Job Networks through College Classmates: Effects of Referrals for Men and Women*

Maria Zhu[†]

Duke University

April 16, 2019

(Most recent version available **here**)

Abstract

This paper analyzes effects of referrals on labor market outcomes, as well as how these effects differ by gender. To do so, I link administrative data on community college student transcripts to matched employer-employee records to study job search through classmates. Using a novel two-step research design, I first identify classroom network effects by exploiting quasi-random variation in section enrollment within courses. Results indicate taking a class with a peer increases the propensity for a student to get a job at a firm where the peer is incumbent. The overall propensity to use classmates in job finding does not differ by gender, although students display an increased propensity to form networks with same-gender peers. In the second step of the research design, I investigate the labor market effects of obtaining a job through a classmate. Consistent with the predictions of a referral-based job search model, workers who obtain jobs through classmates earn more and are less likely to leave the firm, and effects decline with tenure in the firm. Furthermore, while referrals benefit both genders, the earnings premium from referrals for women is less than half the premium for men. From a policy perspective, these findings suggest a key tradeoff between increasing efficiency through referrals and increasing gender equity.

*I am grateful to my advisors, Peter Arcidiacono, Pat Bayer, Rob Garlick, Joe Hotz, and Arnaud Maurel for their guidance and encouragement. I thank Gorkem Bostanci, Hugh Cassidy, David Eil, Paul Goldsmith-Pinkham, Attila Gyetvai, Mike Kofoed, Margaux Luflade, Hugh Macartney, Margie McElroy, John Pepper, Michael Ransom, Tyler Ransom, Seth Sanders, David Selover, John Singleton, Juan Carlos Suárez Serrato, and Xiao Yu Wang for helpful comments. This paper also benefitted from seminar participants at Duke, SOLE, AEFP, Young Economists Symposium, SEA, and WEAI.

[†]maria.zhu@duke.edu

1 Introduction

Approximately 50% of new jobs are found through informal channels, such as employee referrals, rather than formal search methods,[1] and more than half of firms in the U.S. have programs encouraging referral-based hiring (Topa 2011, CareerBuilder 2012). One motivation for these policies is referrals increase efficiency by increasing the number of matches made between workers and firms, as well as the quality of matches.[2] However, referrals do not necessarily benefit all groups equally, and effects may vary due to differences across individuals, the characteristics of their contacts, and the types of firms in which they seek employment (Ioannides and Loury 2004). From a gender perspective, clear systemic differences exist between men and women in the labor market, as evidenced by the persistent gender pay gap in all industrialized countries (Blau and Kahn 2003, 2017). Despite the prevalence of referral-based hiring in the job market, researchers know very little regarding how and whether referrals improve or exacerbate existing gender inequalities in the labor market.

This paper measures labor market returns to classmate referrals, shedding light on the overall benefits of referrals, as well as how these effects differ for men and women. It does so by using transcript data linked to matched employer-employee records for two-year college students in Arkansas. There are two main challenges to the empirical analysis of referral effects: First, referrals and social networks are likely endogenous. For example, it could be the case that high- and low-ability individuals make use of referrals at different rates, or transitory labor market shocks affect both referral uptake and labor market outcomes. By using a large panel data set that follows individuals over multiple employment spells, this analysis can account for individual-specific effects. Furthermore, the identification strategy isolates variation within courses, which are school-, field-, and time-specific. This further controls for the possibility that transitory labor market shocks drive outcomes. While having panel data across an entire labor market allows this paper to address endogeneity rigorously, one tradeoff is that the scale of these data sets generally means they do not contain direct measures of referrals. Thus, the second empirical challenge to this analysis is finding a way

[1] I define formal search methods to include all job search methods that do not involve personal contacts. Following Rees (1966), this includes methods such as job advertisements, state employment services, private employment agencies, and school or college placement services.

[2] For theoretical and empirical motivation, see Simon and Warner 1992, Montgomery (1991), Saloner 1985, Pallais and Sands 2016, Dustmann et al. 2016, and Heath 2018.

to recover information on referrals.

To obtain information on referral relationships, this paper employs a novel two-step research design: In the first step, I provide evidence that students use classmates to find jobs by showing that taking a class with a peer increases the propensity for a student to start working at a firm where the peer is incumbent. One challenge here is to disentangle classmate network effects from other factors correlated with course selection that could also affect job finding. To do so, this study exploits quasi-random variation in section enrollment within college courses. In the second step of the research design, I use the same identifying variation to measure the effects of getting a job through a classmate on labor market outcomes. Results are consistent with testable implications of a referral-based job search model, in which referrals increase information regarding match quality between a worker and firm.

As mentioned above, a key challenge to identification the first step of the research design requires disentangling network effects from other factors correlated with course selection that may also affect job finding. For example, students may be selecting courses based on unobservable characteristics that also affect subsequent course choice. Alternatively, some programs or fields at a school may have high reputations or connections with certain firms, creating a pipeline effect for students in courses into those firms. A course in this setting consists of a set of classes within a term and school that provide the same course content and offer the same course credit. Sections within a course represent the actual classroom setting, and all sections within a course list the same course title and description. While selection into courses may be correlated with factors affecting job finding, enrollment into sections within a course does not suffer from the same issue. Thus, to isolate classmate network effects, I exploit variation in section enrollment within courses. I first examine the validity of this identification strategy by testing whether students sort into sections within a course either systemically or along pre-existing relationships and provide evidence that sorting does not significantly affect job-finding outcomes. Next, within a course, I compare the propensity for a student to start working at a firm where a peer from the same section works, as opposed to a peer from a different section of the course. An increased propensity to work at a firm where a peer from the same section works would indicate students use classmate networks to find jobs.

Results indicate taking an additional class with a peer sizably increases the probability a student will begin working at the same firm as the peer in the years following the last course taken together by the pair. I follow student pairs for up to six years, and the effect is largest immediately after taking the courses: Taking an additional section increases the probability a student will work at a firm where a classmate was incumbent at time of hire one year after their last course together by 8.6% from the baseline propensity of working together. This effect fades slightly over time, and six years after the course, this effect is 5.0%. Overall, these findings indicate students use classmates significantly to help find jobs. Additionally, I find that while men and women do not differ in the propensity to use college classmates in job-finding, women are significantly more likely to form networks with female peers and vice versa for men. This gender homophily in network formation is consistent with the referral patterns Zeltzer (2017) finds in physician referral networks, as well as a well-established literature in sociology (McPherson et al. 2001).

In the second step of the research design, I measure the effects of working in a job obtained through a classmate on earnings and job separation rates. One key advantage of my data is that I follow individuals through multiple employment spells, and there is within-student variation on whether a job was obtained through a classmate. Since informal networks are inherently endogenous, panel observations allow me to control for the possibility that networks are correlated with ability or other unobserved characteristics that affect labor market outcomes. Results indicate getting a job through an incumbent classmate increases initial earnings by 13.6% compared to getting hired without a classmate's help. This effect decreases by 1.6 percentage points per quarter with tenure in the firm. Classmate networks decrease turnover rate by 3.1 percentage points initially, and these effects increase by 0.3 percentage points per quarter with tenure. These findings are consistent with the empirical implications of referral-based job search models from Dustmann et al. (2016) and Simon and Warner (1992), in which referrals improve the efficiency of potential jobs by reducing signal noise on the match quality between a worker and job.

Next, I analyze how the effects of referrals differ for men and women. I find that while both men and women benefit from referrals in terms of earnings, women get a significantly lower earnings premium. Men hired through referrals get a 16.8% earnings increase compared to their earnings at the firm if hired independently of referrals, while women get a 10.3 percentage point lower earnings increase. Additionally, the earnings premium fades for men at a faster rate at 2.2 percentage points, compared to 0.7 percentage points for women, per quarter. Effects of referrals on job separation rate do not differ by gender. I also provide evidence that gender differences in returns to referrals cannot be explained by women working in industries with low returns to referrals or women working in lower paying jobs within a firm. Additionally, in looking at referral effects across firms, I find that in addition to the firm-specific premium mentioned above, referrals increase earnings for men by 3% through getting jobs at higher paying firms. However, referrals actually move women to lower-paying firms on average, and the net effect of referral on earnings for women is zero, even though they have higher within-firm earnings from referrals. This indicates there may be a compensating differential motivating women to use referrals. Further, I show that differences in returns to referrals cannot be explained by gender differences in sorting across industries or jobs within firms. I provide suggestive evidence that differences may arise due to gender differences in bargaining power over the surplus generated by referrals. Overall, from a policy perspective, these findings indicate a tension between increasing overall welfare through referrals and increasing gender equity.

This paper relates to three branches of literature. First, it adds to the empirical work looking at networks in educational settings. These papers are nested in body of work identifying job search networks in a variety of social settings.[3] Existing studies on education networks have overwhelmingly focused on the role of networks at elite institutions. Kramarz and Thesmar (2013) show that French CEOs from elite colleges are more likely to hire board members from the same institution, and Zimmerman (2018) finds similar results in management hiring patterns for graduates from the same degree cohort in elite Chilean schools. Using survey data, Marmaros and Sacerdote (2002) also find that students at Dartmouth College use alumni for job-finding after graduation. My paper contributes to these studies in two ways: First, I establish the importance of peer networks for students at non-elite institutions. Given that these students generally look for very different types of jobs in a different part of the labor market compared to “elite” peers, it is not clear a priori that networks operate in the same way in non-elite settings. Second, this study is the first to identify networks at the classroom level, compared to cohort- or institution-level of prior studies. The granularity of my identification strategy allows me to provide novel estimates of the labor market returns of obtaining a job through an education network at the individual

[3] Residential neighbors: Topa (2001), Bayer et al. (2008), and Hellerstein et al. (2011). Former coworkers: Cingano and Rosolia (2012) and Hensvik and Skans (2016). Ethnic groups: Dustmann et al. (2016) and Beaman (2012). Family members: Kramarz and Skans (2014).

level.

This paper also addresses a recent literature analyzing the effects of referrals on labor market outcomes. Prior studies on this topic have used personnel (Burks et al. 2015, Brown et al. 2015) or survey (Heath 2018, Dustmann et al. 2016) data for analysis. These papers have the advantage of observing information on referrals directly in the data, which is non-trivial to obtain. My paper takes a different approach in measuring referrals indirectly by exploiting quasi-random variation in student enrollment into sections within college courses. I contribute to prior studies with the use of a unique method for measuring referrals that allows me to address endogeneity concerns rigorously. One advantage of administrative data over survey or personnel data is the increased potential scale and scope of data collection. In using a panel data set that spans an entire labor market, this study follows individuals over multiple employment spells, allowing for the inclusion of individual-specific effects that may be correlated with referral use. Additionally, since the identification isolates variation narrowly within school-, time-, and field-specific courses, this strategy controls for transitory labor market shocks that may affect both referral uptake and labor market outcomes. Finally, this data allows me to analyze effects of referrals across an entire labor market. This is especially helpful for analyzing aggregate systemic differences in referral effects across groups, in this case gender.

Third, this study relates to a sizable body of research analyzing factors contributing to the gender pay gap. Studies have looked at the role of factors such as taste-based or statistical discrimination, the interaction of firm and family demands, preferences, and social norms/biases.[4] However, few studies have examined the role of referrals on gender differences in the labor market. In their field experiment on referrals in Malawi, Beaman et al. (2017) show that women sustain a disadvantage from networks because men prefer not to refer women, even when they know women capable of performing the tasks at hand. However, they cannot analyze the effects of referrals on earnings in this setting since the researchers set wages in the experiment. This paper contributes to prior studies on gender differences in the labor market as the first analysis of the role of referrals on labor market outcomes. Results from the analysis also relate to Card et al. (2016), which finds that women both work in firms paying lower premiums and receive a lower portion of the pay premium within

[4]For literature on each topic, see Becker (1971), Fang and Moro (2011), Bertrand et al. (2010), Goldin (2014), Hotz et al. (2017), Bursztyn et al. (2017), Bertrand et al. (2015), and Sarsons (2017).

a given firm compared to male counterparts. I find that referrals contribute to both pay gap margins discussed in the study, to the extent that they shift women to lower-paying firms and exacerbate within-firm gender pay differentials.

In the remainder of the paper, Section 2 describes the data used in the study and provides motivating descriptive statistics. Section 3 describes the empirical strategy and findings for the first part of the research design, which focuses on detecting whether students use classmates to help in job-finding. Next, Section 4 provides a conceptual framework for thinking about the mechanisms by which classmates help each other find jobs. Section 5 presents the empirical strategy and findings for the second part of the research design, which analyzes the effects of getting a job through a referral on labor market outcomes. In Section 6, I first explore alternative explanations for gender differences in effects of referrals before showing results to be consistent with women having lower bargaining power in referral negotiations. Finally, Section 7 concludes.

2 Data and Empirical Setting

This section introduces the empirical setting of the paper. I first describe the data used for the analysis and then present summary statistics. Next, I explain how network links are constructed in this setting.

2.1 Data

Data for this paper come from the Arkansas Department of Higher Education. The main dataset includes enrollment records for all students who attended a public college between 2003 and 2012. For each student, the data contain information on background characteristics including gender, part-time vs. full-time attendance status, high school information, and college degree attainment. Additionally, transcript records provide information on the courses students took, what section they were in, the instructor, and credits earned for the course.

I link students to matched employer-employee labor market records using data collected from Arkansas Unemployment Insurance records.[5] Here, I observe quarterly observations on

[5] The variables in these data are comparable to that found in the Longitudinal Employer-Household Dynamics (LEHD) dataset.

all individuals working for Arkansas from 2001-2012, minus uncovered sectors.[6] The panel contains information on earnings, a six-digit industry code,[7] and a firm identifier number. If a student works multiple jobs in a quarter, the records only report the job with the highest earnings in the quarter. Finally, if an individual does not appear in the labor market files in a given quarter, I cannot disentangle whether she is unemployed, employed outside of Arkansas, or employed in an uncovered sector.[8]

This paper focuses on two-year college students. Arkansas has a total of 22 two-year colleges, and 35% of students enrolled in a public college in Arkansas in the data attend a two-year school.[9] This is slightly lower than the 42% of college students enrolled in a community college nation-wide (Ma and Baum 2016). I analyze two-year, as opposed to four-year college students for a few reasons: First, I look at classmate networks in this study, and classmates represent a relatively more prominent source of student interaction at two-year colleges. Four-year colleges often have a much stronger presence of potentially confounding sources of student networking such as residential dorms, Greek life, on-campus clubs and organizations, etc. Second, two-year college students are more likely to search for jobs in a local labor market, making them more tractable in my data set, as well as potentially increasing the salience of peer networks. Finally, two-year colleges have smaller classes on average than four-year counterparts and typically do not involve large lecture classes: median section size for two-year colleges in this dataset is 14 students, and fewer than 5% of sections contain more than 30 students. This is important for the analysis because different sections actually represent different classes, as opposed to representing discussion sessions within a large lecture class.

