Gender and Mentorship in Entrepreneurship *

Chuan Chen† Barton H. Hamilton‡

Prasanthi Ramakrishnan§

August 22, 2022

Abstract

We study the role informal mentoring can play in bridging the gender gap in venture financing. Mentors can play two roles: financial and role model. Using data from Global Entrepreneurship Research Network (GERN), we document the characteristics of the matched pairs, their funding patterns and then estimate the relative 'value' of a match using Fox (2018) model. Female-Female matches benefit more from the role-model aspect, compared to Male-Male matches. Close to 20 percent of the gains from homophily is from financial aspect for males; this number is <1 percent for females.

---

*We thank Global Entrepreneurship Research Network for providing us with the data. All analysis, interpretations, and conclusions drawn from the data are entirely and solely of our own. We thank Center for Research in Economics and Strategy (CRES), in the Olin Business School, Washington University in St. Louis for their support. We thank Costas Meghir and participants at the 14th Economics Graduate Student Conference, 2020 Australian Gender Economics Workshop, 2020 Academy of Management Conference and Work, Family, Public Policy Seminar at Washington University in St. Louis for their comments. We thank Adam Yu and Alex Martyn for their excellent research assistance.

†Wisconsin School of Business, University of Wisconsin-Madison. Email: chuan.chen@wisc.edu.

‡Olin Business School, Washington University in St. Louis, Campus Box 1133, One Brookings Drive, St. Louis, MO 63130. Email: hamiltonb@wustl.edu.

§Department of Economics, Washington University in St. Louis, Campus Box 1208, One Brookings Drive, St. Louis, MO 63130. Email: p.ramakrishnan@wustl.edu.

1 Introduction

Women face barriers to success in many occupations – these obstacles are particularly acute in entrepreneurship. In 2018, 40 percent of all firms were owned by women, but female-founded startups raised only 2.4 percent of total venture capital (VC) funding invested that year (Pitchbook, 2019).[1] Recent literature suggests that financing frictions associated with new venture fundraising are a key reason for this gap (Howell and Nanda, 2019). This is related to gender differences in access to certain institutional or social networks (Greene et al., 2001; Jackson, 2021; Snellman and Solal, 2022).

In this paper, we examine whether mentors, who are experienced entrepreneurs and can bring knowledge and resources for first-time founders, can mitigate the gender gap in entrepreneurship [2]. While existing literature has focused primarily on formal mentoring[3], where mentors are assigned to one other based on some application process, there is a need to differentiate amongst different mentoring relationships (Mullen and Klimaitis, 2021). We focus on informal mentoring, which is based on the mutual identification and fulfillment of career needs, and is often more effective in providing mentoring functions than formal ones (Ragins and Cotton, 1999). Moreover, many papers highlight that gender composition influenced mentoring functions and outcomes [4].

---

[1] Including companies with both male and female founders raises this number to 10.4 percent (Pitchbook, 2019).

[2] See Athey, Avery and Zemsky (2000); Blau et al. (2010); Carrell, Page and West (2010); Hilmer and Hilmer (2007); Kofoed et al. (2019); Neumark and Gardecki (1998); Porter and Serra (2019) for analysis on gender gap and mentoring

[3] See Blau et al. (2010); Carrell, Page and West (2010); Falk, Kosse and Pinger (2020); Kofoed et al. (2019) for effects of formal mentoring programs.

[4] See Ibarra (1992); McGinn and Milkman (2013); McPherson, Smith-Lovin and Cook (2001); Ragins and Cotton (1999) for a preview of the literature on homophily.

We condense the different functions that mentor can play into 2 broad categories[5]. First, they aid in career development by providing access to networks - in our case, this translates into access or introductions to start-up financing networks and will be referred to as the financial aspect of mentoring. Second, they provide counseling, support, inspiration and advice by virtue of having 'been there, done that', which we refer to as the role-model effect. At the same time, mentors too gain from matching with founders through recognition for developing talent, informational gains about the latest technology or advancements in their field, personal satisfaction, and higher compensation/promotion (Hart, 2009; Ramaswami and Dreher, 2007).

Because the venture market is heavily men dominated (AllRaise, 2019; Kerby, 2018; Robb, Coleman and Stangler, 2014), women may have higher challenges to find female role models (Noe, 1988), facing a trade-off when it comes to these two roles played by mentors. Further, there is literature that the returns of these functions might vary by gender, as a cause or a consequence [6]. Thus, we also investigate the tradeoffs faced in informal mentoring relationships by comparing the relative importance of financial versus role-model effects.

We use unique data on Mentorship and Investment from Global Entrepreneurship Research Network (GERN) covering the New York Tech Industry, collected between March 2013 and March 2014. We supplement this data with data on funding and demographics. Focusing on matches

[5]See Bosma et al. (2012); Hill (1991); Kram and Isabella (1985); O'Brien et al. (2010) for an overview on the different roles that a mentor can provide.

[6]While Howell and Nanda (2019) point out that men are twice as more likely as women to reach out to VC judges after the competition, Snellman and Solal (2022) find that firms with female founders who received funding from female rather than male VCs are two times less likely to raise additional financing.

formed between 2010 and 2013, we have 357 unique founders and 429 unique mentors, of which 19 and 13 percent are female, respectively. Female founders and mentors are more likely to have a MBA degree and less likely to have an engineering undergrad degree as compared to their male counterparts. On the other hand, male mentors are more likely to have founded before.

Our data contains 648 mentor-founder matches, where Female-Female matches accounts for 7.4 percent. Conditional on being matched to a male mentor, male founders are significantly more likely to have an engineering undergrad degree compared to a female mentor. However, there are no statistically significant differences in the characteristics of a male mentor, irrespective of the gender of the founder he is matched to. On the other hand, we find that Female-Female matches have mentors who are more likely to have graduated from a top 20 university as well as are older, in comparison to matches with female founders and male mentors. Using a two-part funding model as well as multinomial logit regressions, we analyze the effects of a mentor on the startup fund raising. We find that startups with female founders tend to have lower probabilities of getting funded, especially when a female founder matches with a male mentor [7].

To understand the potential gender differences in mentorship formation, we study the relative importance of financial and role-model effects across the observed matches. Gender and workplace network theories suggest that men might prefer men for both career objectives and social support, women prefer women for social support and might prefer men for

[7]A common finding amongst the existing literature is that the lower probabilities of funding for females arises from gendered preferences (Ewens and Townsend, 2019; Guzman and Kacperczyk, 2019; Kanze et al., 2018; Malmström, Johansson and Wincent, 2017).

advancing career objectives. Thus, indicating that men and women form homophilous relationships for different reasons (Ibarra, 1992). Further, gender homophily may result in more effective mentoring possibly due to within-group social cohesion (McGinn and Milkman, 2013). We explicitly control for the market competition and unobserved sorting by modeling the mentorship market as a two-sided matching game (Chiappori and Salanié, 2016). Estimating our model using Fox (2018), we find the gender homophily has a positive effect on match surplus, consistent with a vast literature on homophily. The match value is larger for a Female-Female match relative to the Male-Male. In other words, women consider gender homophily to be more important then men [8].

Further, disentangling the value of homophily into the financial and role-model aspects, while men gain close to 18 percent from homophily for financial aspects, this number is only 1 percent for women. Our evidence suggests that Female-Female matches gain more from the role-model aspect as compared to Male-Male matches [9].

The paper proceeds as follows. Section 2 details the GERN data. Section 3 documents the characteristics of mentors and founders in the New York tech startups, with Section 3.1 exploring the type of matching occuring in tech-startups and Section 3.2 exploring the variation in funding by match. Section 4 then presents the setup of the model, estimation procedure along with the results. Section 5 concludes.

[8] Using administrative data from a student-alumni networking platform, Gallen and Wasserman (2021) find that females students strongly prefer female mentors and are 20 percentage points more like to reach out to a female mentor, as compared to male students, with the cited reason for this as more friendly and could give more relevant advice.

