Diversity and Performance in Entrepreneurial Teams*

Sophie Calder-Wang, *The Wharton School, University of Pennsylvania*
Paul A. Gompers, *Harvard Business School, and NBER*
Kanyuan (Kevin) Huang, *The Chinese University of Hong Kong, Shenzhen*

February 3, 2023

Abstract

We study how diversity affects the performance of entrepreneurial teams by exploiting a unique experimental setting in which over 3,000 MBA students participated in a business course to build startups. First, we quantify how selection based on shared personal characteristics contributes to the lack of diversity. Next, when teams are formed through random assignment, we estimate that greater team diversity leads to poorer performances. However, when teams are formed voluntarily, the negative performance effect of diversity becomes greatly alleviated. Lastly, teams with more female members perform better when their faculty advisor is female. These findings suggest that policy interventions to improve diversity should consider the process by which teams are formed, as well as the role of mentoring, to achieve its intended performance goals.

JEL No. J1, J15, J16, L26, M12

* We are grateful to Kari Limmer at Harvard Business School MBA Student and Academic Services for providing us with extensive course information. This paper benefited from conversations with seminar and conference participants at Harvard Business School, Stanford Graduate School of Business, Carnegie Mellon University (Tepper), University of Utah (Eccles), NBER Summer Institute, Kenan Institute Frontiers of Entrepreneurship Conference, and Workshop on Entrepreneurial Finance and Innovation. Patrick Sweeney and Will Levinson provided excellent research assistance. Support for this research was provided by the Division of Research at HBS. All errors are our own.

1. Introduction

Improving diversity, equity, and inclusion has increasingly become a key objective for many corporations, affecting hiring and promotion decisions at all levels in the workplace (Pedulla, 2020; Temple-West and Edgecliffe-Johnson, 2020). A growing number of regulations have made explicit mandates on diversity on corporate boards as a government policy, such as the gender quota in Norway (Ahern and Dittmar, 2012; Masta and Miller, 2013; Bertrand, Black, and Jensen, 2019), and, more recently, the diversity quota in California for under-represented minorities.[1] The SEC has also approved Nasdaq's request to mandate diversity disclosures.[2] Yet, the economic implication of diversity on firm performance is difficult to estimate empirically due to the inherently endogenous team formation process.

In this paper, we focus on the issue of diversity in the entrepreneurial setting. Despite being the backbone of innovation and economic growth (Gornall and Strebulaev 2015), the entrepreneurial ecosystem suffers a striking lack of diversity among start-up founders. Women make up for less than 15% of start-up founders, even though over 40% of science and engineering master's degrees and MBA degrees are awarded to women; Hispanics and Blacks make up for fewer than 5% and 1% of all venture-capital backed founders, respectively (Calder-Wang and Gompers 2017; Calder-Wang, Gompers, and Sweeney 2021). Why are entrepreneurial teams so homogenous? Does team diversity lead to better entrepreneurial outcomes? How would policy intervention, such as mandated gender and racial diversity, affect performance?

To answer these questions, we exploit a unique series of quasi-random variations in a mandatory first-year course to build start-up companies. The team-based course was taken by over 3,000 MBA students from the Classes of 2013 through 2016 at Harvard Business School (HBS). Employing rich individual-level data, our paper fills an empirical gap by not only quantifying the various factors that shape team diversity, but, more crucially, how horizontal (within-team) diversity and vertical (supervisor-team) diversity affect entrepreneurial performance.

[1] As of January 1, 2021, public companies in California are required to meet the minimum requirements for female directors and directors from underrepresented communities on their boards as required respectively by Women on Boards (Senate Bill 826) and Underrepresented Communities on Boards (Assembly Bill 979). However, both laws have been ruled as unconstitutional and the state is appealing the rulings (Gupta, 2022; Fortt, Huber, and Vaseghi, 2022).
[2] As of August 6, 2021, Nasdaq's Board Diversity Rule requires companies listed on its exchange to publicly disclose board-level diversity statistics; and have or explain why they do not have at least two diverse directors. (See Securities Exchange Act Release No. 34-92590).

First, we quantify factors that contribute to a lack of diversity in entrepreneurial teams. For the 2014-2016 cohorts, students were free to select their co-founders from a given set of students. From these cohorts, we estimate the extent to which students form teams with those they share demographic characteristics and personal backgrounds, despite not having had social relationships with potential team members prior to starting business school. Specifically, we find that team formation based upon shared endowed demographic characteristics (e.g., gender, race, and ethnicity) is stronger than team formation based upon acquired characteristics (e.g., industry or education). Individuals are 25% more likely to form teams with people of the same race/ethnicity or gender relative to randomly matching. School ties and shared work experience increase the probability of co-founding a start-up business by 17% and 11%, respectively.

Next, we estimate the causal relationship between entrepreneurial team diversity and performance. By exploiting the quasi-random variation in diversity for the 2013 cohort, for which a computer algorithm was used to assign students to teams, we find that racial/ethnic diversity leads to a significant *decline* in team performance. In other words, randomly-assigned homogenous teams performed better than randomly-assigned teams with greater racial/ethnic diversity. When we look at the intersection of gender and race/ethnicity, we find that teams with both greater gender and racial/ethnic homogeneity performed the best. Because team assignment was dictated by a computer algorithm for the 2013 cohort, the identity of one's teammates is independent of unobservable preferences or characteristics, providing us with a powerful setting to causally estimate the impact of diversity on performance.

Interestingly, although we find that team diversity created by exogenous random assignment reduces performance in the 2013 cohort, diverse teams which are formed endogenously through voluntary matching for the 2014-2016 student cohorts do not suffer as much performance degradation. The negative performance gradient of racial and ethnic diversity in the 2013 cohort is alleviated by about 60% in the 2014-2016 cohorts when teams were formed voluntarily, suggesting endogenous team formation mitigates the detrimental effects of forced diversity.

A unique feature of our empirical setting which differentiates our study from the prior literature is the presence of both exogenous and endogenous team assignment mechanisms: it allows us to estimate the causal impact of the diversity intervention –namely, creating diversity through the exogenous assignment of teams relative to no intervention—on the performance gradient of diversity. In other words, our paper, in fact, contains two parallel experimental settings: one setting with team diversity created by random assignment for the 2013 cohort, which yields a negative causal estimate of the diversity gradient; and another setting with team diversity created through voluntary formation for the 2014-2016 cohort, which does not provide a causal estimate because teams are formed endogenously. Nevertheless, we can still credibly *compare the difference* in the performance gradient on diversity between these two team assignment mechanisms, because it is unlikely that student matriculation decisions are affected by the team assignment mechanism for one specific first-year class. As a result, it allows us to causally estimate a negative impact of the diversity intervention on the performance gradient. Namely, a forced process to create team diversity, relative to voluntary formation, results in worse performances as team diversity increases. In addition, we provide some suggestive evidence that, during voluntary formation, student pairs are more likely to match on unobservable characteristics (e.g., shared career interests) when they do not match on demographic characteristics, which offers a potential channel that could explain why voluntary team formation alleviates the negative performance impact of diversity on performance.

Finally, our paper causally estimates the performance effect of diversity in the vertical relationship between potential capital allocators and entrepreneurial team members. Exploiting the random assignment of faculty advisors and external judges to students, we find that for women, vertical gender ties improve the performance of entrepreneurial teams. Because faculty advisors are randomly assigned, students had no control over who their faculty advisors were, which means these gender ties are exogenous. When the faculty advisor is a woman, the performance of a team with a one-standard-deviation higher fraction of women increases significantly by 15%. We do not find a performance effect of gender ties for male faculty advisors.

Because of the uniqueness of our setting, we are able to differentiate the role of advising separately from the role of evaluation in the vertical relationship, which is novel compared to the existing literature.[3] Specifically, team performance in the course is evaluated by both a faculty advisor and a panel of external judges, all of whom are randomly assigned. While we find significant positive effects of female faculty advisors on female student teams, we do not find any performance effects of female judges who serve as one-time evaluators.

Taken together, it suggests that the positive performance impact of gender ties between female

[3] Existing literature does not disfferentiate the role of advisors from evaluators and yields mixed results (e.g., positive effects of female professors on female students in Carrell, Page, and West 2010; negative effects of co-ethnic supervisors on workers in Marx, Pons, and Suri 2021; and no effect of female evaluators on female candidates in Bagues, Sylos-Labini, and Zinovyeva, 2017).

faculty advisors and female students is likely a result of more effective advising and mentorship throughout the semester when the faculty advisor oversees the course and interacts with their students. Our results highlight the importance of mentorship and the positive performance impact women experience when mentored by women (Athey et al. 2000, Carrell et al. 2010, Matsa and Miller 2011).

Our paper contributes to several strands of literature. First, we contribute to the small but growing body of empirical research on the causal impact of diversity on productivity. Theoretical work on the performance implication of team diversity focuses on the trade-off between knowledge complementarity and communication costs (Lazear 1999; Prat 2002; Alesina and La Ferrara 2005; Lyons 2017). Heterogeneous teams benefit from more diverse pools of skill and knowledge (Hong and Page, 2004). Still, at the same time, differences in gender, race/ethnicity, culture, and native language may increase the cost of efficient communication and/or reduce knowledge sharing among team members, thus lowering productivity (Coffman, 2014; Sandvik, Saouma, Seegert, Stanton, 2020). A few existing studies find a negative impact of diversity on productivity when performing a set of well-defined tasks (Hjort 2014, Marx, Pons, and Suri 2021, Aman-Rana et al. 2022).[4] Our work builds on these previous studies by examining the diversity implication of performing the considerably more complex task of starting an entrepreneurial venture, where, theoretically, the benefit of team diversity through knowledge complementarity could be greater. [5] Yet, the random assignments of afforded in our empirical setting allows us to show that team diversity still leads to performance declines even for such complex tasks.

Our paper is unique in the literature where we observe outcomes in both voluntary and randomized assignment mechanisms in otherwise the same setting to study the performance impact of diversity. Indeed, existing empirical methods either perform observational studies of endogenously formed teams (van Knippenberg and Schippers, 2007) or conduct field experiments to perturb team diversity (Bandiera, Barankay, Rasul, 2013; Hoogendoorn, Oosterbeek, Van Praag, 2013). We highlight a material improvement in the performance-diversity gradient when the team formation mechanism is changed from exogenous assignment to voluntary formation. Our finding

[4] Hjort (2014) and Marx, Pons, and Suri (2021) study the effect of diversity on labor productivity on flower distribution and voter registration in Kenya. They find horizontal diversity in ethnicity reduces factory worker productivity because workers in heterogeneous teams are more likely to complain about their teammates. Aman-Rana et al. (2022) also find that randomly assigning a male co-worker to a female worker reduces the female worker's productivity in a lab experiment of taking a half-hour multiple choice question test.

[5] Yet, the need for efficient communication to perform more complex tasks is also likely higher. Therefore, theory does not necessarily predict whether the net impact will be positive or negative.

suggests that policy interventions targeting greater diversity should consider the process by which diverse teams are formed to prevent the potential negative performance impact of diversity.

Second, compared to the previous literature, we provide an appropriate setting to examine the performance impact of diversity in entrepreneurship specifically. Indeed, our study is particularly relevant to the U.S. entrepreneurial ecosystem: we focus on MBA students at Harvard Business School, where the main criteria for evaluating the start-up businesses were business viability, as judged by experienced venture capital investors and entrepreneurs. Each founding team received sizable initial funding, which resemable seed money from angel investors. Some projects received subsequent VC funding after the field course.[6] Further, among the 3,864 MBA students in our sample, over 20% of them work in venture capital or technology-related startups after graduation, representing a sizable labor inflow to the entrepreneurial ecosystem. In fact, Harvard Business School MBAs represent 22% of the venture capitalists and 16% of founders in the US who have MBAs, by far the largest source of investors and founders in the venture capital ecosystem (Calder-Wang and Gompers 2017). In other words, compared to existing work ((Hoogendoorn et al., 2013; Hoogendoorn and van Praag 2012),[7] our setting is more representative of the typical entrepreneurial environment in the US than existing studies.

Our study are differentiated from prior studies in entrepreneurial finance in that we focus on entrepreneurial teams. The vast majority of the prior literature has been on individual entrepreneurs, i.e., how founders' characteristics, such as gender, ethnicity, and prior experience, affect start-up performance or VC fundings (Ewens and Townsend 2020, Gompers et al., 2005; Gompers et al., 2010; Hebert, 2020; Cahn et al., 2021; Gottlieb et al., 2021; Hacamo and Kleiner 2022), much less research has been conducted on the dynamics of entrepreneurial team formation and how team composition affects entrepreneurial success. The richness of our data allows us to conduct a systematic analysis of the role of homophily along various demographic characteristics and personal backgrounds in the entrepreneurial team formation process. Compared to the broader literature on

[6] For instance, HourlyNerd (now Catalant), an online platform that matches freelance consultants to companies, was started during the class and raised $145 million in subsequent VC funding.

[7] One set of studies that examine the causal link between diversity and entrepreneurial team performance are Hoogendoorn et al. (2013) and Hoogendoorn and van Praag (2012), where the authors run a field experiment with freshmen from a Dutch college. Specifically, they vary the share of women and non-Dutch students in the team. However, contrary to our findings, they do not find consistent results with respect to diversity in ethnicity (i.e., Dutch vs. non-Dutch), likely due to differences in the setting with different racial and ethnic groups.

homophily,[8] our paper can rigorously quantify the strengths of homophilous preference on matching because we can observe the entire choice set of one's potential business partners and that these potential partners almost universally had no connection prior to starting business school. In contrast, naive estimates of homophily using observational data tend to be upward biased because one's choice set is unobserved, and ties can only be formed based on one's pre-existing network. Thus, our study quantifies the impact of homophily that is not biased by unobserved personal networks. More importantly, the random assignments of within entrepreneurial teams, as well as between teams and faculty advisors, allow us to examine the causal implication of horizontal and vertical diversity on entrepreneurial teams in a setting that is pertinent to the U.S. entrepreneurial ecosystem.

Lastly, we contribute to empirical studies of the performance implication of vertical diversity. Existing literature on the impact of vertical relationships is mixed (e.g., positive effects of female professors on female students in Carrell, Page, and West 2010; negative effects of co-ethnic supervisors on workers in Marx, Pons, and Suri 2021; and no effect of female evaluators on female candidates in Bagues, Sylos-Labini, and Zinovyeva, 2017). In our study, we leverage the random assignments of both faculty advisors and external judges to identify the causal impact of mentorship from female advisors, separately from female evaluators, on the success of women. In other words, our setting allows us to highlight that the nature of the vertical relationship (i.e., mentorship v.s. evaluation) matters for its performance implication.

In our setting, the relationship between faculty advisors and student teams resembles that between venture capitalists and entrepreneurial teams, where venture capitalists may provide guidance to entrepreneurs over an extensive period of time (Hellman and Puri, 2002). Our results also suggest that in capital markets, diverse investors could also play an important role in mentoring and advising minority entrepreneurs (Ewens and Townsend 2020). More generally, our finding suggests that vertical ties in the form of mentorship could be an effective channel to improve the performance of minorities.

