Investors and the Entrepreneurship Gender Gap

Michael Ewens and Richard R. Townsend*

March 14, 2017

Abstract

It has been well-documented that women are underrepresented among entrepreneurs funded by early stage equity investors and these investors are predominantly men. In this paper, we explore whether investors play a role in explaining the gender gap in entrepreneurship. To do so we use a unique dataset obtained from AngelList, which allows us to observe detailed investor-founder interactions for a large sample of fundraising startups, some of which succeed in raising capital and some of which fail. We find that, even before investors get involved, only 16% of those seeking capital are women, suggesting that non-finance factors likely play a significant role in explaining the entrepreneurship gender gap. Among those who do seek capital, we find that female founders are roughly as successful as observably similar male founders in garnering attention and funding from investors. However, when we decompose our results by investor gender, we find evidence of segmentation: female (male) founders garner less interest from male (female) investors than male (female) founders, but this is offset by the fact that they garner more interest from female (male) investors. Additional tests suggest that the gender segmentation we observe is due to homopholistic preferences rather than within-gender screening/monitoring advantages. Regardless of the mechanism, our results suggest that an increase in female investors is likely necessary to support an increase in female entrepreneurship.

*Ewens: California Institute of Technology, mewens@caltech.edu. Townsend: University of California, San Diego, Rady School of Business, rrtownsend@ucsd.edu. The authors thank Kevin Laws of AngelList for graciously providing the data. The authors thank David Neumark and Sabrina Howell for comments. All errors are our own.

1 Introduction

It is well known that there is a significant gender gap in entrepreneurship. Recent studies of high-growth startup activity in the U.S. find that only roughly 10-15% of startups are founded by women (Tracy, 2011; Brush et al., 2014; Gompers and Wang, 2017). Many explanations for this phenomenon have been offered, including gender differences in entry into technical disciplines as well as differences in risk aversion.[1] According to such explanations, women drop off from entrepreneurial career paths long before they reach the point of seeking financing from a venture capitalist (VC) or angel investor. On the other hand, many have speculated that much of the gender gap may in fact be due to a lower propensity for investors to fund female entrepreneurs seeking capital. This view largely stems from the fact that over 90% of VCs are men (Gompers et al., 2014). For example, a recent article in the New York Times states that, "venture capitalists are, in a way, the gatekeepers to Silicon Valley, and if they are a group of white men [...] it is no wonder that most of the entrepreneurs fit the same mold."[2] In this paper, we examine whether female founders do in fact have more difficulty garnering interest and raising capital from investors than their male counterparts.

Examining this question rigorously has been difficult for several reasons. First and foremost, standard data sources only provide information on startups that have successfully raised capital, as it is challenging to systematically identify companies in the pre-financing stage. From these data, it is evident that women are dramatically under-represented among *funded* entrepreneurs. However, this under-representation does not necessarily point toward differential treatment of women by investors. In particular, it may be that women are just as under-representated in the entire pool of those seeking capital. Some have also found that, among funded entrepreneurs, female founders are

[1]See Marianne (2011) and Croson and Gneezy (2009) for surveys of empirical and experimental evidence of differences in risk attitudes by gender. For example, Bonin et al. (2007) find that risk preferences predict occupational sorting.

[2]https://www.nytimes.com/2015/04/02/business/dealbook/female-run-venture-funds-alter-the-status-quo.html?_r=0

more likely to pair with female investors (e.g., Marom, Robb, and Sade, 2016). However, this also does not necessarily indicate that male investors are reluctant to fund women. It may be that male investors see fewer companies with female founders due to the nature of their networks, but are no less likely to fund the female founders that they do see. In addition, investment is a two-sided decision in that it must both be offered by an investor and accepted by an entrepreneur. It may be that female founders garner equal interest from male and female investors, but are more likely to accept funding from female investors.

A second challenge is that female-led companies may differ from male-led companies in ways that make them less favorable investments on average. To the extent that such investment characteristics are unobservable in the data, but are observable to investors, it may appear that investors are reluctant to invest in women when in fact they are screening on other attributes. Moreover, if investors cannot observe these characteristics, but know that they correlate with gender, they may statistically discriminate against female entrepreneurs, which is distinct from taste-based discrimination.

In order to address these challenges, we use a proprietary dataset obtained from AngelList, a popular online platform started in 2010 that connects investors with seed stage startups. Companies create profiles on AngelList describing their businesses and founding teams. They can then start a fundraising campaign wherein they specify the amount of capital they are trying to raise along with other desired deal terms. Accredited investors—both angels and VCs—can register on the platform and subsequently connect with companies seeking funds. The site is widely used, even among high quality startups. By 2013, over 60% of companies raising a seed round had an AngelList profile and more than half of those firms attempted to raise capital through the site (Bernstein, Korteweg, and Laws, 2016). Many well-known companies, such as Uber and Pinterest, have raised capital through AngelList.

There are several advantages of this setting for studying the impact of gender on entrepreneurial fundraising. First, unlike much of the past work on this topic, we are not limited to studying companies that successfully raised capital. Instead, we observe a large set of companies that are trying to raise capital—some of which succeed and some of which fail. This allows us to characterize the population of founders seeking financing in a way that has not previously been possible, and to more directly examine whether gender appears to be an important determinant of fundraising success. Second, we not only observe whether an investor ultimately invests in a company, we also observe other investor actions. In particular, we see when an investor decides to "share" a company profile with someone else or "request an introduction" to the founders. As noted earlier, investment is a two-sided decision, but many of the other outcomes we are able to study are expressions of interest that only involve an action on the part of the investor. These actions also precede any personal interactions with founders that may differ across investors and thus complicate the analysis. Third, because of the nature of the platform, all investors have "access" to all deals in the sense that they can see the exact same information about the same set of companies and are free to take action on any company. Therefore, each investor's opportunity set is the same, at least with regard to the one-sided actions discussed above. Finally, we are also able to observe the gender of both the founders and the investors very accurately based on their name and profile picture. This feature of the data means that we can benchmark the behavior of male investors relative to that of female investors for the same set of companies. For example, if female-led companies tend to have unfavorable investment characteristics, one might expect both male and female investors to respond similarly to these characteristics.

We begin by characterizing the set of founders seeking capital on AngelList. We find that only 16% of the roughly 31,000 founders in our sample are female. It is worth noting that the barriers to posting an AngelList fundraising campaign are lower than almost any other form of fundraising. We view this as a strength of the data in the sense that it is unlikely that our sample is selected against including women. We also do not face survivorship bias, as we continue to observe companies even if they subsequently fail and remove their profile from the AngelList site. Thus, if anything, 16% likely represents an upper bound on the percent of innovation-driven entrepreneurs seeking capital that are women. This number is difficult to calculate from standard data sources, which track only companies that successfully raise capital.[3]

The fact that the 16% female share is significantly below the approximate 50% of women in the population, suggests that much of the entrepreneurship gender gap is present before any investor involvement. This would seem to point toward non-finance explanations as the primary drivers of the gender gap. As mentioned earlier, these explanations include differences in entry into technical disciplines as well as differences in risk aversion—both of which may arise from differential treatment of women by society at large or differences in education or experience (e.g., Gompers and Wang, 2017). However, it is also possible that many potential female entrepreneurs do not seek financing primarily because they are discouraged by what they understand to be the attitude of investors toward them as a group. Therefore, further analysis is necessary to understand the role of investors.

To that end, we next examine whether women seeking capital on AngelList do garner less interest from investors than men, controlling for observable characteristics. Perhaps surprisingly, we find small and statistically insignificant differences in the propensity of investors to share, request an introduction, or ultimately fund female-led companies as opposed to male-led companies. In contrast, other characteristics such as founder education and team size correlate strongly with investor interest in the direction one would expect.

Again, a major advantage of our setting is that we are able to observe both founder and investor gender. This feature of the data allows us to further examine whether female and male

[3] The only other potential sources of data on pre-finance stage companies would likely be incubators/accelerators or business plan competitions. However, these would likely provide less representative samples, as a layer of selection is applied prior to a company being accepted into either (or even prior to being eligible/invited to apply).

investors evaluate the same set of female entrepreneurs differently relative to the same set of male entrepreneurs. Interestingly, we find that female founders do garner less interest and funding from male investors. This is offset, however, by the fact that they garner more interest and funding from female investors. That is, there appears to be some degree of segmentation, wherein male investors are more likely to show interest and ultimately invest in male-led startups, while female investors are more likely to show interest and invest in female-led startups. Because the ratio of male entrepreneurs to male investors is roughly similar to the ratio of female entrepreneurs to female investors, it is possible for there to be no difference on average.

