replies to first contacts and do not measure discrimination in either warm introductions or the subsequent stages of the investment process. The economic significance of our results depends on how they generalize to warm introductions and later stages of the investment process. A 9% higher rate of interest on cold emails is significant for entrepreneurs trying to raise money through this process, but this approach is not the dominant method of fundraising. If female entrepreneurs get 9% more interest on all forms of initial contact and there is no other bias in the process, they might be 9% more likely to get their ventures funded—clearly a large effect. On the other hand, if female entrepreneurs are less able to get warm introductions due to being excluded from networks (e.g., Abraham (2019)) or are facing discrimination later in the investment process (e.g., Brooks et al. (2014)), they might face significant discrimination overall despite favorable treatment in cold emails. Our results should not be interpreted as showing a lack of discrimination against females overall or at other steps of the investor decision-making process.

Regardless, our result is both important and surprising given widely-held beliefs about gender discrimination in entrepreneurship. The stark underrepresentation of females in high-impact entrepreneurship and the volume of anecdotal evidence on discrimination led us to expect non-trivial discrimination against female entrepreneurs. Therefore, the most important finding of this study is not that we found a statistically significant, though only marginally economically relevant, bias in favor of females. Rather, it is that our results lead us to largely rule out discrimination *against* females by VCs and angels as can be discerned from their reply rates to unsolicited email pitches.

Although we report all of our results based on sample splits, the randomized nature of the experiment means that the results are extremely robust to different specifications. Randomization means that control variables such as pitch fixed effects or investor characteristics have virtually no effect on the coefficients of the variable of interest. Similarly, logit or probit specifications change the coefficients in the expected way but have no effect on the significance or interpretation. We also get similar results when we redo this analysis at the investor level (e.g., by comparing the fraction of investors who have a higher reply rate to females than to males with the fraction of investors with the opposite preference). Further, note that all of the tests in this subsection and the following subsection were preregistered, with the exception of the industry-group tests and the experienced versus inexperienced angels tests.[25]

### 3.2 Is There a Race Gap in Investor Interest?

**Figure 3: Illustration of Racial Bias in Different Samples.** This figure presents the ratio of the interested reply rate of fictitious Asian entrepreneurs to that of fictitious White entrepreneurs. Ratios are reported for the overall sample and subsamples. Experienced angels are angels with three or more listed investments; inexperienced angels are angels with fewer than three. B2B denotes business-to-business and B2C denotes business-to-consumer. Point estimates are presented as squares; for example, 1.1 would describe Asians getting 10% more interested replies than Whites. The vertical line denotes zero bias. 90% confidence intervals are presented as horizontal lines.

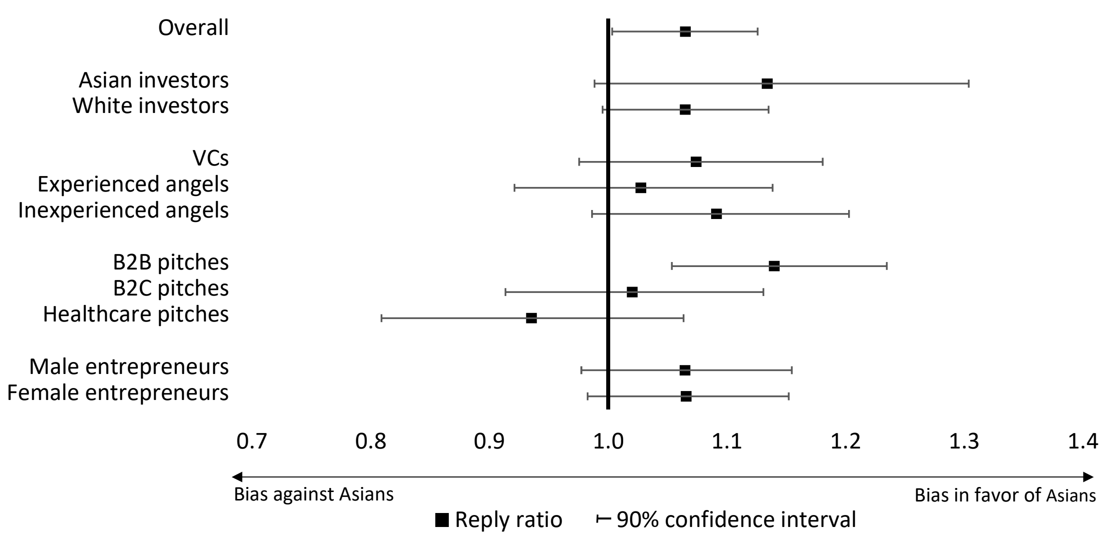



Table 5 and Figure 3 present our results on investor interest in fictitious Asian and White entrepreneurs. Emails that were randomly assigned to be sent by an Asian entrepreneur had a 4.40% chance of

[25]Specifically, the trial is registered in the AEA RCT Registry under AEARCTR-0003277. The gender (and race, in the following subsection) interested reply rate differences were primary outcomes. The tests of homophily, gender congruence, and VCs versus angels, as well as the gender/race interaction reported in the next subsection, were secondary outcomes.

**Table 5: Racial Bias in Different Samples.** This table reports the rate of interested replies to email pitches randomly assigned to be emailed from fictitious White and Asian entrepreneurs. The rows correspond to the subsample of investors receiving the pitches. The Emails sent column lists the number of emails we sent. The Interest column reports the fraction of our emails that generated interested replies. The Ratio column reports the ratio of the interested reply rate Asians received to the interested reply rate Whites received, with rates above one indicating a bias in favor of Asians. The Difference in interest column reports the difference between the interested reply rate for Asians and the interested reply rate for Whites. Experienced Angels are angel investors we were able to link to three or more investments; Inexperienced Angels are angels with fewer than three linked investments. B2B denotes business-to-business and B2C denotes business-to-consumer. Standard errors are given in parentheses. Tests of the null of no discrimination are reported with * denoting significance at the 10% level and *** denoting significance at the 1% level.

| | White entrepreneur | | Asian entrepreneur | | | |
|---|---|---|---|---|---|---|
| | Emails sent | Interest % | Emails sent | Interest % | Ratio | Difference in interest % |
| Overall | 39,736 | 4.15 | 39,941 | 4.40 | 1.06 | 0.25* |
| | | (0.10) | | (0.10) | | (0.14) |
| Asian investors | 6,325 | 4.41 | 6,342 | 4.98 | 1.13 | 0.57 |
| | | (0.26) | | (0.27) | | (0.38) |
| White investors | 31,105 | 4.04 | 31,291 | 4.29 | 1.06 | 0.24 |
| | | (0.11) | | (0.11) | | (0.16) |
| VCs | 11,126 | 5.46 | 11,154 | 5.85 | 1.07 | 0.38 |
| | | (0.22) | | (0.22) | | (0.31) |
| Angels | 28,610 | 3.64 | 28,787 | 3.84 | 1.06 | 0.20 |
| | | (0.11) | | (0.11) | | (0.16) |
| Experienced angels | 11,143 | 4.47 | 11,175 | 4.57 | 1.02 | 0.10 |
| | | (0.20) | | (0.20) | | (0.28) |
| Inexperienced angels | 17,467 | 3.11 | 17,612 | 3.38 | 1.09 | 0.27 |
| | | (0.13) | | (0.14) | | (0.19) |
| B2B pitches | 21,977 | 3.83 | 22,141 | 4.34 | 1.14 | 0.52*** |
| | | (0.13) | | (0.14) | | (0.19) |
| B2C pitches | 12,311 | 3.99 | 12,317 | 4.05 | 1.02 | 0.06 |
| | | (0.18) | | (0.18) | | (0.25) |
| Healthcare pitches | 5,448 | 5.82 | 5,483 | 5.42 | 0.93 | -0.40 |
| | | (0.32) | | (0.31) | | (0.44) |
| Female entrepreneurs | 19,904 | 4.34 | 19,950 | 4.60 | 1.06 | 0.27 |
| | | (0.14) | | (0.15) | | (0.21) |
| Male entrepreneurs | 19,832 | 3.96 | 19,991 | 4.20 | 1.06 | 0.24 |
| | | (0.14) | | (0.14) | | (0.20) |

receiving an interested reply, while those by a White entrepreneur had a 4.15% chance. This represents an absolute difference of 0.25% or a relative difference of 6.1% in favor of the fictitious Asian entrepreneurs, statistically significant at the 10% level. To the extent that race perception affects investor decisions to follow up on unsolicited pitch emails, Asian entrepreneurs are favored by our sample of investors.