The remainder of the paper is organized as follows. Section 2 describes the empirical setting, and

[8] The existence of homophily is widely documented in various settings, e.g., in marriages (Kalmijn 1998; Fiore and Donath 2005), close friendships (Marsden, 1987, 1988; Currarini, Jackson and Pin 2009), professional networks (Gompers, Mukharlyamov, and Xuan 2016; Kleinbaum, Stuart, and Tushman, 2013; Ruef et al. 2003; Reagans 2011; Sorenson and Stuart 2001) and acquaintances (Hampton and Wellman 2000). However, one limitation of these studies is that the observed homophily in these settings cannot differentiate the impact of homophilous preference from pre-existing personal networks.

section 3 describes the data. In section 4, we present results on homophily in team formation. We present our main results on the performance implication of horizontal team diversity in section 5 and our results on vertical diversity in section 6. Section 7 concludes.

2. Empirical Setting

First-year MBA students at Harvard Business School were required to take FIELD (Field Immersion Experience for Leadership Development) from 2012 through 2015 (MBA class years 2013-2016). In the third module of the course (i.e., FIELD 3), students were required to work in small teams to develop and launch a start-up business. FIELD 3 was designed to allow MBA students to "hone their collaborative skills while experiencing the challenge and excitement of being an entrepreneur."

The general structure of the course was as follows:

i. At the beginning of the spring semester, students formed a team of 5 to 7 with members all from the same section.[9] Each team is endowed with seed funding of $3,500 to $5,000.[10]
ii. Throughout the semester, these student teams worked together to develop and build their start-up businesses. They were required to develop a business idea, gather market feedback, create the product/service, manage external resources and vendors, and market and sell the product/service.
iii. After three months, on "Launch Day," each team made a presentation, and those that did not have a product ready to sell were moved to the "Failed Business Track" at the discretion of the faculty member leading the section.
iv. Finally, at the end of the semester on "IPO Day," the surviving teams proceeded to present their projects to a panel of judges from the academic, corporate, entrepreneurship, and venture capital industries. The judges ranked the teams by determining whether they had demonstrated product demand, whether the business was viable (defined as positive cash flow in a five-year period), and their ability to create the most value.

[9] Each year, approximately 900 students matriculate at the Harvard Business School. They are divided into 10 sections of approximately 90 students each. All students take the same first-year required courses with the same section-mates throughout the year.

[10] In 2013 and 2014, the initial funding was $5,000. In 2015 and 2016, the initial fund was $3,500, but students are allowed to request additional funding later on during the semester. The maximum funding should not exceed $6,500.

Importantly, FIELD 3 was taught to 4 cohorts of students using the same course structure but with one critical difference in terms of the team formation mechanism. Team membership was randomly assigned for the Class of 2013 by a computer algorithm developed by the HBS administration. For the Class of 2013, the assignment algorithm randomly formed teams conditional on their observed characteristics. Concretely, the algorithm was developed to ensure that the composition of each team created by the computer approximately reflected the overall composition of the entire section in terms of gender and whether a student was from the US or international. In contrast, teams were formed voluntarily by students for the Class of 2014, 2015, and 2016.

The structure of the Harvard Business School curriculum and the set-up of the FIELD 3 course offer us a unique and powerful setting, allowing for credible identification of the impact of diversity on performance. First, because the assignment of the team members was exogenous for the Class of 2013, the identity of one's teammates and their characteristics were independent of student characteristics or unobservable student preferences. From the student's perspective, the identity of their teammates was exogenous.

Second, the choice of the team formation mechanism (i.e., computer assignment vs. voluntary formation) is also orthogonal to student characteristics, allowing for credible inference on the impact of such diversity intervention. We believe it is rather implausible that a sizable portion of the students deferred their matriculation decisions based on the likely team assignment process for one of the three modules in one out of over ten required first-year courses. In other words, the random assignment of the team-formation mechanism across different cohorts allows us to identify the performance impact of the diversity intervention created by exogenous assignment relative to the control cohort with no intervention.

Third, because HBS utilizes a sectioning algorithm that divides the entire student body into ten equal-sized sections and teaching is performed in parallel across all sections, the assignment of the faculty advisor who supervised all FIELD 3 teams within a given section was also exogenous to student characteristics, allowing us to obtain credible inference on the performance implications of vertical diversity. These students had no ability to choose which section they joined. Each section reflected the student composition of the entire class. Shue (2013) and Lerner and Malmendier (2013) utilize the randomization in the HBS sectioning algorithm to derive the causal impact of peer effects. It is important to note that neither the balanced section assignment in Shue (2013) and Lerner and Malmendier (2013) nor the balanced team assignment in our setting create a problem for the validity of the empirical methodology: From an individual student's perspective in the Class of 2013 cohort, who their faculty advisor will be and who their team members will be for the course is exogenous to their own preferences.

3. Data

We collect an extensive set of information from the Office of MBA Student and Academic Services at Harvard Business School. It includes anonymized student characteristics and background information, the team membership of each student, the team performance, and the characteristics of the panel of judges at the IPO Day.

Student Characteristics and Background:

We obtain anonymous data on student characteristics from the Office of MBA Student and Academic Services. We observe the gender, race/ethnicity, home country, undergraduate institution, past employers, and past industry experience of each MBA student from the class years 2013 to 2016. We were not provided with students' actual names.

Table I reports summary statistics for the 3,684 MBA students in our sample. Women make up 41% of the total student population. Approximately 38% of the students are white Americans, 12% are Asian Americans, 5% are Black, 4% are Hispanics, and 35% are international students. India, Canada, and China represent the top three origin countries for international students, as shown in Appendix Table I. In terms of past work experience, roughly half of the students worked in finance or consulting before business school. Not surprisingly, the big three consulting firms (McKinsey, Bain, and BCG) and bulge bracket investment banks (Goldman and Morgan Stanley) are the top five past employers for Harvard MBA students, as shown in Table II. On average, 11% of students had experience in the technology industry. 25% of the MBA students graduated from Ivy League schools; Harvard, Stanford, and the University of Pennsylvania are the top 3 undergraduate institutions (Table II).

Team Membership and Validation of Assignment Mechanism:

We also obtain the composition of all teams. From 2013 to 2015, there were 150 teams in each class year, and the average team size was 6. In 2016, the average team size was changed to 5, and there were 180 teams.

We provide some initial graphical evidence showing that the computer assignment of teams for the Class of 2013 sought to create balanced teams in terms of gender and international status. Figure

I Panel A shows the distribution of the number of female students on each team for both team assignment mechanisms. For the Class of 2013, the distribution of the gender composition indicates that the computer algorithm created gender-balanced teams. Out of 150 six-person teams, 62% had two women, and 38% had three women. There were zero teams with more than three women. There were also zero teams with no women. In sharp contrast, when students formed teams voluntarily in 2014 and 2015, out of 300 six-person teams, 12% of the teams had no women, 12% of the teams had one woman, 53% of the teams had 2 or 3 women, and another 19% of the teams had four or more women. Figure I Panel B shows the distribution of international students across the team assignment mechanism. While there were only two teams with no international students in 2013 when teams were randomly assigned by the algorithm, 16% had no international students in 2014 and 2015 with voluntary team formation. Similarly, only 4% of the teams had more than four international students in 2013 with computer assignment, but 20% had over four international students with voluntary formation.

Next, we provide a framework to validate the conditional random assignment used in 2013 more formally. The framework also forms the basis for identifying and quantifying the strength of homophily in team formation when students formed teams endogenously. Specifically, we construct student-student pairs by matching each student to every other student within the same section and year. This process creates 335,686 potential pairs, with 81,368 potential pairs for the Class of 2013 and 254,318 for the Class of 2014-2016. We then create a dependent variable *real_match*, which equals one if the two students are members of the same team and 0 otherwise. The independent variable gender (race/ethnicity, school, industry) tie equals one if two students belong to the same gender (race/ethnicity, school, industry) group.[11] Our data construction method is similar to Louch (2000).

To illustrate, consider the following example: James Brown is a student in Section A, which has 90 students, and he needs to form a team of six. We create 89 student-student pairs by matching Mr. Brown to all his section mates, where each pair is a potential teammate with whom Mr. Brown could form a team with. If the match happened randomly, Mr. Brown would pair with an arbitrary teammate with a probability of approximately 5.6% (5 out of 89). The outcome variable *real_match*

[11] Race/ethnicity ties are defined as follows: For domestic students, the race and ethnicity groups are Whites, Asian Americans, Latino, and Blacks. For international students, the race and ethnicity groups instead measure the broad regions that the students come from: Europe, Latin America, South Asia, East Asia, the Middle East, and Africa. Two students share the race/ethnicity tie when they belong to the same racial or ethnic group if they are domestic, or if they hail from the same region if they are international.

equals 1 for the five pairs with whom Mr. Brown is matched in the same team. To measure factors that affect team formation, we compare the probability of being in the same team (*real_match*) when a student pair share the same characteristics (e.g., gender, race/ethnicity, school, and industry) to the probability of matching when a student pair have different characteristics.

Indeed, Table V validates that, once controlling for gender and international status, the matching probabilities generated by the computer algorithm in 2013 are *independent* of shared demographic characteristics or personal backgrounds. In column (1), shared gender or shared international status negatively predicts matching probabilities, indicating that the computer algorithm sought to create balanced teams on these two characteristics. However, other coefficients such as race/ethnicity ties, school ties, and industry ties have no predictive power on whether two students will match in a team. Columns (2) to (3) show that the same patterns hold true with subsamples of student pairs that either share the same gender or belong to different gender groups. Consistently, columns (3) to (6) show that other personal backgrounds do not matter to team assignment in the subsamples of student pairs that share/do not share their international status. As we will describe in the next section, the matching coefficients for teams created by the computer assignment stand in sharp contrast to the matching coefficients for the 2014 to 2016 cohorts when teams are formed voluntarily.

Team Performance and Judge Characteristics:

Beyond student characteristics and team membership, we also collect information about team performance from the MBA program office, summarized in Table III. Specifically, we code team outcomes into four binary indicators:

1) IPO Day: We observe whether a team progresses to the IPO Day. Approximately 75% of the teams were determined to be sufficiently developed to present on IPO Day. Otherwise, it was placed in the "Failed Business Track."
2) Viable: A team that presented on IPO Day was deemed by judges to be a viable business if they believe the business could be cash flow positive over a five-year period. Roughly 50% of all projects were deemed "viable".
3) Section Top 3: A project was ranked in the top 3 of their section by a panel of judges. Because a typical section has approximately 15 teams, about 20% of all projects were ranked as Section Top 3.

4) Class Top 3: A project was ranked as top 3 in the entire class. Since there were 150 teams in each cohort in 2013-2015 and 180 teams in 2016, approximately 2% of all projects were ranked as Class Top 3.

Correspondingly, we construct a composite performance score based on the median of the quantile of the team's outcome. For those who did not progress to IPO Day, their performance score is set to 0.125 because 25% of the teams do not progress and the median quantile of this group is 0.125. Analogously, among those who advanced to the IPO Day, but the project was deemed not viable, their performance score is set to 0.375 because they fall between the 25th and 50th percentile of the class. Among those who had a viable project but not Section Top 3, their score is 0.65, as they fall between the 50th and 80th percentile of the class. Section Top 3 teams are assigned a score of 0.89, as they fall between the 80th and 98th percentile. Finally, Class Top 3 teams are assigned a score of 0.99, as they fall in the top 2 percentile. Our performance measure is increasing in the project outcome.

In addition, for every section in 2014-2016, we obtained the demographic characteristics (e.g., gender, race, and ethnicity) of the panel of judges at the IPO Day. Each panel comprises a faculty advisor who led instructional sessions throughout the semester and external judges from the industry, as shown in Table IV. There are 21 unique faculty advisors, out of which 5 of them are women (24%). There are 76 unique external judges who participated in the IPO Day, 26% of whom are women.

4. Homophily in Team Formation

In this section, we show that homophily along the dimension of gender, race/ethnicity, educational background, and past work experience all play a significant role in team formation when teams are formed voluntarily. We also quantify the relative strengths of different dimensions, showing that demographic ties (i.e., shared gender or race/ethnicity) play a larger role than ties on acquired characteristics (i.e., shared educational institution or past industry background).

Although the homophily phenomenon of "birds of a feather flock together" is well-documented in both sociology (McPherson, Smith-Lovin, and Cook 2001) and economics (Jackson 2014; Bertrand and Duflo 2017), one advantage of our setting is that we can precisely quantify the relative strengths of different components of personal characteristics, because we have the ability to precisely measure the characteristics of all those in one's choice set, namely everyone in the same section. It is much better measured than in typical social settings such as friendship or professional networks, where the potential choice set is challenging to define.

Using the pseudo pair specification described in the previous section, Table VI column (1) presents the regression results for matching from 2014 to 2016 when students were allowed to choose their own teams. Race/ethnicity ties increase the probability of matching by 1.4 percentage points. Given the base rate of matching is 5.6%, this represents a 25% increase from the baseline probability of randomly matching with a student from the same race/ethnicity. Similarly, we find shared gender increases the probability of matching by 1.3%, corresponding to a 23% increase relative to the baseline. Attending the same undergraduate institution increases the probability of matching by 0.85%, a 15% increase from the baseline. Having industry experience in the same sector increases the matching rate by 0.62%, an 11% increase from the baseline. All results are significant and economically meaningful. Table VI column (2) reports the regression result using the 2013 subsample. Given that teams were randomly assigned, the coefficients on race/ethnicity tie, school tie, and industry tie are statistically not different from zero. The negative matching coefficient in front of gender reflects HBS's gender-balanced assignment mechanism.

Table VI column (3) tests whether homophily based on endowed demographic characteristics is stronger than homophily based on acquired characteristics. The variable *Endowed Demographic Match* is an indicator variable equaling one if race/ethnicity tie or gender tie equals one. *Acquired Characteristics Match* equals one if school tie or industry tie equals one. In the 2014-2016 subsample, the coefficient on Endowed Demographic Match is more than twice as large as the coefficient on Acquired Characteristics Match, and the difference is statistically significant at the 1% level.

The main homophily results found here are robust to alternative statistical inference procedures. In the main tables, the standard errors are clustered at the student level. Namely, every 89 pseudo pairs for every given student is considered a cluster. Such clustering procedure is commonly used in the literature (Sufi 2007, Corwin and Schults 2005). Alternatively, we still find all the results remain statistically significant using a randomization inference procedure. Specifically, we create 1,000 random permutations of alternative team formations and re-run the matching regression on the generated data. In Appendix Table III, we find that all of the actual matching coefficients belong to the extreme right tail of the simulated coefficients, admitting a p-value smaller than 5%. Moreover, we find that the standard errors of the simulated coefficients fall closely in line with the analytical standard errors (e.g., 0.0012 vs. 0.0011 for the race/ethnicity tie standard errors), suggesting that asymptotic normality is a reasonable assumption for our setting.

4.1 Breakdown of Homophily by Subgroups

Because we have detailed demographic and personal backgrounds of all students, we further investigate the relative strengths of homophily by subgroups.

By Race / Ethnicity / Nationality:

Along racial and ethnic lines, the propensity to match is strongest among international students hailing from the same region. Among domestic students, we find that both Asian Americans and White Americans have strong tendencies to form teams within their group.

Table VII, column (1) shows that the probability of matching based upon shared race/ethnicity increases by 1.2% and 1.4% among White American and Asian American MBA students, respectively, translating to over 20% increase relative to the baseline of random matching. The coefficient for Black students is 1.3 %, but we lack statistical power likely because Black students only make up 5% of the student body. Latino Americans seem no more likely to match with other Latino Americans.