[9] Similar to mentor-mentee setup, Rocha and Van Praag (2020) find that female founders have a strong influence on their female employee's entrepreneurial choice and propose that is best explained by female founders acting as role-models for their female employees.

2 Data

There are two sources of data that are used in this paper: data collected by the Global Entrepreneurship Network (GERN) and supplemental data collected by us.

2.1 GERN Data

We use data from Global Entrepreneurship Network (GERN) which covers the New York Tech Industry [10]. This data was collected between March 2013 and March 2014 using primary interviews with entrepreneurs and publicly available data from Crunchbase, AngelList, and LinkedIn. Although the sampling methodology is non-random, the final aim was to create a representative pool of founders.

They covered five core questions in their analysis covering the areas of Inspiration, Investment, Mentorship, Serial Entrepreneurship, and Spinoffs. We primarily use data on Investment and Mentorship. Mentorship, as defined in the GERN dataset, refers to a relationship between a mentor and a founder where the mentor advises the founder "on critical business issues at least three times for periods longer than 30 minutes."

2.2 Additional Data

We supplement the GERN data by hand-collecting data on funding (from 2010 to 2018), as well as mentor and founder demographics. The funding data is primarily accessed through Crunchbase and we focus on the following variables: funding in each year from 2009 to 2018[11], whether the company is active in 2018, if the company is not active, then whether

[10]Tech companies are defined as those that are either actively developing a new information technology or those whose businesses are Internet-enabled (excluding financial technology, green technology, and life sciences companies).

[11]If the company received any funding before 2009, this is clubbed in the year 2009.

the company closed or was acquired by someone, and finally, whether or not the company is for-profit. For the demographics data, this is primarily collected through LinkedIn. Appendix A details the process of collecting the data and any restrictions imposed.

2.3 Relevant Variables and Data Exclusions

The variables used in the paper can be grouped into one of three types: demographics, quality, and funding. Demographics include include gender and age of the individual[12].

Quality variables include whether or not the individual graduated from a top 20 university, whether or not the individual has a MBA degree and/or graduate degree, and finally, whether or not the individual had engineering or computer science as their major in their undergraduate degree. These variables serve as a proxy for access to the relevant networks in terms of mentoring or financing. We also focus on their undergraduate major as our data is on high-tech startups. Wadhwa, Freeman and Rissing (2010) find that founders with Ivy League degrees establish startups that produce more revenue and employ more workers. Moreover, founders having a MBA degree established companies faster than others. However, education alone can not be used to say an individual is of high-quality and therefore, we also include if the individual has a founded a company before. This is an important factor, especially for a mentor, as it provides experience in starting up a company - either in terms of what not to do (if that company failed) or what to keep repeating (if it succeeded).

Finally, funding variables include whether the founder received any

[12]While race is an important attribute as well, we ignore it in this analysis. See Hamilton et al. (2022*b*) for racial gap in entrepreneurship and Hamilton et al. (2022*a*) for a short article on the same.

funding by the year 2015. We also include variables as to the amount of funding received by the year 2015. We choose the year 2015 as it indicates short-term funding, given our sample selection.

We limit to those mentorship connections which are formed between the years of 2010 and 2013. This is to maximize the accuracy of the answers since all the questions are collected retrospectively. We also drop companies that are founded before 2004 and after 2013.

3 Who are the Mentors and Founders?

The dataset contains 429 unique mentors and 357 unique founders, of which 13 percent and 19 percent are female, respectively (Figure 1). While this is not a high number, it is representative of the industry. Table 1 presents the descriptive statistics of the dataset, as well as for unique founders and mentors. We notice that the numbers are similar across unique observations and the entire dataset, except for founded before. Around a third of both founders and mentors receive a MBA degree, while a fifth to a quarter received an engineering undergrad degree. While a third of the founders have graduated from a top 20 university and founded before, close to 40 percent of the mentors have done the same.

As the focus of this paper is on gender, we analyze the characteristics of the mentor and founder by gender (Table 2). We present the average values by gender, as well as the t-statistic of the difference. Focusing on founders, female founders are more likely to have received a MBA degree, in comparison to male founders. This could suggest that female founders use MBA as a signal of entrepreneurial will or desire, to counter the gender biases and stereotypes that exist in the industry. However, they're less likely to have graduated with an engineering undergrad degree. This is, however, not surprising as although females have surpassed males in terms of college graduation, they still lag in STEM degrees (Ceci et al., 2014). Although females are more likely to have graduated from a top 20 university and less likely to have founded before, these differences are not statistically significant.

For mentors, we see that the same trend of undergraduate engineering holds as well, with only 6 percent of female mentors having an undergrad engineering degree. Male mentors are more likely to have founded before - this could suggest that the male mentors are more experienced; however, it could be the case the female mentors have successfully run one company for a longer period of time. There is no difference between male and female mentors on receiving a MBA degree. While male mentors are more likely to have graduated from a top 20 university, this is not statistically significant. Lastly, we see that there is an age gap of 9 years between founders and mentors.

3.1 How does Matching Occur in Tech Startups?

Our dataset contains 648 mentor-founder matches. However, mentors can match with multiple founders and vice versa (although much less likely), as can be seen from the number of unique mentors and founders. We will analyze these as one-to-one matches, that is, take each mentor and founder in a match to be a unique individual from another match.

The first question that arises is that if the matching amongst the tech startups is random. Conditional on being matched to a female mentor, 58.5 percent of the founders are female; whereas conditional on being matched to a male mentor, 86.4 percent of the founders are male. This suggests strong homophily on the part of the mentor, as female mentors are more likely to match with female founders and vice versa. On the other hand, conditional on being matched to a female founder, 38.5 percent of the mentors are female; whereas conditional on being matched to male founder, 93.5 percent of the mentors are male. This suggests that homophily on the part of the founder as well.

With the fact that matching is non-random in mind, we break out the matches by the gender of the founder and mentor (Figure 2). Of these, the majority (75.5 percent) of the matches are Male Mentor and Male Founder matches (Male-Male). Female Mentor and Female Founder (Female-Female) only account for 7.4 percent. The cross-gender matches (Male-Female and Female-Male) are 11.9 and 5.2 percent, respectively. Hereafter, all matches will be referred to as the gender of the mentor first and then the gender of the founder. We now break out the characteristics of the mentor and founder by the type of match, as detailed (Table 3). Columns (1) and (2) present the averages of Male-Male and Male-Female matches, while Column (3) presents the difference and whether or not the t-statistic is significant. Columns (4) and (5) present the averages of Female-Male and Female-Female matches, with the differences and t-statistics in Column (6). Column (7) presents the differences between Male-Male and Female-Matches while Column (8) presents the difference between Male-Female and Female-Female matches. We see that conditional on being matched to a male mentor, male founders are less likely to have received a MBA degree and more likely to have received an undergrad engineering degree. This is not that surprising as these mirror the differences seen for founders by gender. However, irrespective of whether one is a male or a female founder, there are no statistically significant differences amongst the characteristics of the mentor. This suggests that the quality of the male mentor does not vary across matches.

Female-Female matches have mentors who are more likely to have graduated from a top 20 university, and who are older than female mentors in Female-Male matches. If we consider age to be a proxy for experience, Female-Female matches have better quality mentors. However, conditional on being matched with a female mentor, female founders are more likely to have received a MBA. On the other hand, male founders are more likely to have graduated from a top 20 university and founded before, whereas when matched with a male mentor, there were no statistically significant differences for these categories. A potential explanation is that Female-Male matches are formed as a result of competition i.e. male founders who could not find matches with male mentors, end up matching with a female mentor.

Conditional on being matched to a male founder, male mentors are more likely to have an engineering undergrad degree, graduated from a top 20 university, and to have founded before. Thus, Male-Male and Female-Male matches are systematically different. For the female founders, the trend is less clear. Female founders do not differ by the gender of founder they are matched to; however, female mentors matched to female founders are more likely to have a MBA degree while male mentors are more likely to have an engineering undergrad degree and to have founded before.