This preference for Asian entrepreneurs persists for all but one of our subsamples, although the effect is too small to attain statistical significance in almost all cases. Specifically, we find that White investors are 6% more likely to reply with interest to Asian entrepreneurs than White entrepreneurs, Asian investors at 13% more likely, VCs are 7% more likely, angel investors are 6% more likely (with experienced angels being 2% more likely and inexperienced angels being 9% more likely), but none of these sample splits are statistically significant. Pitches from B2B start-ups had a 14% higher interested response rates for Asian entrepreneurs, a level that is significantly different from both an unbiased response rate and from the 2% higher rate for B2C start-ups and the 7% lower rate for healthcare start-ups.

The final two rows of Table 5 present results on the intersection of race and gender. Across our four demographic types, White males received a 3.96% rate of interested replies, White females 4.34%, Asian males 4.20%, and Asian females 4.60%. In terms of relative gender preference, White females received 9.4% more interested replies than White males and Asian females received 9.5% more interested replies than Asian males, both statistically significant. In terms of relative race preference, Asian males received 6.0% more than White males and Asian females received 6.1% more than white females, neither statistically significant. In both cases, the difference between the ratios is very close to zero, which shows our results are not being driven by a single race-gender condition.

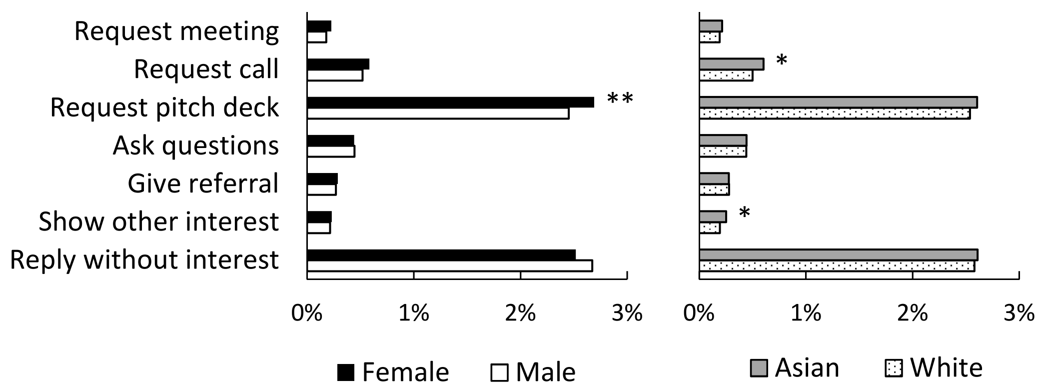
As with our findings on gender, these findings apply only to unsolicited email pitches and are silent on discrimination in the subsequent investment process.

### 3.3 Gender and Racial Differences in the Nature of Replies

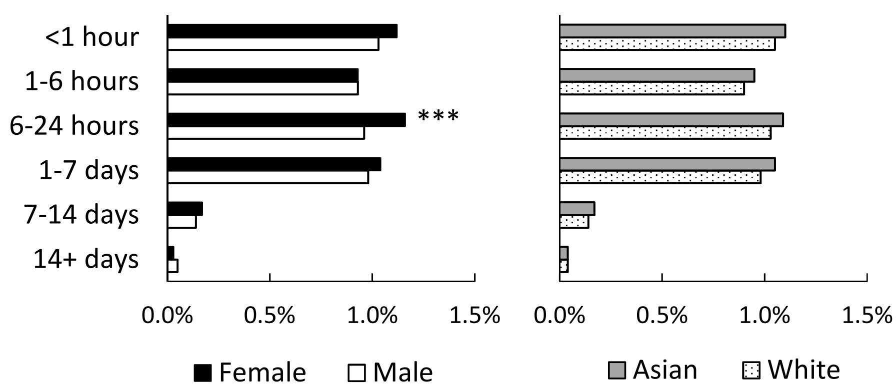
In addition to looking at the rates of interested replies, we looked at whether females or Asians received different types of replies or waited longer to receive replies. Panel A of Figure 4 shows the

**Figure 4: Gender and Racial Differences in the Nature of Replies.** This figure presents the fraction of emails from men, women, Asians, and Whites that got manual replies with different characteristics. Panel A shows the rate of interested replies received within each time period. Panel B shows replies with each level of interest. Each category includes only emails that did not fit in a previous category. For example, if an email got a reply proposing a meeting, asking for a pitch deck, and asking questions about the product, it would be counted only in the "Request meeting" category. In both panels, tests of the null of equal reply rates are reported with * denoting significance at the 10% level and ** denoting significance at the 5% level.

**Panel A:** Differences in type of manual reply




**Panel B:** Differences in time to interested reply



types of manual replies received by different genders and races. We order replies by implied interest, with requesting a meeting being the highest level of interest, requesting a call being the second highest, and so on through requesting a pitch deck, asking questions about the company, providing a referral, other expressions of interest, and finally, replies without interest. Each email was coded based on the highest level of interest it received. Women received significantly more (10%) requests for pitch decks, which drives the difference in interested replies as half of them were requests for pitch decks. Both women and Asians received more requests for in-person meetings (28% and 13%) than their counterparts, although these differences were not significant due to the small number of in-person meeting requests. Similarly, women and Asians received more requests for calls (13% and 21%), with the difference being significant for Asians.[26]

Panel B of Figure 4 shows the time it took replies to arrive. Again, we see little difference between groups—all groups received approximately 25% of their interested replies within one hour and 72% within one day. Kolmogorov-Smirnov tests confirm the distribution of time-to-reply is not significantly different, conditional on an interested reply.

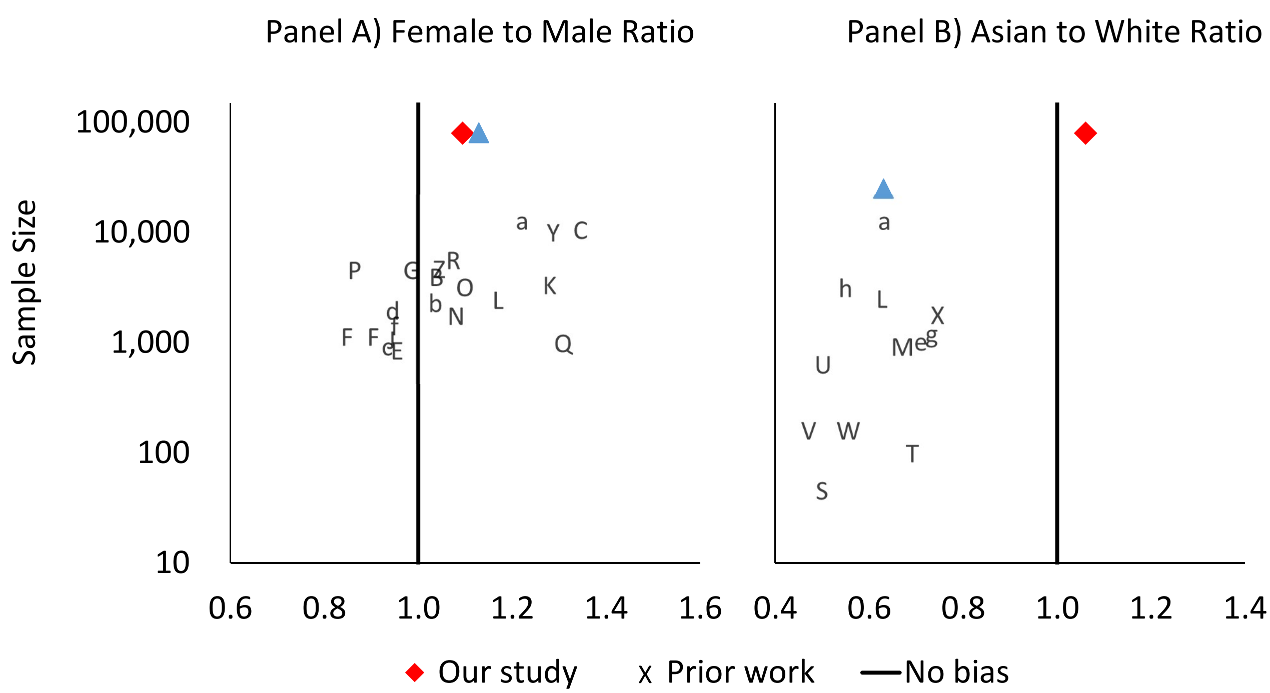
## 4 Comparison to Existing Empirical Evidence