The propensity to match is highest among international students from the same region. An international MBA student is 4.0% more likely to find a teammate from the same region, three times greater than the effect among White and Asian Americans. A detailed breakdown of international students by region in Appendix Table IV shows that the increase is highest among students from East Asia, the Middle East, and Latin America. The coefficients for these groups are around 6%, over 100% larger relative to the baseline and twice as large as the coefficients for European and South Asian students.

By Education and Past Experience:

In Table VIII, we examine the effect of education ties and industry ties on matching in the student teams. In column (1), education ties are much stronger among students from non-Ivy League schools. Specifically, attending the same non-Ivy college increases the matching probability by 1.9%. In column (2), we break down the industry ties by industry sectors. We find the effect strongest among students who worked in non-finance, non-consulting, and non-technology industries, increasing the matching rate by 2.2%. The magnitude of the effects is similar among finance, technology, and consulting industries, which is around 0.45%. Overall, we do not attempt to provide a sociological theory as to why certain subgroups exhibit stronger homophilic tendencies than other subgroups. Still, we quantify the prominent patterns that are present in the data.

4.2 Gender Differences in Homophily

In this section, we provide evidence that men exhibit stronger homophilic tendencies than women for all other non-gender-related demographic and personal characteristics, including race/ethnicity, education, and past experience. In Table IX columns (1) and (2), we find that the coefficients on race/ethnicity, school, and industry tie are statistically larger for men than women. We do not observe that women match with section-mates based on education or industry background at all. These results contrast with Brashears (2008), where homophily in education level is uniform among males and females using 1985 general social survey data. On the other hand, our results here are consistent with the patterns found in real-world venture capital settings. Using the deal-level data from Calder-Wang and Gompers (2021), in Appendix Table VI Panel A, we find that women venture capital partners are much more likely to invest in women entrepreneurs. However, Panel B shows that women VCs are *less* likely to match on the race and ethnicity of the entrepreneurs compared to their male venture capital partner colleagues. Men in venture capital are more likely to invest in entrepreneurs with shared ethnic ties across all ethnicities. Relatedly, similar patterns are also documented among Wall Street analysts in Fang and Huang (2017).

Overall, this section establishes two main results. First, we quantify the relative strength of homophily in both demographic and acquired characteristics. We find that shared characteristics, especially in gender and ethnicity, play economically significant roles in the matching process. Our setting allows us to estimate the strength of various characteristic ties resulting from homophilous preference without being confounded by the pre-existing personal network. Second, we examine the gender difference in homophily, an often-overlooked aspect in prior studies. We document that men exhibit a stronger tendency to match with peers with the same ethnicity/race, education, and industry backgrounds than women in our entrepreneurial class setting. These results contribute to the literature by shedding light on the "black box" of the entrepreneurial team formation process.

5. Performance Implications of Horizontal Diversity

In this section, we analyze the impact of team diversity on performance. We find that increased team diversity created by random assignment is harmful to performance. However, we find diverse teams formed voluntarily exhibit a much weaker penalty on performance. Overall, the finding highlights how diversity is created in an organization is instrumental to whether it generates better outcomes.

5.1 Definition of Diversity Scores

Our unit of performance analysis is at the team level. There are 150-180 teams in each class year, and each team has 5-7 students. We measure team diversity across four different dimensions: race/ethnicity, gender, school, and industry, and construct the diversity measure for each dimension as the following:

$$Diversity\ Score_i = 1 - \frac{\#\ of\ ties\ between\ team\ members\ in\ team_i\ with\ \ shared\ characteristics}{Total\ possible\ ties\ in\ team_i}$$

Or equivalently:

$$Diversity\ Score_i = \frac{\#\ of\ ties\ between\ team\ members\ in\ team_i\ with\ \ different\ characteristics}{Total\ possible\ ties\ in\ the\ team_i}$$

To illustrate our race/ethnicity diversity score, consider a team with six people: three are White, two are Asian Americans, and one is an international student from Latin America. Race/Ethnicity Score in this team will be 1-(3+1)/(5+4+3+2+1)=11/15, as there are three ties among three White team members, one tie between two Asian American students, and fifteen possible ties among all six team members. Gender Score, Education Score, and Industry Score are constructed analogously. Diversity is monotonically increasing in the score. It equals zero if everyone in the team is the same type and equals one if everyone has different characteristics.

Figure II plots the distribution of the race/ethnic diversity scores across different team assignment mechanisms. Panel A plots the probability distribution under the 2013 conditional random assignment compared with voluntary team formation. Notice that the score distribution under voluntary formation has a greater mass among lower-diversity score teams; namely, there are more homogenous teams. When plotted as a cumulative distribution function, Figure II Panel B implies a larger area under the curve for voluntarily formed teams than randomly assigned teams.

The average diversity score on race/ethnicity decreased from 0.76 under random assignment in 2013 to 0.72 under voluntary formation in 2014-2016. The average diversity score on gender also decreased from 0.56 for teams created under random assignment to 0.43 for teams formed voluntarily. The results above are consistent with stronger homophily under voluntary team formation, as documented in the previous section.

5.2 The Impact of Diversity on Performance

As described in the data section, we use the median quantile of the team's project ranking as our performance measure. In this section, we examine the impact of diversity on team performance.

Graphically, Figure III shows the binscatter of team performance on diversity scores. Panel A on the left indicates that among randomly assigned teams (2013), higher diversity scores correspond to poorer team performance. Panel B on the right shows that among voluntarily formed teams (2014-2016), the correlation between performance and diversity is much smaller, with the slope of the binscatter much flatter than the left panel.

The magnitude of the impact of diversity on performance is both statistically and economically significant. Table X column (1) shows the coefficient on the race/ethnicity score is -0.49 for the 2013 cohort with randomly assigned teams. Since the standard deviation of the race/ethnicity score is 0.18,[12] it suggests that a one-standard-deviation increase in racial/ethnic diversity leads to approximately a 9 percentile decline in the performance rankings (e.g., a decline from being ranked at 80th percentile to 71st percentile), or, equivalently, an 18% decline in performance (as the median team is ranked at the 50th percentile.)

In other words, in 2013, where teams were exogenously assigned, relatively more homogenous teams performed better than racially/ethnically more diverse teams. Because the assignment of teams is random in 2013, these students have no ability to select teams based on unobservable student quality or preference. Thus, the negative relationship between team diversity and team performance in 2013 admits a causal interpretation: Higher racial and ethnic diversity levels lead to worse team performance in our entrepreneurial team setting.

One key advantage of our current setting is that the conditional randomized team assignment used in 2013 is like a field experiment run by the HBS administration: Whether a given student will land in a diverse team or homogenous team is completely exogenous to their own unobservable quality or preference. Our setting bypasses the usual empirical challenge associated with interpreting the correlation between diversity and performance. In a typical organizational environment, one may be concerned that more diverse firms are better managed and thus could attract unobservably higher-quality candidates. More diverse firms may also attract candidates who are unobservably better at collaborating with colleagues from different backgrounds. Therefore, our estimate is less

[12] Because the distribution of the diversity scores will vary by team assignment mechanism, we calculate the "true" standard deviation using 1,000 simulations of team formations under pure, unconditional random assignment.

likely to suffer from the usual upward bias associated with such selection on unobservables.

Even though the random assignment successfully removes the scope under which the team diversity may be subject to unobserved student selection, another potential threat to our regression specification is that the diversity score may itself be correlated with other student characteristics that are directly predictive of performance, as we do not observe the same student across different assignment mechanisms.

For instance, more diverse teams may have a higher fraction of Asian Americans because the diversity score is calculated using student race and ethnicity. Meanwhile, we know that in the data, Asian Americans are more likely to have experience in the technology section (Appendix Table VII), which may be predictive of their performance in our entrepreneurship class. In other words, the presence of performance-relevant student characteristics that are also correlated with our diversity score can potentially create bias in our estimate.

To address this concern, in Table X column (2), (5) and (7), we directly control for student characteristics that may be indicative of their quality, but could be correlated with our definition of the diversity score. Specifically, we control for the percentage of students with prior startup experience, the percentage of students who came from top undergraduate institutions,[13] and the percentage of students graduating with MBA honor.[14] These variables are potential proxies for students' abilities. We observe results remain significant, and the magnitude of the coefficient stays similar (0.49 vs. 0.45). Moreover, columns (3), (6), and (9) show that the main results also remain unchanged after adding a variety of other plausible controls, including the fraction of native English speakers, and the fraction of students with work experience in consulting, finance, or technology. We recognize that there will always be a limit as to what we can control for explicitly. Nonetheless, the fact that none of the plausible performance-relevant controls had a substantial impact on the results gave us more confidence in the main specification.

Finally, we also find supportive evidence of the negative impact of diversity along other dimensions, including diversity in industry background and diversity in one's undergraduate institution, as shown in Table X column (1-3). Although the statistical significance is relatively weak,

---

[13] We classify as top universities the Ivy League schools (Brown University, Columbia University, Cornell University, Dartmouth College, Harvard University, Princeton University, University of Pennsylvania, and Yale University) as well as other top U.S. schools (Amherst College, California Institute of Technology, Duke University, MIT, Northwestern University, Stanford University, University of California, Berkeley, University of Chicago, and Williams College).
[14] Students whose grades fall into the top 20% of their section in both their first and second years are awarded the MBA degree with Distinction (i.e., honor). Approximately 12% of MBA students in our sample graduate with honor.

the fact that they are consistently negative across all dimensions and specifications provides us with confidence in the general direction of the result.[15]

5.3 The Impact of "Forced" Diversity on the Performance Gradient

Although we cannot interpret the correlation between diversity and performance causally for the cohorts formed voluntarily (2014-2016) because of the likely selection on unobservables, the presence of these control cohorts provides us the ability to identify the causal impact of the intervention itself (namely, the conditional random assignment utilized by the HBS administration relative to the voluntary team formation) on the performance gradient (e.g., the slope of performance on diversity).

Quantitatively, when HBS changed the way it creates team diversity from random assignment to voluntary formation, the negative performance implication of diversity on performance declined by about 60%. In Table X, comparing the results from the voluntary formation years in column (4) with the random assignment results in column (1), the coefficient in front of the race/ethnicity score is much smaller (-0.18 vs. -0.49). In a full-sample regression, the difference in these two coefficients is the causal impact of the diversity intervention on the performance gradient. In column (7), the coefficient on the interaction term Voluntarily Formed × Race/Ethnicity Score is 0.30, suggesting that changing the team assignment from random assignment to voluntary formation reduces the negative impact of diversity on performance by approximately 60% (0.30/0.49). However, it is also at the cost of creating fewer diverse teams overall.

Our unique empirical setting allows us to interpret the impact of diversity intervention (in the form of random assignment) on the difference in the performance gradient causally.[16] The relevant exogeneity condition is that the treatment assignment (i.e., random team assignment) is orthogonal to any unobservables that would cause the performance gradient to be different. We believe that it is rather implausible that the performance gradient on diversity may fluctuate from cohort to cohort from 2013-2016 in a way that systematically favors the 2013 cohort for some reason. Because the admission process and the school curriculum from 2013 to 2016 remained stable, the use of random assignment in 2013 is independent of any unobservables that may affect the counterfactual

[15]Note that Table X column (1-3) shows no statistically significant coefficients in front of the gender score. It is expected because the 2013 assignment algorithm creates gender-balanced teams by design, leaving little residual variation in the gender diversity of the team for us to exploit.

[16] A few recent papers find that the team formation process affects team performance (Chen 2017, Chen and Gong 2018, Aman-Rana et. al 2022). Different from these studies, this paper examines the effect of diversity on team performance and how it is shaped by the team formation mechanism.

performance gradient for the 2013 cohort.

After all, the presence of two different types of team assignment mechanisms is novel in the literature of field experiments. The ability to compare different assignment mechanisms suggests how team diversity comes about and the type of intervention used to achieve such diversity significantly impact the performance gradient. The random assignment utilized by the HBS program may be the most extreme version of diversity intervention where no information about match-specific qualities is considered at all. Consequently, the outcome of such diversity intervention is that the performance gradient on diversity becomes significantly more negative than a voluntary team formation baseline. To the extent we consider the diversity created by random assignment is "forced," these results suggest that voluntary team formation alleviates as much as 60% of the underperformance of forced diversity.

So, why does the team formation mechanism have such a large impact on the performance gradient of diversity? In other words, why does voluntarily formed diversity result in much smaller performance penalty than randomly assigned diversity? Although we cannot precisely observe every aspect of teamwork to pin down the specific mechanisms, our hypothesis is that a voluntary team formation process allows individuals to select on *unobservable* dimensions (e.g., shared industry interests, career goals, or risk-tolerance, etc.) that can facilitate better collaboration, conditioning on the same level of *observable* racial/ethnic diversity.

To test the hypothesis, in addition to the existing demographic data, we collected additional data that may be correlated with such "unobservable" dimensions from their Class Card profiles: MBA students list their career interests in their Class Card profile, which is an internal portal where students could disclose their career interests, accessible to their fellow classmates.[17] Indeed, when voluntary team formation is allowed, we find that racially/ethnically mismatched student pairs are more likely to match on career interests, as shown in Table XI. Therefore, the evidence on matching on career interests suggest that matching on such unobservables might be an important channel that may explain why the voluntary team formation mechanism can alleviate the negative performance impact of diversity on performance.

5.4 Robustness of the Performance Gradient

In this section, we explore a few additional properties and the robustness of the negative

[17] The most popular self-reported career interests include finance, technology, consulting, entrepreneurship, consumer products, leisure, healthcare, community development, education, and government, as shown in Appendix Table VIII.

performance gradient on diversity. Although we currently cannot isolate the mechanism of the level or the change precisely, some of these findings still shed light on the potential sources of the negative impact.

First, we examine the intersection of gender and ethnicity. Given the negative coefficients on the race/ethnicity score, we find that the performance lift from team homogeneity is driven by the quadrant where students match on *both* gender *and* race/ethnicities. Concretely, we now define a composite race/ethnicity-gender diversity score in which a student pair is considered to share the same characteristics when they share the same gender and the same race/ethnicity. In other words, the score is the lowest when the student pair matches on both dimensions.

Table XII column (1) shows that once we include the race/ethnicity-gender diversity score, the magnitude of the single-dimensioned race/ethnicity score becomes not statistically different from zero. The coefficient in front of the composite race/ethnicity-gender score is large and significant at -0.78. With a standard deviation of 0.1, a one-standard-deviation change in this composite score translates into a change of 7.8 percentiles in the project performance ranking. In column (2), we find that the intersection effect is similar for both men and women, although we have better statistical power for men. Columns (5) and (6) replicate our previous findings on the impact of forced diversity, where voluntary team formation largely alleviates the performance penalty of forced diversity. Overall, this analysis suggests that looking at multiple aspects of diversity at the same time may be important for understanding performance implications.

Second, we also examine the contribution of different racial/ethnic subgroups. In particular, one may be concerned that the negative performance implication is primarily driven by variations in the fraction of American students. To mitigate this concern, we partition the race/ethnicity score by different racial and ethnic subgroups.[18] Table XIII shows that the coefficients in front of the international student score, White American score, and Asian American score all remain negative and statistically significant. The coefficients on the Latino American score and Black score are negative but statistically insignificant, indicating that the fraction of Black (Latino American) and non-Black (non-Latino American) pairs does not affect team performance. However, this may be partly because these students account for a small fraction (5%) of the student body. Overall, these results indicate that the negative performance implication of diversity score is not driven by White

[18] The diversity score for each subgroup is constructed in the same way as the race/ethnicity score. For example, White American score equals 1 - (# of ties between White American team members in team i)/(Total possible ties in team i)

American students.