2.2 Summary Statistics

Table 1 describes the composition of students in the sample. Approximately 60% of students in the two-year college system are female, paralleling national trends.[10] Slightly over half of

[6]Individuals working in uncovered sectors in this data line constitute a very small portion of overall employees in the state. These include self-employed individuals and federal government employees.

[7]Industries are classified using North American Industry Classification System (NAICS) codes.

[8]Analysis from the American Community Survey indicates that in 2000, approximately 12% of individuals in Arkansas with the education level of my sample no longer lived in the state five years later.

[9]Figure 8 in Appendix A shows a map of Arkansas displaying locations of the schools.

[10]In 2012, 57% of two-year college students and 56% of four-year college students nation-wide were female USDOE (2017).

students enroll as part-time, and slightly over half of students are employed in a given term. The average student in the sample is 27 years old, with a median of 23.[11] Students take an average of 2.6 classes a semester. Overall, men and women in this sample do not display large disparities along the characteristics described.

Table 1: Student Characteristics

	All	Male	Female
Female (%)	59.50	–	–
Part-time (%)	55.00	55.43	54.71
Employed (%)	54.00	51.43	55.75
Age	26.71	25.93	27.20
	(9.16)	(8.95)	(9.26)
Number of Courses	2.59	2.49	2.65
	(1.60)	(1.66)	(1.56)
N	389,342	157,674	231,668

Two-Year College Students, 2003-2011. Table displays means (standard deviations in parentheses). Age is imputed from year of high school graduation, assuming students are 18 at the end of high school. Median age is 23.

Next, Table 2 shows labor market statistics for students in the sample. I observe earnings from 2001-2011, and average earnings are $5,097 per quarter, conditional on being employed in the sample.[12] One reason for low earnings is that I observe many students while they are enrolled for college and thus observe part-time earners. Table 2 also shows significant gender differences in earnings: While men earn on average $5,632 per quarter, women earn $4,798, and these values are significantly different at a 95% confidence level. Men and women in the sample have similar separation rates at about 21-22% per quarter.[13]

To explore the types of firms in which men and women work, Table 3 shows the top 10

[11] Nationally, the average of of two-year college students is 28 years old (Kolesnikova and Shimek (2008)).

[12] As an external check, I calculate earnings for a comparable sample of students from the American Community Survey (ACS). The ACS reports similar average quarterly earnings of $4,915 for this demographic, with an average unemployment rate of 6.6% over the sample. Specifically, I look at ACS data for Arkansas residents from 2006-2010 whose highest level of education ranges from some college experience to holding an associate's degree. Additionally, I restrict the sample age to the interquartile range of ages of students in the sample, 19-32.

[13] When I restrict my sample to people who have left school, this rate drops to 13%. Nationally, the BLS reports average annual job separation rates in the US South at 41% in 2013.

Table 2: Labor Market Summary Statistics

	All	Male	Female
Quarterly Earnings ($)	5,097	5,632	4,798
	(5360.41)	(5811.51)	(5066.05)
Job Separation Rate (%)	0.21	0.22	0.21
N	5,684,914	2,039,417	3,645,497

Sample consists of earnings data between 2001-2011 for students enrolled in a two-year college during the sample. Table displays means (standard deviations in parentheses). All earnings in 2010 dollars. ACS matched earnings (all): $4,915. Post-college turnover: 13%. BLS annual job separation rate US South (2013): 41%. Earnings for men vs. women are significantly different at 95% confidence level.

industries of work for students by gender using two-digit North American Industry System (NAICS) codes. The table displays clear gender differences: For example, men are more likely to work in manufacturing and public administration, while women are more likely to work in health care/social assistance and education. Table 12 in Appendix A shows aggregate industry results, and Table 13 shows the full ranking of industries by gender.

Table 3: Concentration of Students by Industry

(a) Male

Industry	(%)
Retail Trade	22.5
Manufacturing	13.2
Accommodation/Food Services	12.8
Health Care/Social Assistance	9.2
Public Administration	6.9
Admin./Waste Mgmt./Remediation	5.9
Construction	4.9
Education	4.5
Wholesale Trade	3.8
Transportation and Warehousing	3.5
Other	15.8
Total	100.0

Industries are classified by two-digit NAICS codes. Each student-quarter employment period represents an observation.

(b) Female

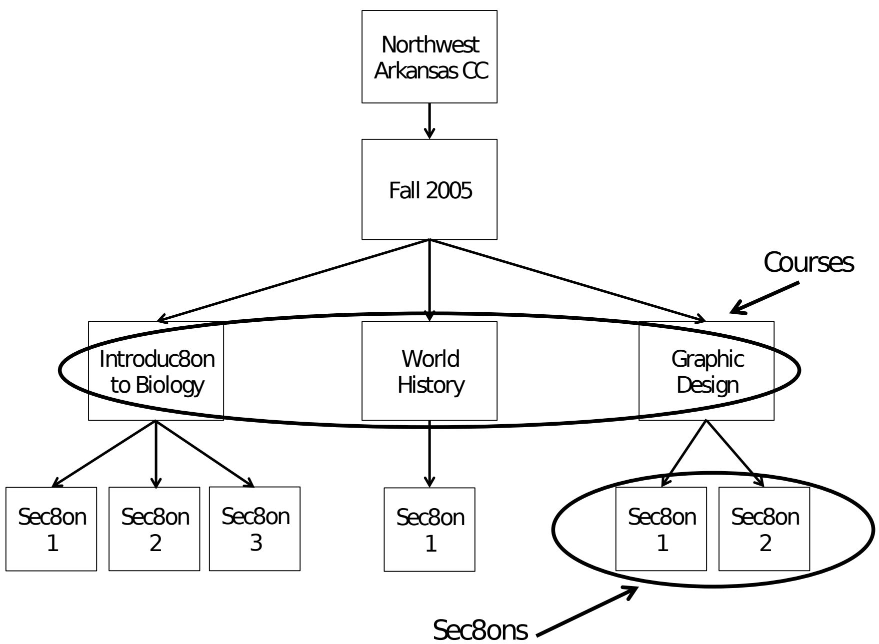
Industry	(%)
Health Care/Social Assistance	30.9
Retail Trade	18.4
Accommodation/Food Services	12.2
Education	9.6
Manufacturing	5.8
Admin./Waste Mgmt./Remediation	4.5
Finance and Insurance	3.8
Prof./Scientific/Technical Services	2.8
Public Administration	2.6
Other Services	1.7
Other	7.7
Total	100.0

Industries are classified by two-digit NAICS codes. Each student-quarter employment period represents an observation.

2.3 Network Links

Since this analysis focuses on classmate network channels, I next formally define courses, sections, and classmates in this setting. Figure 1 shows a diagram of the structure of courses and sections in my data. Each community college offers a wide variety of courses in a given term.[14] A course consists of a set of sections (also known as classes) with the same course name and course number. Sections within a course offer the same course content and provide identical academic credit and represent the actual classroom setting of the course. Thus, "classmates" in this setting refer to a student's peers who enroll in the same section of a course.

Figure 1: Course Layout



Note: Figures displays a simplified example of three courses in a term. Each course consists of a set of sections, or classes, which all contain the same course name and course number.

I restrict my data to analyze students in courses containing multiple sections since identification in this paper comes from variation in section enrollment within a course. This eliminates 31% of sections. To check if these restrictions significantly change the sample, I examine whether the restrictions alter characteristics of the student body in Appendix A. Table 14 indicates restricting the sample does not significantly alter student composition.

[14]Courses are definitionally to be school- and term- specific in this setting.

The final sample consists of 132,625 class sections nested in 29,793 courses. Figures 9 and 10 in the appendix show the distribution of sections per course students per section, respectively. The median number of sections per course is 3, and 90% of courses contain fewer than 10 sections. The median enrollment per class is 14 students, and 90% of classes contain fewer than 25 students.

3 Identifying Network Effects

In the first step of the research design, I measure whether and to what extent students use classmates to find jobs. I first describe the research problem and data setup. I next present the empirical strategy I employ and then show results of the analysis.

3.1 Description

To measure whether students use classmates networks to find job, I examine whether taking a class with a peer increases the propensity for a student to start working at a firm where the peer is incumbent. For this analysis, I construct a dyadic data set consisting of students, i, matched with peers, j. I define a student's peers as the set of students who were enrolled at the same school contemporaneously. For each dyadic observation, I observe in the data whether $\{i, j\}$ were former classmates (defined by whether they shared a section at some point), as well as whether i ends up working firm where j was incumbent at time of hire.

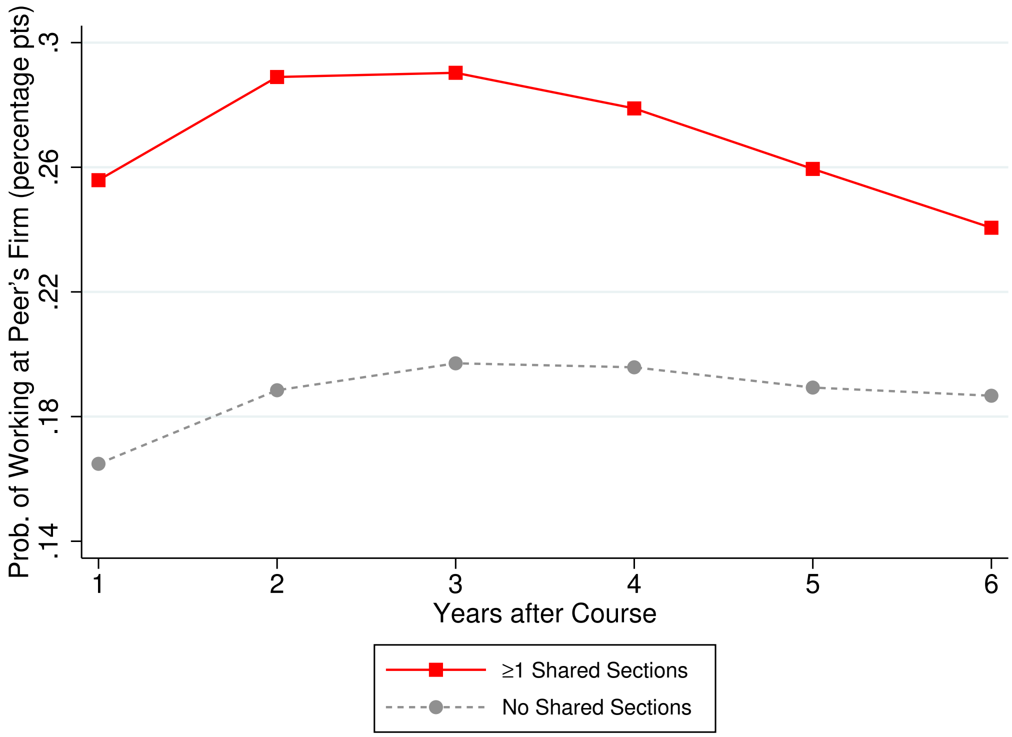
As a starting point, I compare the propensity for i to get a job at a firm where j is incumbent for pairs in which $\{i, j\}$ were former classmates versus pairs in which they were not.[15] Figure 2 displays the probability that x years after their last course together, an individual works at a firm where her peer is incumbent at time of hire.[16] For all years of analysis, the graph indicates that a higher proportion of students start working at a peer's firm when the two shared at least one section together, compared with pairs with no shared sections. A two-sample test of proportions shows that these values are significantly different at a 95% confidence level for each year.

However, this increased propensity to work with former classmates does not necessarily

[15]Note that matched pair observations are not interchangeable in this setting: $\{i, j\} \neq \{j, i\}$; highlighting the inherent asymmetrical nature of network relationships.

[16]For pairs who did not share any sections, I use time since the last term they were both enrolled in school.