Although to our knowledge there have been no audit studies on entrepreneurship or investor gender or racial preferences, there have been many in labor economics. These studies are not directly comparable to ours as none of them had investors as the decision-makers being studied and none of them is related to entrepreneurship. Despite that, prior work helps put our results in context.

In Figure 5, we compare our results with those of studies in the labor economics literature. Panel A (Panel B) shows the findings of 22 (13) correspondence studies with data on gender (race). Each panel provides a scatter plot, with the number of applications in each study on the vertical axis and the bias in favor or against females (Asians) on the horizontal axis. These figures include all of the labor correspondence studies mentioned in Riach and Rich (2002), Bertrand and Duflo (2017), or Baert (2018) that had usable data from the U.S., Canada, or Europe; used real decision-makers; and

[26] Our gender results remain significant if in-person and telephone meeting replies are excluded, while our race results lose statistical significance.

**Figure 5: Male/Female and Asian/White Biases in Labor Correspondence Studies.** This figure shows the relative reply rate of female to male (left) and Asian to European descent White (right) applicants from previous studies in the labor economics literature and compares those rates to the rates we observe. Each study is plotted based on the number of applications that have randomized gender or race assignment (vertical axis) and the ratio of reply rates (horizontal axis, female:male and Asian:White). The vertical line denotes zero bias and the large diamonds are this paper. We include data from correspondence studies in Bertrand and Duflo (2017), Riach and Rich (2002), and Baert (2018) that i) were from the U.S., Canada, or Europe, ii) used randomized assignment of male and female or Asian and European descent White, and iii) were reported in a manner to allow inclusion. The letters denote prior work with A denoting the sample size weighted average of the included studies; B Ahmed et al. (2013); C Albert et al. (2011); D Baert et al. (2016a); E Baert et al. (2016b); F Baert et al. (2016c); G Bailey et al. (2013); H Bartoš et al. (2016); I Berson (2012); J Bertrand and Mullainathan (2004); K Booth and Leigh (2010); L Booth et al. (2012); M Brown and Gay (1994); N Bursell (2007); O Carlsson (2011); P Drydakis (2010); Q Edo et al. (2019); R Eriksson and Rooth (2014); S Esmail and Everington (1993); T Esmail and Everington (1997); U Hubbuck and Carter (1980); V Jowell and Prescott-Clarke (1970); W McGinnity and Lunn (2011); X Midtbøen (2016); Y Neumark et al. (2019); Z Nunley et al. (2014); a Oreopoulos (2011); b Patacchini et al. (2015); c Petit (2007); d Riach and Rich (1987); e Riach and Rich (1991); f Weichselbaumer (2004); g Weichselbaumer (2015); and h Wood et al. (2009). Appendix C contains details on the construction of the figure.



had randomized female and male (Asian and European descent White) applicants.[27]

Our study is one of the largest correspondence studies ever conducted. Prior studies sent 46 to 40,000 inquiries and received between 18 and 6,900 responses, while we sent 80,000 emails and received 3,400 responses.[28] This scale gives us the power to detect small biases. Indeed, our findings of minor bias in favor of females and Asians are statistically significant only due to our large sample size and would not have shown up if our scale had been the same as most previous studies.

Female entrepreneurs getting a 9% higher interested response rate is quantitatively similar to the bias found by prior papers, which have a weighted average bias of 13% in favor of females. In many prior studies, these findings are statistically insignificant due to the aforementioned power issues. Our larger scale makes this effect statistically significant, despite its small size. Every correspondence study is subject to severe limitations, including our own, as discussed in Subsection 2.6. Keeping that in mind, a mild bias in favor of females appears to be a normal result of correspondence studies and our investors are not unusual in this regard.

Our finding of a bias in favor of Asian entrepreneurs, on the other hand, cannot be more different from the literature. Correspondence studies that look at race consistently find a strong bias *against* Asians, who receive 25–60% fewer replies than similarly qualified Whites. For example, Oreopoulos (2011) finds that even their Asian applicants with English first names, Canadian university degrees, and several years of Canadian work experience receive fewer callbacks that similar White job applicants. Against this backdrop, our finding of a small but significant bias *in favor* of Asian entrepreneurs is even more surprising. Thus, it is important to explore potential differences between our setting and the settings of the existing race studies.

[27] In Appendix C, we discuss each study. We use all data that are presented in a usable manner. For example, while the focus of Berson (2012) is ethnic discrimination, we use her data on White applicants of European descent to look at gender discrimination. We restrict our analysis to Whites of European descent because our own White entrepreneurs are in this group and previous research (e.g., see Baert (2018)) shows Whites of North African or Middle Eastern descent also face discrimination. We restrict our focus to the U.S., Canada, or Europe because we want countries similar to the U.S. where Whites of European descent are a majority group.

[28] In Figure 5, we consider the relevant inquiries, not the total number of inquiries. For example, Baert et al. (2016c) sent out 40,000 applications but only randomized gender for 10,000 of them.

## 5 Testing Theories of Discrimination

In this section, we discuss the mechanisms that could explain our results. These explanations fall into two rough groups: taste-based discrimination and statistical discrimination. In a taste-based discrimination framework (Becker, 1957), decision makers have non-financial reasons for preferring a particular group. In our context, investors would show a preference for a particular group of entrepreneurs, despite that group not delivering better financial returns. In a statistical discrimination framework (Phelps, 1972; Arrow et al., 1973; Aigner and Cain, 1977), discrimination arises due to a signal extraction problem. When information is limited, group membership may provide additional information about expected productivity and returns. In our context, investors would be inferring the quality of an investment based on the entrepreneur's race or gender.

In the rest of this section, we explore several plausible mechanisms for our results. We first discuss statistical mechanisms, then taste-based mechanisms, and finally attempt to rule out some alternative channels. Unfortunately, like most past correspondence studies, we struggle to distinguish between statistical and taste-based discrimination.

**Group Membership Perceived to Signal Entrepreneurial Quality.** The simplest statistical discrimination explanation for our results is that females or Asians are simply better entrepreneurs and investors use gender or race as proxy variables to infer information on entrepreneurial ability. This could arise through several channels.

First, these groups might be better entrepreneurs overall. Some tentative evidence points toward this. Ewens and Townsend (2018) find that "... male-led startups that male investors connect with on AngelList are actually less likely to have a successful exit than the female-led startups they connect with." In addition, Fairlie and Robb (2008) find evidence consistent with U.S. Asian-owned businesses being more successful than White-owned businesses.