Even though we use the median percentile as the performance measure, we could also perform our analysis using the binary outcomes that indicate whether a team's performance is above a certain threshold, namely, IPO Day, Viable, Section Top 3, Class Top 3. Appendix Table IX shows that diversity hampers performance at every stage of the project progression for the 2013 cohort. On the left tail, 75% of teams progress to IPO Day with more homogenous teams being more likely to make it. On the right tail, only 2% of the teams are considered Class Top 3 with the most homogenous teams in terms of alignment in both race/ethnicity and gender being most likely to be ranked at the top. Consistent with our main results, such negative performance implications are also greatly alleviated across all levels of outcomes when teams are formed voluntarily.

Given the underlying distribution of the diversity score is skewed to the left, as shown in Figure II, we also provide a robustness test where the diversity score is measured in terms of its percentile in the underlying distribution of diversity scores generated with an unconditional random assignment. In other words, if there were no homophily, the diversity score percentile would admit a perfectly uniform distribution between 0 and 1. To the extent that there may be non-linearity in the raw diversity score, the variable transformation from diversity scores to diversity percentiles addresses this concern. We show in Appendix Table X that all main results on the negative performance impact and the improvement in the performance gradient due to voluntary formation remain robust under the alternative measure of diversity.

Overall, in this section, we provide credible inferences that team diversity negatively impacts performance, leveraging the unique empirical setting of the FIELD 3 course. We also find that the performance penalty of forced team diversity, such as those created by random assignment, becomes greatly alleviated when diversity is created with voluntary team formation. This result has important implications for policies that use gender/ethnicity quotas to promote diversity. Our results suggest that it may be essential to consider match-specific qualities beyond observable demographic characteristics in fostering a well-functioning diverse team.

6. Performance Implications of Vertical Diversity

In addition to the role of diversity at a horizontal level (i.e., among team members), prior work has also examined vertical diversity (i.e., between a supervisor and their subordinates). In our setting, we explore the relevant vertical relationships in FIELD 3, namely, the faculty advisors' role and external judges' role. Specifically, we examine whether a greater overlap between student attributes and the attributes of their faculty advisor and/or judges influences team performance. Because we have both mentors (faculty advisors) and pure evaluators (judges), we are able to disentangle how much of any performance benefit is driven by each factor.

One crucial advantage of our empirical setting is the Harvard Business School's randomized sectioning algorithm allows us to obtain credible causal inferences on the performance impact of vertical diversity. Every year, HBS randomizes every incoming class of over 900 students into ten equal-sized sections, where students in each section sit together to take their entire first-year required courses. These required first-year classes, including FIELD 3, are taught in parallel with the same curriculum content but by different faculty members. Thus, from the student's perspective, the demographic characteristics of their faculty advisor and the external judges are exogenous to their own characteristics.

The outcomes for each team were determined through the development of a start-up business and evaluation of those businesses by a panel of judges on IPO Day. The faculty advisor was a member of the HBS faculty who supervised the section over the entire semester. The faculty advisor was critical to the team's performance because of their role in teaching and advising students throughout the period. Each panel then had an additional four or five external judges from the industry. Because of the different roles played by the faculty advisor, we analyzed the attributes of the faculty advisor and the external judges separately.

Table IV reports the summary statistics on faculty advisors and judges' gender and race/ethnicity from 2014 to 2016. We were not able to obtain data from HBS on faculty advisors and judges for the class of 2013. Among the ten faculty advisors in each class year, there were three women in 2014 and 2015 and two women in 2016. The majority of the faculty advisors are White, with a few Black, Latino, and South Asians. There were no East Asian faculty advisors in our sample. There were more than 40 judges in each class year in our sample. The percentage of female judges increased from 14% in 2014 to 34% in 2016. The percentage of ethnic minority judges varied between 5% and 10% for each subgroup. Because there were so few minority judges, we focus on the gender ties between the faculty advisors or judges and the students.

In Figure IV, we sort all teams into four quartiles based on the percentage of female team members and plot their performances. Conditional on having a female faculty advisor, Panel A shows that team performance increases monotonically as the percentage of women on the team increases. In these sections, the percentages of teams progressing to the IPO day, being rated as viable, and being ranked section top 3 are 53%, 28%, and 8%, respectively, for teams with a low fraction of female members (Quartile 1). These numbers increase to 90%, 76%, and 38%, respectively, for teams with a high female percentage (Quartile 4). The economic magnitude of performance increase is large. For instance, teams with the highest number of female members were four times more likely to be in the section top 3 than teams with few or no female members when the faculty advisor was a female. However, we do not find any relationship between male faculty advisors and the performance of male-dominated teams. Figure IV Panel B shows that team performance does not vary with the percentage of women (or men) in the team in sections with male faculty advisors. The results offer some suggestive evidence that women's performance may be improved with women mentors and supervisors.

To statistically test whether teams with more female members perform better when the faculty advisor is also a female, we run the following regression specification:

$$\begin{aligned} Performance_i = {} & b_1 \times Faculty\ Advisor\ Female_i \times Female\ Team\ Member\ \%_i \\ & + b_2 \times Faculty\ Advisor\ Female_i + b_3 \times Female\ Team\ Member\ \%_i + controls_i \\ & + \epsilon_i \end{aligned}$$

where each observation is a team. The dependent variable is team performance. The key independent variable is the interaction term between being a female faculty advisor and the percentage of women on a team. Because the assignment of faculty advisors to sections is exogenous from the student's perspective, and it is unlikely that the gender of faculty advisors affects how teams are formed in the section, the coefficient on the interaction term, i.e., $\boldsymbol{b_1}$, can be interpreted as the causal impact of female faculty advisors on teams with a high female percentage.[19] Column 1 of Table XIV presents the regression results. Consistent with Figure IV, the coefficient on the interaction term on faculty advisors is positive and statistically significant at the 1% level, indicating teams with a greater number of female members perform significantly better in sections with female faculty advisors. The standard errors are clustered at the section level. The results remain robust if we cluster the results at the faculty level since a few faculty members taught multiple sections across different years.

[19] Since teams are formed at the beginning of the semester, faculty advisors are unlikely to impact how students choose their teammates. Appendix Figure 1 shows that the distribution of female team members does not differ by the faculty advisor's gender.

Meanwhile, when it comes to external judges, we do not find any performance impact of female judges on teams with a greater number of female team members. In the second column of Table XIV, we regress performance on the interaction term between Have Female Judge and the percentage of females on the team. Have Female Judge is a dummy variable that equals one if at least one female judge is on the panel. The coefficient on the interaction term is positive but not statistically significant. The magnitude of the coefficient is also much smaller.

Overall, taken together, we believe that our result does not indicate a more favorable (or stringent) ranking of teams with female team members by female evaluators (e.g., Bagues, Sylos-Labini, Zinovyeva 2017). Instead, the likely channel is that female faculty advisors may have provided better mentorship throughout the year and during the Field 3 course for female students.

7. Conclusion

In this paper, we leverage various sources of randomization unique in our empirical setting to study the impact of homophily on entrepreneurial team formation and the effects of diversity on performance. We quantify the relative strengths of homophily in gender, ethnicity, education, and industry background in entrepreneurial team formation. We find that the effect of endowed demographic attributes (gender and race/ethnicity) is much stronger than team choice based upon acquired characteristics (education and industry). We also find interesting gender differences in homophily. Men exhibit a stronger tendency to match with peers of the same ethnicity/race, education, and industry backgrounds than women. Then, when we examine the effect of horizontal team diversity on performance, we find that for teams in the 2013 cohort when team membership was exogenously assigned, greater diversity across race/ethnicity led to poorer performance than more homogenous teams. When team formation was endogenous, however, such underperformance was much alleviated. Lastly, in terms of vertical diversity, we find that shared gender ties between female team members and their female faculty advisors enhance team performance.

Our results have important real-world implications. Because the goal of the course was to form real start-up businesses, the performance was evaluated by a panel of experienced venture capital investors and entrepreneurs from the industry. In fact, a significant minority of businesses that started during FIELD 3 continued to operate after the course, with some attracting significant outside funding. Moreover, many of these MBA students have chosen to work in startups and the venture capital industry after graduation. Based on the exit surveys, over 20% of the graduating class entered the technology sector or venture capital during this period. Many of them later on progress to leadership positions in their field. It is reasonable to infer that such homophily found in our setting also exists in startup team formation, venture capital investing, and hiring.

Our results on the performance effects of horizontal diversity highlight the need to design and implement diversity policies thoughtfully. Although the conditional random assignment implemented for the 2013 cohort may be thought of as a draconian way to create balanced teams, it exposes the potential harm as we find a strong negative relationship between diversity and performance among these teams. The fact that much of the negative performance gradient is alleviated with voluntary team formation suggests that students could match on other characteristics that are not used by the computer algorithm, which in turn dampens the negative effect of diversity, such as shared career or personal interests, working styles, risk preferences, etc. Although we are not able to extract all of the information unobservable to the computer algorithm that has led to performance improvement under voluntary formation, we find some suggestive evidence that shared career interests seem to matter. Regardless, we offer a lower bound and an upper bound on the performance gradient with respect to two extreme scenarios: we allow for *no* selection on unobservables using the computer assignment in 2013, and we allow for selection on all unobservables using voluntary formation in 2014-2016. The results suggest that diversity interventions may create unintended negative outcomes unless adequate considerations about these match-specific qualities are considered.

In addition, to ensure the benefits of diversity in entrepreneurship, one needs to take into account how subtle treatment effects may dislodge existing biases. To harness the full benefits of diversity, policymakers need to eliminate bias against underrepresented groups. For instance, Calder-Wang and Gompers (2021) show that when venture capitalists have more daughters, they are more likely to hire a female investor, and subsequent firm performance improves after hiring.

Our results for the performance effects of vertical diversity have potentially important implications for female-led startups. The relationship between female teams and female faculty advisors in our setting resembles the relationship between female entrepreneurs and female VCs. Calder-Wang and Gompers (2021) and Gompers et al. (2020) document that female VCs (and entrepreneurs) are underrepresented and under-supported. An effective policy to help women succeed in entrepreneurship needs to take advantage of the superior mentorship that female venture capitalists may be able to provide to female entrepreneurs. It argues for increasing the number of women in venture capital as an effective strategy for greater representation and performance of female entrepreneurs.

Overall, our horizontal and vertical diversity results show that diversity diminishes performance in both of these two dimensions. To minimize the negative effects of horizontal diversity, policy interventions ought to take into account the process by which diverse teams are formed. In the meantime, to leverage the performance effects of vertical diversity, our paper highlights the importance of the role played by minority mentors, who may be particularly effective in improving the outcomes for minority groups. Together, these findings pave a potential path to better achieve both diversity and performance goals.

Reference

Ahern, K. R., & Dittmar, A. K. (2012). The changing of the boards: The impact on firm valuation of mandated female board representation. *Quarterly Journal of Economics*, 127(1), 137-197.

Alesina, A., & Ferrara, E. L. (2005). Ethnic diversity and economic performance. *Journal of Economic Literature*, 43(3), 762-800.

Aman-Rana, S., Minaudier, C., Pereira, B., Chaudry, S. (2022). Gender and Choice over Co-workers: Experimental Evidence. Working paper.

Athey, S., Avery, C., & Zemsky, P. (2000). Mentoring and diversity. *American Economic Review 90*(4), 765-786.

Bagues, M., Sylos-Labini, M., Zinovyeva, N. (2017) Does the gender composition of scientific committees Matter? *American Economic Review* 107(4): 1207-1238.

Bandiera, O., Barankay, I., & Rasul, I. (2013). Team incentives: Evidence from a firm level experiment. *Journal of the European Economic Association*, 11(5), 1079-1114.

Bertrand, M. and Duflo, E. (2017). Field experiments on discrimination. *Handbook of Economic Field Experiments,* 1, 309-393.

Brashears, M. E. (2008). Gender and homophily: Differences in male and female association in Blau space. *Social Science Research*, *37*(2), 400-415.

Cahn, C., Girotti, M. and Landier, A. (2021). Entrepreneurship and information on past failures: A natural experiment. *Journal of Financial Economics*, 141(1), 102-121.

Calder-Wang, S.Q. and Gompers, P. A. (2017). Diversity in innovation. NBER Working Paper, No. 23082.

Calder-Wang, S.Q. and Gompers, P. A. (2021). And the children shall lead: The effect of children's gender on hiring and performance in venture capital. *Journal of Financial Economics,* 142 (1), 1-22.

Calder-Wang, S.Q., Gompers, P. A., and Sweeney, P. (2021). Venture Capital's "Me Too" Moment. NBER Working Paper, No. 28679.

Carrell, S., Page, M., and West, J. (2010). Sex and science: How professor gender perpetuates the gender gap. *Quarterly Journal of Economics*, 125(3), 1101-1144.

Chen, R. (2017). Coordination with endogenous groups. *Journal of Economic Behavior & Organization*, 141, 177-187.

Chen, R. and Gong, J. (2018). Can self selection create high-performing teams?. *Journal of Economic Behavior & Organization,* 148, 20-33.

Coffman, K.B. (2014). Evidence on self-stereotyping and the contribution of ideas. *Quarterly Journal* *of Economics*, 129(4), 1625-1660.

Corwin, S.A. and Schultz, P. (2005). The role of IPO underwriting syndicates: Pricing, information production, and underwriter competition. *Journal of Finance*, 60(1), 443-486.

Currarini, S., Jackson, M. O., & Pin, P. (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, 77(4), 1003-1045.

Ewens, M. and Townsend, R.R. (2020). Are early stage investors biased against women?. *Journal of Financial Economics*, 135(3), 653-677.

Fang, L.H. and Huang, S. (2017). Gender and connections among Wall Street analysts. *Review of Financial Studies*, 30(9), 3305-3335.

Fortt, S., Huber, B., and Vaseghi, M. (2022). California Gender Board Diversity Law Is Held Unconstitutional. *Harvard Law School Forum on Corporate Governmence*.

Gompers, P., Lerner, J. and Scharfstein, D. (2005). Entrepreneurial spawning: Public corporations and the genesis of new ventures, 1986 to 1999. *Journal of Finance*, 60(2), 577-614.

Gompers, P., Kovner, A., Lerner, J. and Scharfstein, D. (2010). Performance persistence in entrepreneurship. *Journal of Financial Economics*, 96(1), 18-32.

Gompers, P. A., Mukharlyamov, V., & Xuan, Y. (2016). The cost of friendship. *Journal of Financial Economics*, 119(3), 626-644.

Gompers, P. A., Mukharlyamov, V., Weisburst, E., and Xuan, Y. (2020). Gender effects in venture capital. Forthcoming *Journal of Finance and Quantitative Analysis.*

Gornall, W., and Strebulaev, I. (2015). The economic impact of venture capital: Evidence from public companies, working paper.

Gottlieb, J.D., Townsend, R.R. and Xu, T. (2021). Does career risk deter potential entrepreneurs? *Review of Financial Studies*, forthcoming.

Hacamo, I. and Kleiner, K. (2022). Forced entrepreneurs. *Journal of Finance*, 77(1), 49-83.

Hebert, C. (2020), Gender stereotypes and entrepreneur financing. Working paper.

Hjort, J. (2014). Ethnic divisions and production in firms. *Quarterly Journal of Economics*, 129(4), 1899-1946.