3.2 How does Funding Vary by Match?

As funding is at the firm level, we take our match level data and calculate the average characteristics of mentors and founders at the firm level. We start with 648 matches across years, and we end up with 434 firm-year observations. Thus, for each observation, we have the *average* characteristics of the founders and mentors. We focus on short-term funding, which is the cumulative funding of the firm by the year 2015. We present the summary statistics of the full firm-level dataset, as well as the dataset we use for regressions in Table 4. For the full dataset, we see that 78 percent of firms receive funding by 2015, and that the average level of funding is $ 12.34 million, with a standard deviation of $28.1 million. We find that the median value of funding is $2 million and this varies by gender mix of the firm. To prevent being biased by extreme values (zero or maximum funding), we also set up a categorical variable - whether the founder received no funding (Zero), received some funding but less than $1.5 million (Low), and received more than $1.5 million (High) [13]. We find that for the entire sample, around 20 percent of the firms receive zero funding, 25 percent receive low funding, and close to 55 percent receive high funding. Figure 3 shows that Male-Male matches have the highest proportion of High funding, with Female-Male matches having a high proportion of Low Funding. Female-Female matches appear to be equally split amongst the three categories. This potentially hints that female entrepreneurs (mentors or founders) might have an inherent bias against them (Ewens and Townsend, 2019; Kanze et al., 2018; Malmström, Johansson and Wincent, 2017) Thus, there is significant variation within the categorical funding variable by match type. On average, there appear to be 2.3 founders per firm (two founders within the same firm can have different mentors). Lastly, 45 percent of the interviewed firms were founded between 2004 and 2010.

[13] We chose $1.5 million as the cutoff as we wanted equal distribution across categories, at least for females due to data limitations.

Methodology We examine the effects of mentor on the startup fund raising performance using two approaches.

First, we estimate a multinomial logit [14] We set it up in the following manner:

$$P(Y_i^{2015} = k) = \frac{\exp(\beta_0^k + X_i\beta_{11}^k + X_j\beta_{12}^k + X_{ij}\beta_{13}^k)}{1 + \sum_{m=1}^{2}[\exp(\beta_0^m + X_i\beta_{11}^m + X_j\beta_{12}^j + X_{ij}\beta_{13}^m)]} \quad k = 2, 3 \tag{1}$$

where i refers to a firm. k refers to the three categories of the funding variables as defined earlier - $k = 1$ refers to zero or no funding (the base category), $k = 2$ refers to low funding or funding less \$1.5million, and $k = 3$ refers to high funding or funding >\$1 million, by the year 2015 . The other variables are as defined as before. All mentor and/or founder characteristics are fractions, or averages, in the case of age.

Further, we present a two-part model (Cragg, 1971), where the first part estimates the probability of funding and the second part estimates the log amount of funding received, if funded, as shown below. [15]

$$\mathbb{E}(y|\mathbf{X}) = \mathbb{P}(d = 1|\mathbf{X})\mathbb{E}(y|d = 1, \mathbf{X}) \tag{2}$$

[14] Our results are similar if we were to employ a generalized ordered logit model. See Williams (2006) for further details.

[15] The two-part model is commonly used to address situations where the dependent variable has many zeros and positive values. We do not use the Tobit type I model because we are interested on both the extensive margin and intensive margin of the mentorship impact on founder's fundraising. We do not use the Tobit type II (heckit) model because the exclusion restriction is hard to find in our setting.

We can define the funding regression in the following manner:

$$y_{ij} = \beta_0^y + X_i\beta_{11}^y + X_j\beta_{12}^y + X_{ij}\beta_{13}^y + \zeta_{ij} \tag{3}$$

$$\text{Prob}(d_{ij} = 1) = \frac{1}{1 + \exp\{\beta_0^d + X_i\beta_{11}^d + X_j\beta_{12}^d + X_{ij}\beta_{13}^d\}} \tag{4}$$

where, ζ_{ij} iid, d_{ij} is an indicator on whether the founder j raised any fund, and y_{ij} is the log of funding raised by founder j when matched with mentor i. X_i, X_j, X_{ij} are the characteristics of the mentor, founder and match that affect start-up capital but not the non-pecuniary benefits.

The results are presented in Table 5. Columns 1, 2, and 3 present the marginal effects from the multinomial logit. We find relative to the Male-Male match, all other matches have a lower probability of high funding. However, we also see that matches with a Female mentor have a higher probability of getting low funding. Columns 3 and 4 present the two-part model - marginal effects from logit regression, as well as the linear regression of log funding. We find that matches with female founders tend to have lower probabilities of getting funded, although this is statistically significant only for matches with Male Mentors. However, matches with Females Mentors have statistically significantly lower levels of funding when compared with Male-Male matches. Other mentor controls do not appear to have any significant effect on the probability or level of funding. In terms of founder controls, we find that graduating from a top university and having a larger founding team leads to a higher probability of getting funded. For the level of funding, top university as well as whether it is an older company affects the level. Time controls are not statistically significant.

We compare histograms of actual funding in 2015 with the expected funding from the Two-Part Model in Figure 4 and the summary statistics are in Table 6. Our model fits fairly well, except some under prediction at the higher end.

4 Value of Match

In this section, we study how the gender homophily affects the mentorship formation. To do so, we examine the revealed preference given observed mentor-mentee matches.

In the informal mentorship market, both mentors and founders choose each other so that they are both able to benefit from the relationship. Further, they are constrained by time and thus, cannot match with everyone. Therefore, market players may not always get their best choices due to market competition. This makes discrete choice models, like logit or probit, infeasible because they assume each decision maker is independent and can *always* get their best choices. Therefore, we model the mentorship market as a two-sided matching game (Roth and Sotomayor, 1992) to explicitly capture the sorting due to market competition.

We apply the two-sided matching framework as proposed by Fox (2018). This framework encompasses a many-to-many matching market, in which players on both sides of the market can be marketed with multiple players on the other side of the market.[16] We assume the capacity of each player is exogenous while most of the mentors in our data have only one founder and vice versa.[17] We focus on transferable utility two-sided matching model. In this framework, players on two sides of the market en-

[16]We do not use Choo and Siow (2006) because it is more suitable large markets.

[17]In order to guarantee the existence of of market equilibrium, we abstract away from potential complementarity among potential partners for the same agent. In other words, the preference on potential partners is independent. This is a common assumption in empirical matching papers.

dogenously decide how they share the total match value. This assumption makes intuitive sense in our dataset focuses on informal mentoring. We have already detailed the reasons why the founders would want to match with mentors. On the other hand, mentors benefit from mentoring through recognition for developing talent, informational gains about the latest technology or advancements in their field, and personal satisfaction. Formal mentors, in addition, could also gain from higher compensation and promotion (Hart, 2009; Ramaswami and Dreher, 2007). A meta-analysis study by Ghosh and Reio Jr (2013) finds that mentoring is reciprocal and collaborative. Thus, we believe there are transfers from both sides of the market and thus, we could justify it as transferable utility.

4.1 Model Setup

There are two types of agents: mentors and founders. Each agent obtain some payoffs from potential matches with agents on the other side of the market.

There are $i \in I$ types of mentors and $j \in J$ type of founders. $\boldsymbol{z_i}$ and $\boldsymbol{z_j}$ refers to the characteristics of founders and mentors, respectively. $\boldsymbol{z_{ij}}$ are the interactions of both founder and mentor characteristics. $\hat{y}_{ij}$ is the predicted amount of financing raised by founder j when matched with mentor i, as predicted by the two-part model detailed in Section 3.2. For each match, there is a transfer τ_{ij} (can be positive or negative) from the founder to the mentor.