There are several potential explanations for gender segmentation in startup financing. One explanation is that female investors have a screening and/or monitoring advantage when it comes to female-led companies. For example, female-led companies may tend to operate in industries that are geared toward female customers, and female investors may have more expertise in these industries. A second possibility is that, setting screening and monitoring advantages aside, people simply tend to gravitate toward others who are similar to themselves. This tendency, known as "homophily," has been shown to exist in a variety of network settings. Here homophily would encompass outright sexism (in both directions), but also more subtle things like a desire on the part of investors to mentor founders that remind them of themselves.

Under any of the above explanations, our segmentation results are important in that they suggest that, in order to support an increase in female entrepreneurship, more female investors are likely to be necessary. That is, pre-fundraising interventions such as increased technical training for women may be important in closing the gender gap in entrepreneurship. However, our results suggest that the effect of such interventions may be limited without a simultaneous increase in female investors to fund the additional entrants.[4] This will be true regardless of whether investors tend to invest

[4]Coleman and Robb (2016) form similar conclusions in their review of the literature. Our work is the first large-scale evidence directly consistent with this explanation.

in entrepreneurs of the same gender due to a screening/monitoring advantage, or homopholistic preferences.

Disentangling these potential sources of segmentation empirically is challenging. Nonetheless, we attempt to do so to the extent possible with our data. We start by preforming a variety of tests to try to account for the possibility that female-led companies may have more female-oriented products.[5] First, we exclude startups in consumer-related industries, as well as startups that have self-assigned keyword tags associated with over 50% female founders (e.g. cosmetics, wedding, and childcare). We find similar segmentation results in both cases. We also find similar results when we control directly for the way a company describes itself on AngelList by including fixed effects for over 1,800 keyword tags. As a final way of excluding female-oriented companies, we also limit the sample to companies that male investors requested an introduction to. We again find that, even within this subsample, female investors show more interest in female founders. This last test also helps to account for a broader set of potential unobservable characteristics, including the possibility that female investors and founders are more likely to match because of similarity in risk preferences.

Finally, to further differentiate among different potential reasons for gender segmentation we examine ex-post company outcomes. If investors prefer to invest in founders of the same gender due to screening or monitoring advantages, we should expect to see same gender pairs outperform mixed gender pairs (Fisman, Paravisini, and Vig, 2015). On the other hand, if segmentation occurs due to homopholistic preferences, we should expect to see the opposite. In this case, investors essentially view it as more costly to invest in the opposite gender and therefore will only do so for the most promising companies. As a result, same gender pairs would underperform mixed gender pairs. In our data we find that companies associated with same gender pairs appear to underperform. They are more likely to fail and less like have an IPO or acquisition. Thus, our results overall seem to

[5]Marom, Robb, and Sade (2016) find that products and services available on Kickstarter posted by women are often in a few narrow categories.

point in the direction of homopholistic preferences rather than within-gender screening/monitoring advantages.

This paper contributes to a growing literature on gender and entrepreneurship. Many studies shown that women are extremely underpreresented among venture-backed entrepreneurs. Gompers and Wang (2017) find that just 10.7% of venture-backed founders were women from 2010–2015. Brush et al. (2014) estimate the number to be 15% using data from 2011-2013. Defining entrepreneurship more broadly, Tracy (2011) finds that 12.4% of "high-impact" firms with less than 20 employees in 2004–2008 were owned by women. In contrast to these papers, we are able estimate the female share both among those who successfully raise capital but also among all those seeking funding.

Other studies have shown that women are also underrepresented on the investor side. For example, Gompers et al. (2014) find that just 6.1% of VCs are women. This naturally begs the question of whether finance plays a role in the underpresentation of women among venture-backed entrepreneurs. Coleman and Robb (2009, 2016) find that women who do become entrepreneurs use less external equity financing and, possibly as a result, hire fewer employees and have slower businesses grow. Brooks et al. (2014) conduct a lab experiment in which the same entrepreneurial pitch is delivered by a woman and a man and then evaluated by non-investor experiment participants. They find that participants are significantly more likely to make mock investments in male entrepreneurs than female entrepreneurs delivering the same pitch. Our paper differs in that we study real investors making equity investments in real companies. In addition, we investigate whether the gender of the investor plays a role in how the gender of the participant is evaluated. Marom, Robb, and Sade (2016) study fundraising campaigns on the crowdfunding site Kickstarter. They find evidence that men are significantly less likely than women to back women-led projects.[6] Our analysis differs in

[6]Greenberg and Mollick (2016) provide an explanation using both lab and observational data for why women might perform better on these platforms.

that we study equity financing by angel and VC investors rather than rewards-based crowdfunding. Thus, we seek to understand the extent to which differential treatment of women by traditional investors plays a role in explaining the previously documented entrepreneurship gender gap. In contrast, Marom, Robb, and Sade (2016) seek to understand the extent to which the advent of crowdfunding may help to democratize access to capital by dramatically changing the composition and incentives of capital providers.

The rest of the paper proceeds as follows. Section 2 provides background about AngelList. Section 3 describes our data. Section 4 discusses our results. Section 5 concludes.

2 The AngelList Platform

Traditionally, seed-stage startup financing has been largely done through personal networks. Founders often seek capital from potential investors that they either know directly or indirectly through a mutual acquaintance. AngelList was founded in 2010 with the goal of making it easier for founders and investors to connect. Since launching, the platform has attracted much attention and grown rapidly in popularity, becoming an important part of the startup ecosystem. By 2013, over 60% of companies raising a seed round had an AngelList profile and more than half of these firms attempted to raise capital through the site (Bernstein, Korteweg, and Laws, 2016). Many well-known companies, such as Uber and Pinterest, have raised capital through AngelList.

The website allows founders to create startup profiles describing their idea, progress thus far, and personal/professional background. They can then start a fundraising campaign wherein they specify the amount of capital they are trying to raise along with other desired deal terms. Accredited investors—both angels and VCs—can register on the platform and subsequently connect with companies seeking funds. There are a variety of ways that an investor can interact with a startup. First, an investor can "follow" a startup, which adds the startup's updates to the investor's newsfeed. This allows the investor to easily keep track of a company of interest. Second, an investor can "share" a startup profile with someone else—either another AngelList user (through a private message) or someone off the platform (through an email with an embedded link). Investors often share deals with others that they know may be interested. Since multiple investors are frequently involved in a round of financing, sharing a deal also does not necessarily preclude the sharer from investing as well. Third, an investor can request an "introduction" to a startup. If the request is accepted, the investor can communicate directly with the founders and view confidential documents such as pitch decks, financials, or in depth business plans. Absent an introduction, communication is not possible, nor is full data access. Importantly, intro requests can only be made to startups with an active fundraising campaign. Thus, a request for an introduction can be viewed as a direct precursor to investment. Finally, an investor can "fund" a startup. This last step, however, happens offline, although founders can self-report consummated financing rounds in the funding section of their startup profile. In recent years, AngelList has also begun facilitating financings directly through the platform with equity crowdfunding syndicates. However, as of the time we obtained our data from AngelList, syndicates were still a fairly nascent addition to the site. Thus, we focus exclusively on the original "social network for startups" part of the platform as described above.

The only other paper we are aware of that uses AngelList's proprietary data is Bernstein, Korteweg, and Laws (2016). They examine how the likelihood of an investor visiting a startup's profile is affected by the inclusion or omission of certain categories of information from an email sent to investors highlighting the startup. The three categories of information they consider are the startup's founding team, its performance metrics, and its existing investors. They find that that omitting information about the founding team has the biggest negative impact on investor click-through rates from emails. Given that investors on AngelList find it important to see information about the founding team, it is plausible that characteristics like founder gender may play an important role in their decision-making. In contrast to Bernstein, Korteweg, and Laws (2016), we use the full AngelList dataset rather than focusing on the small set of companies featured in emails. We also study a broader set of investor actions that are more closely tied to actually making an investment than email clicks.

3 Data

In this section we describe our key variables, data sources, and sample conditions.

3.1 Investor-startup interactions

As described in Section 2, investors on AngelList can interact with startups in several ways that signal interest. We focus on investor sharing, requests for introductions, and investment. Data on sharing and intro requests come directly from AngelList. However, as described above, actual investments occur offline. Therefore, AngelList's data on investment is user-entered and somewhat incomplete.