Second, the specific hypothetical entrepreneurs we created could come across as stronger due to their race or gender. In all of our emails, the senders claimed to be graduate students at highly ranked universities and the co-founders and (explicitly or implicitly) CEOs of their start-ups. To the extent that group members face bias in achieving those goals, investors may take a positive signal from group membership, as in Fryer Jr (2007) or Bohren, Imas, and Rosenberg (2018). For example, investors may perceive it to be more difficult for females or Asians to get admitted into technical graduate programs or attract a team if others are biased.[29]

Third, if group members get less utility from entrepreneurship, they require higher thresholds on expected returns and overall expected benefits from entrepreneurship to become founders of start-ups. For example, if women perceive high-impact entrepreneurship to be hostile towards them, the only women who will pursue entrepreneurship will be those with the best opportunities. Investors might expect that and anticipate that the women who pursue ventures are of higher quality. Self-selection into entrepreneurship should thus lead to differential outcomes and lead investors to consider group membership as a signal of differentiated quality. Similar mechanisms could come from cultural pressures (e.g., Fernandez and Fogli (2009)) or taste differences (e.g., Buttner and Moore (1997)). Of course, a perception of differentiated characteristics may in itself lead to larger self-selection through self-fulfilling prophecies.

Finally, it is possible that some groups are shut out of traditional investor networks and that changes how unsolicited emails are perceived. Sending an unsolicited pitch email may suggest to investors that the sender was unable to find an interest for the company within their network, which would be a signal of low quality. However, for groups that are less likely to have access to strong professional networks, the signal might be weaker and less informative. Thus, receiving an unsolicited email pitch from somebody who may not have access to traditional networks sends less of a negative signal. This would be consistent with studies by Aldrich (1989), Renzulli, Aldrich, and Moody (2000), Shaw, Carter, and Brierton (2001), Abraham (2019), and others that suggest female business owners may have less diverse networks or more difficulty accessing them. More specifically, Orser, Riding, and Manley (2006) argue that females are less likely to seek external equity capital, Kwapisz and Hechavarria (2017) that they are less likely to ask for start-up financing more generally, and Howell and Nanda (2019) that they may find networking frictions a key barrier.

Importantly, all of these channels work equally well if investors' decision-making is driven by stereotypes

[29] This channel seems especially relevant for Asians, given that the timing of our study coincided with widely reported court proceedings against Harvard University for racial bias against Asians in college admissions. Case timeline taken from www.townandcountrymag.com/society/money-and-power/a24561452/harvard-lawsuit-affirmative-action-timeline/, accessed March 20,2019.

or incorrect beliefs, in which case we would generally classify them as taste-based discrimination. This may be an explanation for our result that female entrepreneurs get better results when pitching companies perceived to be more feminine. However, we observe a preference for female founders even when pitching companies perceived to be more masculine, which seems unlikely to be due to stereotypes, as past research suggests stereotypes in male-dominated, technical fields do not favor females (Bordalo, Coffman, Gennaioli, and Shleifer, 2016; Herbert, 2018).

**Groups Membership Perceived to Signal Other Characteristics.** Even if VCs do not interpret gender or race as signals of entrepreneurial quality, statistical discrimination can arise if they are interpreted as signals of other important characteristics. For example, investors may believe that other investors practice taste-based discrimination against some groups. Such discrimination would reduce the competition among investors for quality start-ups founded by members of those groups, which could lead to better negotiated outcomes for investors. As detailed in Section 1, there is currently a strong perception of bias against females in high-impact entrepreneurship, which makes this mechanism plausible for female founders. A related explanation would be that some groups are perceived to be worse negotiators. Mazei, Hüffmeier, Freund, Stuhlmacher, Bilke, and Hertel (2015) survey the literature and report that males achieve better economic outcomes in most types of negotiations than females, Hellmann et al. (2019) find that female teams on crowdfunding sites ask for too little; and Amatucci and Sohl (2004) report that many female entrepreneurs felt they should have asked for more in negotiations.

Another example is that investors may believe it is easier to screen, monitor, or work with start-ups whose founders are members of certain groups. For example, Shane, Dolmans, Jankowski, Reymen, and Romme (2012) find experimental evidence in the context of technology transfer offices of universities that licensing officers are positively disposed to Asian inventors, who are perceived as easy to work with, and negatively disposed to female inventors.

**Investors Seek to Support Group Members.** Bias in favor of females could arise from a sincerely-held goal of addressing gender inequality. Investors may be interested in supporting traditionally disadvantaged groups, such as females. In fact, many of the replies we received from both VCs and angels explicitly mention an interest in supporting female founders. A pro-female bias could also result from external pressure. For example, the Institutional Limited Partners Association recommends that investors in VC funds ask about female and minority representation as a part of the due diligence process.[30]

**Investors Are Acting Out of Romantic or Sexual Interest.** Although our initial emails were strictly business-oriented, some investors could be replying out of romantic interest or even the intent to engage in sexual coercion.[31] As the overall reply rate is low, even a small number of investors with these motivations could bias our results. It is difficult to rule out this hypothesis due to the lack of systematic data about who engages in this behavior or how they do so. However, the tests we can run do not provide support for this channel. First, none of the replies we received implied sexual or romantic intent. In fact, as Figure 4 shows, our results are driven by female senders being more likely to receive a request for a pitch deck (without asking for a call or a meeting). Asking for a pitch deck (rather than setting up a call or meeting) does not seem like a natural way to initiate sexual or romantic contact. Second, female investors replying to females are a large driver of the gender gap we observe. Finally, we find a significant preference (and in fact non-significantly stronger preference) for female entrepreneurs among out-of-state investors (e.g., an investor in New York messaged by an entrepreneur at a California-based university). Presumably, out-of-state investors would see less likelihood of romantic or sexual contact. Also note that we did not provide any images of our founders, which presumably limits the role of attraction.

**Homophily** Homophily refers to the tendency for people to seek out or be attracted to those who are similar to themselves. For example, male investors may be more interested in investing in male founders. Homophily could be a form of either statistical discrimination (e.g., male investors are better able to evaluate male founders) or taste-based discrimination (e.g., male investors prefer to spend time with male founders). Anecdotal evidence supports homophily. For example, Huang et al. (2017) find that "female investors place great importance on the gender of the founders they are considering investing in. 51% of women consider the founders' gender to be highly important, while only 6% of their male counterparts considered the founder's gender to be highly important." Unfortunately, and in common with previous research in venture capital, we have little power to

[30] Specifically, their Due Diligence Questionnaire asks "For investments made by the Firm during the last five years, what is the average percentage of female board members per company? Average percentage of minorities?" See ilpa.org/due-diligence-questionnaire, accessed March 20, 2019.

[31] Many female founders report harassment or discrimination. See, for example, fastcompany.com/90235322/special-report-2018-female-founders-speak-out or blog.ycombinator.com/survey-of-yc-female-founders-on-sexual-harassment-and-coercion-by-angel-and-vc-investors, both accessed June 26, 2019.

effectively test for homophily as our sample of investors is overwhelmingly White males. However, the coefficients we observe are suggestive. Female investors are 11% more interested in female entrepreneurs than male investors and Asian investors are 7% more interested in Asian entrepreneurs than White investors. Note that homophily cannot be the full story for our gender results, as male investors show a preference for female entrepreneurs.

## 6 Conclusion

In this paper, we provide the first experimental evidence on the gender and racial preferences of real investors. We employed a correspondence study design where fictitious entrepreneurs with randomly assigned gender and race sent 80,000 emails to 28,000 venture capitalists and angels. Emails sent by entrepreneurs with female names had 9% more interested replies than those with male names. Emails sent by entrepreneurs with Asian-sounding names had 6% more interested replies than those with White-sounding names. Both differences are statistically significant.

The measured 9% bias in favor of female entrepreneurs is economically significant; however, it is impossible to determine how it generalizes to warm introductions or the later stages of the investment process. Discrimination against females later in the investment process (as shown by Brooks et al. (2014)) would reverse this effect. Thus, our results should not be interpreted as showing females are not discriminated against overall in the startup fundraising process.

Despite this, our results are important given the stark underrepresentation of females in high-impact entrepreneurship and the anecdotal evidence on discrimination. In that context, our most important finding is perhaps not that we found a statistically significant bias in favor of females. Instead, it is that we can largely rule out discrimination *against* females by VCs and angels at a specific stage of the investment process, as can be discerned from their reply rates to unsolicited email pitches.

## References


Abraham, M., 2019. Gender-role incongruity and audience-based gender bias: An examination of networking among entrepreneurs. Administrative Science Quarterly p. 0001839219832813.