Hoogendoorn, S., Oosterbeek, H., & Van Praag, M. (2013). The impact of gender diversity on the performance of business teams: Evidence from a field experiment. *Management Science, 59*(7), 1514-1528.

Hoogendoorn, S. and Van Praag, M. (2012). Ethnic diversity and team performance: A field experiment, working paper.

Hong, L. and Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences* 101(46), 16385–16389.

Jackson, M. O. (2014). Networks in the understanding of economic behaviors. *Journal of Economic Perspectives,* 28, 3–22.

Kleinbaum, A. M., Stuart, T. E., & Tushman, M. L. (2013). Discretion within constraint: Homophily and structure in a formal organization. *Organization Science, 24*(5), 1316-1336.

Lazear, E. P. (1999). Globalisation and the market for teammates. *The Economic Journal*, 109(454), 15-40.

Lerner, J. and Malmendier, U., (2013). With a little help from my (random) friends: Success and failure in post-business school entrepreneurship. *Review of Financial Studies*, 26(10), 2411-2452.

Louch, H. (2000). Personal network integration: transitivity and homophily in strong-tie relations. *Social Networks*, *22*(1), 45-64.

Lyons, E. (2017). Team production in international labor markets: Experimental evidence from the field. *American Economic Journal: Applied Economics*, 9(3),70-104.

Marx, B., Pons, V., & Suri, T. (2021). Diversity and team performance in a Kenyan organization. *Journal of Public Economics*, 197, 104332

Matsa, D. A., & Miller, A. R. (2011). Chipping away at the glass ceiling: Gender spillovers in corporate leadership. *American Economic Review*, 101(3), 635-39.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), 415-444.

Pedulla, D. (2020). Diversity and inclusion efforts that really work. *Harvard Business Review*.

Prat, A. (2002). Should a Team be Homogeneous? *European Economic Review, 46*, 1187-1207.

Sandvik, J. J., Saouma, R. E., Seegert, N. T., & Stanton, C. T. (2020). Workplace knowledge flows. *Quarterly Journal of Economics*, 135(3), 1635-1680.

Shue, K., (2013). Executive networks and firm policies: Evidence from the random assignment of MBA peers. *Review of Financial Studies*, 26(6),1401-1442.

Sufi, A. (2007). Information asymmetry and financing arrangements: Evidence from syndicated loans. *Journal of Finance*, 62(2), 629-668.

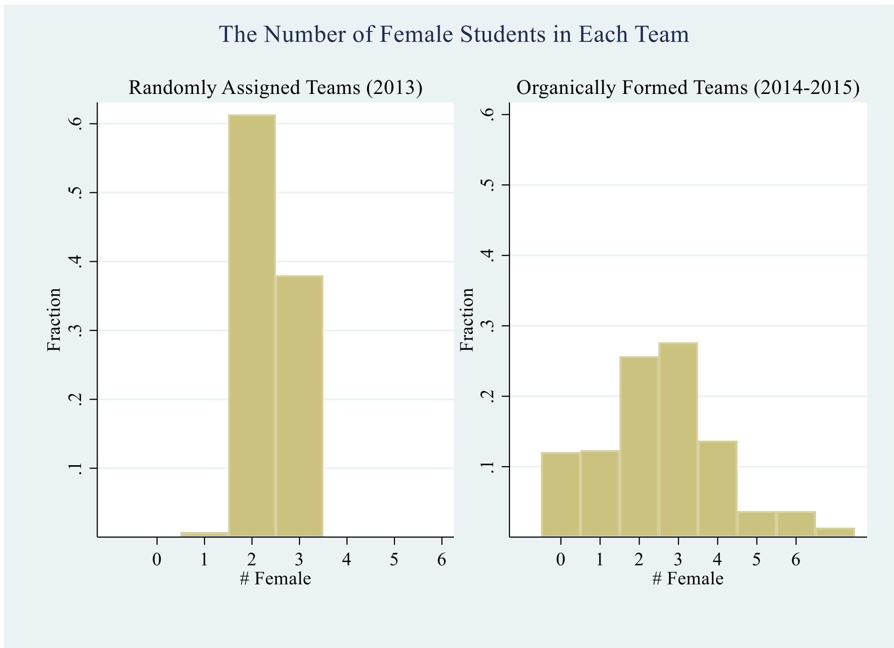
Temple-West, P. and Edgecliffe-Johnson, A. (2020). Push for diversity brings rush of business for executive headhunters. *Financial Times*.

Van Knippenberg, D. and Schippers, M. (2007). Work group diversity. *Annual Review of Psychology*, 58: 515-541.

Figures and Tables

Figure I. Distribution of Student Characteristics across Team Assignment Mechanisms

Panel A. The Distribution of Female Student Counts under Computer Assignment vs. Voluntary Formation



Panel B. The Distribution of International Student Counts under Computer Assignment vs. Voluntary Formation

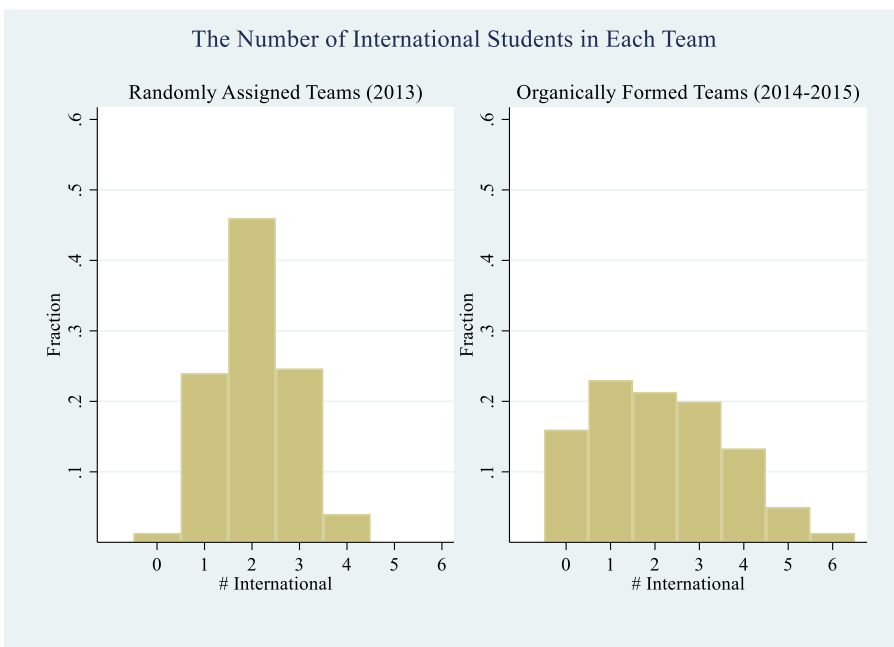
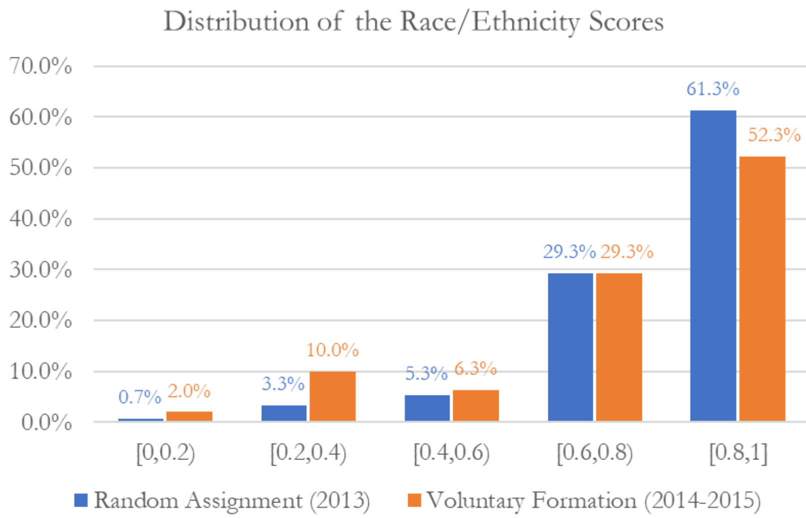


Figure II. Distribution of Diversity Scores across Team Assignment Mechanisms

Panel A. The PDF of Race/Ethnicity Scores under Random Assignment vs. Voluntary Formation



Panel B. The CDF of Race/Ethnicity Scores under Random Assignment vs. Voluntary Formation

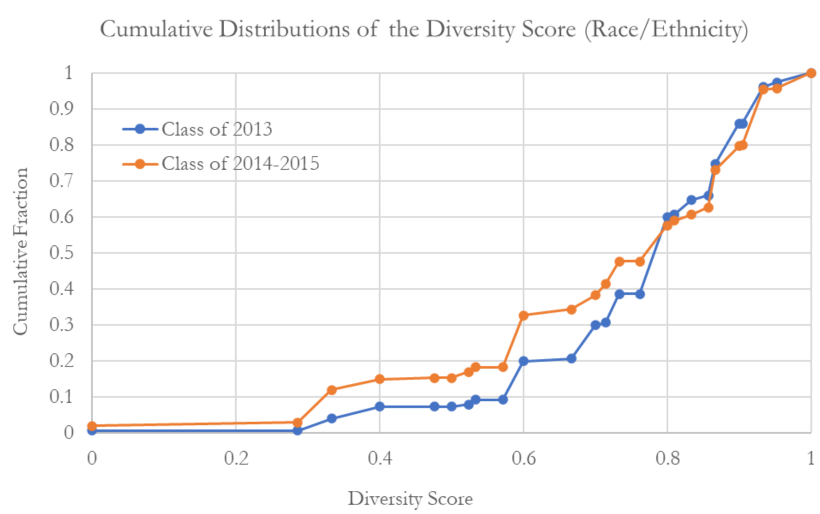
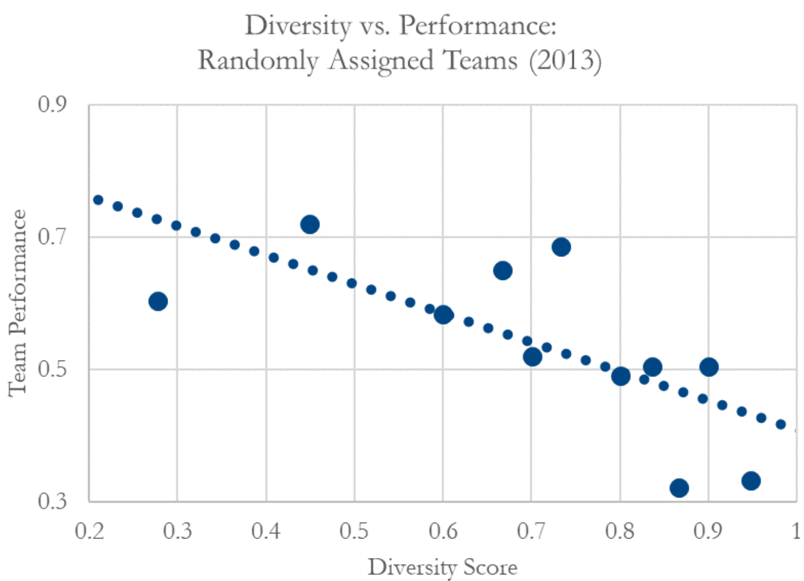


Figure III. Horizontal Team Diversity and Team Performance

The figures below plot the binscatter of team performance by race/ethnicity diversity scores. The Y Axis is the team performance, measured as the median of the quantile of the team's outcome. The X-Axis is the race/ethnicity score of the team. Larger scores imply a more diverse team. The left panel plots teams assigned randomly in 2013; the right panel plots teams formed voluntarily in 2014-2016.

Panel A. Diversity vs. Performance with Random Team Assignment



Panel B. Diversity vs. Performance with Voluntary Team Formation

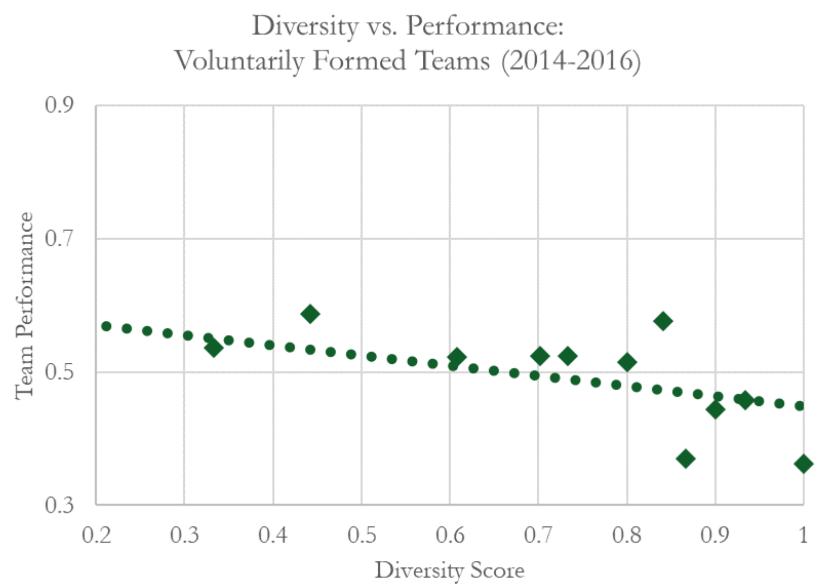
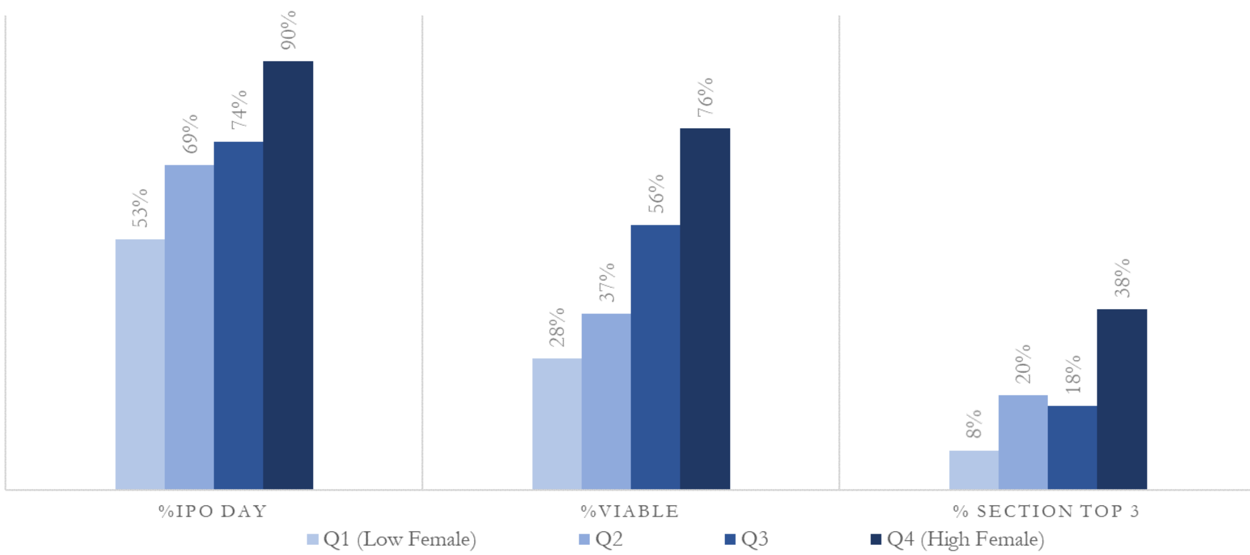


Figure IV. Team Performance Conditional on Judge's Gender (2014-2016)

This figure plots the team performance conditional on judge's gender and female percent in the team. Team performance measures (Y Axis) are the percentage of teams in section top 3, viable and IPO day. Teams are sorted into four quantiles by percent female in the team. The sample includes all teams in 2014-2016. Teams in 2013 are excluded as HBS does not have judge information for that year.