Given this, we supplement AngelList's data with three additional sources. First, we match our sample with startups in Dow Jones' VentureSource database. This allows us to identify companies in our sample that eventually raised money from VCs. Second, we also match our sample to startups that report raising capital on Crunchbase. Crunchbase began as a wiki-style website to track IT startups primarily, however, it has recently evolved into a for-profit data provider for a larger set of entrepreneurial firms. Crunchbase's coverage is likely to be better than VenureSource for seed rounds with no institutional investor. Finally, to further ensure that we capture seed rounds as well a possible, we also match our sample with fundraising data gathered directly from SEC Form D filings. In principle, these filings are required for all private equity financings.[7]

[7] Matching with VentureSource and Crunchbase is based on a cleaned version of a startup's web domain. Matching with Form D filings is based on location, founding date, and company name.

Overall, the additional investment data sources do improve our coverage of fundraising outcomes. AngelList alone gives us a 7.8% investment rate. Using the additional data sources we find a 13.2% investment rate.[8] Nonetheless, for some of our analysis we will still rely mainly on the AngelList data. This is because it is easier to determine the gender of investors on AngelList. When the investor is listed as a firm rather than an individual—as is often the case in the other datasets—it is unclear how to assign gender. In addition, financings reported by founders on AngelList are more likely to be directly connected to the AngelList fundraising campaign. Finally, we supplement the investor gender variable for those startups that raised VC through their board members listed in VentureSource using the algorithm discussed below.

3.2 Startup outcomes

We focus on two measures of startup outcomes following a fundraising campaign. The first is an indicator equal to one if a startup has failed based on whether its website is no longer active as of November 2016. We deem a website as inactive if it fails to load and/or if its domain is available for purchase. The second measure of startup outcomes we consider is an indicator equal to one if a startup has had a successful exit via IPO or acquisition according to VentureSource or Crunchbase. Successful exits are quite rare in our sample. Some 5% of firms that successfully raised capital in our sample had a successful exit by November 2016. This is likely due to the fact that AngelList is relatively new, so even the high performing companies that originally raised capital through the site have not had enough time to have an IPO or acquisition.

[8] While this investment rate may sound low, it is considerably higher than the percent of companies that surveyed VCs typically report funding out of their opportunity set (e.g., Gompers et al., 2016). Of course, we are examining the probabiliy of raising capital from any investor rather than a particular one. We are not aware of any well-accepted benchmark for this probability.

3.3 Identifying gender

We assign a gender to founders and investors in our sample based on their name and profile picture. In particular, we run all first names through the genderize.io API, which gives the probability a first name is female based on a large sample.[9] For all individuals where the returned probability is less than 100%, we determine gender based on the user's profile picture. To do this, we use Crowdflower, which is a service like Amazon Mechanical Turk with additional quality controls. In particular, "test pictures" for which the correct answer has already been determined by us are randomly mixed in with pictures that have not been categorized. Crowdflower contributors who fail too many test questions are excluded, and the work of less trusted contributors is double-checked by more trusted contributors.

Many of the startups in our sample have a single founder, in which case it is straightforward to categorize a startup as "female-led" or "male-led" based on the gender of that founder. However, some of the startups in our sample have multiple founders. In these cases we categorize startups based on the gender of the founder who is also the CEO.

3.4 Non-gender founder characteristics

A founder's AngelList profile can include a short bio with information on their education and past work experience. However, often founders provide only sparse information about themselves on AngelList and instead use the option to link their AngelList profile to their LinkedIn profile. In addition, for some of the founders who do not link the two profiles, we are still able to find their LinkedIn profile manually by searching LinkedIn for their name along with the name of their startup. Therefore, for a large fraction our sample, we are able to obtain rich data on founders' educational

[9] http://genderize.io

and professional background from public LinkedIn profiles.[10]

In terms of education, we can observe the schools a founder attended, degrees obtained, and years of graduation. When we observe the year of college graduation, this provides a fairly accurate proxy for age. We crudely categorize founders as having attended an "elite" school if they hold a degree from a top-10 university according to the U.S. News & World Report rankings. In terms of work experience, we can observe the number of jobs held, past job titles, and number of years in the work force. We categorize founders as "serial entrepreneurs" if they held the title of founder at a different company prior to their AngelList fundraising campaign. Finally, we also observe the number of connections a founder has with other users on LinkedIn, censored at 500 (i.e. "500+"). Table 1 provides a full listing of these background variables.

3.5 Industry and location classification

Startups on AngelList describe themselves in part through various categories of keyword "tags." There are 1,805 distinct industry tags and companies can use multiple tags in combination to describe themselves. We map these tag combinations into VentureSource industry categorizations using the subsample of AngelList startups that also appear in VentureSource. For startups in the overlapping sample we already have both AngelList tags and VentureSource industries. For startups that are not in the overlapping sample (i.e. only in AngelList) we identify the nearest neihbors in the overlapping sample.[11] Based on these nearest neighbors we compute a probability distribution for each company over the seven major VentureSource industries.[12] We then categorize a company according to its most probable VentureSource industry. We also do the same using VentureSource 18-industry and 43-industry categorization schemes.

[10]Public profiles were searched and downloaded manually by an RA not an automated "scraper," in accordance with LinkedIn's terms of service.

[11]Nearest neihbors are startups with the highest number of common AngelList tags.

[12]Our results are similar whether we control directly for the industry probabilities, or assign according to the most probable.

Starups use 5,841 distinct location tags. We geocode these using the google maps API and then categorize them according to the 19-region scheme used by the National Venture Captial Association (NVCA). The NVCA regions are coarse where there are few startups and more granular where there are many. For example, there is one region in the Southwest, but four regions in California.

3.6 Final sample

The final sample of founders and startups satisfies several conditions that help to minimize measurement error and captures a representative set of startups seeking capital in our sample period. The sample begins with all first-time fundraising events for startups started between 2010 and November 2015. We remove non-US startups as our ability to track fundraising success is quite difficult in Europe or other regions. Next, we require that AngelList have a founding team where we could confidently identify the gender of each founder. Any startup that raised venture capital before our sample period is excluded to ensure we study first-time financings. The startup's fundraising campaign must also have a non-missing capital sought. Finally, we require that the startup maps to a VentureSource industry and NVCA region based on its tags. In the end we have 31,846 startups in the sample.

4 Results

We begin in Table 2 by examining the gender composition of entrepreneurs at different stages. As mentioned before, standard datasets do not cover those who have yet to successfully raise capital. This makes it impossible to assess the extent to which the gender gap that has been documented previously among *funded* entrepreneurs is also present among the pool of those seeking funding. In our data we can observe a large sample of entrepreneurs seeking funding and can follow them through the fundraising processes and ultimately to exit/failure. This allows us to get a sense of the point in the entrepreneurial process where women appear to drop out. We view these simple summary statistics as an important contribution in and of themselves. Overall, we find that only 16.2% of those who try to raise capital on AngelList are women. This suggests that, in fact, much of the gender gap is already present before investors even get involved. It should also be noted that the barriers to fundraising on AngelList are arguably lower than the barriers to any other type of fundraising. Therefore, we think this number likely represents an upper bound. That is, women are likely even more under-represented among those approaching investors in the traditional manner. This large pre-funding gender gap suggests that non-finance factors may account for much of the overall entrepreneurship gender gap (e.g., Gompers and Wang, 2017). However, it remains possible that many women would be interested in raising capital for an entrepreneurial venture but are discouraged by their perception that the fundraising environment for women is difficult and so do not even try.

It does appear that women who attempt to raise capital on AngelList are somewhat less successful in doing so than men. Women constitute approximately 13.5% of those who receive an intro request from an investor, those who have their profile shared by an investor, and those who ultimately raise capital. Thus, they are somewhat more under-represented among those groups than among the fundraising population, reflecting a lower success rate. In terms of company outcomes, very few AngelList companies have had an exit via IPO or acquisition at this point, partly due to the fact that the site remains relatively new. Of those few companies that have had a successful exit, 11.1% are female-led. Among those companies that remain in operation (i.e., have not failed) 15.9% are female-led. Overall, women appear to drop off at each stage but by far the biggest drop off seems to occur prior to attempting to raise capital. In other words, we see a decline from approximately 50% to 16.2% in the fundraising stage, but then much smaller declines from 16.2% thereafter.

Table 3 presents summary statistics for our sample of founders. Again, approximately 16% are female. Approximately 86% do not have a co-founder. The average (median) fundraising target is \$670,000 (\$300,000). The median company on the platform posted its first AngelList fundraising campaign fairly recently, in 2014. Most companies that post a fundraising campaign appear to generate little interest from investors. Only 8.6% have their company profile page shared by an investor, 10% receive a request for an introduction, and 9% are successful in raising capital from any source, with 5% raising capital from an AngelList user. While these numbers are low, they are consistent with common claims of investors that they do not seriously consider most of the deals they could potentially invest in (e.g. Gompers et al. (2016)). Some 43% of the founders in our sample report holding a bachelor's degree, 7% report holding an MBA, and 3% report holding some other advanced degree. These numbers are based on the information founders post on AngelList as well as LinkedIn. It is possible that actual educational attainment in our sample is higher than reported educational attainment if some founders that hold degrees choose to omit this information from their online profiles. Nonetheless, we interpret these variables as reflecting the information that was available to investors online at the time of the fundraising campaign. This is likely the information upon which investors decided to share or request an introduction to a company and thus is the appropriate information to control for in regressions where those are the outcome variables. In the process of actually funding a company, investors likely learn additional information from conversations with the founders. Thus, when fundraising success is our outcome variable, our ability to control for the information that investors had is more limited.