Agan, A., Starr, S., 2017. Ban the box, criminal records, and racial discrimination: A field experiment. The Quarterly Journal of Economics 133, 191–235.

Ahmed, A. M., Andersson, L., Hammarstedt, M., 2013. Are gay men and lesbians discriminated against in the hiring process? Southern Economic Journal 79, 565–585.

Aigner, D. J., Cain, G. G., 1977. Statistical theories of discrimination in labor markets. ILR Review 30, 175–187.

Albert, R., Escot, L., Fernández-Cornejo, J. A., 2011. A field experiment to study sex and age discrimination in the Madrid labour market. The International Journal of Human Resource Management 22, 351–375.

Aldrich, H., 1989. Networking among women entrepreneurs. Women-owned businesses 103, 132.

Alesina, A. F., Lotti, F., Mistrulli, P. E., 2013. Do women pay more for credit? Evidence from Italy. Journal of the European Economic Association 11, 45–66.

Amatucci, F. M., Sohl, J. E., 2004. Women entrepreneurs securing business angel financing: tales from the field. Venture Capital 6, 181–196.

Arrow, K., et al., 1973. The theory of discrimination. Discrimination in labor markets 3, 3–33.

Baert, S., 2018. Hiring discrimination: An overview of (almost) all correspondence experiments since 2005. In: *Audit studies: Behind the scenes with theory, method, and nuance*, Springer, pp. 63–77.

Baert, S., De Pauw, A.-S., Deschacht, N., 2016a. Do employer preferences contribute to sticky floors? ILR Review 69, 714–736.

Baert, S., De Visschere, S., Schoors, K., Vandenberghe, D., Omey, E., 2016b. First depressed, then discriminated against? Social Science & Medicine 170, 247–254.

Baert, S., Norga, J., Thuy, Y., Van Hecke, M., 2016c. Getting grey hairs in the labour market. An alternative experiment on age discrimination. Journal of Economic Psychology 57, 86–101.

Bagues, M. F., Esteve-Volart, B., 2010. Can gender parity break the glass ceiling? Evidence from a repeated randomized experiment. The Review of Economic Studies 77, 1301–1328.

Bailey, J., Wallace, M., Wright, B., 2013. Are gay men and lesbians discriminated against when applying for jobs? A four-city, internet-based field experiment. Journal of Homosexuality 60, 873–894.

Bartoš, V., Bauer, M., Chytilová, J., Matějka, F., 2016. Attention discrimination: Theory and field experiments with monitoring information acquisition. American Economic Review 106, 1437–75.

Becker, G., 1957. The Economics of Discrimination. University of Chicago Press.

Bernstein, S., Giroud, X., Townsend, R. R., 2016. The impact of venture capital monitoring. The Journal of Finance 71, 1591–1622.

Bernstein, S., Korteweg, A., Laws, K., 2017. Attracting early-stage investors: Evidence from a randomized field experiment. The Journal of Finance 72, 509–538.

Berson, C., 2012. Does competition induce hiring equity?, working Paper, Banque de France.

Bertrand, M., Duflo, E., 2017. Field experiments on discrimination. In: *Handbook of Economic Field Experiments*, Elsevier, vol. 1, pp. 309–393.

Bertrand, M., Goldin, C., Katz, L. F., 2010. Dynamics of the gender gap for young professionals in the financial and corporate sectors. American Economic Journal: Applied Economics 2, 228–55.

Bertrand, M., Mullainathan, S., 2004. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. American Economic Review 94, 991–1013.

Bohren, J. A., Imas, A., Rosenberg, M., 2018. The dynamics of discrimination: Theory and evidence, Working Paper, Carnegie Mellon University.

Booth, A., Leigh, A., 2010. Do employers discriminate by gender? A field experiment in female-dominated occupations. Economics Letters 107, 236–238.

Booth, A. L., Leigh, A., Varganova, E., 2012. Does ethnic discrimination vary across minority groups? Evidence from a field experiment. Oxford Bulletin of Economics and Statistics 74, 547–573.

Bordalo, P., Coffman, K., Gennaioli, N., Shleifer, A., 2016. Stereotypes. The Quarterly Journal of Economics 131, 1753–1794.

Brooks, A. W., Huang, L., Kearney, S. W., Murray, F. E., 2014. Investors prefer entrepreneurial ventures pitched by attractive men. Proceedings of the National Academy of Sciences 111, 4427–4431.

Brown, C., Gay, P., 1994. Racial discrimination: 17 years after the act. Equal employment opportunity: Labor market discrimination and public policy pp. 315–28.

Bursell, M., 2007. What's in a name?-A field experiment test for the existence of ethnic discrimination in the hiring process.

Buttner, E. H., Moore, D. P., 1997. Women's organizational exodus to entrepreneurship: self-reported motivations and correlates with success. Journal of Small Business Management 35, 34–46.

Carlsson, M., 2011. Does hiring discrimination cause gender segregation in the Swedish labor market? Feminist Economics 17, 71–102.

Darolia, R., Koedel, C., Martorell, P., Wilson, K., Perez-Arce, F., 2015. Do employers prefer workers who attend for-profit colleges? Evidence from a field experiment. Journal of Policy Analysis and Management 34, 881–903.

Drydakis, N., 2010. Labour discrimination as a symptom of HIV: Experimental evaluation-The Greek case. Journal of industrial Relations 52, 201–217.

Edelman, B., Luca, M., Svirsky, D., 2017. Racial discrimination in the sharing economy: Evidence from a field experiment. American Economic Journal: Applied Economics 9, 1–22.

Edo, A., Jacquemet, N., Yannelis, C., 2019. Language skills and homophilous hiring discrimination: Evidence from gender and racially differentiated applications. Review of Economics of the Household 17, 349–376.

Egan, M. L., Matvos, G., Seru, A., 2019. When harry fired sally: The double standard in punishing misconduct. Working Paper, Harvard Business School.

Eriksson, S., Rooth, D.-O., 2014. Do employers use unemployment as a sorting criterion when hiring? Evidence from a field experiment. American Economic Review 104, 1014–39.

Esmail, A., Everington, S., 1993. Racial discrimination against doctors from ethnic minorities. BMJ: British Medical Journal 306, 691.

Esmail, A., Everington, S., 1997. Asian doctors are still being discriminated against. BMJ 314, 1619.

Ewens, M., Townsend, R., 2018. Are early stage investors biased against women?, Working Paper, UCSD.

Fairlie, R. W., Robb, A. M., 2008. Race and entrepreneurial success: Black-, Asian-, and White-owned businesses in the United States. MIT Press Books 1.

Fernandez, R., Fogli, A., 2009. Culture: An empirical investigation of beliefs, work, and fertility. American Economic Journal: Macroeconomics 1, 146–77.

Fryer, R. G., Levitt, S. D., 2004. The causes and consequences of distinctively black names. The Quarterly Journal of Economics 119, 767–805.

Fryer Jr, R. G., 2007. Belief flipping in a dynamic model of statistical discrimination. Journal of Public Economics 91, 1151–1166.

Goldin, C., Rouse, C., 2000. Orchestrating impartiality: The impact of "blind" auditions on female musicians. American Economic Review 90, 715–741.

Gompers, P., Gornall, W., Kaplan, S. N., Strebulaev, I. A., 2018. How do venture capitalists make decisions? Journal of Financial Economics Forthcoming.

Gompers, P. A., Mukharlyamov, V., Weisburst, E., Xuan, Y., 2014. Gender effects in venture capital, Working Paper, Harvard Business School.

Gompers, P. A., Wang, S. Q., 2017. Diversity in innovation, Working Paper, Harvard Business School.

Hellmann, T., Mostipan, I., Vulkan, N., 2019. Be careful what you ask for: Fundraising strategies in equity crowdfunding, Working Paper, Oxford University.

Herbert, C., 2018. Mind the gap: Gender stereotypes and entrepreneurs' financing, Working Paper, Tillburg University.