Founder's Match Value Let v_{ij}^F refer to the match value for a founder.

$$v_{ij}^F = \alpha^F E(y_{ij}) + z_i \gamma_1^F + z_j \gamma_2^F + z_{ij} \gamma_3^F + \epsilon_{ij} - \tau_{ij} \tag{5}$$

where ϵ_{ij} refers to gains to the founder from unobservable characteristics of the match. The $E(y_{ij})$ is fitted value using the two-part model estimates in the previous section. It captures the expected founder fundraising performance in the future during the mentorship formation process.

Mentor's Match Value Let v_{ij}^M refer to the match value for a mentor.

$$v_{ij}^M = \alpha^M E(y_{ij}) + z_i\gamma_1^M + z_j\gamma_2^M + z_{ij}\gamma_3^M + \eta_{ij} + \tau_{ij} \tag{6}$$

where η_{ij} refers to gains to the mentor from unobservable characteristics of the match.

Total Match Value The total match value is composed of the founder's value and the mentor's value. Each of this can be further divided into two components: non-pecuniary benefits and pecuniary benefits from a match.

$$U^{ij} = v_{ij}^F + v_{ij}^M \tag{7}$$

$$= \underbrace{(\alpha^F + \alpha^M)E(y_{ij})}_{\text{pecuniary}} + \underbrace{z_i(\gamma_1^F + \gamma_1^M) + z_j(\gamma_2^F + \gamma_2^M) + z_{ij}(\gamma_3^F + \gamma_3^M)}_{\text{non-pecuniary}} + \epsilon_{ij} + \eta_{ij} \tag{8}$$

$$= \alpha_1 z_i + \alpha_2 z_j + \alpha_3 E(y_{ij}) + \alpha_4 z_{ij} + \varepsilon_{ij} \tag{9}$$

The transfer is cancelled out in the combined utility for a given match.

The market equilibrium is unique in terms of who-match-with-whom. As in Fox (2018), the necessary and sufficient condition for the equilibrium is the so-called pairwise stability in which there is no feasible deviation for any player in any two pairs of observed matches. In our setting, the equilibrium condition implies that the total match value for for any two pairs of observed matches, founder i with mentor j and founder i' with mentor j', is larger than the total match value of the two pairs if they switch partners. Mathematically, it is

$$U^{ij} + U^{i'j'} \geq U^{ij'} + U^{i'j} \tag{10}$$

4.2 Model Identification and Estimation

The maximum score estimator proposed by Fox (2018) builds on the market equilibrium condition. In our setting, it compares pairs of observed matches with the alternative but unrealized matches for all potential matches between mentors and founders. Mathematically, we obtain the parameter vector that maximize the following score:

$$\sum \mathbb{1}[U^{ij} + U^{i'j'} \geq U^{ij'} + U^{i'j}] \tag{11}$$

Identification follows as detailed in Fox (2010). Since the identification is based on the differences between pairs of matches, factors from only one side of the market are unidentified. In other words, we can only identify α_3 and α_4 from (9). Further, this model is identified up to scale, we normalize the norm of the parameter vector to 1, ie., $||\alpha|| = 1$. For each model setting, we run the estimator twice with the parameter of expected funding to be positive and negative respectively. We pick the estimation that generates the highest score.

We estimate two models. Model 1 has the gender interaction of founder and mentor along with the expected funding, while Model 2 includes the quality variables of both mentor and founder. Both of these models are relative to the Male-Male case.

We follow Akkus, Cookson and Hortacsu (2015) to obtain subsampling and confidence intervals for the maximum score estimator - we run 100 bootstraps for estimating the confidence intervals.

4.3 Results

4.3.1 Parameter Estimates

We present the results from the matching model in Table 7. The main result is that the value of a match is the maximum for a Female-Female match, relative to the base case of Male-Male - it is 0.635 points higher for a Female-Female match as compared to a Male-Male match. This is a strong indication that the female mentors and female founders *choose* to match with each other, and thus, there is a strong preference for homophily amongst females.

On the other hand, cross-gender matches have a lower match surplus than the matches in which homophily exists. Male-Female match provides the lowest value of a match, amongst the match-gender types. This provides a direct argument and rationale for the existence of Female-Female matches through the role-model effect. This also raises the issue of the cross-gender matches as they do not seem to benefit from either financial or role-model aspect. This could imply that one possible explanation is that cross-gender matches are formed as a result of competition in the market where these founders were unable to find mentors of their own gender. Moreover, these effects are statistically significant at 5 percent level of significance. We find a positive effect of expected funding on the match value - if the expected funding increases by $ 1 million then the match value rises by $ 0.023 million. These effects hold for Model 1 and 2.

Now, focusing on the quality variables, we see that if both mentor and founder have a MBA, then the value of a match is higher by 0.541 points relative to a match where either mentor or founder does not have MBA. Similar results are seen for having an engineering degree and top 20 university - 0.414 points and 0.295 points, respectively. All of these are significant at 5 percent level of significance. However, we see a lower value of match if both mentor and founder have founded before, relative to either mentor or founder having not founded before. An explanation for this might be that the match may not be as useful as the founder may already have the role-model aspect covered and may need more help on the financial side. However, this is insignificant as well. We also see a strong positive effect of the interaction of age - if both mentor and founder are older, then the match has a higher surplus. The fit of this model is 66.1 percent - implying that 66.1 percent of the matches are correctly identified by the model. Adding a founder gender interaction with expected funding or removing age interaction from the specification does not change the key result.

4.3.2 Value from Homophily

As we have expected funding as one of the variables in the matching model, we can normalize all our variables with respect to it to get a dollar value of the marginal effect to the match value. In Model 1, keeping everything else constant, a Female-Female match adds $27 million relative to a Male-Male match. This rises to $73 million in Model 2. However, both cross-gender matches will take away from the match surplus, relative to a Male-Male match.

The key question of this paper has been what is the value of homophily and we can use our matching model to answer this question. Using average expected funding for each type of match, we can calculate the relative value of a match for each type. Using this, we can then ask: conditional on being a male mentor, what is the value of homophily? We define this as the difference in the match value from being matched to the same gender founder (male founder) versus an opposite gender founder (female founder). We repeat this for both mentors and founders, male and female. We present these results in Figure 5 using the parameter estimates from Model 1, for ease of exposition. For each gender and type of individual, we present three columns. The first is the value of homophily as defined above. The second and third columns are the decomposition of this value of homophily into fraction role model and fraction financial, respectively.

First, we find that on average, females gain more from homophily than males. Second, amongst males, mentors gain more from homophily, while we find the opposite for females. Third, while males gain close to 18 percent from homophily due to financial aspects, females gain less than 1 percent from the same.

5 Conclusion

We study whether we can bridge the gap in entrepreneurship through mentoring. We focus on informal mentoring, where both mentors and founders form a match based on 'mutual identification and fulfillment of career needs'. While there are multiple roles a mentor can play, we condense them into two aspects - one, through career development by providing access to networks, and thus, helping in financing and other observable outcomes (financial aspect); two, by providing advice and support in dealing with obstacles of entrepreneurship (role-model aspect).

We find that conditional on being matched to a male mentor, there are no statistically significant differences across Male-Male and Male-Female matches. On the other hand, conditional on being matched to a female mentor, Female-Female matches have older mentors and mentors who have graduated from a top 20 university, as compared to Female-Male matches. Further, we find that Male-Male matches have the highest probability of receiving funding of more than $1.5mn, relative to any other match type.

To disentangle the role model and financial effects, we estimate a two-sided transferable utility matching model (á la Fox (2018)) to understand the parameters that affect the value of a match. We find the role-model aspects of the mentor-founder relationship has a significant impact among Female-Female match, and the gender homophily plays an important role during the mentorship formation. Using the expected funding to normalize the parameters, we find that a Female-Female match adds $27 million to the match value, relative to a Male-Male match. Moreover, on decomposing the match value into role-model and financial effect, we find females gain more from homophily than males and most of this is from the role-model effects.

References

Akkus, Oktay, J Anthony Cookson, and Ali Hortacsu. 2015. "The determinants of bank mergers: A revealed preference analysis." *Management Science*, 62(8): 2241–2258.