Table 4 presents further summary statistics, now separated by male and female founders. We find that male and female founders are similar on many dimensions e.g. number of co-founders, age, years of work experience, startup founding experience, number of LinkedIn connections, and education. The biggest difference appears to be that women on AngelList set lower fundraising targets than men. This finding is consistent with previous research (Marom, Robb, and Sade (2016) and Becker-Blease and Sohl (2007)). In terms of outcomes, female founders appear to be slightly less successful than men in garnering interest from investors in terms of profile shares, intro requests, and raised capital.

4.1 Interactions between investors and founders

We now explore whether founder gender correlates with investor interest and fundraising success in a regression framework. We begin by using investor sharing of a company profile as a proxy for interest. Despite the low cost of sharing on the platform, less than 10% of startups had an investor share their profile with another AngelList user. Our focus on the set of startups raising capital ensures that these sharing events coincide with information sharing about the capital raising event. In column 1 of Table 5, we include minimal controls. Specifically, we include fixed effects for the year the startup joined AngelList and the year it posted its first fundraising campaign. These fixed effects account for the fact that older companies have had more time to generate interest among investors. We find that, on average, female-led companies tend to be shared less by investors. In terms of magnitudes, the coefficient in column (1) suggests the the probability of a company being shared at least once by an investor is .68% lower if the company has a female founder. In column (2), we control for the amount of capital sought as well as team size, industry, and location fixed effects. With the inclusion of these controls, the coefficient on female flips signs and becomes statistically indistinguishable from zero. This change would suggest that women may tend to start companies that have less desirable characteristics to investors in terms of size, industry, and location; however, once these characteristics are taken into account, investors are just as likely to share a female-led company as a male-led company. The coefficient remains insignificant and reduces further in magnitude as we add additional controls for education and experience in column (3). Note that the education and experience variables go in the direction one would expect. Startups founded by college graduates are more likely to be shared, as are startups founded by individuals who hold a degree from an “elite” university, and startups founded by serial entrepreneurs. Thus, it does not seem that the lack of a significant coefficient on the female indicator is due to issues of power, as our sample is large enough to detect significant effects for other founder characteristics. One form of differential treatment across genders would be a differential response by investors to the same credentials for men and women. For example, one could imagine that women benefit less than men from having attended an elite university in terms of generating investor interest. Such differential treatment would be along the lines of the Bertrand and Mullainathan (2004) finding that employers are less responsive to resume quality for job applicants with African-American sounding names. To investigate whether a similar pattern holds in our setting, we allow the education and experience variables to interact with the female indicator in column (4). None of the estimated coefficients on these interaction terms are statistically significant. Thus, there is no evidence that educational credentials or experience are discounted for women. Finally, to ensure that our insignificant results are not driven by noise introduced by the presence of co-founders, in column (5) we limit our sample to only companies with a single founder. We find very similar results in this case.

While the sharing behavior of investors is interesting to examine, the way in which sharing relates to investment is unclear. It may be the case that observably similar female- and male-led companies are equally likely to be shared, but when it comes to actually raising capital, female-led companies do worse. To move one step closer to the funding stage, Table 6 examines investor requests for introductions. As discussed before, such requests are a direct precursor to funding, as investors need to request an introduction in order to communicate with a startup’s founder(s). We find very similar results to those in Table 5. Female-led companies are about 1.3% less likely to receive a request for an introduction, but once controls are added to the regression this effect goes away. Again, companies led by founders with a college degree, founders who attended an elite university are serial entrepreneurs are more likely to receive requests for introductions and there is no evidence that such credentials are discounted for women.

The lack of differential outcomes in sharing and introduction requests between observably similar female- and male-led companies does not preclude the possibility that women still have more trouble actually raising capital. While investment is perhaps more important than investor sharing or introduction requests, it is also more complex. Investment involves communication that is unobservable to the researcher, making it difficult to control for investors' information set. Investment is also a two-sided decision where an investor must make an offer and a founder must accept it. This means that investment could partially reflect the preferences of founders rather than investors. Nonetheless, in Table 7 we examine actual fundraising outcomes. We consider a startup to have successfully raised capital if subsequent to posting its AngelList fundraising campaign it raised a round according to their AngelList, CrunchBase, SEC filings or VentureSource profile. Again, we find results that are qualitatively similar to before. After controlling for observable firm, founder and financing characteristics, there is little difference—statistically and economically—between male and female-founded firms in fundraising success. Thus, the previous results were not driven by the preliminary or lower stakes nature of investor sharing and intro requests relative to actual investment.

4.2 Investor gender

Thus far it appears that women and men with similar characteristics are similarly successful in garnering interest and funding from investors. Up until this point, however, we have ignored investor gender. A major advantage of our setting is that we are able to observe gender on both sides of the potential transaction. It is natural to think that female and male investors may evaluate female entrepreneurs differently. In order to begin examining whether investor gender matters, we compare the characteristics of companies that garner interest from male and female investors. In Table 8 we divide the introductions made by male and female investors.[13] Overall, we find that companies that garner interest from men and women are similar in terms of fundraising success, team size, capital sought, and industry. Importantly, there is no large economic difference in startup outcomes by either of the measures at the bottom of the table, suggesting that male and female investors differ little in the ability to identify quality. The primary difference that emerges is that companies garnering interest from female investors are significantly more likely to have female founders.

Motivated by this observation we repeat the analysis of Tables 5–7, now decomposing the dependent variables by gender of the investor who made the share, introduction or provided the capital. The results are shown in Table 9. Differential treatment emerges with this decomposition. In particular, in the odd numbered columns, female-led companies are in fact significantly less likely be shared, to receive an introduction request, or to receive funding from male investors. Interestingly, however, they are significantly *more* likely to be shared, to have an intro request, or to receive funding from female investors. The magnitudes are approximately equal, which is why they more or less cancel out when the outcomes variables are not decomposed by gender. For example, female-led companies are approximately 4% less likely to be shared by a male investor and 4% more likely to be shared by a female investor. Compared to a baseline sharing probability of approximately 8.6% overall, these effects are economically large. Similarly, female-led companies are approximately 1% less likely to be receive an intro request from a male investor and 1% less likely to receive an intro request from a female investor, with a baseline intro request probably of approximately 10%. Finally, female-led companies are approximately 0.5% less likely to receive funding from a male investor and 0.3% less likely to receive funding from a female investor, with a baseline funding probability of 5%. It should be noted that our definition of funding in this case is fairly restrictive as we only look at rounds that were raised from AngelList users and that were subsequently reported on AngelList.

[13] If a company garners interest from both male and female investors they are included in both samples.

It is fairly common for companies to successfully raise capital subsequent to posting a fundraising campaign, but not to update their AngelList profile with this information (or disaggregate the individual investors or partners). Such funding rounds may be raised from Angels that are not on the AngelList platform or from VCs. We ignore these rounds for this analysis as it is more difficult to assign a gender to these non-AngelList investors, which are VC funds rather than individuals. These rounds are also less likely to be directly connected to the AngelList fundraising campaign.

4.3 Exploring mechanisms

The results in Table 9 suggest some degree of segmentation in startup financing. Male investors are more likely to show interest in and ultimately invest in male-led startups, while female investors are more likely to show interest and invest in female-led startups. These patterns cannot be explained by differences in the personal networks of male and female founders, because all investors on AngelList have access to all companies and presumably those who interact through the platform have no preexisting relationship. However, there are several other potential explanations. One explanation is that female investors have a screening and/or monitoring advantage when it comes to female-led companies. For example, female-led companies may tend to operate in industries that are geared toward female customers, and female investors may have more expertise in these industries. A second possibility is that, setting screening and monitoring advantages aside, people simply tend to gravitate toward others who are similar to themselves. This tendency, known as "homophily," has been shown to exist in a variety of network settings. Here homophily would encompass outright sexism (in both directions), but also more subtle things like a desire on the part of investors to mentor founders that remind them of themselves.