Howell, S., 2017. Reducing information frictions in venture capital: The role of new venture competitions, Working Paper, NYU Stern.

Howell, S., Nanda, R., 2019. Networking frictions in venture capital, and the gender gap in entrepreneurship, Working Paper, Harvard Business School.

Huang, L., Wu, A., Lee, M. J., Bao, J., Hudson, M., Bolle, E., 2017. The American angel: The first in-depth report on the demographics and investing activity of individual American angel investors, Working Paper, University of Pennsylvania.

Hubbuck, J., Carter, S., 1980. Half a Chance?: A Report on Job Discrimination Against Young Blacks in Nottingham. Commission for Racial Equality London.

Jowell, R., Prescott-Clarke, P., 1970. Racial discrimination and white-collar workers in Britain. Race 11, 397–417.

Kwapisz, A., Hechavarria, D. M., 2017. External financing challenges in prosocial entrepreneurship. In: *Academy of Management Proceedings*, Academy of Management Briarcliff Manor, NY 10510, vol. 2017, p. 12697.

Lee, M., Huang, L., 2018. Gender bias, social impact framing, and evaluation of entrepreneurial ventures. Organization Science 29, 1–16.

Mazei, J., Hüffmeier, J., Freund, P. A., Stuhlmacher, A. F., Bilke, L., Hertel, G., 2015. A meta-analysis on gender differences in negotiation outcomes and their moderators. Psychological Bulletin 141, 85.

McGinnity, F., Lunn, P. D., 2011. Measuring discrimination facing ethnic minority job applicants: An Irish experiment. Work, Employment and Society 25, 693–708.

Midtbøen, A. H., 2016. Discrimination of the second generation: Evidence from a field experiment in Norway. Journal of International Migration and Integration 17, 253–272.

Milkman, K. L., Akinola, M., Chugh, D., 2012. Temporal distance and discrimination: An audit study in academia. Psychological Science 23, 710–717.

Neumark, D., Bank, R. J., Van Nort, K. D., 1996. Sex discrimination in restaurant hiring: An audit study. The Quarterly Journal of Economics 111, 915–941.

Neumark, D., Burn, I., Button, P., 2019. Is it harder for older workers to find jobs? New and improved evidence from a field experiment. Journal of Political Economy 127, 922–970.

Nunley, J. M., Pugh, A., Romero, N., Seals, R. A., 2014. An examination of racial discrimination in the labor market for recent college graduates: Estimates from the field. Auburn University Department of Economics Working Paper Series 6.

Oreopoulos, P., 2011. Why do skilled immigrants struggle in the labor market? A field experiment with thirteen thousand resumes. American Economic Journal: Economic Policy 3, 148–71.

Orser, B. J., Riding, A. L., Manley, K., 2006. Women entrepreneurs and financial capital. Entrepreneurship Theory and Practice 30, 643–665.

Patacchini, E., Ragusa, G., Zenou, Y., 2015. Unexplored dimensions of discrimination in Europe: Homosexuality and physical appearance. Journal of Population Economics 28, 1045–1073.

Petit, P., 2007. The effects of age and family constraints on gender hiring discrimination: A field experiment in the French financial sector. Labour Economics 14, 371–391.

Phelps, E. S., 1972. The statistical theory of racism and sexism. American Economic Review 62, 659–661.

Raina, S., 2017. VC financing and the entrepreneurship gender gap, Working Paper, University of Alberta.

Renzulli, L. A., Aldrich, H., Moody, J., 2000. Family matters: Gender, networks, and entrepreneurial outcomes. Social forces 79, 523–546.

Riach, P. A., Rich, J., 1987. Testing for sexual discrimination in the labour market. Australian Economic Papers 26, 165–178.

Riach, P. A., Rich, J., 1991. Testing for racial discrimination in the labour market. Cambridge Journal of Economics 15, 239–256.

Riach, P. A., Rich, J., 2002. Field experiments of discrimination in the market place. The Economic Journal 112, F480–F518.

Robb, A. M., Robinson, D. T., 2014. The capital structure decisions of new firms. The Review of Financial Studies 27, 153–179.

Shane, S., Dolmans, S., Jankowski, J., Reymen, I., Romme, G., 2012. Which inventors do technology licensing officers favor for start-ups? Frontiers of Entrepreneurship Research 32, 1.

Shaw, E., Carter, S. L., Brierton, J., 2001. Unequal entrepreneurs: Why female enterprise is an uphill business. Work Foundation.

Tinkler, J. E., Whittington, K. B., Ku, M. C., Davies, A. R., 2015. Gender and venture capital decision-making: The effects of technical background and social capital on entrepreneurial evaluations. Social Science Research 51, 1–16.

Weichselbaumer, D., 2004. Is it sex or personality? The impact of sex stereotypes on discrimination in applicant selection. Eastern Economic Journal 30, 159–186.

Weichselbaumer, D., 2015. Discrimination against migrants in Austria: An experimental study, IZA Discussion Paper.

Wood, M., Hales, J., Purdon, S., Sejersen, T., Hayllar, O., 2009. A test for racial discrimination in recruitment practice in British cities. Department for Work and Pensions Research Report 607.

Ye, J., Han, S., Hu, Y., Coskun, B., Liu, M., Qin, H., Skiena, S., 2017. Nationality classification using name embeddings. In: *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, ACM, pp. 1897–1906.