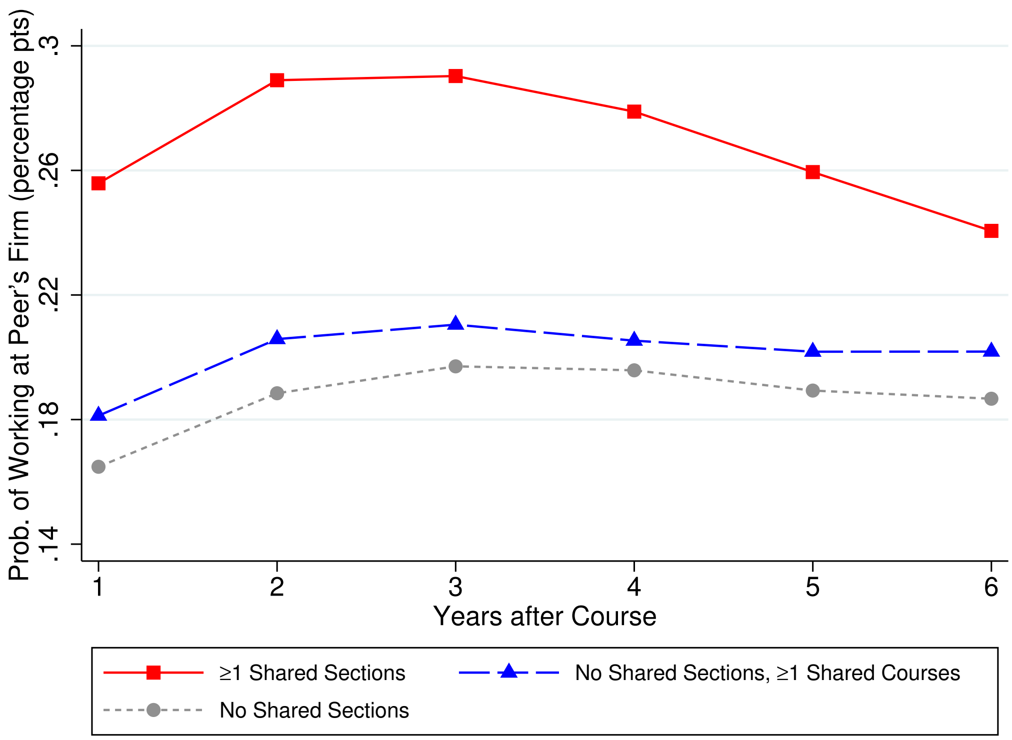
Figure 2: Probability of Working at Firm with Incumbent Peer



Note: Observations denote matched pairs $\{i, j\}$. For peers with shared sections/courses, I use the full sample of within-course pair-level observations, $N \approx 304$ million. For peers with no shared sections, I take a 10% sample of pairs, $N \approx 168$ million. For peers with no sections, the x-axis measures number of years after last potential class together. Point estimates for a given year are all significantly different at a 95% confidence level.

arise due to network effects through classroom exposure; other factors correlated with course selection could also affect job finding outcomes. First, since students select courses non-randomly, it may be that unobserved factors affecting course choice also affect the firms at which students seek employment. Second, some jobs may have certain degree or accreditation requirements, and consequently firms hiring in those jobs may disproportionately draw from courses associated with those degrees or programs. Finally, if certain programs or tracks of study have strong reputations or connections with certain firms, this could lead to a pipeline hiring effect in the relevant courses. To control for these other factors correlated with course enrollment, I include a matched pair sample consisting of students matched with peers with whom they shared at least one course, but no sections, in Figure 3. Intuitively, comparing the propensity of working at a peer's firm between section-mates and course-mates (who did not share any sections) isolates the effects of differences in classroom exposure between peers.

Figure 3: Probability of Working at Firm with Incumbent Peer



Note: Observations denote matched pairs $\{i, j\}$. For peers with shared sections/courses, I use the full sample of within-course pair-level observations, $N \approx 304$ million. For peers with no shared sections, I take a 10% sample of pairs, $N \approx 168$ million. For peers with no sections, the x-axis measures number of years after last potential class together. Point estimates for a given year are all significantly different at a 95% confidence level.

From Figure 3, for all years of analysis there exists a gap in the propensity of working at a peer's firm between pairs who shared at least one course and pairs with no shared sections. This gap represents the proportion of the increased probability of working at a pair's firm that arises due to factors correlated with course selection. Thus, moving forward, I restrict my analysis to pairs of individuals who share at least one course together, using variation in section enrollment to identify the effects of classroom exposure on job finding. Figure 3 indicates that students have a higher propensity to work at a firm where a former classmate was incumbent, compared to a former course-mate who was not a classmate, for all years of analysis. A two-sample test of proportions shows these values are significantly different at a 95% confidence level for each year, providing suggestive evidence of networks occuring at the classroom level. To uncover causal relationships, the next section explains the empirical strategy, which isolates comparisons to variation in section exposure within pairs of students who take the same set of courses together. Additionally, it restricts comparisons to sections of a course taught by the same instructor, in order to ensure instructor effects do not drive job-finding outcomes. Finally, the methodology allows for heterogeneous effects across the number of courses a pair takes together.[17]

3.2 Empirical Strategy

This section describes the empirical strategy used to analyze the effects of taking an additional section with a peer on the propensity for a student to get a job at a firm where the peer is incumbent. As mentioned above, observations consist of matched pairs of students with former course-mates to examine the propensity for a student to work at a firm where a peer works. I run the following pair-level regression:

$$F_{ij\tau} = \rho_{c\tau} + \gamma_\tau N_{ij} + \epsilon_{ij\tau} \tag{1}$$

where i and j represent two individuals who take a bundle of courses, c, together. The outcome of interest, $F_{ij\tau}$ is an indicator variable that equals one if τ years after a pair's last course taken together, i works at a firm where j was incumbent when i was hired and zero otherwise. The inclusion of a course bundle fixed effect, $\rho_{c\tau}$, restricts comparison of observation $\{i, j\}$ to the set of pairs that took exactly the same bundle of courses together as i and j. The course bundle fixed effect captures the baseline propensity for student to start working at a firm where a peer with whom she shares a set of courses c is incumbent, independent of classroom interaction effects. To capture the effect of classroom interaction, N_{ij} measures the number of sections i and j shared, of the total number of courses they took together.[18] The coefficient of interest, γ_τ, captures how much taking an additional section together increases the propensity for i to work at a firm where j is incumbent. A positive estimated value γ_τ indicates classmates serve as a channel through which students are finding jobs, while an estimated coefficient not statistically from zero indicates no networking

[17]Over 70% of pairs take one course together, and less than 1% of pairs take take more than five courses together. Figure 11 in Appendix A shows the distribution of number of courses a given pair takes together.

[18]To test for non-linearities in the effects of section exposure, I run a specification that includes a quadratic term on N_{ij}. Table 16 in Appendix C displays estimation results from this specification. The effects on the quadratic term are small and statistically insignificant for most year of analysis, so I use Equation 1 as the primary specification.

occurring through this channel.

The key assumption in this approach requires that while individuals sort non-randomly into the courses they choose to take, they do not sort into sections within a course in a way that also affects subsequent firm placement. This assumption is motivated by two considerations. First, within a course, each section has the same course title and course description and offers the same number and type of credits, which eliminates much of the heterogeneity that drives selection across courses. Second, as indicated in Table 1, 54% percent of students hold jobs while in school and students take on average 2.6 courses per semester. Thus, even if students would like to select into sections, work or other academic obligations may place constraints on how much choice they have. To control for the the possibility students sort into sections based on instructor, I restrict the analysis to comparisons of pairs taught by the same instructor. Additionally, I control for the attendance status (part-time vs. full-time) of both i and j to account for the fact that certain types of students are more constrained to courses that take place at certain times or days of the week, especially since over half of students work while in school.[19]

Potential Obstacles to Identification

As mentioned above, identification requires that students do not sort into sections within a course in a way that also affects the outcome of interest. To test the plausibility of this assumption, I assess two types of sorting: systematic and pair-wise. First, students may sort systematically into different sections of a course along unobservable characteristics. If these unobservable characteristics also affect firm choice later on, this could upwardly bias estimates of network effects. Second, a student and peer may purposefully try to enroll in the same sections due to pre-existing friendships. If friends tend to share unobservable characteristics that make it more likely they will end up at the same firm, independently of network effects, this would also upwardly bias network estimates. In Appendix B, I provide evidence that sorting on these margins does not drive results.

[19] I do not observe the time of day or day of week of a section in the data.

Alternative Specification

The linear probability model denoted in Equation 1 presents a straightforward approach to measuring the effects of classroom exposure with intuitive coefficients. However, one concern is that this approach may not fit the data well since $F_{ij\tau}$ represents a binary outcome that takes on many zero values. Thus, I also estimate effects of classroom exposure on job finding using a logit specification:

$$F_{ij\tau} = \mathbb{1}\left[\tilde{\rho}_{c\tau} + \tilde{\gamma}_{\tau} N_{ij} + \tilde{\epsilon}_{ij\tau} > 0\right] \tag{2}$$

Parameter estimates from Equation 2 cannot be as easily interpreted directly, but the intuition for what the variables capture remains the same as in Equation 1.

3.3 Results

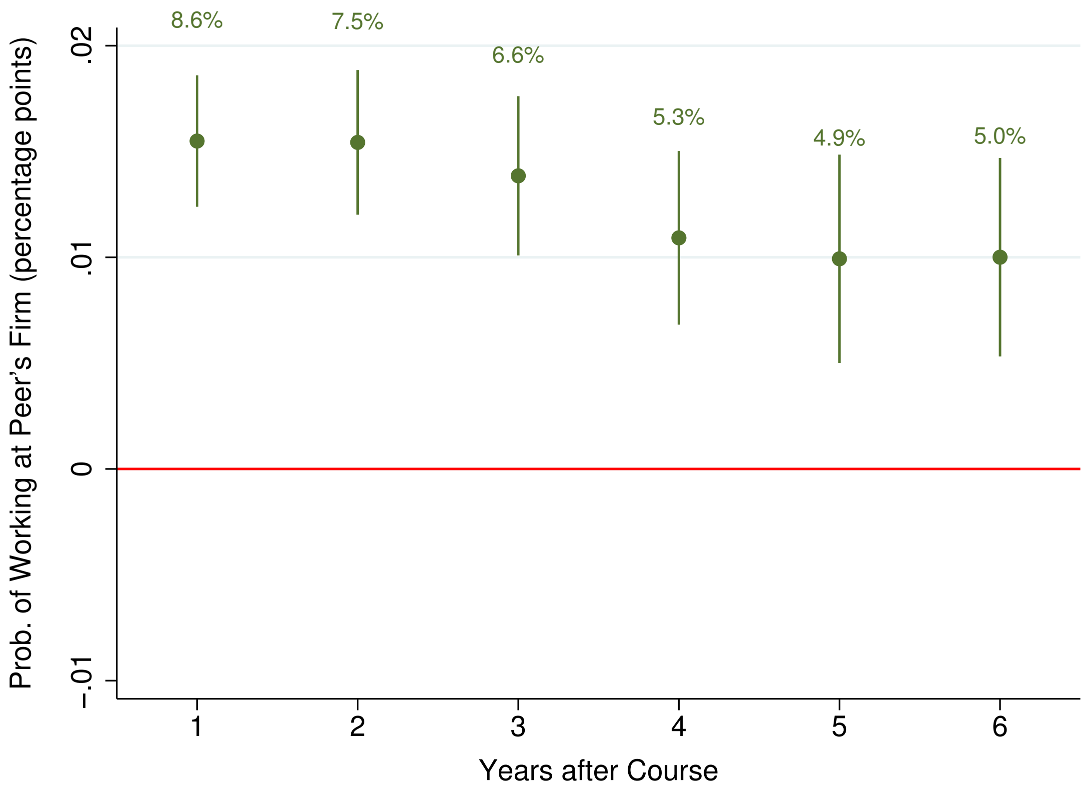
Main Findings

I first show estimation results from Equation 1, which analyzes the effect of taking an additional section with a peer on the propensity for a student to start working at a firm where the peer is incumbent. Figure 4 shows results for up to six years after the last course taken together. The x-axis denotes years after the pair takes their last course together, and I estimate coefficients separately for each year. The outcome is measured in percentage points, and I cluster standard errors by course bundle. I display estimates using 95% confidence intervals, and numbers above point estimates represent the percent increase of the effect from baseline values.

Results indicate that taking an additional section with a peer significantly increases the propensity to work at a firm where the peer was incumbent at time of hire for each of the first six years after taking courses together. For example, taking an additional course together increases the propensity for an individual to work at a firm where the peer is incumbent three years after the course by .014 percentage points. This represents a 6.6% increase in the baseline propensity to work at a firm with an incumbent peer after three years, .210 percentage points.[20] Thus, while the estimated effects of classroom exposure on job-finding at a peer's firm are small in magnitude, they constitute a significant increase from the

[20] Baseline values are shown in Table 15.

Figure 4: Effects of an Additional Section Together on Probability of Working at Peer's Firm



Note: Coefficient estimates have been multiplied by 100 to reflect percentage changes, and standard errors are clustered by course bundle. Estimates are displayed with 95% confidence intervals. Numbers above point estimates represent the percent increase of the effect from baseline values. Results are displayed numerically with sample sizes in Table 15 in Appendix A.

baseline probability of getting a job at a firm with an incumbent peer. Table 15 in Appendix A displays these results numerically.

Since estimates are positive and statistically significant for all years of analysis, I aggregate the years of analysis in Table 4 to interpret the magnitude of findings. Specifically, I look at the effects of taking a section with a peer on the propensity of getting a job at a firm where the peer is incumbent sometime within the first three years of after the last course taken together.[21] The first column of Table 4 displays estimates from the linear probability model in Equation 1, and the second column displays estimates from the logit specification in Equation 2.[22] For the logit specification, I also include average marginal effects.

[21] I choose to restrict the analysis to the first three years after the course since I only observe a limited number of pairs for all six years of analysis, as shown in Table 15.

[22] The logit specification contains fewer observations because course bundles with no variation in the

Table 4: Aggregate Effects of Taking Additional Section on Probability of Working at Peer's Firm

	Linear (γ)	Logit ($\tilde{\gamma}$)
Baseline	0.58	0.58
Number of Sections	0.0279***	6.0458***
	(0.0029)	(0.4594)
Course Bundle FE	X	X
Average Marginal Effect	–	0.0382
N	31,857,553	25,247,770

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by course bundle. Observations represent matched pairs of students with peers, $\{i, j\}$. Baseline and coefficient estimates multiplied by 100 so linear estimate and AME represent percentages. Outcome: whether student starts working at firm where peer is incumbent within three years of last shared course.

Results from both specifications indicate that taking an additional section with a peer has a positive and statistically significant effect on getting a job at the peer's firm within three years after the last course together. The baseline propensity for an individual to start working at a firm where a peer from the same set of courses, independent of network effects, is 0.58%. Thus, taking an additional section with a classmate increase the propensity for a student to start working at the classmate's firm within three years by 4.8-6.6% from the baseline, depending on whether the linear or logit estimates are used. Furthermore, given that students take an average of 2.59 classes per semester with an average of 15 students per section, these effects translate into approximately a 1.7-2.3 percentage point increase in the propensity for a student to get a job at a firm with at least one classmate over two years of college.[23]

Gender Heterogeneity

Next, I analyze gender differences in classmate network formation. First, I look at how the propensity to use classmates for job-finding differs between men and women in Column (1) of Table 5. As in Table 4, I aggregate job finding outcomes for the first three years after the

outcome are dropped from analysis. As a robustness check, I have run the linear specification without perfectly predicted course bundles as well, and results do not significantly change.