Under any of the above explanations, our segmentation results are important in that they suggest that, in order to support an increase in female entrepreneurship, more female investors are likely to be necessary. In particular, we showed earlier that most of the drop-off in female entrepreneurship occurs before the fundraising stage. This suggests that pre-fundraising interventions such as increased technical training for women may be important in closing the gender gap in entrepreneurship. However, our results suggest that the effect of such interventions may be limited without a simultaneous increase in female investors to fund the additional entrants. This will be true regardless of whether investors tend to invest in entrepreneurs of the same gender due to a screening/monitoring advantage, or homophilistic preferences.

Disentangling the above stories empirically is of course challenging. Nonetheless, we attempt to do so to the extent possible with our data. We begin by examining whether the gender segmentation we observe is due to a tendency for female-led companies to be ones that cater to female customers, resulting in a screening/monitoring advantage of female investors. Our baseline regressions include industry fixed effects, but it could be argued that our 7-industry classification scheme is too crude to entirely account for subtle differences in the products of companies with male and female founders. Thus, in Table 10, we make various attempts to limit our analysis to companies that lack any gender component in their products. We begin in Panel A by limiting our sample only to companies in industries that we classify as Non-Consumer by excluding those companies in "Consumer Services" or "Consumer Products." Presumably, it is less likely that these non-consumer companies are geared toward a specific gender. Nonetheless we find similar results in this subsample. Next, in Panel B we take advantage of the fact that we observe granular keyword tags describing a company's market. There are over 1,800 such tags and companies typically describe themselves using some combination of these. We exclude from the sample any company with a tag that we classify as female based on the fact that over 50% of companies using that tag have a female founder. This excludes companies that use tags such as "women," "cosmetics," "wedding," and "childcare." Again we find similar results using this restricted sample. We also find similar results even if we lower the female tag threshold to well-below 50% or if we classify tags manually using our own judgement. Thus, even among companies that do not appear to have a strong gender component in their products, female-led (male-led) companies garner more interest and funding from female (male) investors. Finally, in Panel C, we make use of all of the information contained in the keyword tags by using our whole sample and including a full set of tag fixed effects. Again we find similar results, meaning that segmentation seems to occur, even among companies that describe themselves similarly on AngelList.

It remains possible that, even among companies that lack an obvious gender component, or that describe themselves similarly, those with a female founder are still different in some way that generates interest from female investors. Importantly, since founders and investors do not interact personally prior to an intro request, there is no soft information in our setting. Any differences between male- and female-led companies that investors respond to would have to be ones that appear on the company profiles but which we have not controlled for in our regressions. One possible characteristic along these lines might be risk. If female-led companies somehow appear less risky to investors, and female investors prefer lower risk companies, that might at least partially explain the segmentation we observe. To examine stories of this nature, we limit our sample to only include companies that received an introduction request from a male investor. Again, most companies in our sample receive no such requests, so this is a very stringent sample restriction. To the extent that male investors avoid companies with certain kinds of unobservable characteristics that correlate with female founders (e.g., female oriented product, low growth potential, low risk, etc.), the companies in this subsample should be less likely to have those characteristics. If our baseline results are driven by female investors gravitating toward companies with unobservable female characteristics, we should find no effect of having a female founder in this subsample of companies that do not have such characteristics. However, as shown in Table 11, within this subsample we once again find that female investors are more likely to share or request an introduction to a female-led company, all else equal. The results for investment are weaker (likely due to power) but go in the same direction. Overall, these findings provide further evidence against our results being driven by difference in unobservable company characteristics.

4.3.1 Startup outcomes

Finally, to further differentiate among the different potential reasons for gender segmentation we examine company outcomes. If investors prefer to invest in founders of the same gender due to screening or monitoring advantages, we should expect to see same gender pairs outperform mixed gender pairs (Fisman, Paravisini, and Vig, 2015). On the other hand, if segmentation occurs due to homphilistic preferences, we should expect to see the opposite. In this case, investors essentially view it as costly to invest in the opposite gender and therefore will only do so for the most promising companies. As a result, same gender pairs would underperform mixed gender pairs. We face a few challenges in implementing this test. First, very few companies in our sample actually raise funding from an AngelList user and report the fundraising event on the platform. Therefore, we have a small sample of funded companies where we can identify gender on both sides of the transaction. To get around this issue, we define realized pairs more broadly based on introduction requests rather than only using investment. A second challenge is that very few companies in our sample have had a successful exit at this point in time, likely due to the recency of the platform. Recall that the median company in our sample first joined on AngelList in 2014. While early successes are relatively uncommon, early failures are not. Therefore, we focus primarily on startup failure as out outcome of interest.

The results are shown in Table 12. Observations are now at the startup-investor pair level and there is one observation for every realized pair based on intro requests. Standard errors are clustered by startup. The variable of interest "Same gender" is equal to one if the founder and investor are of the same gender. Each panel of the table reports a different company outcome variable. In Panel A the outcome of interest is startup failure. Column (1) pools all investor-company pairs and includes investor fixed effects. The coefficient on "Same gender" is positive and significant, suggesting same-gender pairs are more likely to fail. The investor fixed effects show that the same investor does worse when they investor is of the same gender as the founder. One concern might be that these results are driven my female investors requesting introductions to low quality female entrepreneurs, essentially as a form of charity. Given that there are relatively few women in the startup world, they may feel a greater responsibility than men to help each other out. However, columns 2–3 show that, if anything, the underperformance of same gender pairs is mainly driven by male-male pairs rather than female-female pairs. These columns consider the sample of male and female investors separately. Panel B repeats the same analysis using an IPO/Acquisition indicator variable as a measure of success. We find qualitatively similar results in this case, although the coefficients are not statistically significant. That is, the point estimates suggest that same gender pairs are less likely to be associated with an IPO or Acquisition. Overall, this evidence further points in the direction of homopholistic preferences as opposed to within gender screening or monitoring advantages.

5 Conclusion

Using a novel database of startup companies seeking capital, we find an under-representation of female-founded firms but little evidence of differential outcomes by male and female founders. The lack of differential treatment on average disappears when we consider the gender of the investors evaluating the startups. Here, male (female) investors exhibit preference for their same-gender counterparts across a variety of interactions. We explore reasons for this behavior including advantages in ex-ante screening, matching on risk preferences and post-investment interaction advantages. None of these explanations completely diminish the patterns observed. The results imply some form of homopholistic preferences among investors and suggest that interventions that seek to increase the pool of female entrepreneurs must coincide with an increase in the supply of female investors. Fortunately, the platform we study – AngelList – and others growing after the passing of the JOBs Act in 2012 have increased the "democratization" of early-stage capital. AngelList has few, if any, barriers to entry on both the company and investor-side which is only possible due to changes in the rules on crowdfunding and solicitation regulations.

References

Becker-Blease, John R and Jeffrey E Sohl. 2007. "Do women-owned businesses have equal access to angel capital?" *Journal of Business Venturing* 22 (4):503–521.

Bernstein, Shai, Arthur Korteweg, and Kevin Laws. 2016. "Attracting Early-Stage Investors: Evidence from a Randomized Field Experiment." *The Journal of Finance* .

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

Bonin, Holger, Thomas Dohmen, Armin Falk, David Huffman, and Uwe Sunde. 2007. "Cross-sectional earnings risk and occupational sorting: The role of risk attitudes." *Labour Economics* 14 (6):926–937.

Brooks, Alison Wood, Laura Huang, Sarah Wood Kearney, and Fiona E Murray. 2014. "Investors prefer entrepreneurial ventures pitched by attractive men." *Proceedings of the National Academy of Sciences* 111 (12):4427–4431.

Brush, Candida, Patricia G Greene, Lakshmi Balachandra, and Amy E Davis. 2014. "Diana Report: Women Entrepreneurs 2014–Bridging the Gender Gap in Venture Capital." *Arthur M. Blank Center for Entrepreneurship Babson College* .

Coleman, Susan and Alicia Robb. 2009. "A comparison of new firm financing by gender: evidence from the Kauffman Firm Survey data." *Small Business Economics* 33 (4):397–411.

———. 2016. *The Next Wave: Financing Women's Growth-Oriented Firms*. Stanford University Press.

Croson, Rachel and Uri Gneezy. 2009. "Gender differences in preferences." *Journal of Economic literature* 47 (2):448–474.

Fisman, Raymond, Daniel Paravisini, and Vikrant Vig. 2015. "Cultural proximity and loan outcomes." .

Gompers, Paul, William Gornall, Steven N Kaplan, and Ilya A Strebulaev. 2016. "How Do Venture Capitalists Make Decisions?" *NBER Working Paper* .

Gompers, Paul A, Vladimir Mukharlyamov, Emily Weisburst, and Yuhai Xuan. 2014. "Gender effects in venture capital." *Working paper* .