Panel A. Performance by % Female in a Team when Faculty Section Leaders are Women



Panel B. Performance by % Female in a Team when Faculty Section Leaders are Men

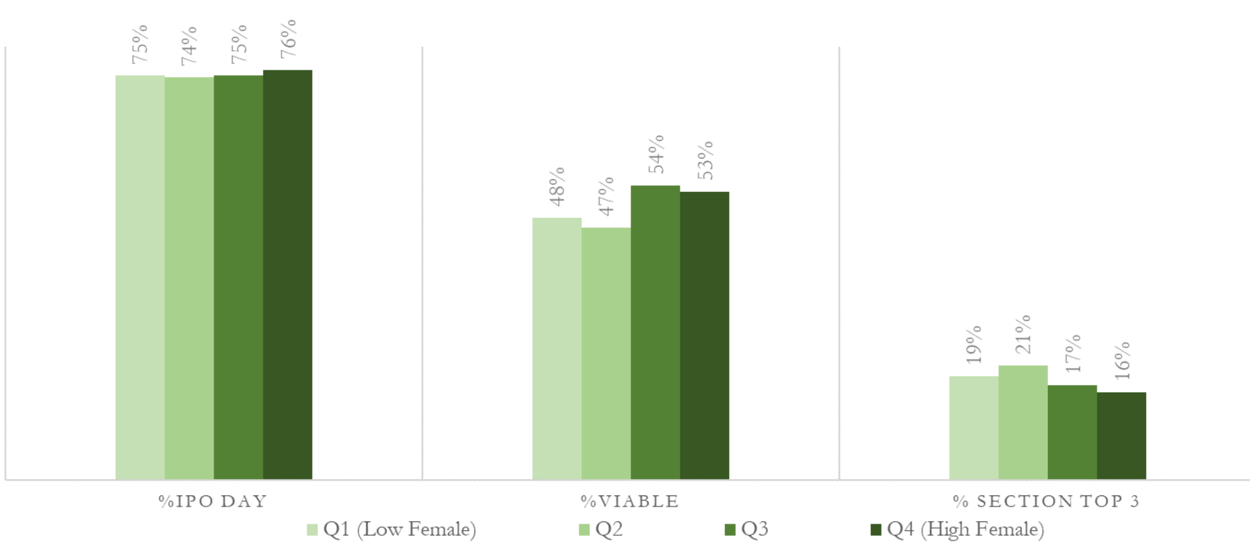


Table I. Summary Statistics of MBA Backgrounds

The table below presents the summary statistics of the demographic and employment backgrounds of Harvard Business School MBA Class of 2013, 2014, 2015, and 2016.

	2013	**2014**	**2015**	**2016**	**Total**
# of Students	907	915	931	931	3684
Team Size	6.06	6.13	6.25	5.2	5.91
Age	28.89	29.1	29.07	29.21	29.06
% of Female	39.25%	40.44%	41.14%	41.35%	40.55%
% of White American	37.16%	39.45%	37.70%	39.53%	38.46%
% of Asian American	14.33%	11.80%	11.92%	11.82%	12.46%
% of Black	4.52%	5.68%	5.59%	5.80%	5.40%
% of Latinos American	3.75%	4.26%	4.83%	3.65%	4.13%
% International	34.07%	34.32%	34.59%	37.06%	35.02%
Employment Background					
% Finance Background	29.66%	29.29%	33.83%	36.84%	32.44%
% Consulting	21.94%	20.55%	20.62%	25.13%	22.07%
% Technology	9.04%	9.84%	10.85%	13.96%	10.94%
% Healthcare	8.16%	7.87%	6.34%	8.92%	7.82%
Education Background					
% Ivy League	26.90%	25.03%	23.63%	22.99%	24.62%
% Top School	41.23%	37.92%	38.35%	34.26%	37.92%

Table II. Past Employment and Education Background

This table summarizes the employment and education background of HBS MBAs.

Rank	Company	Obs	Percent
1	McKinsey & Company	308	8.40%
2	Bain & Company	184	5.02%
3	Boston Consulting Group	173	4.72%
4	Goldman Sachs	166	4.53%
5	Morgan Stanley	138	3.77%
6	Google	78	2.13%
7	Credit Suisse	54	1.47%
8	J.P. Morgan	47	1.28%
9	Deloitte Consulting	45	1.23%
10	Booz & Company	44	1.20%
11	UBS Investment Bank	42	1.15%
12	Bank of America Merrill Lynch	38	1.04%
13	Bain Capital	32	0.87%
14	United States Marine Corps	29	0.79%
15	Accenture	26	0.71%
15	Citigroup	26	0.71%
15	Barclays Capital	25	0.68%
15	Oliver Wyman	25	0.68%
15	The Blackstone Group	25	0.68%
20	Deutsche Bank	24	0.65%
20	The Carlyle Group	24	0.65%
	Top 20 Total	1553	42.37%
	Sample Total	3,665	

Rank	School	Obs	Percent
1	Harvard University	286	8.17%
2	Stanford University	157	4.49%
3	University of Pennsylvania	151	4.31%
4	Yale University	124	3.54%
5	Princeton University	102	2.91%
6	Duke University	81	2.31%
7	MIT	72	2.06%
8	United States Military Academy	70	2.00%
9	Dartmouth College	67	1.91%
10	University of California	64	1.83%
11	Cornell University	63	1.80%
12	Georgetown University	60	1.71%
13	Brown University	57	1.63%
13	Columbia University	57	1.63%
15	Northwestern University	56	1.60%
16	University of Virginia	52	1.49%
17	Indian Institute of Technology	50	1.43%
18	University of Texas	45	1.29%
19	University of Michigan	38	1.09%
20	Brigham Young University	37	1.06%
	Top 20 Total	1689	48.26%
	Sample Total	3,500	

Table III. Summary Statistics on Team Performance Measures

This table reports our performance measure and the percentage of teams presented on IPO day, and ranked viable, section top 3, or class top 3.

Class Year	Freq.	IPO Day	Viable	Section Top 3	Class Top 3	Performance	SD
2013	150	78.67%	46.67%	20.00%	2.67%	0.502	0.275
2014	150	70.00%	39.33%	20.00%	2.00%	0.46	0.29
2015	150	73.33%	55.33%	20.00%	2.00%	0.512	0.287
2016	180	76.11%	52.78%	16.67%	2.22%	0.504	0.272
Total	630	74.60%	48.73%	19.05%	2.22%	0.495	0.281

Table IV. IPO Day Judge Characteristics

This table reports summary statistics on judges' gender and race/ethnicity. Each section has one faculty advisor, who is a faculty member from HBS, and 3-4 other judges from the industry.

Class Year	# Judges	% Female	% Black	% Latino	% East Asian	% South Asian	% White
Faculty Advisor							
2014	10	30%	20%	10%	0%	10%	60%
2015	10	30%	10%	0%	0%	20%	70%
2016	10	20%	20%	20%	0%	0%	60%
All Judges							
2014	49	14.29%	6.12%	4.08%	6.12%	6.12%	77.55%
2015	43	27.91%	6.98%	0.00%	9.30%	9.30%	74.42%
2016	44	34.09%	11.36%	4.55%	4.55%	6.82%	68.18%

Table V. Matching Properties of Computer-Assigned Teams for Class of 2013

This table reports the regression results of matching on race/ethnicity (gender, school, industry) ties. Each observation is a student-student pair. The dependent variable Real Match equals one if the pair is in the same team. The independent variables race/ethnicity (gender, school, industry) tie equals one if the pair has the same race/ethnicity (gender, school, industry). In addition, Both Non-US Citizens is an indicator variable equal to one if the student pairs are non-US citizens. Race/ethnicity Tie (US) is an indicator variable equal to one if the student pairs are both US citizens with the same race/ethnicity. Race/ethnicity Tie (International) is an indicator variable equal to one if the student pairs are both non-US citizens from the same region. Robust standard errors are clustered at the student level. Statistical significance at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.

	Dependent Variable: Real Match					
	(1)	(2)	(3)	(4)	(5)	(6)
Subsample	Full Sample	Same Gender	Different Gender	Both US	Both Non-US	US, Non-US Pairs
Gender Tie	-0.017***			-0.016***	-0.019***	-0.017***
	(0.001)			(0.002)	(0.004)	(0.002)
Both Non-US Citizens	-0.0094***	-0.010***	-0.0084**			
	(0.002)	(0.003)	(0.004)			
Race/ethnicity Tie (US)	-0.00053	0.0012	-0.0025	0.0029		
	(0.002)	(0.003)	(0.003)	(0.002)		
Race/ethnicity Tie (International)	-0.0063	-0.0058	-0.0066		-0.0069	
	(0.006)	(0.007)	(0.010)		(0.006)	
School Tie	-0.0025	0.0030	-0.0087	0.0019	0.0082	-0.011
	(0.006)	(0.008)	(0.009)	(0.007)	(0.018)	(0.011)
Industry Tie	-0.000011	-0.0013	0.0013	-0.0010	0.0092	-0.0014
	(0.002)	(0.003)	(0.003)	(0.003)	(0.006)	(0.003)
Team Member Count	0.011***	0.010***	0.011***	0.015***	0.015***	0.0057***
	(0.000)	(0.001)	(0.001)	(0.002)	(0.005)	(0.002)
Constant	0.0017	-0.013**	-0.00022	-0.032***	-0.034	0.035***
	(0.002)	(0.006)	(0.006)	(0.011)	(0.032)	(0.013)
Observations	81,368	42,140	39,228	35,228	6,764	39,376
R-squared	0.002	0.000	0.000	0.002	0.003	0.001
Year FE	YES	YES	YES	YES	YES	YES

Table VI. Matching Regression

This table reports the regression results of matching on race/ethnicity (gender, education, industry) ties. Each observation is a student-student pair. The dependent variable *Real Match* equals one if the pair is in the same team. In columns 1 and 3, the independent variables *race/ethnicity (gender, education, industry) tie* equals one if the pair has the same race/ethnicity (gender, education, industry). In columns 2 and 4, the independent variable *Endowed Demographic Match* (*Acquired Characteristics Match*) equals one if the pair has the same race/ethnicity or gender (education or industry background). Robust standard errors are clustered at the student level (i.e., one student is matched to 89 potential matches, and they are treated as one cluster). Statistical significance at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.

	Dependent variable: Real Match			
	(1) Voluntarily Formed (2014-2016)	(2) Randomly Assigned (2013)	(3) Voluntarily Formed (2014-2016)	(4) Randomly Assigned (2013)
Race/ethnicity Tie	0.014***	-0.00084		
	(0.001)	(0.002)		
Gender Tie	0.013***	-0.017***		
	(0.001)	(0.001)		
School Tie	0.0085**	-0.0028		
	(0.004)	(0.006)		
Industry Tie	0.0062***	-0.00027		
	(0.001)	(0.002)		
Endowed Demographic Match			0.015***	-0.015***
			(0.001)	(0.001)
Acquired Characteristics Match			0.0066***	-0.00039
			(0.001)	(0.002)
Team Member Count	0.011***	0.011***	0.011***	0.011***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.019***	0.00088	-0.019***	0.00076
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	254,318	81,368	254,318	81,368
R-squared	0.003	0.001	0.002	0.001
Year FE	YES	YES	YES	YES

Table VII. Detailed Matching Regression: On Ethnicity and Gender Groups

This table reports the regression results of the probability of match on race/ethnicity ties and gender ties. Each observation is a student-student pair. The dependent variable real match equals one if the students are teammates. The independent variables are race/ethnicity, or gender ties equal one if both students share the same race/ethnicity or gender. Robust standard errors are clustered at the student level. Statistical significance at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.

	Dependent variable: Real Match			
	(1)	(2)	(3)	(4)
	Voluntarily Formed (2014-2016)		Randomly Assigned (2013)	
Both White	0.012***		-0.00024	
	(0.001)		(0.002)	
Both Asian American	0.014***		0.0019	
	(0.004)		(0.005)	
Both Latino American	0.0031		0.0041	
	(0.012)		(0.022)	
Both Black	0.013		-0.000034	
	(0.009)		(0.018)	
Both International (same region)	0.040***		-0.016***	
	(0.005)		(0.005)	
Both Male		0.012***		-0.014***
		(0.001)		(0.001)
Both Female		0.017***		-0.022***
		(0.002)		(0.001)
Team Member Count	0.011***	0.011***	0.011***	0.011***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.012***	-0.016***	-0.0077***	0.00064
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	254,318	254,318	81,368	81,368
R-squared	0.002	0.002	0.000	0.002
Year FE	YES	YES	YES	YES

Table VIII. Detailed Matching Regression: On Education and Industry Backgrounds

This table reports the regression results of the probability of match on education ties and industry ties. Each observation is a student-student pair. The dependent variable real match equals one if the students are teammates. The independent variables are industry or education ties equals one if both students share the same education or industry background. Robust standard errors are clustered at the student level. Statistical significance at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.

	Dependent variable: Real Match			
	(1)	(2)	(3)	(4)
	Voluntarily Formed (2014-2016)		Randomly Assigned (2013)	
Both Ivy School	0.0023		0.0061	
	(0.005)		(0.009)	
Both Non-Ivy School	0.019***		-0.014*	
	(0.006)		(0.008)	
Both Finance Industry		0.0042***		-0.00084
		(0.001)		(0.003)
Both Tech Industry		0.0046		0.021**
		(0.004)		(0.010)
Both Consulting Industry		0.0043**		-0.0052
		(0.002)		(0.004)
Both Other Industries		0.022***		0.0068
		(0.004)		(0.005)
Team Member Count	0.011***	0.011***	0.011***	0.011***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.0096***	-0.010***	-0.0079***	-0.0082***
	(0.000)	(0.000)	(0.000)	(0.001)
Observations	254,318	254,318	81,368	81,368
R-squared	0.001	0.001	0.000	0.000
Year FE	YES	YES	YES	YES

Table IX. Match Regression: Differences by Gender

This table reports the regression results of the probability of match on race/ethnicity ties, education ties, and industry ties by gender. Each observation is a student-student pair. The dependent variable real match equals one if the students are teammates. Columns 1 and 3 report the results of the male student subsample. Columns 2 and 4 report the results of the female student subsample. Robust standard errors are clustered at the student level. Statistical significance at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.

	Dependent variable: Real Match			
	(1)	(2)	(3)	(4)
	Voluntarily Formed (2014-2016)		Randomly Assigned (2013)	
	Male	Female	Male	Female
Race/ethnicity Tie	0.015***	0.011***	-0.0024	0.00088
	(0.002)	(0.002)	(0.002)	(0.003)
School Tie	0.015***	0.00018	-0.0031	-0.0025
	(0.005)	(0.005)	(0.008)	(0.009)
Industry Tie	0.0087***	0.0031*	-0.0033	0.0039
	(0.002)	(0.002)	(0.003)	(0.003)
Team Member Count	0.010***	0.011***	0.011***	0.011***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.011***	-0.011***	0.0063***	0.0080***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	150,093	104,225	49,434	31,934
R-squared	0.002	0.002	0.000	0.000
Year FE	YES	YES	N/A	N/A

Table X. Impact of Team Diversity on Performance

This table regresses team performance on team diversity scores. The dependent variable *Performance* is the median of the quantile of the team's outcome. The independent variables are diversity scores described in the paper. *Voluntarily Formed* is an indicator variable equal to one if the team is in 2014-2016 subsample. Robust standard error is clustered at year-section level. Statistical significance at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.