AllRaise. 2019. "Where are the numbers on women funders and founders?" `https://www.allraise.org/data/`, [Online; accessed 12-August-2019].

Athey, Susan, Christopher Avery, and Peter Zemsky. 2000. "Mentoring and diversity." *American Economic Review*, 90(4): 765–786.

Blau, Francine D, Janet M Currie, Rachel TA Croson, and Donna K Ginther. 2010. "Can mentoring help female assistant professors? Interim results from a randomized trial." *American Economic Review*, 100(2): 348–52.

Bosma, Niels, Jolanda Hessels, Veronique Schutjens, Mirjam Van Praag, and Ingrid Verheul. 2012. "Entrepreneurship and role models." *Journal of economic psychology*, 33(2): 410–424.

Carrell, Scott E, Marianne E Page, and James E West. 2010. "Sex and science: How professor gender perpetuates the gender gap." *The Quarterly Journal of Economics*, 125(3): 1101–1144.

Ceci, Stephen J, Donna K Ginther, Shulamit Kahn, and Wendy M Williams. 2014. "Women in academic science: A changing landscape." *Psychological Science in the Public Interest*, 15(3): 75–141.

Chiappori, Pierre-André, and Bernard Salanié. 2016. "The econometrics of matching models." *Journal of Economic Literature*, 54(3): 832–61.

Choo, Eugene, and Aloysius Siow. 2006. "Who marries whom and why." *Journal of political Economy*, 114(1): 175–201.

Cragg, John G. 1971. "Some statistical models for limited dependent variables with application to the demand for durable goods." *Econometrica: Journal of the Econometric Society*, 829–844.

Ewens, Michael, and Richard R Townsend. 2019. "Are early stage investors biased against women?" *Journal of Financial Economics*.

Falk, Armin, Fabian Kosse, and Pia Pinger. 2020. "Mentoring and schooling decisions: Causal evidence."

Fox, Jeremy T. 2010. "Identification in matching games." *Quantitative Economics*, 1(2): 203–254.

Fox, Jeremy T. 2018. "Estimating matching games with transfers." *Quantitative Economics*, 9(1): 1–38.

Gallen, Yana, and Melanie Wasserman. 2021. "Do Male and Female Students Use Networks Differently?" Vol. 111, 154–58.

Ghosh, Rajashi, and Thomas G Reio Jr. 2013. "Career benefits associated with mentoring for mentors: A meta-analysis." *Journal of Vocational Behavior*, 83(1): 106–116.

Greene, Patricia G, Candida G Brush, Myra M Hart, and Patrick Saparito. 2001. "Patterns of venture capital funding: is gender a factor?" *Venture Capital: An international journal of entrepreneurial finance*, 3(1): 63–83.

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

Hamilton, Barton, Andrés Hincapié, Maggie Isaacson, Prasanthi Ramakrishnan, and Siddhartha Sanghi. 2022*a*. "What Can Be Done to Promote Black Entrepreneurship?" *The Regional Economist*.

Hamilton, Barton H, Andrés Hincapié, Prasanthi Ramakrishnan, and Siddhartha Sanghi. 2022*b*. "Entrepreneurship in Black and White."

Hart, E Wayne. 2009. "In focus/mentoring—Nurturing relationships provide many benefits." *Leadership in Action: A Publication of the Center for Creative Leadership and Jossey-Bass*, 29(1): 17–20.

Hill, Linda A. 1991. "Beyond the myth of the perfect mentor: Building a network of developmental relationships."

Hilmer, Christiana, and Michael Hilmer. 2007. "Women helping women, men helping women? Same-gender mentoring, initial job placements, and early career publishing success for economics PhDs." *American Economic Review*, 97(2): 422–426.

Howell, Sabrina, and Ramana Nanda. 2019. "Networking Frictions in Venture Capital, and the Gender Gap in Entrepreneurship." *Harvard Business School Entrepreneurial Management Working Paper*, , (19-105).

Ibarra, Herminia. 1992. "Homophily and differential returns: Sex differences in network structure and access in an advertising firm." *Administrative science quarterly*, 422–447.

Jackson, Matthew O. 2021. "Inequality's Economic and Social Roots: The Role of Social Networks and Homophily." *Available at SSRN 3795626*.

Kanze, Dana, Laura Huang, Mark A Conley, and E Tory Higgins. 2018. "We ask men to win and women not to lose: Closing the gender gap in startup funding." *Academy of Management Journal*, 61(2): 586–614.

Kerby, Richard. 2018. "Where Did you Go to School?" https://blog.usejournal.com/where-did-you-go-to-school-bde54d846188, [Online; accessed 31-May-2019].

Kofoed, Michael S, et al. 2019. "The Effect of Same-Gender or Same-Race Role Models on Occupation Choice Evidence from Randomly Assigned Mentors at West Point." *Journal of Human Resources*, 54(2): 430–467.

Kram, Kathy E, and Lynn A Isabella. 1985. "Mentoring alternatives: The role of peer relationships in career development." *Academy of management Journal*, 28(1): 110–132.

Malmström, Malin, Jeaneth Johansson, and Joakim Wincent. 2017. "We recorded VCs' conversations and analyzed how differently they talk about female entrepreneurs." *Harvard Business Review*.

McGinn, Kathleen L, and Katherine L Milkman. 2013. "Looking up and looking out: Career mobility effects of demographic similarity among professionals." *Organization Science*, 24(4): 1041–1060.

McPherson, Miller, Lynn Smith-Lovin, and James M Cook. 2001. "Birds of a feather: Homophily in social networks." *Annual review of sociology*, 415–444.

Mullen, Carol A, and Cindy C Klimaitis. 2021. "Defining mentoring: a literature review of issues, types, and applications." *Annals of the New York Academy of Sciences*, 1483(1): 19–35.

Neumark, David, and Rosella Gardecki. 1998. "Women Helping Women? Role Model and Mentoring Effects on Female Ph. D. Students in Economics." *Journal of Human Resources*, 220–246.

Noe, Raymond A. 1988. "Women and mentoring: A review and research agenda." *Academy of management review*, 13(1): 65–78.

O'Brien, Kimberly E, Andrew Biga, Stacey R Kessler, and Tammy D Allen. 2010. "A meta-analytic investigation of gender differences in mentoring." *Journal of Management*, 36(2): 537–554.

Pitchbook. 2019. "The VC Female Founders Dashboard." *https://pitchbook.com/news/articles/the-vc-female-founders-dashboard*, [Online; accessed 12-August-2019].

Porter, Catherine, and Danila Serra. 2019. "Gender differences in the choice of major: The importance of female role models." *American Economic Journal: Applied Economics*.

Ragins, Belle Rose, and John L Cotton. 1999. "Mentor functions and outcomes: A comparison of men and women in formal and informal mentoring relationships." *Journal of applied psychology*, 84(4): 529.

Ramaswami, Aarti, and George F Dreher. 2007. "The benefits associated with workplace mentoring relationships." *The Blackwell handbook of mentoring: A multiple perspectives approach*, 211–231.

Robb, Alicia, Susan Coleman, and Dane Stangler. 2014. "Sources of economic hope: Women's entrepreneurship." *Available at SSRN 2529094*.

Rocha, Vera, and Mirjam Van Praag. 2020. "Mind the gap: The role of gender in entrepreneurial career choice and social influence by founders." *Strategic Management Journal*, 41(5): 841–866.

Roth, Alvin E, and Marilda Sotomayor. 1992. "Two-sided matching." *Handbook of game theory with economic applications*, 1: 485–541.

Snellman, Kaisa, and Isabelle Solal. 2022. "Does investor gender matter for the success of female entrepreneurs? Gender homophily and the stigma of incompetence in entrepreneurial finance." *Organization Science*.

Wadhwa, Vivek, Richard Freeman, and Ben Rissing. 2010. "Education and tech entrepreneurship." *Innovations: Technology, Governance, Globalization*, 5(2): 141–153.