## Race, Gender, and Entrepreneurship: A Randomized Field Experiment On Venture Capitalists and Angels

## ONLINE APPENDIX

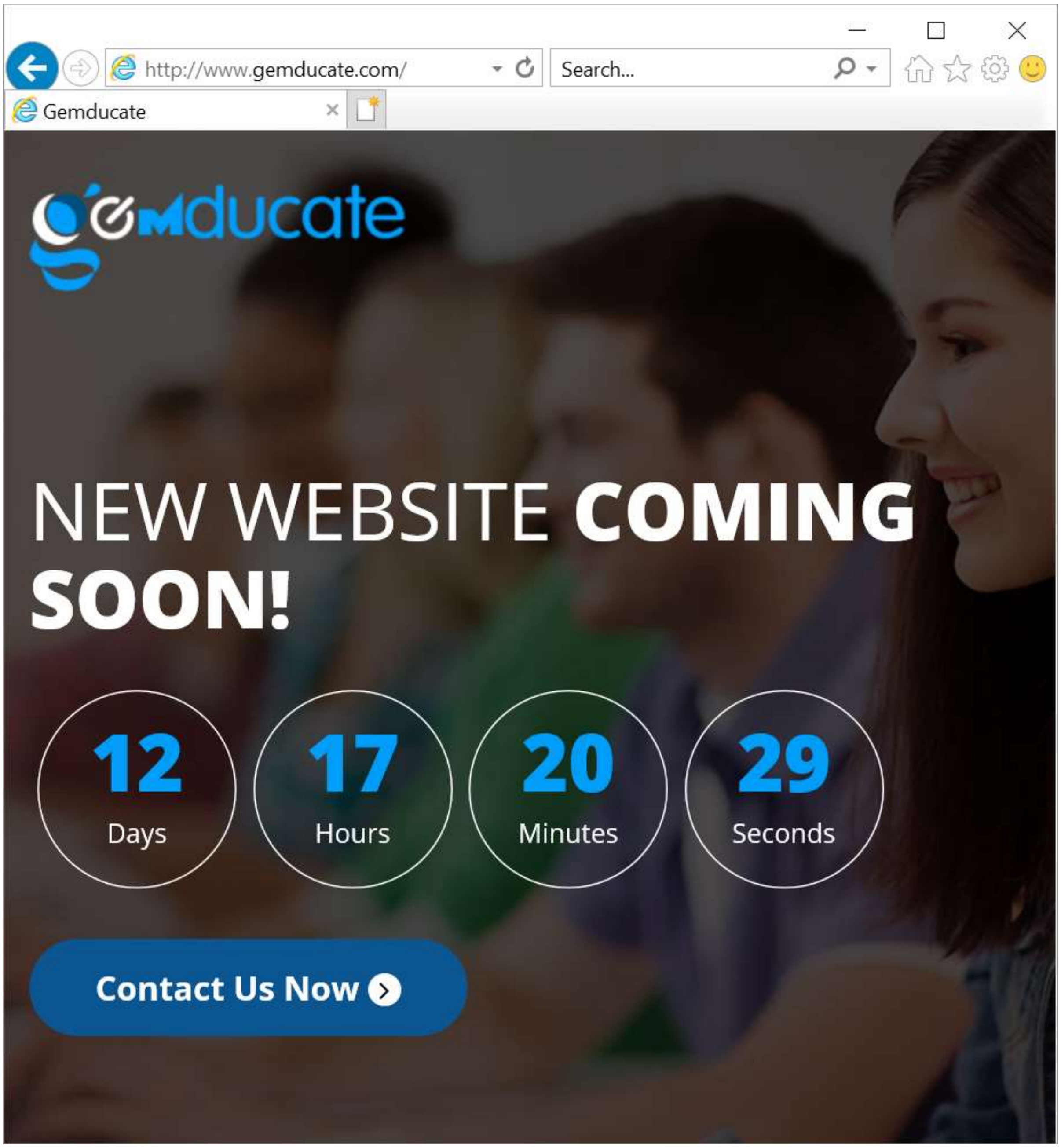
August 2019

**Will Gornall, Ilya A. Strebulaev**

## A Start-up and Entrepreneur Names Used

| Company | First Name | | Last Name | |
|---|---|---|---|---|
| | Male | Female | White | Asian |
| A-Assist | Richard | Erica | Hansen | Choi |
| AI.Boost | Vincent | Abigail | Welch | Li |
| Arcdatum Analytics | William | Danielle | Peterson | Ho |
| Asolution Encryption | Jack | Natalie | Hoffman | Chan |
| BE012 | Luke | Christine | Meyer | Chung |
| BetterCarbon Lab | Michael | Jessica | Schmidt | Pham |
| Breathe Freely | Stephen | Angela | Burke | Lin |
| Cartnnect | Jason | Michelle | Beck | Ho |
| Chatack | Evan | Nicole | Peters | Chen |
| Conveks Studio | Joseph | Catherine | Walsh | Chan |
| Dropfresh | Eric | Mary | Carpenter | Vang |
| DtoZ Workspace | Adam | Jennifer | Jensen | Liu |
| EIO Nail Solution | Patrick | Jacqueline | Keller | Yu |
| ERTech | Shane | Kathleen | Snyder | Nguyen |
| EVagle | Mark | Rachael | Myers | Vang |
| Fantasleep | Sean | Katie | Stone | Chung |
| Forawall | Jonathan | Patricia | Hoffman | Vang |
| Gemducate | Christopher | Samantha | Cohen | Huynh |
| GeneShop | Peter | Veronica | Schultz | Chang |
| Immunolyx Labs | Robert | Brenda | Barker | Lin |
| Innerlib | James | Melissa | Jensen | Le |
| Ituz Solution | Edward | Kathryn | Becker | Lin |
| Mantatrack | Ryan | Samantha | Burke | Nguyen |
| Marnificent | Steven | Stephanie | Schwartz | Huang |
| Min-D | Thomas | Margaret | Larson | Li |
| Mortar Technology | Brian | Amanada | Larson | Pham |
| MyHomeLab | Scott | Hayley | Schwartz | Choi |
| NLP Blic | Jeremy | Madeline | Miller | Huynh |
| Oldays Software | Bryan | Molly | Peterson | Le |
| Optzana | Justin | Vanessa | Cohen | Wu |
| Pandoor Lab | Kevin | Allison | Weaver | Wong |
| Penpall | Michael | Jessica | Hansen | Huynh |
| Pixtailor | Matthew | Emily | Myers | Chang |
| QmA Communication | Jeffrey | Katherine | Schultz | Kim |
| Quan Technology | Christopher | Lisa | Snyder | Chan |
| R2V | David | Megan | Beck | Pham |
| ReadMe | Samuel | Rachael | Miller | Wang |
| Riplex Marketplace | Ian | Grace | Sullivan | Wong |
| SereneTrip | Paul | Melanie | Keller | Huang |
| Simp Technology | George | Heather | Schmidt | Yu |
| Solopure Diagnostics | Marcus | Cynthia | Barker | Yang |
| SoundMap | Nicholas | Julie | Hansen | Liu |
| Soupred | Gregory | Rebecca | Walsh | Wang |
| Superior Kinetics | Nicholas | Julie | Carpenter | Yu |
| Syllo | Luke | Christine | Cohen | Huang |
| Tellanect | Trevor | Valerie | Carroll | Chen |
| Terach Property | Anthony | Caroline | Peters | Yang |
| Toolism | Andrew | Christina | Meyer | Wu |
| Vbun Systems | Derek | Karen | Stone | Chang |
| Videa Suite | Charles | Amy | Becker | Pham |

## B Example Company Website



# C Details on Included Correspondence Studies

This appendix details the construction of Figure 5. Our intent was not to conduct a formal meta-study but merely to compare our results to the prior literature. In Tables A.1 and A.2, we outline all of the data points used in the figure.

The general procedure was to collect callback ratios of female versus male or Asian versus European descent White applicants in past correspondence studies. We considered all studies referenced in Riach and Rich (2002), Bertrand and Duflo (2017), or Baert (2018). We used only data from Asian or European descent White applicants to jobs in the United States, Canada, or Europe. For example, Berson (2012) tested for labor market discrimination against Moroccans in France using four types of applicants: Moroccan males, Moroccan females, ethnic French males, and ethnic French females, and found respective reply rates of 2.40%, 3.92%, 11.28%, and 13.68% for those groups (refer to her Table 4). We use the ratio of the ethnic French female to ethnic French male reply rate for our gender plot ($1.21 = 13.68\%/11.28\%$) and do not add this to our race plot. We exclude Whites of non-European descent such as Moroccan or Turkish individuals because these groups also experience labor market discrimination and we do not want to confound our estimates. This excludes several studies that look at discrimination against North African and Middle Eastern individuals. We restrict our focus to the U.S., Canada, and Europe as our focus is on European descent Whites as a majority group and we want countries comparable to the U.S.

Many studies were not relevant for our charts. For example, studies such as Agan and Starr (2017) that only used one gender could not be used for our gender plot and the many studies that did not use Asian applicants could not be considered for our race plot. Similarly, data generated by sending male and female applicants to different types of jobs could not be used for our gender plot (e.g., Riach and Rich (1991)), unless there was randomization within certain job types. For example, Neumark et al. (2019) sent male and female applications to different sets of jobs, but had both apply to sales jobs, so we restrict our focus to sales jobs.

Not all studies included in the three literature reviews focus on ethnicity or gender discrimination; for example, some consider discrimination against gay or lesbian (Ahmed et al., 2013), older (Baert et al., 2016c), or long-term unemployed (Eriksson and Rooth, 2014) individuals. For these studies, we consider all applicants who are male or female and Asian or European descent White when calculating the callback ratios. For example, we add up the number of homosexual and heterosexual applicants and callbacks for females and males separately and calculate the callback ratios in Ahmed et al. (2013).

We cannot include studies that do not present data that allow us to infer the callback rates for our groups of interest. For example, Milkman et al. (2012) report their data in a way that makes it difficult to back out the underlying reply rates. We do include studies that present their data in the form of regressions (e.g., Nunley et al. (2014)), provided we can infer overall callback ratios from the regression coefficients.