	Dependent variable: Performance								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Subsamples	Randomly Assigned (2013)			Organically Formed (2014-2016)			Full Sample		
Race/ethnicity Score	-0.49***	-0.45***	-0.54**	-0.18***	-0.14**	-0.057	-0.49***	-0.46***	-0.44***
	(0.105)	(0.088)	(0.173)	(0.056)	(0.059)	(0.074)	(0.103)	(0.091)	(0.103)
Gender Score	-0.025	-0.18	-0.13	-0.043	-0.054	-0.054	-0.00037	-0.14	-0.062
	(0.658)	(0.598)	(0.592)	(0.059)	(0.059)	(0.057)	(0.630)	(0.582)	(0.573)
School Score	-0.96*	-0.75	-0.77	0.20	0.33	0.34	-0.97**	-0.77	-0.78
	(0.476)	(0.528)	(0.498)	(0.295)	(0.310)	(0.348)	(0.455)	(0.485)	(0.466)
Industry Score	-0.084	-0.048	-0.16	0.12	0.12	0.093	-0.085	-0.048	-0.11
	(0.165)	(0.165)	(0.190)	(0.082)	(0.084)	(0.137)	(0.156)	(0.153)	(0.167)
Voluntarily Formed ×Race/ethnicity Score							0.30**	0.32***	0.36***
							(0.117)	(0.107)	(0.108)
Voluntarily Formed × Gender Score							-0.043	0.083	0.0072
							(0.633)	(0.584)	(0.577)
Voluntarily Formed × School Score							1.16**	1.10*	1.12*
							(0.541)	(0.562)	(0.565)
Voluntarily Formed × Industry Score							0.20	0.16	0.20
							(0.177)	(0.176)	(0.173)
Top School Ratio		0.060	0.074		0.064	0.0089		0.063	0.028
		(0.067)	(0.072)		(0.061)	(0.063)		(0.050)	(0.051)
Start-up Ratio		0.46	0.50		0.34**	0.26*		0.36***	0.29**
		(0.365)	(0.408)		(0.136)	(0.138)		(0.125)	(0.126)
Honor Student Ratio		0.31	0.33*		0.26***	0.27***		0.27***	0.28***
		(0.182)	(0.179)		(0.086)	(0.095)		(0.077)	(0.082)
English Speaker%			-0.18			0.15**			0.11
			(0.201)			(0.071)			(0.066)
Consulting%			-0.038			-0.032			-0.032
			(0.211)			(0.112)			(0.099)

Finance%			-0.14			-0.020			-0.032
			(0.195)			(0.110)			(0.095)
Technology%			0.12			0.19*			0.17*
			(0.127)			(0.109)			(0.087)
Team Member Count	0.054	0.036	0.039	0.094***	0.078***	0.080***	0.090***	0.074***	0.075***
	(0.057)	(0.061)	(0.058)	(0.027)	(0.028)	(0.029)	(0.025)	(0.025)	(0.026)
Constant	1.58*	1.43	1.72	-0.26	-0.39	-0.53	1.36*	1.21*	1.14*
	(0.844)	(0.891)	(1.099)	(0.313)	(0.301)	(0.370)	(0.682)	(0.638)	(0.643)
Observations	150	150	150	480	480	480	630	630	630
R-squared	0.107	0.145	0.157	0.056	0.091	0.111	0.067	0.103	0.116
Year FE	N/A	N/A	N/A	YES	YES	YES	YES	YES	YES

Table XI Matching on Career Interests and Language

This table reports regression results on matching on communication cost and shared career interests. Each observation is a focal student-teammate pair (real match=1). In column 1, the dependent variable is the number of career interests shared by the focal student and the teammate. The independent variable *Different Race/Ethnicity* is an indicator variable equal to one if the focal student and the teammate have different races or ethnicities. Robust standard errors are clustered at the student level. Statistical significance at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.

Dependent Variable	(1) Number of Shared Career Interests
Different Race/ Ethnicity * Voluntarily Formed	0.17***
	(0.063)
	-0.12**
Different Race / Ethnicity	(0.057)
	0.038
	(0.032)
Team Member Count	0.17***
	(0.063)
Constant	0.90***
	(0.191)
Observations	11,896
R-squared	0.013
Year FE	YES

Table XII. Impact of the Intersection of Gender and Race/Ethnicity on Performance

This table regresses team performance on team Ethnicity-Gender scores. The dependent variable *Performance* is the median of the quantile of the team's outcome. The independent variables are *Ethnicity-Gender scores* measure the fraction of student pairs who do not match on either gender or race/ethnicity. *Voluntarily Formed* is an indicator variable equal to one if the team is in 2014-2016 subsample. Control variables are *top school ratio, start-up ratio, honor student ratio* and *team member count*. Robust standard error is clustered at year-section level. Statistical significance at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.

	Dependent variable: Performance					
	(1)	(2)	(3)	(4)	(5)	(6)
Subsamples	Randomly Assigned (2013)		Organically Formed (2014-2016)		Full Sample	
Race/Ethnicity Score	-0.083	-0.015	-0.13	-0.10	-0.12	-0.041
	(0.137)	(0.187)	(0.120)	(0.123)	(0.140)	(0.186)
Race/Ethnicity-Gender Score	-0.78***		-0.0078		-0.73***	
	(0.233)		(0.157)		(0.226)	
Race/Ethnicity-Gender Score (Male)		-0.80***		-0.0038		-0.75***
		(0.239)		(0.155)		(0.236)
Race/Ethnicity-Gender Score (Female)		-1.29*		-0.20		-1.28*
		(0.669)		(0.220)		(0.644)
Race/Ethnicity Score * Voluntarily Formed					-0.0020	-0.058
					(0.185)	(0.225)
Race/Ethnicity-Gender Score * Voluntarily Formed					0.72**	
					(0.278)	
Race/Ethnicity-Gender Score (Male) * Voluntarily Formed						0.75**
						(0.285)
Race/Ethnicity-Gender Score (Female) * Voluntarily Formed						1.08
						(0.680)
Gender Score	-0.086	-0.27	-0.040	-0.040	-0.0076	-0.20
	(0.587)	(0.630)	(0.073)	(0.073)	(0.576)	(0.614)
Gender Score * Voluntarily Formed					-0.036	0.15
					(0.579)	(0.616)
Observations	150	150	480	480	630	630
R-squared	0.163	0.168	0.084	0.087	0.100	0.104
Control Variables	YES	YES	YES	YES	YES	YES
Year FE	N/A	N/A	YES	YES	YES	YES

Table XIII. Impact of Team Diversity on Performance by Subgroups

This table regresses team performance on team diversity scores. The dependent variable *Performance* is the median of the quantile of the team's outcome. The independent variables are diversity scores that measure the fraction of student pairs in each team that does not belong to the same racial or ethnic group. Robust standard error is clustered at year-section level. Statistical significance at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.

	Dependent Variable: Performance	
	(1)	(2)
	Randomly Assigned	Voluntarily Formed
Subsamples	(2013)	(2014-2016)
White American Score	-0.50***	-0.17***
	(0.097)	(0.053)
Asian American Score	-1.03***	-0.020
	(0.302)	(0.222)
Black Score	-0.36	-0.10
	(1.687)	(0.727)
Latino American Score	-2.59	-1.28
	(1.913)	(0.988)
International Student Score	-1.69***	0.10
	(0.440)	(0.170)
Top School Ratio	0.090	0.044
	(0.082)	(0.051)
Startup Ratio	0.42	0.38***
	(0.375)	(0.127)
Honor Student Ratio	0.26	0.25***
	(0.173)	(0.080)
Team Member Count	0.057	0.067***
	(0.070)	(0.024)
Constant	6.11	1.47
	(4.091)	(1.359)
Observations	150	630
R-squared	0.168	0.092
Year FE	N/A	YES

Table XIV. The Effect of Judge Gender on Performance

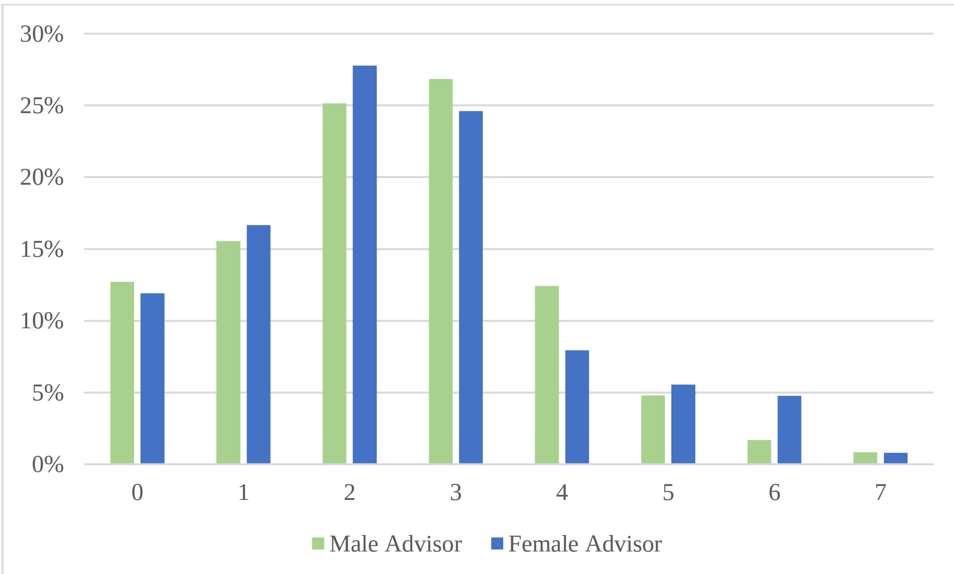
This table regresses team performance on judge's gender interacted with percent of female in the team. The dependent variable *Performance* is the median of the quantile of the team's outcome. The independent variable *Faculty Advisor Female* is an indicator variable equals to 1 if the female judge is a faculty advisor. *Have Female Judge* equals 1 if there is at least one female judge in the section. *Female Team Member%* is the percent of females in the team. Robust standard error is clustered at year-section level. Statistical significance at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.

	Dependent Variable: Performance		
	(1)	(2)	(3)
		2014-2016	
Faculty Advisor Female * Female Team Member %	0.31***		0.31***
	(0.094)		(0.103)
Faculty Advisor Female	-0.15***		-0.15***
	(0.045)		(0.047)
Female Team Member %	0.032	0.062	0.063
	(0.063)	(0.093)	(0.094)
Have Female Judge * Female Team Member %		0.065	-0.038
		(0.112)	(0.122)
Have Female Judge		-0.036	0.020
		(0.054)	(0.056)
Top School Ratio	0.048	0.043	0.049
	(0.060)	(0.062)	(0.061)
Start-up Ratio	0.35**	0.34**	0.35**
	(0.134)	(0.135)	(0.134)
Honor Student Ratio	0.30***	0.31***	0.30***
	(0.093)	(0.093)	(0.093)
Team Member Count	0.076***	0.078***	0.075***
	(0.027)	(0.027)	(0.027)
Constant	-0.076	-0.10	-0.087
	(0.158)	(0.157)	(0.158)
Observations	480	480	480
R-squared	0.101	0.085	0.101
Year FE	YES	YES	YES

Online Appendix

Appendix Figure 1. The Distribution of the Female Team Members by the Faculty Advisor's Gender

This figure plots the distribution of the number of female team members by the faculty advisor's gender. The X-axis is the number of female team members in a team. The Y-axis is the conditional probability density (i.e., % teams with a certain number of female members). The green (blue) bar shows teams with a male (female) faculty advisor. A chi-square test of homogeneity of the two distributions is insignificant (χ^2=5.8, the critical value for the 10% significance is 12).



Appendix Table I. Home Country of HBS MBA Students

This table reports the top 20 home countries of HBS MBA students in our sample.

	Country	Freq.	Percent
1	USA	2,394	64.98%
2	India	172	4.67%
3	Canada	125	3.39%
4	China	76	2.06%
5	UK	59	1.60%
6	Brazil	52	1.41%
7	Australia	46	1.25%
8	France	45	1.22%
9	Germany	45	1.22%
10	Israel	33	0.90%
11	Korea	30	0.81%
12	Japan	28	0.76%
13	Mexico	28	0.76%
14	Turkey	28	0.76%
15	Argentina	27	0.73%
16	Lebanon	25	0.68%
17	Russia	25	0.68%
18	Spain	24	0.65%
19	Nigeria	23	0.62%
20	Chile	19	0.52%
	Total	3,684	100.00%

Appendix Table II. HBS Section Assignment in 2013-2016

This table reports the regression results of matching on race/ethnicity (gender, education, industry) ties within sections. Each observation is a student-student pair in the section. The dependent variable Real Match equals one if the pair is in the same section. The independent variables race/ethnicity (gender, education, industry) tie equals one if the pair has the same race/ethnicity (gender, education, industry). In addition, Both Non-US Citizens is an indicator variable equal to one if the student pairs are non-US citizens. Robust standard errors are clustered at the student level.