Williams, Richard. 2006. "Generalized ordered logit/partial proportional odds models for ordinal dependent variables." *The Stata Journal*, 6(1): 58–82.

6 Tables and Figures

Table 1: Descriptive Statistics of Dataset

	Entire Dataset		Unique Observations	
	Founder	Mentor	Founder	Mentor
Female	0.19	0.13	0.19	0.13
	[0.39]	[0.33]	[0.39]	[0.33]
Age	30.56	39.84	31.71	40.09
	[6.06]	[9.70]	[6.58]	[10.37]
Received a MBA degree	0.29	0.32	0.29	0.31
	[0.45]	[0.47]	[0.46]	[0.46]
Engineering Undergrad Degree	0.21	0.22	0.24	0.21
	[0.41]	[0.42]	[0.43]	[0.41]
Graduated from a Top 20 University	0.36	0.42	0.34	0.38
	[0.48]	[0.49]	[0.47]	[0.49]
Founded Before	0.28	0.45	0.34	0.43
	[0.45]	[0.50]	[0.48]	[0.50]
Observations	648	648	357	429

Source: Authors' calculations from the GERN Dataset
Note: Standard deviations in parentheses. Please refer to Section 2 and Appendix A for further details.

Table 2: Summary Statistics by Gender of Founder and Mentor

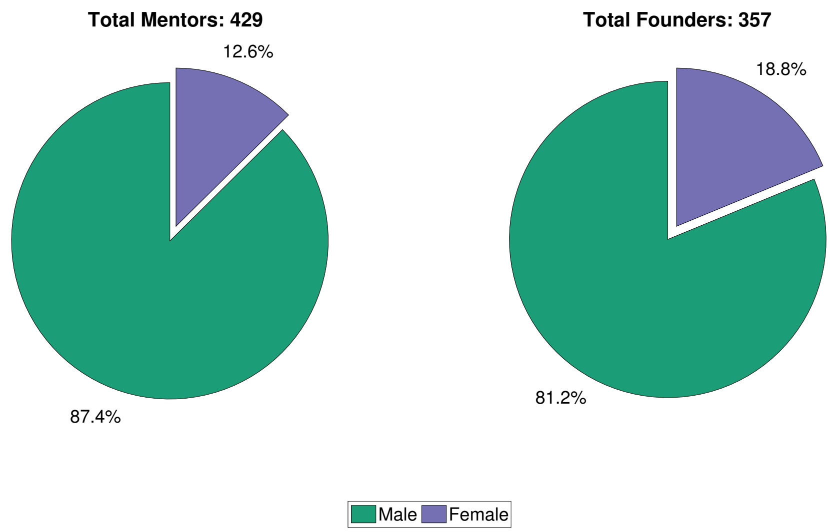
	Founders			Mentors		
	Mean		Diff.	Mean		Diff.
	Male	Female		Male	Female	
Age	31.81	31.25	0.56	40.05	40.41	-0.36
Received a MBA degree	0.26	0.43	-0.17***	0.31	0.31	-0.00
Graduated from a Top 20 University	0.32	0.40	-0.08	0.39	0.30	0.10
Engineering Undergrad Degree	0.27	0.10	0.17***	0.23	0.06	0.18***
Founded Before	0.36	0.27	0.09	0.45	0.31	0.14*
Observations	290	67	357	375	54	429

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' calculations from the GERN Dataset

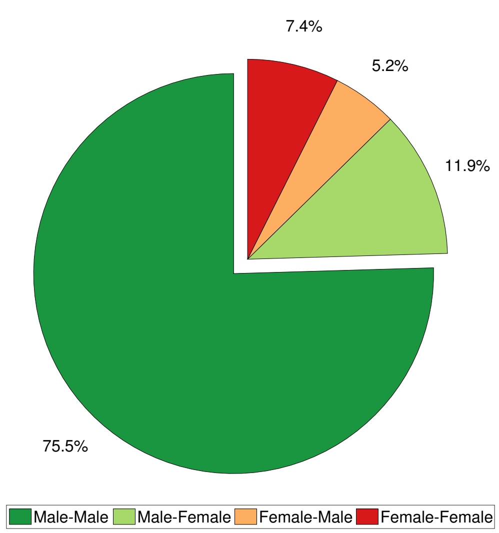
Note: The significance stars on the Difference column refer to whether or not the t-statistic of the difference is significant. Please refer to Section 2 and Appendix A for further details.

Figure 1: Gender Mix of Mentors and Founders



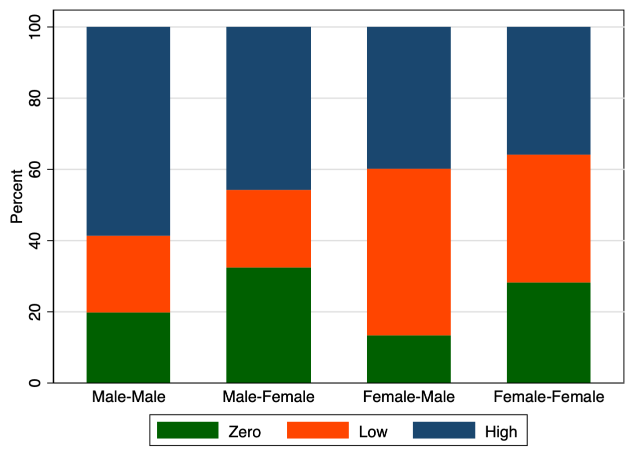
Source: Authors' calculations from the GERN Dataset
Notes: 1. There are a total of 648 connections.

Figure 2: Summary of Connections in Dataset



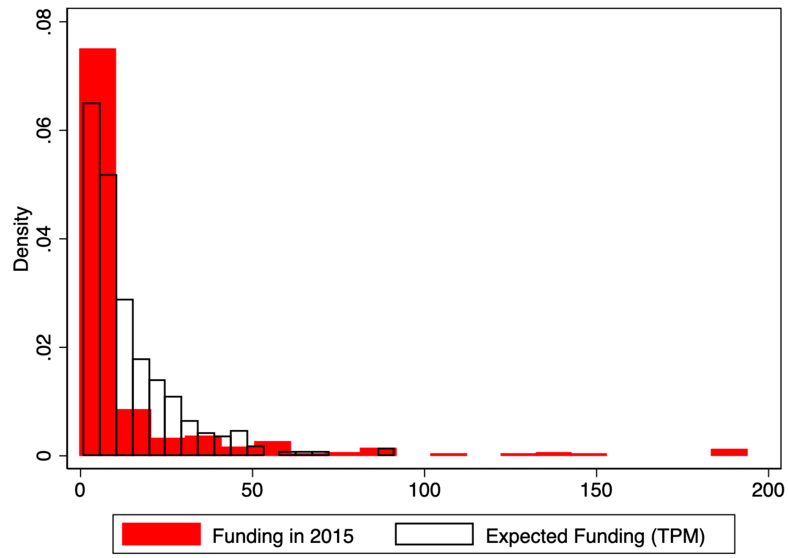
Source: Authors' calculations from the GERN Dataset
Notes: 1. There are a total of 648 connections.