**Table A.1: Details of Studies Whose Data on Male-Female Reply Differences were Reported in Figure 5**. This table presents details on each data point in Figure 5. The Code column presents the letter associated with each study; Study presents the citation for the study; Apps and Replies are the number of applications and replies that are relevant to our tests; Ratio is the ratio of reply rates of females and males among Asian and European descent White applicants; Setting gives the country and year of the experiment; Notes provides comments on our interpretation of the paper's data.

| Code | Study | Apps | Replies | Ratio | Setting | Notes |
|---|---|---|---|---|---|---|
| B | Ahmed et al. (2013) | 3,990 | 1,136 | 1.04 | Sweden, 2010 | Paper focuses on sexual orientation |
| C | Albert et al. (2011) | 10,620 | 931 | 1.34 | Spain, 2005-06 | |
| D | Baert et al. (2016a) | 1,152 | 120 | 0.90 | Belgium, 2013-14 | |
| E | Baert et al. (2016b) | 864 | 88 | 0.96 | Belgium, 2015 | Paper focuses on depression |
| F | Baert et al. (2016c) | 1,152 | 148 | 0.85 | Belgium, 2014-15 | Paper focuses on age |
| G | Bailey et al. (2013) | 4,608 | 575 | 0.99 | U.S., 2010 | Paper focuses on sexual orientation |
| I | Berson (2012) | 312 | 39 | 1.21 | France, 2011 | Paper focuses on ethnicity |
| J | Bertrand and Mullainathan (2004) | 1,077 | 93 | 0.94 | U.S., 2001-02 | Paper focuses on ethnicity. Only sales jobs used as other jobs did not have random gender assignment |
| K | Booth and Leigh (2010) | 3,365 | 962 | 1.28 | Australia, 2007 | |
| L | Booth et al. (2012) | 2,473 | 741 | 1.17 | Australia, 2007 | Paper focuses on ethnicity |
| N | Bursell (2007) | 1,776 | 684 | 1.08 | Sweden, 2006-07 | Paper focuses on ethnicity. Ratio of male to female respondents inferred from linear probability model in Table 6. |
| O | Carlsson (2011) | 3,228 | 968 | 1.10 | Sweden, 2005-07 | |
| P | Drydakis (2010) | 4,608 | 1,041 | 0.87 | Greece, 2007-08 | Paper focuses on HIV status |
| Q | Edo et al. (2019) | 1,008 | 173 | 1.31 | France, 2011-12 | Paper focuses on ethnicity |
| R | Eriksson and Rooth (2014) | 5,662 | 1,585 | 1.08 | Sweden, 2007 | Paper focuses on unemployment |
| Y | Neumark et al. (2019) | 10,055 | 2,155 | 1.29 | U.S., 2015 | Paper focuses on age. Only sales jobs used as other jobs did not have random gender assignment |
| Z | Nunley et al. (2014) | 4,698 | 846 | 1.05 | U.S., 2013 | Paper focuses on ethnicity. Ratio of male to female respondents inferred from linear probability model in Table 2. |
| a | Oreopoulos (2011) | 12,910 | 1,284 | 1.22 | Canada, 2009 | Paper focuses on ethnicity |
| b | Patacchini et al. (2015) | 2,320 | 256 | 1.04 | Italy, 2012 | Paper focuses on sexual orientation |
| c | Petit (2007) | 942 | 218 | 0.93 | France, 2002 | |
| d | Riach and Rich (1987) | 1,982 | 870 | 0.95 | Australia, 1983-86 | |
| f | Weichselbaumer (2004) | 1,431 | 699 | 0.95 | Austria, 1998-99 | |

**Table A.2: Details of Studies Whose Data on Asian-White Reply Differences were Reported in Figure 5**. This table presents details on each data point in Figure 5. The Code column presents the letter associated with each study; Study presents the citation for the study; Apps and Replies are the number of applications and replies that are relevant to our tests; Ratio is the ratio of the reply rates of Asian and European descent White applicants; Setting gives the country and year of the experiment; Notes provides comments on our interpretation of the paper's data.

| Code | Study | Apps | Replies | Ratio | Setting | Notes |
|---|---|---|---|---|---|---|
| H | Bartoš et al. (2016) | 183 | 55 | 0.40 | Czechia, 2009-10 | |
| L | Booth et al. (2012) | 2,517 | 738 | 0.63 | Australia, 2007 | |
| M | Brown and Gay (1994) | 932 | 411 | 0.67 | U.K., 1984-85 | Data from Riach and Rich (2002) |
| S | Esmail and Everington (1993) | 46 | 18 | 0.50 | U.K., 1992 | |
| T | Esmail and Everington (1997) | 100 | 44 | 0.69 | U.K., 1996 | |
| U | Hubbuck and Carter (1980) | 644 | 293 | 0.50 | U.K., 1977-79 | Data from Riach and Rich (2002) |
| V | Jowell and Prescott-Clarke (1970) | 160 | 106 | 0.47 | U.K., 1969 | |
| W | McGinnity and Lunn (2011) | 160 | 42 | 0.56 | Ireland, 2008 | |
| X | Midtbøen (2016) | 1,800 | 672 | 0.75 | Norway, 2009-10 | |
| a | Oreopoulos (2011) | 12,910 | 1,284 | 0.63 | Canada, 2009 | |
| e | Riach and Rich (1991) | 1,038 | 253 | 0.71 | Australia, 1984-88 | |
| g | Weichselbaumer (2015) | 1,222 | 421 | 0.73 | Austria, 2012-13 | |
| h | Wood et al. (2009) | 3,149 | 284 | 0.55 | U.K., 2008-09 | |

# D Additional Tests

**Table A.3: Additional Tests**. This table presents additional checks on the data. Interested reply rate to first email provides the rate of interested replies for the first email sent to each investor by Male, Female, White, and Asian entrepreneurs. The Interested reply rate before rows provide the rate of interested replies based only on emails sent and replies received before the specified dates. The Interested reply rate, in-state (Out-of-state) row provides the rate of interested replies for emails to investors who reside in the same (in a different) state from (to) the college the entrepreneur claims to attend. The Questioning reply rate row provides the percentage of emails sent by each type of investor that received questioning replies, as defined in Subsubsection 2.5.4. The Spam/fraud reply rate row similarly provides the rate of emails receiving allegations of spam or fraud. The Excluded from sample rate row provides the fraction of the total emails sent by each type of entrepreneur that were excluded from the sample (primarily due to bounced emails), as described in Subsection 2.4. This final row uses all emails sent as its denominator, while all other rows exclude emails that were not delivered. Tests of the null of equal reply rates are reported with * denoting significance at the 10% level, ** denoting significance at the 5% level, and *** denoting significance at the 1% level.

| | Gender | | | Race | | |
|---|---|---|---|---|---|---|
| | Male | Female | Difference | White | Asian | Difference |
| Interested reply rate to first email | 5.11 | 5.77 | 0.66 ** | 5.17 | 5.70 | 0.53 * |
| | (0.19) | (0.20) | (0.27) | (0.19) | (0.20) | (0.27) |
| Interested reply rate, before Nov. 15 | 3.79 | 4.18 | 0.40 ** | 3.80 | 4.17 | 0.37 ** |
| | (0.11) | (0.11) | (0.16) | (0.11) | (0.11) | (0.16) |
| Interested reply rate, before Nov. 10 | 4.09 | 4.73 | 0.64 *** | 4.23 | 4.60 | 0.36 * |
| | (0.14) | (0.15) | (0.20) | (0.14) | (0.15) | (0.20) |
| Interested reply rate, in-state | 5.43 | 5.70 | 0.26 | 5.33 | 5.80 | 0.47 |
| | (0.23) | (0.24) | (0.33) | (0.23) | (0.24) | (0.33) |
| Interested reply rate, out-of-state | 3.66 | 4.09 | 0.43 *** | 3.78 | 3.97 | 0.19 |
| | (0.11) | (0.11) | (0.16) | (0.11) | (0.11) | (0.16) |
| Questioning reply rate | 0.67 | 0.66 | -0.01 | 0.66 | 0.68 | 0.02 |
| | (0.04) | (0.04) | (0.06) | (0.04) | (0.04) | (0.06) |
| Spam/fraud reply rate | 0.03 | 0.06 | 0.03 * | 0.05 | 0.04 | -0.01 |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |
| Excluded from sample rate | 7.89 | 7.79 | -0.10 | 7.81 | 7.87 | 0.07 |
| | (0.13) | (0.13) | (0.18) | (0.13) | (0.13) | (0.18) |