[23]To derive these estimates, I use a back of the envelope calculation in which I assume the likelihood of working with each peer is an independent event. Thus, the linear estimate translates into a reported $.017 = (1 - .0058)^{155.4} - [1 - (.0058 + .000279)]^{155.4}$.

last course taken together by the pair. The first line of coefficients on number of sections shared shows the effect of taking an additional section with a peer for male students, and results indicate an increase of .026 percentage points in the propensity to get a job at the peer's firm within three years. The second line shows how this propensity differs for women. Results indicate that women do not differ significantly from men in their propensity to form classmate networks for any years of analysis.

Table 5: Gender Analysis of Effects of Taking Additional Section on Probability of Working at Peer's Firm

	(1) Linear	(2) Linear
Number of Sections	0.0262***	0.0173***
	(0.00366)	(0.00378)
Number of Sections×Female	0.00122	
	(0.00497)	
Female	0.0681***	
	(0.00494)	
Number of Sections×Same Gender		0.0145***
		(0.00489)
Same Gender		0.105***
		(0.00410)
Course Bundle FE	X	X
N	31,228,512	31,228,512

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by course bundle. Coefficient estimates multiplied by 100, so estimates represent percentages. Outcome: whether student starts working at firm where peer is incumbent at some point within three years of last shared course.

Column (2) of Table 5 investigates cross-gender patterns in network formation. The first row of coefficient estimates on number of sections shared represents the effects of taking an additional section together with a peer of the different gender. Results indicate an additional shared section leads to an increase of .017 percentage points in the propensity to get a job at a peer's firm when the peer is a different gender from the student. Next, the estimates on number of sections interacted with an indicator for same gender peer measures how this effect differs when a student's peer has the same gender as the student. The results here indicate that the effect of taking an additional section with a same-gender peer increases the propensity for a student to get a job by .015 percentage points more than the propensity with different-gender peers. These findings provide evidence of a significant gender homophily effect, in which students are almost twice as likely to network with peers of the same gender. One advantage of the pair-level analysis is that these effects account for the fact that men and women likely encounter peers of each gender at different rates due to sorting into courses. These findings indicate that conditional on taking a class with a peer of a certain gender, network effects are higher if the peer is the same gender as the student.

In summary, results indicate that taking a class with a peer increases the propensity for a student to get a job at a firm where her peer is incumbent. The overall propensity to use classmates in job finding does not differ by gender, although students display a significantly higher propensity to form networks with same-gender peers. Overall, these effects indicate classmates play a significant role in helping college students find jobs.

4 Conceptual Framework

This section sets up a conceptual framework for interpreting the network relationships I detect in the Section 3. Specifically, I identified a network relationship in which a classmate helps a student get a job at the classmate's firm. A priori, I do not know the channel through which these networks operate. One possibility may be that classmates help students by providing referrals to employers, but another possibility is that networks operate through classmates providing information about job openings, without referring them to employers. In this section, I first provide a framework for a referral-based job search model in which referrals increase the information on match quality between a worker and firm. Next, I provide testable implications of this model, as well as contrasting implications of a non-referral network mechanisms and other referral mechanisms. In Section 5, I measure the effects of obtaining a through classmates on labor market outcomes and use findings to assess the mechanisms outlined here.

4.1 Setup

The intuition behind the empirical implications of the model is based on Dustmann et al. (2016). This section does not attempt to give a deep theoretical analysis, but rather provide a context for how to conceptualize the economic setting behind the network effects detected in the previous section.

The labor market consists of workers and firms. Firms can hire workers through either the referral or external market. Referrals in this setting operate through classmate channels. A worker's productivity is match specific, and before a worker is hired, firms and workers observe an imperfect noisy signal of true productivity. Referrals reduce signal noise by providing firms and workers with more information regarding the match quality they otherwise would not have. In order to focus on the role of information, I assume the mean of the productivity distribution is the same in the referral and external market. Firms and workers use the signal they observe to form a belief about the worker's true productivity and make wage offers based on this belief. Since the expected value of the match is higher in the referral than in the external market, firms first try to fill a position through referrals. If hired, workers and firms learn the individual's true productivity over time. Firms and workers determine wages by splitting total surplus through a Nash bargaining process.

Sequence of Events

1. Employers observe vacancies, and an employer randomly chooses an employee to ask for a referral for each of the open spots. If the worker that the employee refers is unemployed, then the worker and firms meet; if the worker is employed, they do not meet and employers skip to the external market in Step 3.

    - Since having more classmates at the firm increases the chance a peer gets selected for a referral, increasing the number of incumbent classmates increases the probability a student gets referred.
    - I assume employees are more likely to refer former classmates with whom they have had more interactions, Thus, increasing the number of classes shared with an incumbent increases the propensity of getting referred.

2. The firm and referred worker observe a signal of the worker's match-specific productivity. The employer makes a wage offer, which the worker can either accept or turn down. If the worker turns down the offer, the position will remain vacant for the hiring cycle and the worker remains jobless.

3. Workers who were not referred and vacancies not filled by referrals participate in the

external market. Once again, the firms and workers that meet observe a signal of the worker's match-specific productivity. The employer makes a wage offer, which the worker can either accept or turn down. If the worker turns down the offer, the position remains vacant and the worker remains jobless.

- A student with an incumbent classmate at a firm may still get hired through the external market if the the classmate was not selected to provide a referral, or if the classmate was selected and referred a different peer.

4. In the next period, the firm and worker will learn the worker's true ability with probability α. Employers update their beliefs and subsequently their wage offers accordingly. If the worker turns the offer down, she becomes jobless and the firm has a vacant position.

5. The match ends for exogenous reasons with probability δ.

4.2 Empirical Implications

This model generates testable implications for the relative earnings and turnover levels for workers hired in the referral vs. external market. Since workers in the referral market have a more precise signal of productivity than those in the external market, this leads to a higher reservation match quality in the referral market. Intuitively, larger uncertainty of a worker's true match quality leaves larger opportunity for future wage growth, so workers are more likely to accept lower expected match qualities, a priori. If these matches do not end up being good, workers can always leave the firm. Since reservation match quality is higher in the referral market, workers hired through referrals are initially better matched than counterparts in the external market. As a result, they are more productive and hence earn more and have lower turnover rates, compared to external hires. With tenure, true productivity levels become revealed. Workers with bad matches leave the firm, and only good matches stay, for both referral and external hires. Thus, external and referral hires start looking more similar, and differences in earnings and turnover converge with tenure.

Table 6: Effects of Referral vs. External Hires

	Initial	with Tenure
Earnings	+	-
Turnover	-	+

Empirical Predictions

In Appendix C, I discuss the implications of a non-referral network channel, as well as alternative referral mechanisms. I show that these models generate different testable implications on earnings and turnover.

5 Measuring Effects of Network Jobs

Findings in Section 3 indicate students use classmates to find jobs. Here, I measure the effects of obtaining a job from a classmate on earnings and job separation rates. One key challenge to address in this analysis is I do not observe directly whether an individual works in a job obtained through a classmate. In this section, I first describe the research problem and setup and then explain the empirical strategy. Next, I present results and show that findings are consistent with the testable implications of a referral-based job-search model, as described in Section 4.

5.1 Description

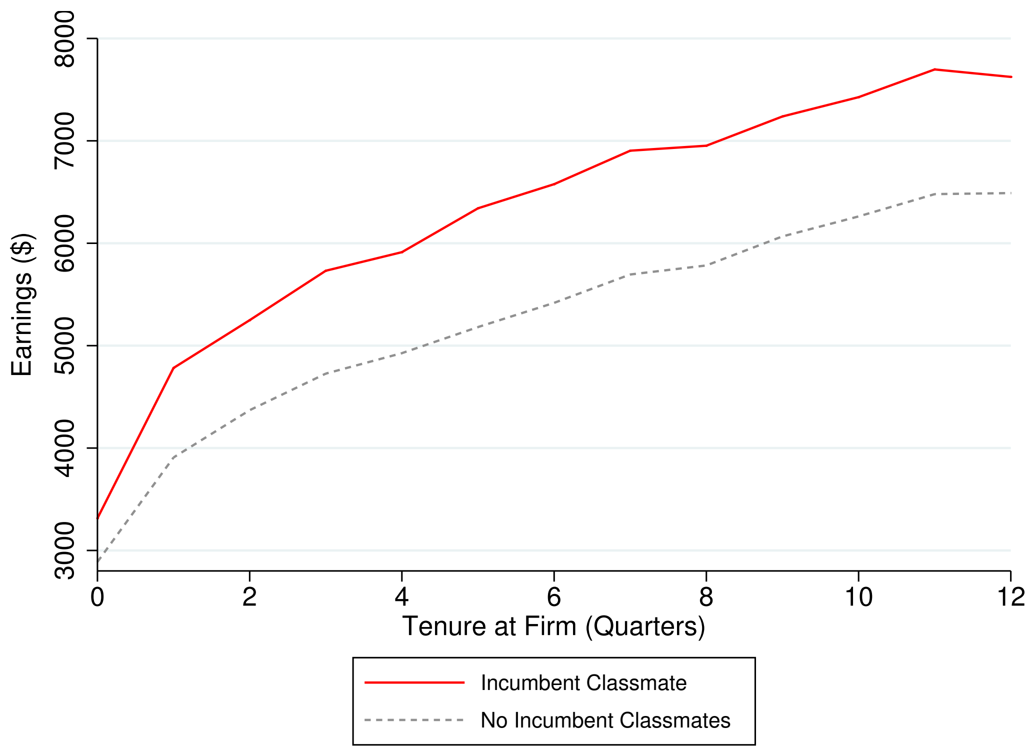
The first step in measuring the effects of obtaining a job through a classmate requires recovering information on whether an individual works in a job obtained through a classmate.[24] While I do not observe this variable directly in the data, I know from the first step of the research design that students do use incumbent classmates to find jobs. Thus, as a starting point, I compare earnings for students with no incumbent classmates at time of hire with that for students who have at least one incumbent classmate in Figure 5.[25] The graph shows that students who have at least one incumbent classmate at the firm consistently earn more

[24]For notational ease, a "classmate" in this paper refers to a peer with whom the student has shared a section.

[25]I do not follow individuals after 12 quarters because approximately 90% of workers no longer work at the firm by the 10th quarter, indicating fairly short employment spells in this context.

than students without any classmates at the firm.

Figure 5: Earnings by Incumbent Classmate Presence

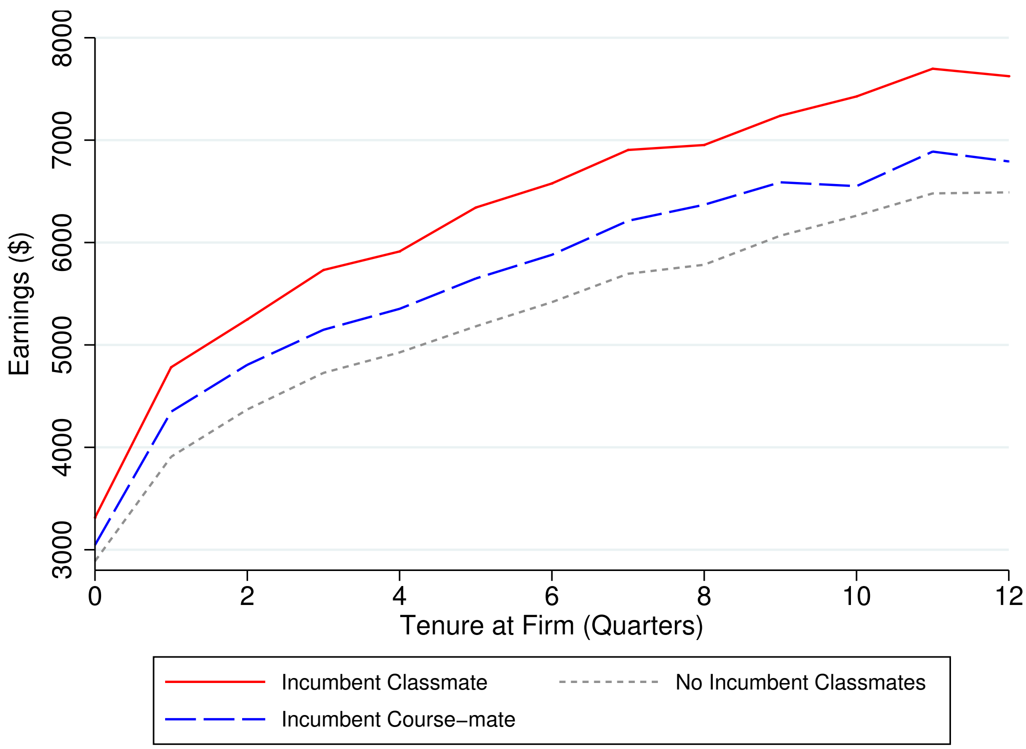


Note: The solid line represents earnings over tenure for individuals who had at least one incumbent classmate at the firm at time of hire, and the short-dashed line represents earnings for individuals with no incumbent classmates at time of hire.

Figure 5 suggests there may be earnings returns to network jobs, but the differences in earnings may be due to other factors correlated with enrollment into certain courses, independent of network effects. For example, it may be that firms disproportionately hire classmates because of a pipeline effect due to the reputations associated with schools, programs, or instructors. Subsequently, the firm may pay a premium for reputation, which is independent of classmate networking. Another example may be that courses that are correlated with higher paying jobs tend to be concentrated at a few firms. If this were the case, extraneous factors such as student sorting into courses or accreditation provided by certain courses may be driving earnings differences between students getting jobs at firms with vs. without incumbent classmates.

To disentangle the effects of extraneous factors correlated with course enrollment from network effects, I look at earnings for individuals who get a job at a firm with a at least one incumbent course-mate, but no incumbent classmates. The difference in earnings between students who get hired at firms with incumbent course-mates vs. incumbent classmates represents the portion of the total differences due to factors correlated with course enrollment, independent of classmate network effects. Figure 6 shows that factors correlated with course enrollment do explain a significant portion of the earnings differences between students who get hired at a firm with an incumbent classmate vs. those who do not. This indicates that in the empirical strategy, it will be important to control for course-level factors that may affect labor market outcomes.