Figure 3: Categorical Funding by Match Type



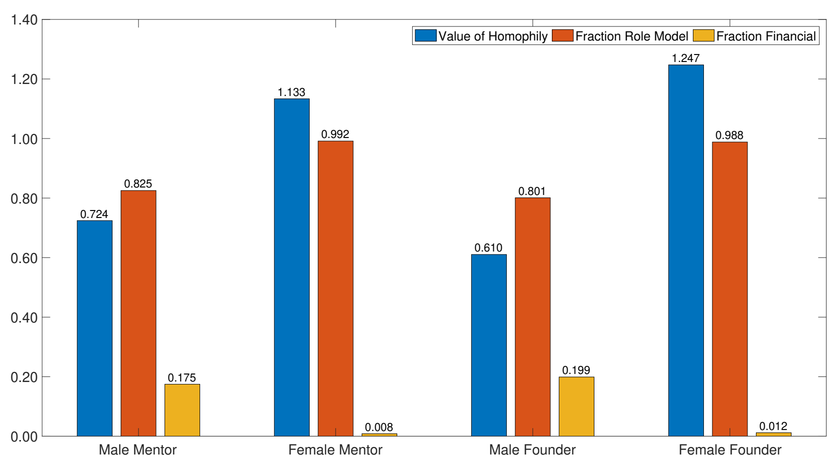
Source: Authors' calculations from the GERN Dataset. *Note:* 1. The type of match is defined as gender of mentor-gender of founder. Therefore, 'Male-Male' refers to a match of a Male Mentor with a Male Founder. 2. The categorical variable for funding is divided into three categories - whether the founder received no funding (Zero), received some funding but less than $1.5 million (Low), and received more than $1.5 million (High), using short-term funding (funding till year 2015).

Figure 4: Comparing Predictions from Two Part Model



Source: Authors' calculations from the GERN Dataset. *Note:* 1. The expected funding is calculated using a smearing estimator for a Two-Part Model, as defined in Section 3.2.

Figure 5: What is the Value of Homophily?



Source: Authors' calculations from the GERN Dataset. *Note:* The gains from homophily are defined as the gains from matching to a mentor or founder of the same gender. For example, the numbers for Male Mentor refers to the difference in match value from a Male-Male match versus a Male-Female match. This match value is estimated from Model 1 as estimated in Section 4. Further details are in Section 4.3.2.

Table 3: Summary Statistics by Type of Connection

	Male Mentors			Female Mentors			Founders	
Gender of Founder →	Male	Female	Diff.	Male	Female	Diff.	Diff.	Diff.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Founder's Characteristics								
Age	30.40	31.19	-0.79	30.62	31.15	-0.53	-0.22	0.05
Received a MBA degree	0.23	0.57	-0.34***	0.18	0.48	-0.30***	0.05	0.09
Engineering Undergrad Degree	0.25	0.09	0.16***	0.15	0.08	0.06	0.10	0.01
Graduated from a Top 20 University	0.35	0.40	-0.05	0.50	0.29	0.21*	-0.15*	0.11
Founded Before	0.29	0.25	0.04	0.32	0.15	0.18*	-0.03	0.10
Mentor's Characteristics								
Age	39.78	40.38	-0.59	37.47	41.21	-3.74**	2.31	-0.83
Received a MBA degree	0.31	0.30	0.01	0.32	0.46	-0.13	-0.01	-0.16*
Engineering Undergrad Degree	0.25	0.21	0.05	0.06	0.02	0.04	0.19**	0.19***
Graduated from a Top 20 University	0.43	0.36	0.07	0.24	0.48	-0.24**	0.20**	-0.12
Founded Before	0.48	0.47	0.01	0.26	0.29	-0.03	0.21**	0.18*
Observations	489	77	566	34	48	82	523	125

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' calculations from the GERN Dataset

Note: Column (3) and (6) refer to difference between Male and Female Founder, conditional on being matched with a Male and Female Mentor, respectively. Column (7) and (8) is the difference between Male and Female Mentors for a Male and Female Founder, respectively. Please refer to Section 2 and Appendix A for further details. for further details.

Table 4: Summary Statistics for Firm Level Dataset

	Full Dataset					Dataset for Regressions				
	Mean	S.D	Min	Max	N	Mean	S.D	Min	Max	N
Funding Variables										
1 if 2015 Funding	0.78	0.41	0.00	1.00	390	0.80	0.40	0.00	1.00	365
2015 Funding (in $ mn)	12.34	28.12	0.00	193.80	390	12.64	28.54	0.00	193.80	365
Log Funding in 2015	1.29	1.96	-3.69	5.27	306	1.27	1.97	-3.69	5.27	293
Funding 2009 or before	0.68	2.94	0.00	25.00	392	0.68	2.91	0.00	25.00	365
Funding 2010-2015	13.27	35.25	0.00	362.60	392	11.96	27.96	0.00	193.80	365
Type of Funding										
Zero	21.54				84	19.73				72
Low	24.62				96	25.21				92
High	53.85				210	55.07				201
Fraction of Match										
Male-Male	0.75	0.42	0.00	1.00	434	0.76	0.42	0.00	1.00	365
Male-Female	0.12	0.32	0.00	1.00	434	0.11	0.30	0.00	1.00	365
Female-Male	0.05	0.20	0.00	1.00	434	0.06	0.22	0.00	1.00	365
Female-Female	0.08	0.26	0.00	1.00	434	0.07	0.25	0.00	1.00	365
Mentor Controls										
Age	40.10	9.12	21.00	81.00	434	39.99	9.15	21.00	81.00	365
MBA	0.33	0.44	0.00	1.00	434	0.33	0.44	0.00	1.00	365
Engineering Undergrad	0.19	0.37	0.00	1.00	434	0.19	0.36	0.00	1.00	365
Top University	0.41	0.46	0.00	1.00	434	0.41	0.46	0.00	1.00	365
Founded Before	0.44	0.46	0.00	1.00	434	0.45	0.46	0.00	1.00	365

Table 4: Summary Statistics for Firm Level Dataset (Continued)

	Full Dataset					Dataset for Regressions				
	Mean	S.D	Min	Max	N	Mean	S.D	Min	Max	N
Founder Controls										
Age	31.16	6.14	18.00	63.00	434	31.25	6.13	18.00	63.00	365
MBA	0.30	0.45	0.00	1.00	434	0.30	0.45	0.00	1.00	365
Engineering Undergrad	0.22	0.41	0.00	1.00	434	0.21	0.40	0.00	1.00	365
Top University	0.35	0.47	0.00	1.00	434	0.35	0.47	0.00	1.00	365
Founded Before	0.32	0.46	0.00	1.00	434	0.32	0.46	0.00	1.00	365
Age Interaction	1262.12	419.35	546.00	3350.00	434	1263.30	425.23	546.00	3350.00	365
Other Variables										
Number of Founders	2.29	1.14	1.00	6.00	416	2.39	1.15	1.00	6.00	365
Indicator if Founded Between 2004 and 2010	0.45	0.50	0.00	1.00	382	0.45	0.50	0.00	1.00	365

Source: Authors' calculations from the GERN Dataset

Note: 1. The categorical variable for funding is divided into three categories - whether the founder received no funding (Zero), received some funding but less than $1.5 million (Low), and received more than $1.5 million (High), using short-term funding (funding till year 2015). 2. We do not have data on all founders, and therefore, the number of founders indicates the number of founders who founded the company, whereas all the other variables are based on the number of founders who have mentor matches.

Table 5: Analysis of Short-Term Funding

	Multinomial Logit Model			Two Part Model	
	Zero	Low	High	Logit	Linear
Fraction of Match (Base: Male-Male)					
Male-Female	0.130**	0.039	-0.169**	-0.133**	-0.367
	[0.0654]	[0.0774]	[0.0839]	[0.0655]	[0.3725]
Female-Male	-0.021	0.224**	-0.203*	0.038	-1.185**
	[0.0976]	[0.1004]	[0.1149]	[0.0978]	[0.5619]
Female-Female	0.113	0.158*	-0.271***	-0.103	-0.658*
	[0.0737]	[0.0836]	[0.1010]	[0.0745]	[0.3967]
Founder Controls					
Age	0.003	0.007	-0.010	-0.003	-0.011
	[0.0207]	[0.0194]	[0.0229]	[0.0205]	[0.0885]
Top University	-0.107**	-0.059	0.166***	0.112**	0.519**
	[0.0535]	[0.0509]	[0.0537]	[0.0532]	[0.2470]
Engineering Undergrad	0.117**	-0.009	-0.108*	-0.118***	-0.204
	[0.0463]	[0.0604]	[0.0640]	[0.0458]	[0.3311]
MBA	-0.003	-0.033	0.036	0.001	0.153
	[0.0471]	[0.0517]	[0.0566]	[0.0468]	[0.2786]
Founded Before	0.040	0.038	-0.078	-0.040	0.017
	[0.0455]	[0.0492]	[0.0542]	[0.0453]	[0.2831]
Age Interaction	0.000	-0.000	0.000	-0.000	0.001
	[0.0005]	[0.0004]	[0.0005]	[0.0005]	[0.0021]
Number of Founders	-0.038*	-0.008	0.047**	0.040*	0.117
	[0.0224]	[0.0206]	[0.0221]	[0.0224]	[0.1049]
Indicator if Founded Between 2004 and 2010=1	-0.069	-0.118**	0.187***	0.069	0.967***
	[0.0445]	[0.0486]	[0.0547]	[0.0444]	[0.2534]
Constant					0.680
					[2.6760]
Mentor Controls	Yes	Yes	Yes	Yes	Yes
Time Controls	Yes	Yes	Yes	Yes	Yes
Observations	365	365	365	365	293
Pseudo $R2^2$†	0.112	0.112	0.112	0.110	0.160