Gompers, Paul A and Sophie Q Wang. 2017. "Diversity in Innovation." *NBER Working Paper* .

Greenberg, Jason and Ethan Mollick. 2016. "Activist Choice Homophily and the Crowdfunding of Female Founders." *Administrative Science Quarterly* :0001839216678847.

Marianne, Bertrand. 2011. "New perspectives on gender." *Handbook of labor economics* 4:1543–1590.

Marom, Dan, Alicia Robb, and Orly Sade. 2016. "Gender Dynamics in Crowdfunding (Kickstarter): Evidence on Entrepreneurs, Investors, Deals and Taste-Based Discrimination." *Investors, Deals and Taste-Based Discrimination (December 6, 2015)* .

Tracy, Spencer L. 2011. "Accelerating job creation in America: The promise of high-impact companies." *SBA Office of Advocacy, Washington* .

6 Figures and Tables

Table 1: Variable definitions

Notes: The table describes the main variables used throughout the analysis.

Variable	**Definition**
Received share	An indicator equal to one if the startup received at least one share from an investor on the AngelList platform. A share allows an investor to point to a startup's profile or fundraising activity to another user on the platform.
Received introduction	An indicator equal to one if the startup received at least one introduction from an investor on the AngelList platform. An introduction is a one-way inquiry by investor to startup firm when the firm has an fundraising event listed on their profile. Direct communication between investor and startup is not possible on the platform without reciprocal action by the startup firm.
Funded (AL, CB, SEC, VS)	An indicator equal to one if the startup was observed as raising capital after their fundraising round on AngelList, Crunchbase, SEC filings or VentureSource.
Funded (AL)	An indicator equal to one if the startup was observed as raising capital after their fundraising round on AngelList.
Startup failed	An indicator equal to one if the startup did not have any online activity as of November 2016 that indicated the firm was still active. A firm is active if their website does not return and error and the returned page has some mention of the startup's name.
Had IPO or Acquisition	An indicator equal to one if the startup firm had an initial public offering or was acquired by November 2016. Such exits are observed in either Crunchbase or VentureSource.
Team size	The number of founders of the startup as listed in the startup's AngelList profile.
Year fundraised	The year the startup's fundraising first began on AngelList.
Capital sought	The original capital amount sought in the startup's first fundraising campaign on AngelList.
Serial founder	An indicator equal to one if the startup founder's LinkedIn profile indicated a past title as "founder" or "co-founder" at a another firm prior to the founding of the current startup.
Bachelors degree	An indicator equal to one if the startup founder's LinkedIn profile indicated that they received a bachelor's degree or equivalent. Such information is also found in the short founder biography on their AngelList profile and/or their Crunchbase profile if available.
MD / PhD / JD and MBA	An indicator equal to one if the founder had a PhD, MD or JD (MBA) in their AngelList or LinkedIn profile.
Elite school (any)	An indicator equal to one if any of the founder's pre-startup degrees were from any of the following universities: MIT, Princeton, UPenn, U. Chicago, Harvard, Yale, Caltech, John Hopkins, Duke, Stanford, Yale, Columbia or Northwestern.
Years experience pre-startup	A count of the number of years from the first observed job date to the founding of the startup as available on the founder's LinkedIn profile.
Number of LI connections	The number of LinkedIn network connections observed as of July 2016 on their LinkedIn profile page. The variable is truncated above by 500.
Age	The age in years of a startup founder as defined by the year of their college graduation (minus 21) observed on their LinkedIn profile or as reported directly on Crunchbase.

Table 2: "Conversion rates" by founder gender

Notes: The table reports the fraction of founders in our main sample—detailed in Section 3.6—that reached each "stage" of the possible stages that we study. The first row reports the initial distribution of founders by gender in our main sample. The percentages displayed report the proportion of founders active in that stage that are of each gender. The stages are defined in Table 1.

Stage	Men	Women
# sought capital	26,681 (83.8%)	5165 (16.2%)
Received share	2357 (86.5%)	366 (13.5%)
Received introduction	2879 (86.5%)	449 (13.5%)
Raised capital	2440 (86.5%)	380 (13.5%)
Startup failed	14,080 (83.4%)	2787 (16.5%)
Had IPO or Acq.	137 (88.9%)	17 (11.1%)

Table 3: Summary statistics for startups and founders

Notes: Table reports summary statistics for main variables used throughout the analysis (defined in Table 1). The unit of observation is founder-CEO by startup.

	mean	sd	min	p25	p50	p75	max	count
Female	0.16	0.37	0	0	0	0	1	31846
Team size (truncated at 4)	1.20	0.57	1	1	1	1	4	31846
Year fundraising start	2013.4	1.35	2009	2012	2014	2015	2015	31846
Solo founder	0.86	0.34	0	1	1	1	1	31846
Capital sought (millions)	0.67	0.84	0.0075	0.100	0.30	1	3	31846
Received share?	0.086	0.28	0	0	0	0	1	31846
Received introduction?	0.10	0.31	0	0	0	0	1	31846
Funded (AL)	0.052	0.22	0	0	0	0	1	31846
Funded (AL, CB, VS, SEC)	0.089	0.28	0	0	0	0	1	31846
Bach. degree	0.43	0.50	0	0	0	1	1	31846
MBA	0.069	0.25	0	0	0	0	1	31846
PhD/MD/JD	0.030	0.17	0	0	0	0	1	31846
Had IPO or Acquisition	0.0048	0.069	0	0	0	0	1	31846
Startup failed	0.53	0.50	0	0	1	1	1	31846

Table 4: Male-led vs female-led startups

Notes: Table reports differences in means, medians and standard deviations for the set of firm and founder variables by whether the founder (founder-CEO if multiple founders) is male or female. Variables are as defined in Table 1. The first panel reports the differences in observables in the full sample of founders. The second panel reports observables gathered from LinkedIn profile pages for the sub-sample of founders where we are able to locate a profile. % Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Male founder				Female founder			
	Obs	Mean	Median	Std dev	Obs	Mean	Median	Std dev
Team size (truncated at 4)	26,681	1.21	1.00	0.59	5,165	1.15	1.00	0.48
Year fundraising start	26,681	2013.39	2014.00	1.35	5,165	2013.55	2014.00	1.32
Solo founder	26,681	0.86	1.00	0.35	5,165	0.89	1.00	0.31
Capital sought (millions)	26,681	0.70	0.35	0.85	5,165	0.53	0.25	0.75
Received share?	26,681	0.09	0.00	0.28	5,165	0.07	0.00	0.26
Received introduction?	26,681	0.11	0.00	0.31	5,165	0.09	0.00	0.28
Funded (AL)	26,681	0.05	0.00	0.23	5,165	0.04	0.00	0.20
Funded (AL, CB, VS, SEC)	26,681	0.09	0.00	0.29	5,165	0.07	0.00	0.26
Bach. degree	26,681	0.43	0.00	0.50	5,165	0.45	0.00	0.50
MBA	26,681	0.07	0.00	0.26	5,165	0.06	0.00	0.24
PhD/MD/JD	26,681	0.03	0.00	0.17	5,165	0.03	0.00	0.16
Had IPO or Acquisition	26,681	0.01	0.00	0.07	5,165	0.00	0.00	0.06
Startup failed	26,681	0.53	1.00	0.50	5,165	0.54	1.00	0.50

LinkedIn Sample statistics

	Male founder				Female founder			
	Obs	Mean	Median	Std dev	Obs	Mean	Median	Std dev
Variable								
Serial founder	12732	0.33	0.00	0.47	2104	0.26	0.00	0.44
Number jobs on LinkedIn	13745	4.48	4.00	3.39	2257	4.45	4.00	3.35
Years experience pre-startup	12538	13.87	12.00	9.01	2043	13.03	12.00	8.24
Number LI connections	15513	300.21	393.00	214.06	2450	321.73	470.00	207.17
Age	5484	35.55	34.00	10.41	856	34.09	33.00	8.92