Figure 6: Earnings by Incumbent Classmate/Course-mate Presence



Note: The solid line represents earnings over tenure for individuals who had at least one incumbent classmate at the firm at time of hire, and the short-dashed line represents earnings for individuals with no incumbent classmates at time of hire. The long-dashed line represents individuals who had at least one incumbent course-mate (who enrolled in a different section) at time of hire.

While differences in earnings between individuals who get hired at firms with incumbent classmates compared to those hired with incumbent course-mates from different sections are suggestive of positive earnings returns to networks, the trends in Figure 6 cannot be interpreted causally. First, these values do not take into account endogeneity in using classmate networks to find jobs. Second, Figure 6 does not say anything about the actual mapping between working at a firm with an incumbent classmate and getting the job from the classmate: Just because a student starts working at a firm with an incumbent classmate does not necessarily imply the classmate helped them get a job. There is also some probability that the student got the job independently of the classmate. I address these issues in turn with the empirical strategy in the next section.

5.2 Empirical Strategy

In this section, I first describe the analysis I would run in an ideal data world, in which I have explicit information on whether a worker works in a job obtained through a classmate. Next, since I do not observe how students obtain jobs, I describe how I recover this information using estimation results from the first step of the research design.

5.2a Ideal Data Setting

In an ideal data world, I would run the following regression:

$$Y_{imt} = \alpha_1 Ref_{imt} + \alpha_2 (Ref \times Tenure)_{imt} + \delta' X_{imt} + \gamma_t + \eta_i + f_m + \epsilon_{imt} \tag{3}$$

where Y_{imt} represents either log quarterly earnings of worker i at time t or an indicator variable for whether the worker leaves the firm in the next quarter. The variable Ref_{imt} denotes an indicator that equals one if i works in a job obtained through an incumbent classmate at time t and zero otherwise.[26] The vector X_{imt} represents a set of time-varying characteristics, including tenure. Finally, I denote π_t, η_i, and f_m as time, individual, and firm fixed effects, respectively. The coefficients of interest, α_1 and α_2, represent the initial effects of referrals on labor market outcomes and how these effects evolve with tenure at firm, respectively. If jobs obtained through classmates actually operate in the manner outlined in Section 4, then the coefficients should be $\alpha_1 >$ and $\alpha_2 < 0$ when Y_{imt} represents earnings and $\alpha_1 < 0$ and $\alpha_2 > 0$ when Y_{imt} represents job separation.

One concern in measuring the effects of referrals is that referral jobs are not randomly assigned to individuals, so the correlation between working in a job obtained through a

[26] I do not formally establish a referral mechanism until Section 4, but I use the variable Ref to denote obtaining a job through an incumbent classmate for ease of notation.

referral and labor market outcomes may be driven by unobserved factors. The benefit of panel data across the labor market is that I can follow individuals over time across multiple employment spells. Individual fixed effects here account for the possibility that low-ability and high-ability workers use referrals at different rates. When measuring turnover rates, I modify Equation 3 to include a control for prior turnover behavior in the three years leading up to getting hired at the firm, rather than an individual fixed effect.[27] Finally, I include firm fixed effects to account for the possibility that low-productivity and high-productivity firms hire through referrals at different rates. This ensures referred workers are compared to non-referred workers within the same firm.

To my knowledge, a data set that allows for the direct estimation of Equation 3 does not exist. Thus, I next describe an estimation strategy to recover the parameters of interest in Equation 3.

5.2b Recovering Network Information

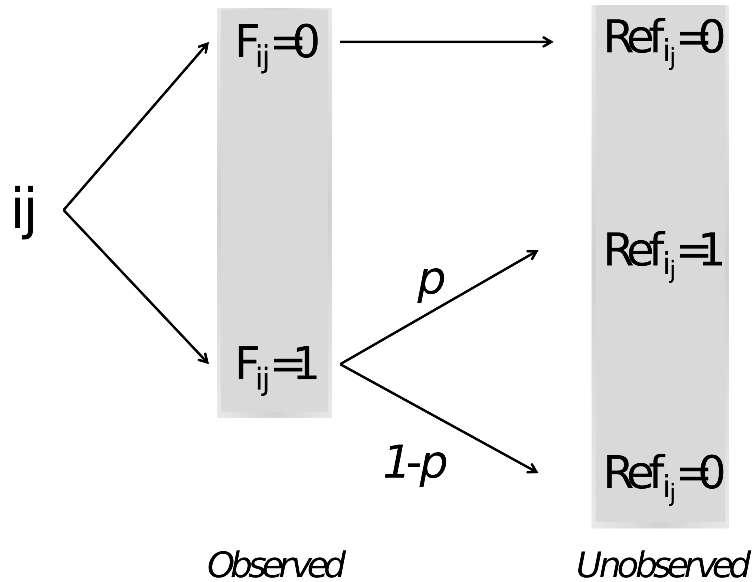
The goal here is to recover information on the unobserved variable Ref_i, denoting whether i works in a job obtained through a classmate at a given time.[28] As a starting point, I use information obtained from the pair-level analysis in the first step of the research design: Consider individual i and her former classmate, j. I denote Ref_{i_j} as an indicator for whether i obtained her job specifically through classmate j. When i gets a job at a firm, I observe in the data whether j is incumbent at time of hire. Following the notation in Equation 1 in Section 3, I denote this as an indicator F_{ij}, in which F_{ij} equals one if j is incumbent and zero otherwise. Figure 7 illustrates these possibilities graphically. If $F_{ij} = 0$, then $Ref_{i_j} = 0$ by definition, since j must be incumbent for Ref_{i_j} to equal one in Equation 3.[29] In the case where $F_{ij} = 1$, i.e. i does work at a firm where j is incumbent at time of hire, Ref_{i_j} can take on a value of either zero or one. With some probability p, i obtained the job through networking with j, and with some probability $1 - p$, i obtained the job independently of j.

[27] The rationale for this is that a priori, it is not clear that turnover is a monotonic function of student type since job separation can be good (e.g. climbing career ladder towards better jobs) or bad (e.g. getting fired). While high-productivity individuals will generally earn more than low-productivity individuals regardless of referrals, referrals may affect turnover in different directions.

[28] For ease of notation, I refer to this variable Ref_i instead of Ref_{imt} in this section.

[29] Technically, i may still have been referred even if j was not incumbent, which is not captured in this analysis. The presence of these situations would bias estimates toward zero, leading to conservative esitmates. Further, these non-captured referrals would not change the direction of effects, which provide key testable implications to understanding network mechanisms in Section 4.

Figure 7: Pair-Level Referral Probability Chart



Calculating p

To calculate p, the probability i obtained her job through j, conditional on getting hired at j's firm, recall Equation 2 from Section 3:[30]

$$F_{ij\tau} = \mathbb{1}\left[\tilde{\rho}_{c\tau} + \tilde{\gamma}_{\tau} N_{ij} + \tilde{\epsilon}_{ij\tau} > 0\right]$$

The probability that i starts working at a firm where j is incumbent ($F_{ij} = 1$) is a function of two components: the set of courses the pair takes together (ρ_c) and the number of sections out of those courses the pair shares (N_{ij}). Intuitively, the probability that i obtained her job from j, is the proportion of the total probability of working together that is due to section interaction, as opposed to other factors affecting why people in the same course may end up at the same firms. To calculate this probability, I exploit the same variation in section enrollment across courses used to detect networks in Section 3. It follows from Equation 2 that:

$$Prob(F_{ij\tau} = 1) = \frac{1}{1 + exp(-(\tilde{\rho}_{c\tau} + \tilde{\gamma}_{\tau} N_{ij}))} \tag{4}$$

[30] I use the logit rather than linear specification because a linear probability model can generate out-of-bounds predicted odds.

Equation 4 uses observable characteristics (N_{ij}) and estimated parameters ($\tilde{\rho}_c, \tilde{\gamma}$) to predict whether i works at a firm where j is incumbent. To avoid the incidental parameters problem in estimating $\tilde{\rho}_c$, I use the bias-corrected unconditional logit estimator proposed by Stammann et al. (2016).

Next, suppose i and j take a bundle of courses, c, together and have the same section for n of those courses. From Equation 4, the probability that i in time t will work at a firm where j was incumbent at time of hire is:[31]

$$Prob(F_{ij} = 1|N = n) = \frac{1}{1 + exp(-(\tilde{\rho}_c + \tilde{\gamma}n))}$$

Suppose that instead of sharing $n > 0$ sections together, i and j share zero sections. In this case, the probability of i working at j's firm becomes:

$$Prob(F_{ij} = 1|N = 0) = \frac{1}{1 + exp(-(\rho_c))}$$

Intuitively, $Prob(F_{ijt} = 1|N = 0)$ represents the baseline probability that i will up at j's firm, independently of classmate referrals. From the law of total probability:

$$\begin{aligned} Prob(F_{ij} = 1|N = n) &= Prob(Ref_{i_j} = 1|N = n) + Prob(F_{ij} = 1 \cap Ref_{i_j} = 0|N = n) \\ &= Prob(Ref_{i_j} = 1|N = n) + Prob(F_{ij} = 1|N = 0) \end{aligned}$$

where the first term after the equality follows since $Ref_{i_j} = 1 \rightarrow F_{ij} = 1$. I obtain the second line because the probability an individual works at her peer's firm without being referred by the peer, conditional on the set of courses they share, is equivalent to the probability the individual works at the peer's firm conditional on taking no classes together. Additionally, as before, I assume $Prob(Ref_{i_j} = 1|N = 0) = 0$. Rearranging terms, I obtain:

$$Prob(Ref_{i_j} = 1|N = n) = Prob(F_{ij} = 1|N = n) - Prob(F_{ij} = 1|N = 0)$$

Intuitively, the probability that i works at a job she obtained through j, conditional on sharing n sections, is the total probability i starts working at j's firm minus the probability that she would have gotten there independent of a referral from j. Using this, I calculate

[31] For ease of notation, I suppress time subscripts for the next few derivations.

p, the probability of obtaining the job through a referral conditional on i working at a firm where j was incumbent at time of hire:

$$\begin{aligned} p = Prob(Ref_{i_j\tau} = 1 | F_{ij\tau} = 1, N = n) &= \frac{Prob(Ref_{i_j\tau} = 1 \cap F_{ij\tau} = 1 | N = n)}{Prob(F_{ij\tau} = 1 | N = n)} \\ &= \frac{Prob(Ref_{i_j\tau} = 1 | N = n)}{Prob(F_{ij\tau} = 1 | N = n)} \\ &= \frac{Prob(F_{ij\tau} = 1 | N = n) - Prob(F_{ij\tau} = 1 | N = 0)}{Prob(F_{ij\tau} = 1 | N = n)} \\ &= 1 - \frac{\frac{1}{1+exp(-(\tilde{\rho}_{c\tau}))}}{\frac{1}{1+exp(-(\tilde{\rho}_{c\tau}+\tilde{\gamma}_{\tau} n))}} \end{aligned}$$

One result from this estimation is that variables in the final line of estimation denote either observable variables (n) or parameters estimated in Equation 2 ($\tilde{\gamma}, \tilde{\rho_c}$), meaning p can be solved for numerically. As illustrated in Figure 7, I now know the probability that i got her job from j for both cases where i starts working at a firm where j is incumbent and where i does not start working at such a firm. I write this expression as:[32]

$$Prob(Ref_{i_j t} = 1 | N = n) = \begin{cases} 1 - \frac{\frac{1}{1+exp(-(\rho_{ct}))}}{\frac{1}{1+exp(-(\rho_{ct}+\gamma_t n_{ij}))}} & F_{ijt} = 1 \\ 0 & F_{ijt} \neq 1 \end{cases} \tag{5}$$

As shown in Table 4, the effects of taking an additional section together on the propensity to work at a firm with an incumbent peer does not differ by gender. However, women have a higher baseline propensity than men to work at a firm with an incumbent peer, independent of classroom interaction effects. This in turn may lead to biased estimates on gender differences in the effects of obtaining a job through a classmate if I do not account for these differences. In Appendix D.1, I show how I extend this analysis to take into account gender differences in the calculation of the probability an individual obtained a job from a peer.

Aggregate Referral Probability

From Equation 5, I obtain $Prob(Ref_{i_j t} = 1)$, the probability an individual i was referred

[32]Equation 2 estimates the probability i works in a job referred by j, τ years after their last course together. To convert the time frame to a particular point in time t, I set $t = t_0 + \tau$, where t_0 denotes the year in which i and j took their last course together.

by a *specific* classmate J, for any given pair of individuals i and j. However, the variable of interest, $Prob(Ref_{it})$ denotes the probability an individual was referred by *any* of her classmates. To get from $Prob(Ref_{i_j t} = 1)$ to $Prob(Ref_{it})$, without loss of generality, suppose i gets hired at a firm where three former classmates, $\{j, k, l\}$, are incumbent. Additionally, I assume pair-level probabilities to be independent of each other (i.e., $Prob(Ref_{i_j t} \perp Ref_{i_k t} \perp Ref_{i_l t})$). Under this assumption, $Prob(Ref_{it})$ can be expressed as:

$$\begin{aligned} Prob(Ref_{it}) &= Prob(Ref_{i_j t} \cup Ref_{i_k t} \cup Ref_{i_l t}) \\ &= 1 - Prob(Ref^c_{i_j t} \cap Ref^c_{i_k t} \cap Ref^c_{i_l t}) \\ &= 1 - (1 - Prob(Ref_{i_j t}))(1 - Prob(Ref_{i_k t}))(1 - Prob(Ref_{i_l t})) \end{aligned} \tag{6}$$

where I use the pair-level independence assumption to arrive at the third line of Equation 6. To test the plausibility of this assumption, I calculate upper and lower bounds for the value of $Prob(Ref_{it})$, given the set of peers incumbent when i gets hired. Subsequently, I estimate the equation of interest using these estimates to see how much results vary within bounds. In Appendix D.2, I describe how I calculate the bounds and show estimation results using these bounds. My findings indicate bounded estimates are fairly narrow and show the same general trends as the main specification, providing support that the independence assumption is reasonable in this context.