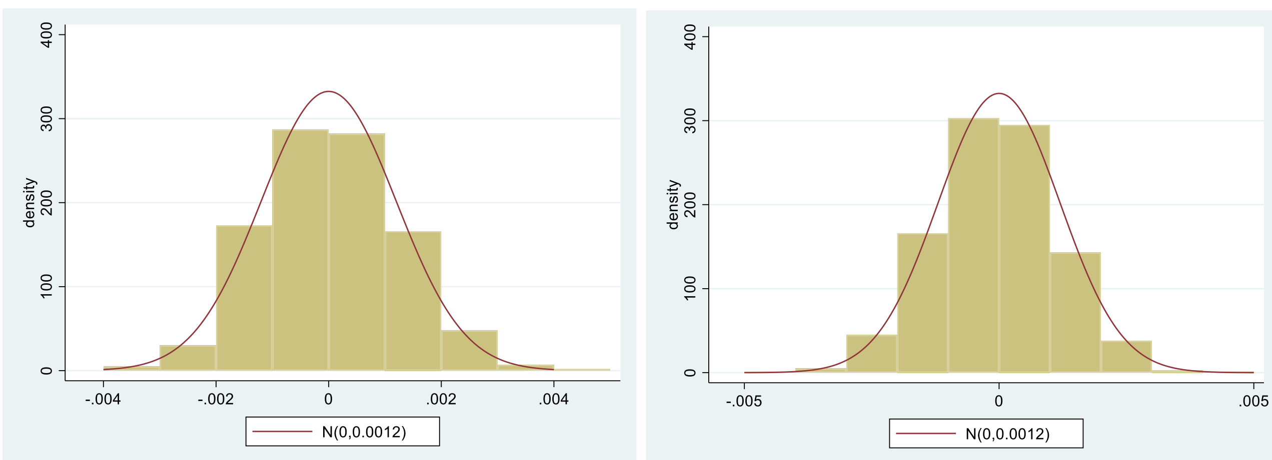
	Dependent Variable: Real Match			
	(1)	(2)	(3)	(4)
Class Year	2013	2014	2015	2016
Gender Tie	-0.0018***	-0.0018***	-0.0018***	-0.0019***
	(0.000)	(0.000)	(0.000)	(0.000)
Both Non-US Citizens	-0.00012	0.00061	-0.00028	-0.00053
	(0.001)	(0.001)	(0.001)	(0.000)
Race/ethnicity Tie (US)	-0.0018***	-0.0021***	-0.0014**	-0.0023***
	(0.001)	(0.000)	(0.001)	(0.000)
Race/ethnicity Tie (International)	-0.0085***	-0.012***	-0.011***	-0.012***
	(0.002)	(0.002)	(0.001)	(0.001)
School Tie	-0.013***	-0.014***	-0.015***	-0.011***
	(0.002)	(0.002)	(0.001)	(0.002)
Industry Tie	-0.0031***	-0.0033***	-0.0030***	-0.0026***
	(0.001)	(0.001)	(0.000)	(0.000)
Team Member Count	0.00079***	0.00032***	0.00075***	0.00037***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.042***	0.044***	-0.039***	0.033***
	(0.005)	(0.005)	(0.011)	(0.011)
Observations	821,742	836,310	865,830	865,830
R-squared	0.000	0.000	0.000	0.000

Appendix Table III. Randomization Test of Homophily

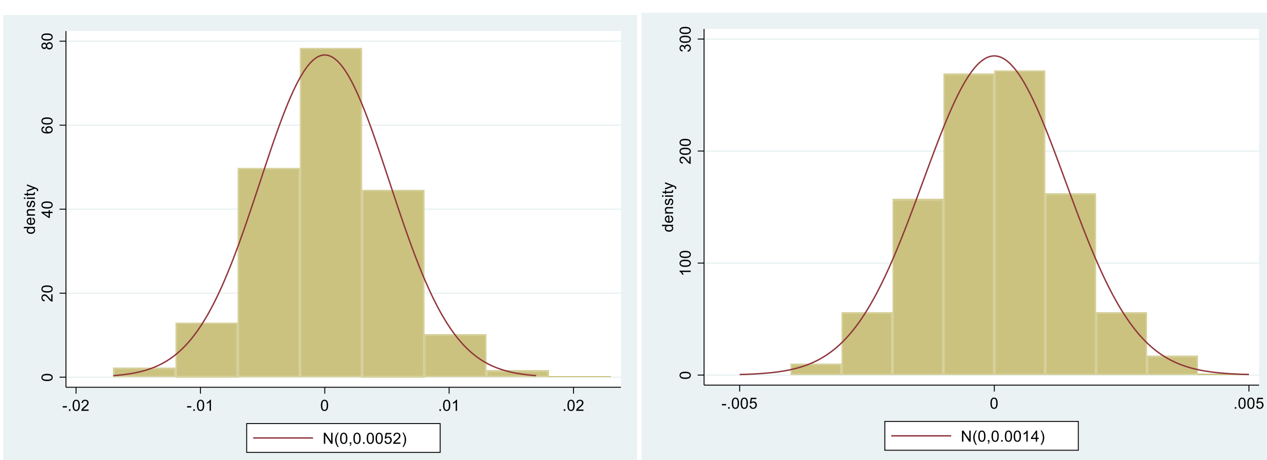
This table compares simulated matching results for 2014-2016 matching with random team assignment to actual regression coefficients. For simulated coefficients, we use 1000 iterations and report the mean and standard deviations of simulated regression coefficients. P-value is the probability of the simulated value greater than the actual matching coefficient.

Subsample: 2014-2016	Simulated Matching Coefficients		Actual Matching Coefficients		
Coefficients	Mean	SD	Value	SE	p-value
Race/Ethnicity Tie	0.0000	0.0012	0.0135	0.0011	<1%
Gender Tie	-0.0001	0.0012	0.0131	0.0011	<1%
School Tie	0.0002	0.0052	0.0085	0.0038	<5%
Industry Tie	0.0001	0.0014	0.0062	0.0012	<1%

Panel A. Distribution of Race/Ethnicity Tie Coefficients (left) and Gender Tie Coefficients (Right)



Panel B. Distribution of Race/Ethnicity Tie Coefficients (left) and Gender Tie Coefficients (Right)



Appendix Table IV. Match between International Students

This table reports the regression results of the probability of match among international students. Each observation is a student-student pair. The dependent variable real match equals to 1 if the students are teammates. The independent variables are ethnicity characteristics, equaling to 1 if both students are from the same region. Robust standard errors are clustered at the student level.

	Dependent Variable: Real Match	
	(1)	(2)
	Voluntarily Formed Teams (2014-2016)	Randomly Assigned Teams (2013)
Both European	0.025***	-0.0089
	(0.007)	(0.008)
Both South Asia	0.029***	-0.027***
	(0.008)	(0.008)
Both East Asia	0.063***	-0.040***
	(0.016)	(0.011)
Both Latin America	0.062***	-0.011
	(0.019)	(0.022)
Both Middle East	0.067***	0.043
	(0.018)	(0.048)
Both African	-0.047***	-0.058***
	(0.001)	(0.001)
Team Member Count	0.011***	0.011***
	(0.000)	(0.000)
Constant	-0.011***	-0.0075***
	(0.000)	(0.001)
Observations	254,318	81,368
R-squared	0.002	0.000
Year FE	YES	YES

Appendix Table V. Detailed Match Regression: Differences by Gender

This table reports the regression results of the probability of match on race/ethnicity ties, education ties, and industry ties by gender. Each observation is a student-student pair. The dependent variable real match equals one if the students are teammates. Columns 1 to 3 report the results of the male student subsample. Columns 4 to 6 report the results of the female student subsample. Robust standard errors are clustered at the student level.

	Dependent variable: Real Match					
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Male Subsample			Female Subsample		
Both White	0.014***			0.0091***		
	(0.002)			(0.002)		
Both Asian American	0.015**			0.014***		
	(0.007)			(0.005)		
Both Latino American	0.016			-0.025		
	(0.016)			(0.019)		
Both Black	0.0033			0.024*		
	(0.011)			(0.014)		
Both International	0.043***			0.036***		
	(0.006)			(0.008)		
Both Ivy School		0.0089			-0.0039	
		(0.007)			(0.007)	
Both Non-Ivy School		0.025***			0.0090	
		(0.008)			(0.009)	
Both Finance Industry			0.0080***			-0.0017
			(0.002)			(0.002)
Both Tech Industry			0.0087*			-0.0022
			(0.005)			(0.006)
Both Consulting Industry			0.0026			0.0056**
			(0.003)			(0.003)
Both Other Industries			0.021***			0.023***
			(0.005)			(0.007)
Team Member Count	0.011***	0.011***	0.011***	0.011***	0.011***	0.011***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.012***	-0.0096***	-0.0100***	-0.012***	-0.0095***	0.011***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)
Observations	150,093	150,093	150,093	104,225	104,225	104,225
R-squared	0.002	0.001	0.001	0.002	0.001	0.001
Year Fixed Effect	Y	Y	Y	Y	Y	Y

Appendix Table VI. Matching Between Venture Capital Investor and Entrepreneurs

In this table, we analyze all venture capital deals in the US between 2010 and 2016, covering 5,731 venture capital investors and 11,471 deals. In the regression tables, each observation is a VC-entrepreneur pair, where each VC-deal is matched to all pseudo-deals in the same year-industry. The independent variables Same Ethnicity (Same School) are binary indicators. Robust standard errors are clustered at the VC-deal level.

Panel A. VC-Entrepreneur Match by Demographic Characteristics

	(1)
Dependent Variable	Real Match
Both Female	0.0045***
	(0.001)
Both Male	0.00098***
	(0.000)
Same Ethnicity	0.0025***
	(0.000)
Same School	0.013***
	(0.001)
Number of Pseudo Deals	-0.000***
	(0.000)
Observations	3,937,141
R-squared	0.005
Year FE	YES

Panel B. VC-Entrepreneur Match by Race / Ethnicity: Gender Differences

	Dependent variable: Real Match	
	(1)	(2)
Subsample	Male	Female
Both White	0.0017***	0.00033
	(0.000)	(0.001)
Both Jewish	0.0034***	0.0010
	(0.001)	(0.002)
Both East Asian	0.022***	0.016***
	(0.002)	(0.004)
Both Indian	0.015***	0.011***
	(0.001)	(0.003)
Both Hispanic	0.013***	0.030
	(0.004)	(0.023)
Both Black	0.14	
	(0.098)	
Observations	3,609,876	291,338
R-squared	0.006	0.006
Year FE	YES	YES

Appendix Table VII. Correlation between Industry Experience and Demographic Characteristics

This table regresses past industry employment on student gender, race, and ethnicity. Each observation is a student. The analysis includes all students from 2013 to 2016.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	Finance	Technology	Consulting	Healthcare	Retail	Top School	Start-up Exp
Female	-0.020	-0.0043	0.094***	5.3e-06	0.0096	0.10***	0.0054
	(0.016)	(0.011)	(0.014)	(0.009)	(0.006)	(0.016)	(0.007)
Asian American	0.046*	0.061***	-0.028	0.032*	0.00017	0.14***	0.0056
	(0.025)	(0.019)	(0.020)	(0.017)	(0.010)	(0.026)	(0.012)
Black	0.052	-0.028	0.010	-0.041**	0.027	0.026	-0.022*
	(0.036)	(0.019)	(0.030)	(0.018)	(0.017)	(0.038)	(0.012)
Hispanic American	-0.021	-0.0051	0.011	-0.017	0.014	-0.047	-0.0067
	(0.039)	(0.025)	(0.032)	(0.023)	(0.018)	(0.041)	(0.017)
International	-0.035**	0.017	0.13***	-0.056***	-0.00052	-0.29***	-0.017**
	(0.017)	(0.012)	(0.016)	(0.009)	(0.007)	(0.016)	(0.007)
Constant	0.31***	0.079***	0.14***	0.099***	0.029***	0.45***	0.031***
	(0.018)	(0.011)	(0.016)	(0.011)	(0.007)	(0.019)	(0.007)
Observations	3,684	3,684	3,684	3,684	3,684	3,684	3,684
R-squared	0.008	0.008	0.039	0.015	0.002	0.121	0.007
Year FE	YES	YES	YES	YES	YES	YES	YES

Appendix Table VIII. Summary Statistics on Career Interests

This table reports summary statistics on the ten most popular career interests listed by HBS students on their Class Card profiles.

	Career Interests	Percent
1	Finance	47.6%
2	Technology	33.9%
3	Consulting	25.8%
4	Entrepreneurship/Startup	24.3%[20]
5	Consumer Products	19.1%
6	Leisure	16.9%
7	Healthcare	14.3%
8	Community Development	12.4%
9	Education	11.6%
10	Government	11.2%
	# of students	3684
	Median # of career interests	3
	% of students listing at least 1 career interest	81%

[20] Entrepreneurship/Startup option only become available after 2013. The 24.3% is the percentage of students choose entrepreneurship as one of her career interests after 2013.

Appendix Table IX. Impact of Team Diversity on Performance: Binary Outcomes

This table reports logit regression results on the effect of Race/ethnicity-Gender Score. Robust standard errors are clustered at year-section level. The dependent variables IPO day/Viable/Section Top 3/Class year top 3 are indicator variables equals 1 if the team presented on IPO day/the project is deemed viable by judges/the team is section top 3/the team is class year top 3. The independent variables are diversity scores described in the paper. Robust standard error is clustered at year-section level. Each coefficient's standard error appears directly below the coefficient estimate. Statistical significance at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.

Panel A. Random Assignment (2013)	(1)	(2)	(3)	(4)
Dependent Variable =	IPO Day	Viable	Section Top 3	Class Top 3
Race/Ethnicity-Gender Score	-10.87***	-7.126***	-3.928**	-6.343***
	(2.903)	(1.951)	(1.655)	(1.545)
Top School Ratio	1.557**	0.982**	1.087	0.736
	(0.785)	(0.486)	(1.057)	(1.482)
Start-up Ratio	2.237	3.140*	3.159	2.772
	(3.173)	(1.900)	(3.315)	(8.376)
Honor Student Ratio	2.502	2.800*	1.633	-2.538
	(2.561)	(1.644)	(1.336)	(5.475)
Team Member Count	0.929	0.285	-0.243	-0.683
	(0.653)	(0.595)	(0.659)	(0.665)
Constant	4.656	3.623	2.721	5.773
	(4.320)	(4.258)	(3.334)	(4.332)
Observations	150	150	150	150
Panel B. Voluntary Formation (2014-2016)	(5)	(6)	(7)	(8)
Dependent Variable =	IPO Day	Viable	Section Top 3	Class Top 3
Race/Ethnicity-Gender Score	-1.737**	-0.851*	-0.980	-2.284
	(0.761)	(0.493)	(0.630)	(1.484)
Top School Ratio	0.655	0.497	-0.240	-0.354
	(0.539)	(0.354)	(0.594)	(1.093)
Start-up Ratio	2.279*	2.789***	3.005***	
	(1.275)	(1.024)	(1.132)	
Honor Student Ratio	1.326	1.715***	2.401***	0.625
	(0.846)	(0.643)	(0.831)	(1.667)
Team Member Count	0.440*	0.702***	0.256	0.745
	(0.253)	(0.248)	(0.229)	(0.632)
Constant	-0.856	-4.543***	-2.515	-6.477
	(1.481)	(1.484)	(1.633)	(4.444)
Year FE	YES	YES	YES	YES
Observations	480	480	480	375

Appendix Table X. Alternative Measure of Race/Ethnicity Diversity

This table regresses team performance on team diversity scores. The dependent variable *Performance* is the median of the quantile of the team's outcome. The independent variables *Race/ethnicity Percentile* is the median percentile of the simulated distribution of the Race/ethnicity score. We obtain the simulated distribution by randomly assigning students to teams in each class year and computing the Race/ethnicity score in each iteration. We iterate this process for 1000 times. *Race/Ethnicity-Gender Percentile* is constructed analogously. Robust standard error is clustered at year-section level. Each coefficient's standard error appears directly below the coefficient estimate. The Statistical significance at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.

	Dependent Variable: Performance			
	(1)	(2)	(3)	(4)
Subsample	2013	2014-2016	Full Sample	Full Sample
Race/ethnicity Percentile	-0.36***	-0.16***	-0.34***	-0.14
	(0.053)	(0.044)	(0.058)	(0.115)
Race/ethnicity Percentile * Voluntary Formation			0.19**	-0.076
			(0.073)	(0.134)
Race/Ethnicity - Gender Percentile				-0.27**
				(0.124)
Race/Ethnicity - Gender Percentile * Voluntary Formation				0.34**
				(0.145)
Gender Score	0.040	-0.049	-0.044	-0.070
	(0.673)	(0.059)	(0.059)	(0.063)
School Score	-1.06*	0.19	-0.085	-0.040
	(0.485)	(0.291)	(0.259)	(0.254)
Industry Score	-0.13	0.13	0.082	0.081
	(0.142)	(0.082)	(0.073)	(0.074)
Team Member Count	0.055	0.089***	0.082***	0.083***
	(0.052)	(0.028)	(0.025)	(0.025)
Constant	1.48*	-0.28	0.22	0.22
	(0.693)	(0.309)	(0.284)	(0.279)
Observations	150	480	630	630
R-squared	0.144	0.064	0.073	0.081
Year FE	N/A	YES	YES	YES