$*p < 0.1, **p < 0.05, ***p < 0.01$, † For the linear regression, R-squared is presented, instead of Pseudo R-squared.

Source: Authors' calculations from the CERN Dataset; *Notes:* 1. The type of match is defined as gender of mentor-gender of founder. 2. Please refer to Section 3.2 for further details on Two-Part Model. 3. Number of founders here refers to the total number of founders of the company, rather than the number of founders who have mentor matches. All other variables are based on founders who have mentor matches.

Table 6: How does Expected Funding for Observed Data look like?

	Mean	S.D	Min	Max	N
Funding in 2015	12.64	28.54	0.00	193.80	365
Expected Funding (Smearing Estimator)	13.96	13.23	1.03	91.31	365
Observations	365				

Source: Authors' calculations from the GERN Dataset. *Note:* 1. The expected funding is calculated using a smearing estimator for a Two-Part Model, as defined in Section 3.2.

Table 7: Parameter Confidence Intervals from Matching Model (100 bootstraps)

Variable	Model 1				Model 2			
	Estimate	95 % CI		in $	Estimate	95 % CI		in $
Type of Match (Base: Male-Male)								
Male-Female	-0.598	-0.837	-0.119	-25.651	-0.197	-0.573	-0.070	-28.885
Female-Male	-0.489	-0.780	-0.050	-20.983	-0.400	-0.580	-0.138	-58.604
Female-Female	0.635	-0.000	0.853	27.239	0.500	0.216	0.628	73.296
Expected Funding (in Million $)	0.023	0.012	0.052	1.000	0.007	0.001	0.012	1.000
Quality Variables								
MBA					0.541	0.352	0.609	79.198
Engg					0.414	0.225	0.525	60.582
Founded Before					-0.014	-0.128	0.031	-1.992
Top 20 University					0.295	0.153	0.429	43.195
Age					0.004	0.002	0.005	0.527
# Inequalities	55193				55193			
% Satisfied	58.6				66.1			

Source: Authors' calculations from the GERN Dataset
Note: 1. The type of match is defined as gender of mentor-gender of founder. 2. This data covers all the connections in the years 2010 to 2013. 3. The confidence intervals are calculated with 100 bootstraps, using Akkus, Cookson and Hortacsu (2015)'s procedure. 4. All quality variables are interactions of Mentor and Founder.

Appendices

A Data Appendix

Using the GERN data, we have 867 mentor-founder matches for the years between 2010 and 2013. In the GERN dataset itself, there is data on the gender, education qualifications (years of graduation as well as the type of degree) as well as the amount invested in the company of the founder. We verify and contribute to this data by manually collecting data from LinkedIn and Crunchbase. All data is accurate as of 23 July 2018.

A.1 Demographics

From the GERN data, we have the name of the founder and/or mentor as well as the company that they are associated with in the dataset. We identify a person only if we can match the name of the person as well as the company listed in the dataset. We then verify or collect the following variables in the following manner:

- *Gender:* Gender is concluded using the picture on the LinkedIn or Crunchbase website (Crunchbase has a column with gender in some cases). We limit gender to Male or Female (although in few cases, the people identified as gender fluid, we categorized them based on their name to male or female). This is further corroborated by the pronouns used in the profile (he/she).

- *Race/Ethnicity:* Race/ethnicity is collected using the picture on the LinkedIn or Crunchbase website. We limit the options to White, Black, Asian, Indian and Other. As this may not be the most reliable source, we do not use this variable.

- *Whether Founded a Company in the Past:* This is a variable that is not provided in the GERN dataset. This variable is collected based on the LinkedIn profile preferably. If enough information is not available, then we consult the Crunchbase bio. An individual is said to have founded a company in the past if there exists a company to which he was listed as Founder and the year of founding was before the company he/she is attached to in the GERN dataset. In some cases, when the individual is older, we also use Bloomberg or Angellist to corroborate this as individuals end up listing on their current affiliations on LinkedIn.

- *Education:* We verify and collect the following data - the university that the individual graduated from, year of graduation for the undergraduate degree, for MBA, and graduate degree (if received). We do not count honorary degrees for this. There are cases where the individual will list the name of the university but not the graduation year on LinkedIn – this is then collected by checking the websites in the following order: Crunchbase, Bloomberg, and Angellist. If none of these have the year of graduation, we use a Google search to find the name of the person along with their graduation (generally school websites for top universities have graduating class details). In some cases, the individual went to college but did not finish and dropped out or did not go to college at all – in this case, we checked to see if we could find the age of the individual and calculate the hypothetical graduation year from there. We tried to ensure that if the graduation year was not listed on the LinkedIn website, then the year is in line with the first job listed in the profile.

- *Whether the Undergrad Major was Computer Science or Engineering Degree:* We verify and collected this variable using the major listed on the LinkedIn profile. We follow the same procedure as above, along with checking their company bio as often the undergrad major is listed there.

We construct the following variables using the above data:

- *Whether Attended a Top 20 University:* Using the information on the undergraduate university, we construct this variable with data from the US News Rankings [18]. The top 20 universities, according to the website, are: Princeton, Harvard, Columbia, MIT, University of Chicago, Yale, Stanford, Duke, University of Pennsylvania, John Hopkins, Northwestern, California Institute of Technology, Dartmouth, Brown, Vanderbilt, Cornell, Rice, University of Notre Dame, UCLA, and Washington University in St. Louis.

- *Age:* The age of mentor and founder are constructed from the graduation year of the individual. If the individual didn't graduate or graduated late, efforts were made to ensure that the graduation year was adjusted to reflect the true age of the individual.

Some additional caveats to the data collected - sometimes, the name of the company changed and therefore, we would first check on Crunchbase to see if the name had changed (they have a variable that is called 'Also Known As' for the company); otherwise, do a google search to check if the company name stayed the same.

[18] https://www.usnews.com/best-colleges/rankings/national-universities, as accessed on June 5, 2019

Further, we do not have data on all founders, and therefore, the number of founders indicates the number of founders who founded the company, whereas all the other variables are based on the number of founders who have mentor matches.

A.2 Funding

As above, the GERN dataset has data on the amount invested in the company. We verify and collect this data using only Crunchbase. Data on funding is collected for every year between 2010 and 2018 (till 23 July 2018). Any funding before 2010 is clubbed together. Using Crunchbase, we also collect data on if the company is still active or not (this may not be a very accurate indicator), and if it is closed, the year of closing. Since many of these are startups, we also gather data on any acquisitions that might have taken place, along with the date and amount (if available). Using this data, we construct the following variables - indicators for whether or not each founder's company had received funding by 2013, 2015, and 2018, the amount of funding received by 2013, 2015, and 2018, and finally, a categorical variable – whether received zero funding, received between 0 and \$1.5 million, and finally, received more than \$1.5 million.