5.2c Implementation

To assess the effects of obtaining a job through a classmate on labor market outcomes, I use the calculated probability that student i obtained her job through a classmate at time t in place of the true unobserved indicator value of the variable from Equation 3:

$$Y_{imt} = \alpha_1 \widehat{Ref}_{imt} + \alpha_2 (\widehat{Ref} \times Tenure)_{imt} + \delta' X_{imt} + \gamma_t + \eta_i + f_m + \epsilon_{imt} \tag{7}$$

where $\widehat{Ref}_{imt} = Prob(Ref_{imt} = 1)$. The intuition for the equation remains the same as in Equation 3: The coefficients of interest, α_1 and α_2, represent the initial effects of referrals on labor market outcomes and how these effects evolve with tenure at firm, respectively. If jobs obtained through classmates actually operate in the manner outlined in Section 4, then the coefficients should be $\alpha_1 >$ and $\alpha_2 < 0$ when Y_{imt} represents earnings and $\alpha_1 < 0$ and $\alpha_2 > 0$ when Y_{imt} represents job separation.

Measurement Error Concerns

In principle, the proxy variable should have the property that $E[\widehat{Ref}_{imt}] = E[Ref_{imt}]$. However, measurement error concerns arise since Ref_{imt} is latent in the analysis. More specifically, I face a case of non-classical measurement error since the unobserved error may be correlated with the latent variable. Thus, I cannot say anything directly about the magnitude or direction of the bias. I address these concerns in two ways: First, I provide an upper bound on the magnitude of the potential bias using the conditions laid out in Lewbel (2007) in Appendix D.3. Next, I discuss an alternate estimation strategy for obtaining estimates of α_1 and α_2 using a mixture model in Appendix D.4. The mixture model approach integrates out over Ref_{imt} in the likelihood function, using the information on wages and turnover in each employment spell to determine the likelihood that that spell was the result of a referral.

5.3 Results

This section presents findings on the effects of obtaining a job through an incumbent classmate on labor market outcomes. I first test key predictions of the referral model in Section 4, which postulates that individuals hired through classmate referrals initially earn more and have a lower propensity to leave their jobs compared to external hires, and these effects decline with tenure at the firm. I show that results are consistent with these predictions, providing support that classmates provide referrals to one another. Next, I explore how effects of referrals differ by gender.

Main Findings

I first look at the aggregate effects of obtaining a job from an incumbent classmate on log earnings and turnover rates in Table 7. For each outcome of interest, I look at results with and without firm fixed effects. The specification without firm fixed effects provides a look at the overall effect of classmate networks, consisting of both the role of networks in sorting people to higher- or lower- paying firms, while the specification with fixed effects isolates the effects of networks within employers. The latter specification plays a key role in assessing the model predictions in Section 4, which focuses on differentiating referral and external market outcomes within employers.

Results on earnings indicate getting a job through an incumbent classmate increases earnings initially by 10.5%, compared to individuals hired independently of classmates, and these effects decrease by 1.9 percentage points per quarter quarter of tenure, not accounting for firm fixed effects. With the inclusion firm fixed effects, incumbent classmate increases initial earnings by 13.6% compared to those at the same firm hired independently of classmate channels, and these effects decrease by 1.6 percentage points per quarter. One concern is that earnings effects to referrals will eventually become negative, after nine quarters at the firm. However, the mean tenure length for workers in this setting is approximately four quarters, and over 90% of workers leave before reaching nine quarters, indicating that referral effects are positive and converging towards zero through the relevant period. The increase in initial earnings effect with the addition of firm fixed effects also suggests that while students who use classmate networks get a sizable overall earnings premium, they may be going into lower paying firms. I explore this effect further in the subsequent analysis looking at gender heterogeneity. Next, columns (1) and (2) of Table 7 show the effects of classmate networks on turnover. I first look at effects without firm fixed effects and find that getting a job through an incumbent classmate decreases initial turnover propensity by 3.0 percentage points compared to those hired independently of incumbent classmates, and these effects fade by .3 percentage points per quarter at the firm. The addition of firm fixed effects does not significantly change outcomes.

Overall, I find that obtaining a job through an incumbent classmate initially increases earnings and decreases turnover, and these effects fade with tenure. These findings are consistent with the predictions of a referral-based job search model, in which referrals provide information about a worker's match quality to employers.

Gender Heterogeneity

Next, I look at differences in the effects of referrals by gender. I first look at overall effects of referrals on earnings and turnover, which incorporates both across- and within-firm components of referrals. Next, I include firm fixed effects in my specification to focus on differences

Table 7: Effects of Classmate Referrals on Labor Market Outcomes

	Log Earnings		Turnover	
	(1)	(2)	(1)	(2)
Referral	0.105***	0.136***	-0.030***	-0.031***
	(0.018)	(0.017)	(0.011)	(0.011)
Ref×Tenure	-0.019***	-0.016***	0.003*	0.003**
	(0.002)	(0.002)	(0.001)	(0.001)
Individual FE	X	X	X	X
Firm FE		X		X
N	3,944,492	3,937,551	3,214,723	3,209,052

*** p<0.01, ** p<0.05, * p<0.1. Specifications with firm fixed effects include interactions of firm effects with tenure.

in returns to referrals for men and women within the same employer. I focus on the latter outcome in particular since it highlights systemic gender discrepancies in how the referral mechanism operates.

The first column for each outcome of interest in Table 8 shows the overall effects of referrals, incorporating both the role of referrals in placing workers in higher- or lower-paying firms, as well as higher- or lower- paying positions within a firm. Findings indicate that the overall differences in earnings and turnover propensity between getting a job in the referral market vs. the external market comes from men using referrals: Men in referral jobs initially earn 20.0% more and have a turnover rate of 7.3 percentage points less than counterparts hired in the external market. These effects disappear completely for women, who have a 20.3 percentage point lower initial earnings than men and a 7.2 percentage point higher turnover propensity. In sum, these results suggest that there may be compensating differentials not tied to earnings and turnover propensity in the types of jobs women use referrals to get.

The second column for each outcome in Table 8 looks at effects of referrals on log earnings and turnover rates with the inclusion of firm fixed effects. I find that while men get an initial earnings premium of 16.8% from referrals, this effect is 10.3 percentage points lower for women. Additionally, while the earnings premium for referred workers fades by by 2.2 percentage points per quarter for men, it fades at a slower rate of .7 percentage points per quarter for women. Referrals decrease initial turnover propensity for men by 3.7 percentage points and this effect fades by .3 percentage points per quarter. Turnover effects do not differ significantly by gender.

Table 8: Gender Heterogeneity in Effects of Classmate Referrals

	Log Earnings		Turnover	
	(1)	(2)	(1)	(2)
Referral	0.199***	0.168***	-0.073***	-0.037**
	(0.025)	(0.023)	(0.015)	(0.015)
Referral×Tenure	-0.026***	-0.022***	0.004***	0.003**
	(0.002)	(0.002)	(0.002)	(0.001)
Female×Referral	-0.203***	-0.103***	0.072***	0.019
	(0.033)	(0.030)	(0.019)	(0.019)
Fem×Ref×Ten	0.011***	0.015***	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)
Individual FE	X	X	X	X
Firm FE		X		X
N	3,944,492	3,937,551	3,214,723	3,209,052

*** p<0.01, ** p<0.05, * p<0.1. Specifications with firm fixed effects include interactions of firm effects with tenure.

Overall, these findings indicate that while referrals benefit both men and women, the earnings premium of getting a job through a referral at a given firm is significantly lower for women. These findings indicate that compared to male colleagues who get jobs through referrals at the same firm, women benefit significantly less from some aspect of the mechanism behind referrals that leads to higher earnings. The next section explores the potential causes for this in more detail.

6 Gender Analysis

In this section, I explore possible causes for gender differences in the effects of referrals on earnings in more detail. First, I analyze the role of gender differences in sorting across industries and sorting across jobs within a firm. I show that these factors cannot explain differences in returns to referrals between men and women. Next, I analyze the role that gender differences in bargaining power may play in leading to gender differences in the returns to referrals. I discuss how empirical results from Section 5.3 are consistent with such a framework and show further empirical explorations of this mechanism.

6.1 Alternative Explanations

Heterogeneity in Referral Effects on Earnings by Industry

One potential explanation for gender differences in effects of referrals on earnings could be that certain industries provide a higher earnings premium to referrals than others, and women are disproportionately likely to get referrals in these industries. The inclusion of firm fixed effects in Equation 3 accounts for the fact that certain firms may be higher or lower paying, but it does not account for the possibility of heterogeneity in returns to referrals, conditional on baseline pay differentials.[33]

To empirically test this hypothesis, I extend Equation 7 to include interactions of Ref_{imt} with the gender of the individual, as well as the gender composition of the individual's industry:

$$\begin{aligned} Y_{imt} = {} & \alpha_1 \widehat{Ref}_{imt} + \alpha_2(\widehat{Ref} \times Tenure)_{imt} + \alpha_3 Prop\_fem_{imt} + \alpha_4(Prop\_Fem \times Tenure)_{imt} \\ & + \alpha_5(Prop\_Fem \times \widehat{Ref})_{imt} + \alpha_6(Prop\_Fem \times \widehat{Ref} \times Tenure)_{imt} + \beta_1 Count_{imt} \\ & + \beta_2(Count \times Tenure)_{imt} + \gamma Tenure_{imt} + \delta' X_{imt} + y_t + \eta_i + f_m + \epsilon_{imt} \end{aligned} \tag{8}$$

where $Prop\_Fem$ measures the proportion of the industry i works in consisting of women. If the lower returns to referrals for women are being driven by women working in industries with lower returns to referrals, I would expect a negative coefficient on α_5, which captures how the effects of referrals on earnings differ for industries with higher concentrations of women.

Table 9 shows results from estimating Equation 8. Column (1) shows results with industry aggregated at a 2-digit NAICS code level, and column (2) shows results with industry aggregated at the full 6-digit NAICS code level. Findings in both columns indicate that the gender composition of an industry does not significantly affect the earnings returns to networks, as indicated by the third line of coefficient estimates. In fact, the fourth line of estimates indicates an increase the female concentration of an industry actually *increases* the returns to network with tenure. Estimation results from Equation 8 indicate differences

[33]Controlling for firm effects automatically controls for industry effects since firms are nested within industries.

in returns to referrals between men and women cannot be explained by women working in industries with lower earnings returns to referrals.

Table 9: Effects of Referrals on Earnings across Industries

	Log Earnings	
	2-digit NAICS	6-digit NAICS
Referral	0.050	0.103***
	(0.038)	(0.035)
Referral×Tenure	-0.046***	-0.041***
	(0.004)	(0.004)
Referral×Prop. Female	0.084	-0.006
	(0.056)	(0.006)
Referral×Tenure×Prop. Female	0.052***	0.049***
	(0.053)	(0.006)
Individual FE	X	X
Firm FE	X	X

*** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses. The outcome of interest is log earnings. Specifications include student fixed effects, firm fixed effects, and interactions of firm effects with tenure.

Heterogeneity in Referral Job Type

A second possible explanation for the differences in effects of networks on earnings between men and women could be due to men and women getting referred to different jobs within a firm. For example, suppose a firm consists of only managers and secretaries, and most men in the firm get hired as managers while women in the firm get hired as secretaries. Further, the pay distribution for managers is higher than that of secretaries. Even if managers of all genders get the same referral premium and secretaries of all genders get the same referral premium as well, this would result in a higher overall referral premium for men since they also more likely get the premium of getting hired as a manager, rather than a secretary within the firm. This effect could further be exacerbated by gender homophily in network formation detected in Section 3 if people tend to refer others into their own jobs.

To assess whether heterogeneity in referral job type drives the referral earnings gap between men and women, I extend Equation 7 to include interactions of firm fixed effects by gender.[34] In doing so, I control for the fact that women may disproportionately work in lower paying jobs and isolate the analysis to the network premium within a firm. Table

[34] I also interact these effects with tenure.

10 presents results of this analysis. Column (1) displays original estimates with firm fixed effects from Table 8, and column (2) includes firm fixed effects interacted with whether a student is female. Results indicate that the differences in returns to networks between men and women persist with the inclusion of gendered firm fixed effects, indicating that gender homophily in job referral patterns does not drive results. Even compared to other women in the firm, women get a significantly lower earnings premium to referrals compared to male counterparts compared with other men in the firm. These findings indicate sorting into job types within firm by gender cannot explain the gap in returns to earnings from referrals.

Table 10: Effects of Referrals on Earnings, Firm×Gender Controls

	Log Earnings	Log Earnings
Referral	0.168***	0.138***
	(0.023)	(0.023)
Referral×Tenure	-0.022***	-0.022***
	(0.002)	(0.002)
Female×Referral	-0.103***	-0.080***
	(0.030)	(0.031)
Fem×Ref×Ten	0.015***	0.015***
	(0.002)	(0.002)
Individual FE	X	X
Firm FE	X	X
Firm×Female FE		X

*** p<0.01, ** p<0.05, * p<0.1. Specifications with firm fixed effects also include interactions of firm effects with tenure.

6.2 Gender Differences in Bargaining Power

In this section, I assess the potential for gender differences in bargaining power to explain discrepancies in the returns to earnings from referrals. Specifically, the main results indicate classmate networks are consistent with a referral-based job search model in which referrals increase information available on the match quality between a worker and firm. Improved match quality subsequently means there exists a higher surplus from the match to be split between the worker and firm. But as results in the Section 5 suggest, the surplus may be greater for men than women. One potential explanation for differences in returns to earnings from referrals is that men are able to bargain for a higher portion of the surplus generated by referrals compared to women.

Previous literature has shown that women generally have a lower propensity to negotiate and are less willing to compete compared to male counterparts (Niederle and Vesterlund (2011), Babcock and Laschevar (2003), Leibbrandt and List (2015)). However, results in Section 5.3 indicate women potentially have worse bargaining power in referral situations than men even compared to existing differences in the overall labor market. One question that arises is why women would have lower bargaining power in referral situations.

One reason for this stems from the literature on gender differences in negotiating. Studies have found that while women in general display a lower willingness to negotiate than men, these effects are especially stark in situations with increased "situational ambiguity" (Babcock and Laschevar 2003, Leibbrandt and List 2015). As defined in Babcock and Laschevar (2003), a setting has a lot of ambiguity if "people have to come up with their own interpretations as to what is the appropriate response", while situations with clear appropriate responses mitigate the amount of situational ambiguity. Referrals may add to the level of situational ambiguity in a negotiation setting since matches are made outside of the conventional job search framework and involve the presence of a third party, the referrer.