Table 5: Sharing and gender of founding team

Note: The table reports the linear probability model where the dependent variable is one received at least one share on AngelList by the end of the sample (11/2015) and were observed raising capital on the platform. A unit of observation is a US-based startup on the platform where we can identify the gender of all the founders and where the capital sought is at least $5000. Columns (1)–(4) include all founders in the sample. Column (5) consider just solo founders. All variables are as defined in Table 1 with the addition of the female indicator variable interacted with founder education variables. "Round year FE" are fixed effects for the year the financing opened. "Firm join year FE" are fixed effects for the year that the startup joined the AngelList platform. "Team size FE" are fixed effects for founding team size, which is truncated at five for teams larger than five. "Industry FE" are industry fixed effects defined in Section 3. "Location FE" are fixed effects. Robust standard errors reported in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Startup received share?				
					Solo
	(1)	(2)	(3)	(4)	(5)
Female	-0.00686*	0.00114	0.00120	-0.00810	-0.00567
	(0.00377)	(0.00372)	(0.00371)	(0.00737)	(0.00722)
Female X Previous founder				-0.000104	-0.0000783
				(0.0000790)	(0.0000771)
Female X Bachelors				0.00859	0.00366
				(0.00785)	(0.00753)
Female X Phd/MD/JD				-0.0215	0.00996
				(0.0270)	(0.0290)
Female X MBA				0.0299	0.0276
				(0.0214)	(0.0223)
Female X Elite school				-0.0266	-0.0306
				(0.0186)	(0.0187)
Previous founder			0.0248***	0.0248***	0.0211***
			(0.00543)	(0.00544)	(0.00552)
Bach. degree			0.0209***	0.0196***	0.0146***
			(0.00326)	(0.00362)	(0.00347)
PhD/MD/JD			-0.0171*	-0.0137	-0.00648
			(0.0104)	(0.0113)	(0.0115)
MBA			0.00699	0.00302	0.0121
			(0.00746)	(0.00804)	(0.00828)
Elite school (any)			0.0221***	0.0261***	0.0291***
			(0.00767)	(0.00858)	(0.00879)
Log capital sought		0.00377***	0.00206**	0.00208**	0.00217**
		(0.000847)	(0.000885)	(0.000884)	(0.000846)
Constant	0.519***	0.463***	0.450***	0.451***	0.468***
	(0.0549)	(0.0639)	(0.0635)	(0.0636)	(0.0736)
Observations	31846	31846	31846	31846	27459
R^2	0.0968	0.155	0.159	0.159	0.0860
Mean dep. var.	0.0855	0.0855	0.0855	0.0855	0.0595
Round year FE?	Y	Y	Y	Y	Y
Firm join year FE?	Y	Y	Y	Y	Y
Team size FE?	N	Y	Y	Y	Y
Industry FE?	N	Y	Y	Y	Y
Location FE?	N	Y	Y	Y	Y

Table 6: Introduction and gender of founding team

Note: The table reports the linear probability model where the dependent variable is one if the startup raising capital on AngelList received at least one introduction by the end of the sample period (11/2015). A unit of observation is a US-based startup on the platform where we can identify the gender of all the founders and where the capital sought is at least $5000. Variables and FEs are as defined in Table 1. Robust standard errors reported in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Startup received introduction?				
					Solo
	(1)	(2)	(3)	(4)	(5)
Female	-0.0134***	0.000573	0.000297	-0.00580	-0.0114
	(0.00428)	(0.00409)	(0.00407)	(0.00826)	(0.00801)
Female X Previous founder				-0.0000219	-0.0000534
				(0.0000887)	(0.0000854)
Female X Bachelors				0.00594	-0.000355
				(0.00871)	(0.00875)
Female X Phd/MD/JD				0.00722	0.0163
				(0.0341)	(0.0280)
Female X MBA				0.0400*	0.0471**
				(0.0243)	(0.0188)
Female X Elite school				-0.00861	-0.000706
				(0.0226)	(0.0182)
Previous founder			0.0306***	0.0306***	0.0252***
			(0.00607)	(0.00607)	(0.00497)
Bach. degree			0.0312***	0.0304***	0.0223***
			(0.00364)	(0.00406)	(0.00370)
PhD/MD/JD			-0.00298	-0.00375	0.0118
			(0.0122)	(0.0132)	(0.0106)
MBA			0.00545	-0.000189	0.00325
			(0.00827)	(0.00886)	(0.00728)
Elite school (any)			0.0415***	0.0421***	0.0342***
			(0.00877)	(0.00968)	(0.00748)
Log capital sought		0.00940***	0.00699***	0.00696***	0.00617***
		(0.000895)	(0.000938)	(0.000938)	(0.00101)
Constant	0.519***	0.411***	0.391***	0.392***	0.413***
	(0.0549)	(0.0633)	(0.0630)	(0.0630)	(0.0461)
Observations	31846	31846	31846	31846	27459
R^2	0.0408	0.149	0.155	0.155	0.0689
Mean dep. var.	0.105	0.105	0.105	0.105	0.0694
Round year FE?	Y	Y	Y	Y	Y
Firm join year FE?	Y	Y	Y	Y	Y
Team size FE?	N	Y	Y	Y	Y
Industry FE?	N	Y	Y	Y	Y
Location FE?	N	Y	Y	Y	Y

Table 7: Funding success and gender of founder

Note: The table reports the linear probability model where the dependent variable is one if the founder was observed successfully raising their first observed round of financing on AngelList, Crunchbase or VentureSource by the end of the sample (11/2015). A unit of observation is a US-based startup on the platform where we can identify the gender of all the founders and where the capital sought is at least \$5000. Variables are as defined in Table 1. All the "FE" are as defined in Table 5. Robust standard errors reported in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Startup raised round?				
	(1)	(2)	(3)	(4)	(5) Solo
Female	-0.0138***	0.00240	0.00238	-0.00385	-0.00355
	(0.00401)	(0.00383)	(0.00382)	(0.00805)	(0.00748)
Female X Previous founder				-0.0000813	-0.0000435
				(0.0000846)	(0.0000784)
Female X Bachelors				-0.000942	-0.00451
				(0.00819)	(0.00739)
Female X Phd/MD/JD				-0.0334	-0.0265
				(0.0332)	(0.0328)
Female X MBA				0.0365	0.0636***
				(0.0236)	(0.0245)
Female X Elite school				0.00804	0.00574
				(0.0216)	(0.0214)
Previous founder			0.0233***	0.0231***	0.0174***
			(0.00580)	(0.00581)	(0.00561)
Bach. degree			0.0171***	0.0173***	0.0133***
			(0.00341)	(0.00379)	(0.00348)
PhD/MD/JD			0.0337***	0.0390***	0.0407***
			(0.0124)	(0.0136)	(0.0140)
MBA			0.0182**	0.0127	0.00644
			(0.00804)	(0.00861)	(0.00824)
Elite school (any)			0.0228***	0.0207**	0.0205**
			(0.00824)	(0.00906)	(0.00886)
Log capital sought		0.0149***	0.0129***	0.0129***	0.0110***
		(0.000858)	(0.000894)	(0.000894)	(0.000824)
Constant	0.278***	0.219***	0.206***	0.206***	0.179***
	(0.0491)	(0.0592)	(0.0594)	(0.0593)	(0.0601)
Observations	31846	31846	31846	31846	27459
R^2	0.0149	0.124	0.129	0.129	0.0359
Mean dep. var.	0.0886	0.0886	0.0886	0.0886	0.0540
Round year FE?	Y	Y	Y	Y	Y
Firm join year FE?	Y	Y	Y	Y	Y
Team size FE?	N	Y	Y	Y	Y
Industry FE?	N	Y	Y	Y	Y
Location FE?	N	Y	Y	Y	Y

Table 8: Differences in firms that receive introductions from male and female investors

Notes: Table reports differences in means for the set of firm-level variables for all the introductions made by male and female investors. Variables are as defined in Table 1. The column "Diff/s.e." reports the difference between the startup and founder characteristics between these two sample of introductions. The stars report the p-value of the two-side t-test for differences in means between the two samples. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Male investor	Female investor	Diff/s.e.
At least one female founder	0.197	0.291	-0.0942***
			0.0146
# female founders	0.224	0.344	-0.120***
			0.0178
Received share?	0.758	0.735	0.0230
			0.0156
Funded (AL, CB, VS, SEC)	0.661	0.647	0.0140
			0.0172
Team size (truncated at 4)	2.009	2.026	-0.0168
			0.0353
Solo founder	0.370	0.360	0.00978
			0.0176
Capital sought (millions)	0.988	0.992	-0.00408
			0.0295
Information technology	0.347	0.300	0.0475**
			0.0173
Consumer goods/services	0.346	0.377	-0.0307
			0.0173
Startup failed	0.238	0.246	-0.00748
			0.0155
Startup had IPO/Acq.	0.0684	0.0679	0.000531
			0.00919
Number of introductions made	13,759	981	