A second reason women may have lower bargaining power than men in referral settings compared to non-referral settings is that classmates who help students get jobs may also provide salary information. Women on average earn less than men, and results in Section 3 find that women tend to use female peers to find jobs and vice versa for men, Thus, to the extent that workers may anchor salary negotiation expectations to the earnings of the peers who referred them, this could lead to women bargaining for lower wages in referred jobs. To empirically assess this possibility, Table 11 shows how the earnings of the peer referrer affect the earnings premium in a referred job.[35] Column (1) shows that an increase in real quarterly earnings of the referrer by one standard deviation increases the initial earnings premium of network jobs by 41%. However, this effect may be driven by the fact that referrals sort people to higher paying firms, rather than the role of the referrer's earnings. Next, I add in firm fixed effects in column (2) to isolate the effect of referrer's earnings from firm-level pay premiums. Here, an increase in real quarterly earnings of the referrer by one standard deviation increases the initial earnings premium of network jobs by 14% within a firm. These findings indicate the effects of referrals on an individual's earnings are linked to the earnings of the referrer.

[35] I use the average wage of incumbent classmates at the firm as a proxy for referrer earnings.

Table 11: Effects of Referrals on Earnings, Interacted with Peer Referrer Earnings

	Log Earnings (1)	Log Earnings (2)
Referral	-.0391***	0.088***
	(0.018)	(0.017)
Referral×Referrer Earn	0.413***	0.143***
	(.011)	(0.010)
Referral×Tenure	-0.015***	-0.015***
	(0.002)	(0.002)
Referral×Tenure×Referrer Earn	-0.016***	-0.005***
	(0.001)	(0.001)
Individual FE	X	X
Firm FE		X
N	3,933,151	3,933,151

*** p<0.01, ** p<0.05, * p<0.1. Standard Errors in parentheses. Specifications with firm fixed effects also include interactions of firm effects with tenure. Incumbent peer earnings measured in standard deviations from mean.

If women do indeed capture a lower proportion of the surplus generated from referrals compared to men, this leads to lower initial earnings for women. Since women still get a non-negative portion of the surplus generated, this does not alter the match quality margin at which offers are accepted or rejected in the referral market in Section 4. After getting hired, true productivity is revealed over time for everyone, so both men and women in the referral market converge in earnings toward their counterparts in the external market with tenure in firm. Following the framework of Dustmann et al. (2016), bargaining does not enter a worker's turnover decision, which is purely a function of match quality. Thus, differences in bargaining power between men and women in the referral market do not lead to gender differences in turnover. Thus far, results from the analysis in Section 5 are consistent with the mechanism described above. These results provide suggestive evidence that gender differences in bargaining power lead to differential earnings returns between men and women, although I have not ruled out all alternative channels.

7 Conclusion

This paper studies the effects of referrals on labor market outcomes, paying particular attention to how these effects differ by gender. I employ a novel two-step research design to study classmate networks at community colleges. I first exploit quasi-random variation in section enrollment within college courses to show that taking a class with a peer significantly increases the propensity a student will get a job at a firm where the peer is incumbent. Next, I use estimation results from the first step of the research design to calculate the probability a student works in a job obtained through a classmate network and subsequently to calculate the effects of working in a network job on labor market outcomes. Obtaining a job through a classmate increases earnings and decreases turnover initially, and these effects fade over tenure at firm. These findings are consistent with the predictions of a referral-based job search model. However, while both genders benefit from referrals, the earnings premium for women is less than half of that for men. Over tenure at firm, effects of referrals converge between men and women for those who remain at the firm, and overall turnover rates do not differ by gender.

This paper has two main contributions: First, it uses a novel research design, which corrects for endogeneity concerns in referrals and assesses referral effects across an entire labor market. Second, it is the first study to analyze gender differences in the effects of referrals, highlighting an important tension between increasing labor market efficiency through referrals and increasing gender equity. From a policy perspective, the findings in this study have important implications for how we think about the role of referrals in the labor market. While referrals improve overall labor market efficiency and benefit both employers and workers, these benefits vary significantly across groups. Namely, for men and women with otherwise equal earning potential, referrals disproportionately benefit male workers. This finding is especially striking, given that the majority of firms in the United States employ programs that explicitly encourage referrals from incumbent workers (CareerBuilder 2012). Thus, while referrals improve information transmission and match quality between workers and firms in an economy, these effects come with a cost of exacerbating gender gaps in the labor market.

In the next steps of this paper, I plan to look further into the mechanism driving gender differences in returns to referrals. I would like to explore in more detail the possibility that women have lower bargaining power than men over the surplus generated by referrals, as well as alternative stories that may be contributing to this gap. The implications of these findings are potentially important from a policy perspective in understanding how to address gender disparities from referrals.

References

Altonji, J. G., Elder, T. E., and Taber, C. R. Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *journal of political economy*, page 34.

Babcock, L. and Laschevar, S. (2003). *Women Don't Ask: Negotiation and the Gender Divide*. Princeton University Press, stu - student edition edition.

Bayer, P., Ross, S., and Topa, G. (2008). Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. *Journal of Political Economy*, 116(6):1150–1196.

Beaman, L., Keleher, N., and Magruder, J. (2017). Do Job Networks Disadvantage Women? Evidence from a Recruitment Experiment in Malawi. *Journal of Labor Economics*, 36(1):121–157.

Beaman, L. A. (2012). Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the U.S. *The Review of Economic Studies*, 79(1):128–161.

Becker, G. S. (1971). *The Economics of Discrimination*. University of Chicago Press.

Bertrand, M., Goldin, C., and Katz, L. F. (2010). Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors. *American Economic Journal: Applied Economics*, 2(3):228–255.

Bertrand, M., Kamenica, E., and Pan, J. (2015). Gender Identity and Relative Income within Households. *The Quarterly Journal of Economics*, 130(2):571–614.

Blau, F. and Kahn, L. (2003). Understanding International Differences in the Gender Pay Gap. *Journal of Labor Economics*, 21(1):106–144.

Blau, F. D. and Kahn, L. M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55(3):789–865.

Brown, M., Setren, E., and Topa, G. (2015). Do Informal Referrals Lead to Better Matches? Evidence from a Firms Employee Referral System. *Journal of Labor Economics*, 34(1):161–209.

Burks, S. V., Cowgill, B., Hoffman, M., and Housman, M. (2015). The Value of Hiring through Employee Referrals. *The Quarterly Journal of Economics*, 130(2):805–839.

Bursztyn, L., Fujiwara, T., and Pallais, A. (2017). 'Acting Wife': Marriage Market Incentives and Labor Market Investments. *American Economic Review*, 107(11):3288–3319.

Card, D., Cardoso, A. R., and Kline, P. (2016). Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women. *The Quarterly Journal of Economics*, 131(2):633–686.

CareerBuilder (2012). *Referral Madness: How Employee Referral Programs Turn Good Employees into Great Recruiters and Grow Your Bottom Line*. CareerBuilder E-Book.

Cingano, F. and Rosolia, A. (2012). People I Know: Job Search and Social Networks. *Journal of Labor Economics*, 30(2):291–332.

Dustmann, C., Glitz, A., Schnberg, U., and Brcker, H. (2016). Referral-based Job Search Networks. *The Review of Economic Studies*, 83(2):514–546.

Fang, H. and Moro, A. (2011). Theories of Statistical Discrimination and Armative Action: A Survey. *Handbook of Social Economics*, 1a:78.

Glitz, A. (2017). Coworker networks in the labour market. *Labour Economics*, 44:218–230.

Goldin, C. (2014). A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, 104(4):1091–1119.

Heath, R. (2018). Why Do Firms Hire Using Referrals? Evidence from Bangladeshi Garment Factories. *Journal of Political Economy*, 126(4):1691–1746.

Hellerstein, J. K., McInerney, M., and Neumark, D. (2011). Neighbors and Coworkers: The Importance of Residential Labor Market Networks. *Journal of Labor Economics*, 29(4):659–695.

Hensvik, L. and Skans, O. N. (2016). Social Networks, Employee Selection, and Labor Market Outcomes. *Journal of Labor Economics*, 34(4):825–867.

Hotz, V. J., Johansson, P., and Karimi, A. (2017). Parenthood, Family Friendly Firms, and the Gender Gaps in Early Work Careers. Working Paper 24173, National Bureau of Economic Research.

Ioannides, Y. M. and Loury, L. D. (2004). Job Information Networks, Neighborhood Effects, and Inequality. *Journal of Economic Literature*, 42(4):1056–1093.

Kolesnikova, N. and Shimek, L. M. (2008). Community Colleges: Not So Junior Anymore. *The Regional Economist*, (Oct):6–11.

Kramarz, F. and Skans, O. (2014). When Strong Ties are Strong: Networks and Youth Labour Market Entry. *The Review of Economic Studies*, 81(3):1164–1200.

Kramarz, F. and Thesmar, D. (2013). Social Networks in the Boardroom. *Journal of the European Economic Association*, 11(4):780–807.

Leibbrandt, A. and List, J. A. (2015). Do Women Avoid Salary Negotiations? Evidence from a Large-Scale Natural Field Experiment. *Management Science*, 61(9):2016–2024.

Lewbel, A. (2007). Estimation of Average Treatment Effects with Misclassification. *Econometrica*, 75(2):537–551.

Ma, J. and Baum, S. (2016). Trends in community colleges: Enrollment, prices, student debt, and completion. *College Board Research Brief*.

Marmaros, D. and Sacerdote, B. (2002). Peer and social networks in job search. *European Economic Review*, 46(4-5):870–879.

McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27(1):415–444.

Montgomery, J. D. (1991). Social Networks and Labor-Market Outcomes: Toward an Economic Analysis. *The American Economic Review*, 81(5):1408–1418.

Niederle, M. and Vesterlund, L. (2011). Gender and Competition. *Annual Review of Economics*, 3(1):601–630.

Pallais, A. and Sands, E. G. (2016). Why the Referential Treatment? Evidence from Field Experiments on Referrals. *Journal of Political Economy*, 124(6):1793–1828.

Rees, A. (1966). Information Networks in Labor Markets. *The American Economic Review*, 56(1/2):559–566.

Saloner, G. (1985). Old Boy Networks as Screening Mechanisms. *Journal of Labor Economics*, 3(3):255–267.

Sarsons, H. (2017). Interpreting Signals in the Labor Market: Evidence from Medical Referrals. *Working Paper*.

Schmutte, I. M. (2015). Job Referral Networks and the Determination of Earnings in Local Labor Markets. *Journal of Labor Economics*, 33(1):1–32.

Simon, C. J. and Warner, J. T. (1992). Matchmaker, Matchmaker: The Effect of Old Boy Networks on Job Match Quality, Earnings, and Tenure. *Journal of Labor Economics*, 10(3):306–330.

Stammann, A., Hei, F., and McFadden, D. (2016). Estimating Fixed Effects Logit Models with Large Panel Data.

Topa, G. (2001). Social Interactions, Local Spillovers and Unemployment. *The Review of Economic Studies*, 68(2):261–295.

Topa, G. (2011). Labor Markets and Referrals. *Handbook of Social Economics*, 1:1193–1221.

USDOE (2017). Total fall enrollment in degree-granting postsecondary institutions, by level of enrollment, sex of student, and other selected characteristics: 2016. Technical Report Table 303.60, US Department of Education.

Zeltzer, D. (2017). Gender Homophily in Referral Networks: Consequences for the Medicare Physician Earnings Gap. *AEJ: Applied Economics (forthcoming)*.

Zimmerman, S. (2018). Elite Colleges and Upward Mobility to Top Jobs and Top Incomes. *American Economic Review (forthcoming)*.

Appendix A Additional Tables and Figures

Figure 8: Community Colleges in Arkansas

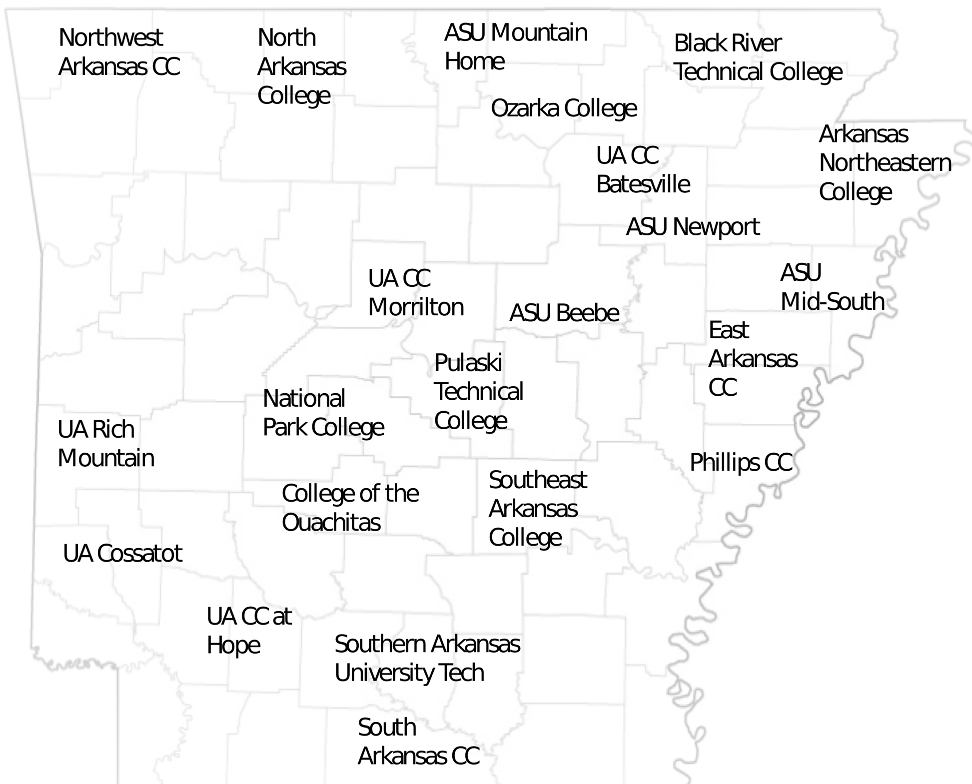


Table 12: Concentration of Students by Industry

Industry	Concentration
Health Care/Social Assistance	23.1
Retail Trade	19.9
Accomodation/Food Services	12.4
Manufacturing	8.4
Education	7.8
Admin./Waste Mgmt./Remediation	5.0
Public Administration	4.2
Finance and Insurance	3.1
Professional/Scientific/Technical Services	2.7
Wholesale Trade	2.2
Construction	2.2
Transportation and Warehousing	2.1
Other Services	1.8
Information	1.6
Arts/Entertainment/Recreation	1.2
Real Estate/Rental and Leasing	1.0
Utilities	0.3
Mining/Quarrying/Oil and Gas	0.3
Agriculture/Forestry/Fishing/Hunting	0.3
Management of Companies/Enterprises	0.3
Total	100.0

Industries are classified by two-digit NAICS codes. Each student-quarter employment spell represents an observation the calculation.