Table 9: Shares, introductions and funds raised from female and male investors

Note: The table reports estimates from linear probability models where the dependent variable is one if the founder's firms received a share, introduction or capital by the end of sample. The dependent variable is split into an indicator for whether the interaction or capital came from at least one male or female investor. For the capital raised outcome, gender is assigned from the list of investors on their startup's AngelList profile page. A unit of observation is a US-based startup on the platform where we can identify the gender of all the founders and where the capital sought is at least \$5000. Variables are as defined in Table 1. All the "FE" are as defined in Table 5. Standard errors clustered at the startup reported in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Received share		Received intro.		Raised round	
	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.0411***	0.0456***	-0.00740*	0.0113***	-0.00440***	0.00366***
	(0.00266)	(0.00308)	(0.00385)	(0.00231)	(0.00133)	(0.00111)
Previous founder	0.0212***	0.00712***	0.0294***	0.0130***	0.00700***	0.00232
	(0.00517)	(0.00265)	(0.00594)	(0.00329)	(0.00266)	(0.00143)
Bach. degree	0.0189***	0.00381**	0.0296***	0.00724***	0.00515***	-0.0000659
	(0.00305)	(0.00149)	(0.00355)	(0.00179)	(0.00149)	(0.000737)
PhD/MD/JD	-0.0204**	0.00548	-0.00613	-0.000366	-0.000845	-0.00278
	(0.00973)	(0.00561)	(0.0119)	(0.00686)	(0.00546)	(0.00232)
MBA	0.00240	-0.00101	0.00517	-0.00207	-0.00117	0.00167
	(0.00704)	(0.00342)	(0.00808)	(0.00446)	(0.00364)	(0.00203)
Elite school (any)	0.0222***	0.0000659	0.0422***	0.0150***	0.00356	0.00130
	(0.00732)	(0.00351)	(0.00860)	(0.00497)	(0.00380)	(0.00196)
Log capital sought	0.00156*	0.000168	0.00656***	0.00254***	0.00132***	0.000503***
	(0.000809)	(0.000393)	(0.000915)	(0.000406)	(0.000343)	(0.000193)
Constant	0.436***	0.0629*	0.392***	0.170***	0.0495*	-0.00965***
	(0.0638)	(0.0346)	(0.0629)	(0.0464)	(0.0298)	(0.00171)
Observations	31846	31846	31846	31846	31846	31846
R^2	0.145	0.0495	0.152	0.0651	0.0330	0.0105
Round year FE?	Y	Y	Y	Y	Y	Y
Firm join year FE?	Y	Y	Y	Y	Y	Y
Team size FE?	Y	Y	Y	Y	Y	Y
Industry FE?	Y	Y	Y	Y	Y	Y
Location FE?	Y	Y	Y	Y	Y	Y

Table 10: Shares, intros and funds raised from female and male investors: Industry differences and robustness

Note: The table repeats the estimation from Table 9 for sub-samples of entrepreneurial firms split by industry classification or alternative fixed effect specifications. In Panel A, we consider all firms that are not in the “Consumer Products” or “Consumer Services” industry categories. Panel B presents the subset of firms that do not have one of the tags most commonly associated with female-founded firms (at least 50% of the tag has a female founder): lingerie, material science, mom, women, pet sitting, hair extensions, bridal community, women’s apparel and accessories, childcare, women-focused, mothers, beauty, hairsalon, natural skin care, child care, wedding, sex, natural food grocers, ethical manufacturing, organic, sex industry, behavioral therapy, aba therapy, autism, psychology, pet care, health foods, cosmetics, spas, elderly, flowers, gifts, vintage clothing, jewelry, teaching stem concepts, baby and kids, babies, natural food manufacturers, deep learning, patient management, underserved children, female 18–34, homeless shelter, open data, happiness, minorities, and bakeries. Panel C presents the main specification where the industry fixed effects are replaced by fixed effects for all the tags used by the startup firm on their AngelList profile. Startups can have more than one tag. All regressions include the controls found in Table 9 with the exception of Panel B that replaces industry FE with tag FE. Robust standard errors reported in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Panel A: Non-Consumer industries					
	Received share		Received intro.		Received capital	
	Female	Male	Female	Male	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.0589***	-0.0568***	0.0124***	-0.0127	0.00131	-0.00812***
	(0.00626)	(0.00498)	(0.00475)	(0.00775)	(0.00195)	(0.00279)
Observations	13406	13406	13406	13406	13406	13406
R^2	0.0563	0.158	0.0733	0.155	0.0103	0.0351
Mean dep. var.	0.0169	0.0935	0.0299	0.135	0.00470	0.0199
	Panel B: Excluding most popular tags of female founded firms (see note)					
	Received share		Received intro.		Received capital	
	Female	Male	Female	Male	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.0569***	-0.0511***	0.0132***	-0.00934*	0.00461***	-0.00582***
	(0.00414)	(0.00356)	(0.00311)	(0.00514)	(0.00151)	(0.00178)
Observations	24944	24944	24944	24944	24944	24944
R^2	0.0542	0.151	0.0664	0.146	0.0106	0.0333
Mean dep. var.	0.0167	0.0885	0.0262	0.121	0.00425	0.0166
	Panel C: Inclusion of industry tags as controls					
	Received share		Received intro.		Received capital	
	Female	Male	Female	Male	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.0470***	-0.0437***	0.0149***	-0.00528	0.00313**	-0.00710***
	(0.00365)	(0.00358)	(0.00298)	(0.00495)	(0.00135)	(0.00180)
Observations	26355	26355	26355	26355	26355	26355
R^2	0.209	0.246	0.211	0.239	0.134	0.186
Founders	25750	25750	25750	25750	25750	25750
Mean dep. var.	0.0170	0.0852	0.0255	0.117	0.00372	0.0161
FE?	Y	Y	Y	Y	Y	Y

Table 11: Female investor interest and funding among startups with at least one male investor introduction or share

Notes: This table reports the linear probability estimates of female investor interest for a the subsample of startups that received at least one introduction or share from a male investor. A unit of observation is a US-based startup on the platform where we can identify the gender of all the founders and where the capital sought is at least $5000. Variables and fixed effects are as defined in Table 5. Robust standard errors reported in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Received from a female investor:		
	Share	Intro	Funding
	(1)	(2)	(3)
Female	0.243***	0.0484**	0.0197**
	(0.0213)	(0.0189)	(0.00901)
Previous founder	0.0185*	0.0305**	0.00515
	(0.0101)	(0.0137)	(0.00591)
Bach. degree	0.00901	0.0143	-0.00158
	(0.00773)	(0.0111)	(0.00449)
PhD/MD/JD	0.0424*	-0.0118	-0.0152*
	(0.0224)	(0.0276)	(0.00847)
MBA	-0.00682	-0.0208	0.00773
	(0.0125)	(0.0178)	(0.00866)
Elite school (any)	-0.0215*	0.0410**	0.00820
	(0.0123)	(0.0186)	(0.00846)
Log capital sought	0.0000296	0.0224***	0.00316*
	(0.00254)	(0.00373)	(0.00177)
Constant	0.185	0.213*	-0.0397***
	(0.135)	(0.128)	(0.00941)
Observations	4249	4249	4249
R^2	0.130	0.0851	0.00987
Mean dep. var.	0.0631	0.133	0.0184
Round year FE?	Y	Y	Y
Firm join year FE?	Y	Y	Y
Team size FE?	Y	Y	Y
Industry FE?	Y	Y	Y
Location FE?	Y	Y	Y

Table 12: Differences in outcomes by investor gender: introductions

Notes: The table reports linear probability model estimates for the dependent variables defined in Table 1. Here a unit of observation is a pair: investor requesting an introduction and startup founder. That is, for each observed introduction, the variable "Same gender" is equal to one if the founder and investor have the same gender. Column (1) in each panel considers all observed pairs, while columns (2) and (3) consider the male and female investor sample separately. Each panel uses the dependent variables measuring startup success as defined in Table 1. Robust standard errors clustered at the startup are reported in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Panel A: Startup failed		
		Investor gender	
	All	Male	Female
	(1)	(2)	(3)
Same gender	0.0664**	0.0729**	-0.0208
	(0.0306)	(0.0358)	(0.0681)
Constant	0.174	0.182	0.548
	(0.211)	(0.241)	(0.333)
Observations	11872	11062	810
R^2	0.0730	0.0767	0.0444
Mean dep. var.	0.761	0.762	0.754
	Panel B: Startup had IPO/Acq.		
	All	Male	Female
	(1)	(2)	(3)
Same gender	-0.0138	-0.0160	-0.00234
	(0.0322)	(0.0375)	(0.0390)
Constant	0.0972	0.104	-0.0115
	(0.104)	(0.102)	(0.231)
Observations	11872	11062	810
R^2	0.0795	0.0795	0.0678
Mean dep. var.	0.0684	0.0684	0.0679
Investor FE?	Y	Y	Y
Year FE ?	Y	Y	Y
Industry FE ?	Y	Y	Y
Location FE ?	Y	Y	Y