Table 13: Concentration of Students by Industry by Gender

(a) Male

Industry	Concentration
Retail Trade	22.5
Manufacturing	13.2
Accommodation/Food Services	12.8
Health Care/Social Assistance	9.2
Public Administration	6.9
Admin./Waste Mgmt./Remediation	5.9
Construction	4.9
Education	4.5
Wholesale Trade	3.8
Transportation and Warehousing	3.5
Professional/Scientific/Technical Services	2.5
Other Services	2.0
Finance and Insurance	1.8
Information	1.7
Arts/Entertainment/Recreation	1.5
Real Estate/Rental and Leasing	1.2
Mining/Quarrying/Oil and Gas	0.7
Utilities	0.7
Agriculture/Forestry/Fishing/Hunting	0.6
Management of Companies/Enterprises	0.2
Total	100.0

Industries are classified by two-digit NAICS codes. Each student-quarter employment spell represents an observation the calculation.

(b) Female

Industry	Concentration
Health Care/Social Assistance	30.9
Retail Trade	18.4
Accommodation/Food Services	12.2
Education	9.6
Manufacturing	5.8
Admin./Waste Mgmt./Remediation	4.5
Finance and Insurance	3.8
Professional/Scientific/Technical Services	2.8
Public Administration	2.6
Other Services	1.7
Information	1.5
Wholesale Trade	1.4
Transportation and Warehousing	1.3
Arts/Entertainment/Recreation	1.0
Real Estate/Rental and Leasing	1.0
Construction	0.7
Management of Companies/Enterprises	0.3
Utilities	0.2
Agriculture/Forestry/Fishing/Hunting	0.1
Mining/Quarrying/Oil and Gas	0.1
Total	100.0

Industries are classified by two-digit NAICS codes. Each student-quarter employment spell represents an observation the calculation.

Student Characteristics Before and After Sample Restriction

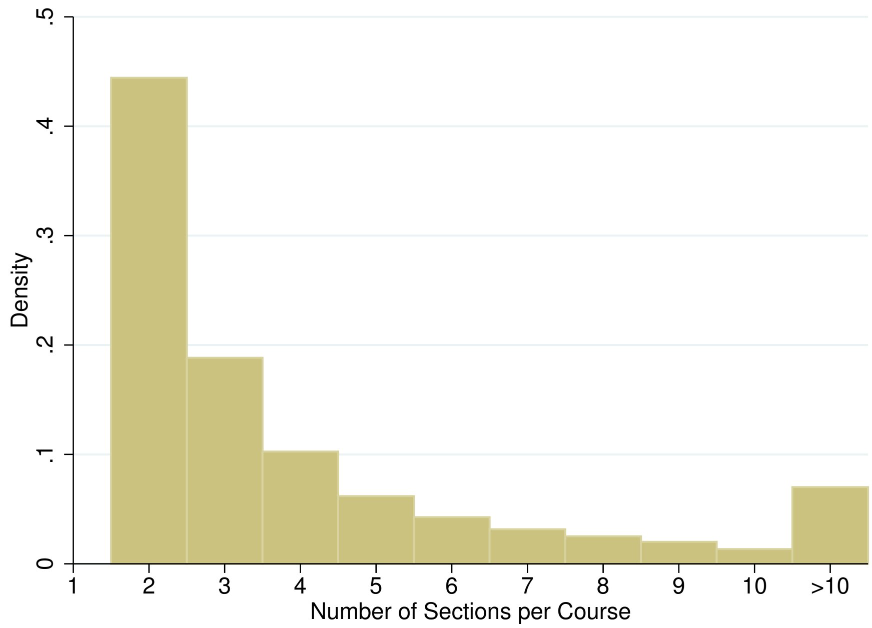
Table 14 displays student characteristics before and after restricting the sample. The first three columns show the final restricted sample from 1, and the last three columns show the original sample. Most of the values in the final sample are statistically different from corresponding values in the original sample due in part to the large sample size of individuals. However, the economic significance in the difference in magnitudes between the values is fairly small, and none of the gender trends in the data change.

Table 14: Student Characteristics Before and After Sample Restriction

	Final Sample (Restricted)			Original Sample		
	All	Male	Female	All	Male	Female
Female (%)	59.50	–	–	59.06	–	–
Part-time (%)	54.98	55.38	54.71	57.36	58.02	56.89
Employed (%)	54.00	51.43	55.75	53.75	51.92	55.02
Age	26.71	25.93	27.20	27.05	26.44	27.44
	(9.16)	(8.95)	(9.26)	(9.36)	(9.24)	(9.42)
Number of Courses	2.59	2.49	2.65	2.94	2.88	2.97
	(1.60)	(1.66)	(1.56)	(1.92)	(2.04)	(1.83)
N	389,342	157,674	231,668	445,789	182,505	263,284

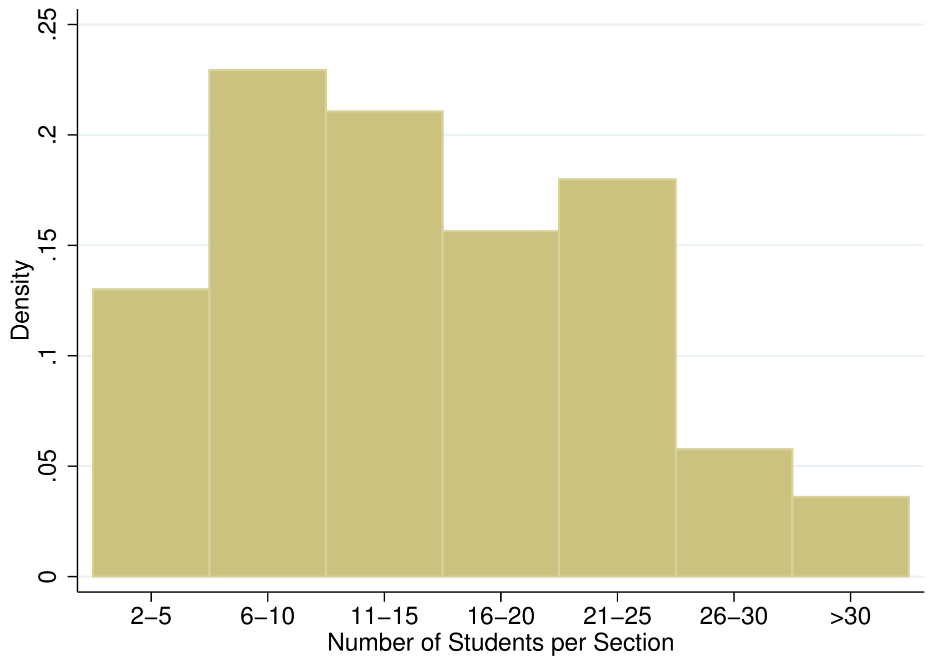
Two-Year College Students, 2003-2011. Table displays means (standard deviations in parentheses). Age is imputed from year of high school graduation, assuming students are 18 at the time they finish high school.

Figure 9: Distribution of Sections Per Course



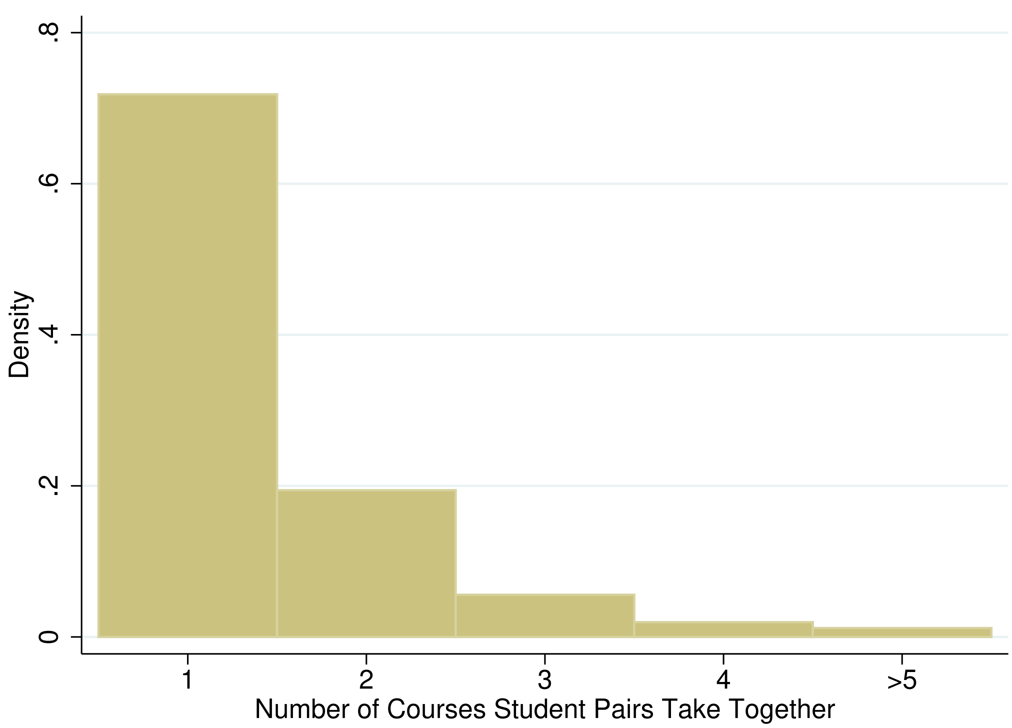
Note: Observations in histograms represent courses in the sample. This figure does not include sections from courses with only one section since these courses are dropped from the analysis.

Figure 10: Distribution of Students Per Section



Note: Observations in histogram represent sections in the sample. This figure does not include sections from courses with only one section since these courses are dropped from the analysis.

Figure 11: Distribution of Courses Per Pair



Identifying Network Effects: Main Results

This section displays results from Figure 4 in Section3 in table form. All coefficients in Table 15 represent estimates from the linear probability model in Equation 1. Coefficients can be interpreted as the effect of taking an additional section with a peer on the probability a student works at a firm where their peer was incumbent at time of hire, x years after their last course together. Columns represent the number of years after the last course, and the sample size decreases across the years because I am able to follow fewer pairs as years increase due to data constraints.

Table 15: Effects of Taking Additional Section on Probability of Working at Peer's Firm

	1yr.	2yrs.	3yrs.	4yrs.	5yrs.	6yrs.
Baseline	0.181	0.206	0.210	0.205	0.202	0.202
Number of Sections	0.0155***	0.0154***	0.0139***	0.0109***	0.00993***	0.0100***
	(0.00159)	(0.00174)	(0.00192)	(0.00209)	(0.00251)	(0.00239)
Course Bundle FE	X	X	X	X	X	X
N	42,977,504	38,239,214	31,831,440	25,028,626	18,906,500	13,283,296

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by course bundle. Coefficient estimates and baseline values multiplied by 100, so linear estimates represent percentages. Outcome: whether student works at firm where peer is incumbent x years after last course together.

Identifying Network Effects: Alternative Specification

Table 16: Effects of Taking Additional Section on Probability of Working at Peer's Firm

	1yr.	2yrs.	3yrs.	4yrs.	5yrs.	6yrs.
Baseline	0.181	0.206	0.210	0.205	0.202	0.202
Number of Sections	0.0145***	0.0142***	0.0108***	0.00881***	0.00465	0.00691**
	(0.00209)	(0.00239)	(0.00262)	(0.00283)	(0.00325)	(0.00351)
(Number of Sections)^2	0.000325	0.000376	0.000898*	0.000621	0.00151**	0.000868
	(0.000449)	(0.000430)	(0.000510)	(0.000440)	(0.000744)	(0.000576)
Course Bundle FE	X	X	X	X	X	X
N	42,977,504	38,239,214	31,831,440	25,028,626	18,906,500	13,283,296

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by course bundle. Coefficient estimates and baseline values multiplied by 100, so linear estimates represent percentages. Outcome: whether student works at firm where peer is incumbent x years after last course together.

Appendix B Checking for Sorting into Sections

I empirically test the validity of the identification assumption that students do not sort into sections within a course along factors that subsequently affect where they work. Specifically, I test for student sorting along two dimensions: First, I assess whether students sort systematically into certain sections of a course along characteristics that affect future job finding. Second, I look at whether students who have pre-existing relationships sort in sections together in ways that ways that affect subsequent subsequent propensity to work together.

Appendix B.1 Testing for Systematic Sorting into Sections

Assessment of Sorting on Observable Characteristics

Following the strategy of Bayer et al. (2008), I first analyze the correlation on observable characteristics between students and peers in their section, conditioning on overall course demographics.[36] For each section in the sample, I randomly select one student. Next, for each observable characteristic of interest in the data, I calculate the mean of the characteristic of the non-selected students in the section, in order to construct a measure of average characteristics of the section.[37]

Table 17 shows correlations along observable characteristics. Column 1 shows unconditional correlations, and column 2 shows correlations conditioned on course fixed effects. To obtain values in column 2, I first regress both individual and block measures separately on course fixed effects and then report the correlation between the residuals. These findings isolate the portion of the overall correlation to sorting patterns across sections within a course. As expected, results in column 1 show a significant degree of sorting across courses. Reassuringly, column 2 shows that with the addition course fixed effects, correlation on observable characteristics falls significantly. For example, the correlation between whether an individual is female and the proportion of peers in her section who are female is .28 overall, but this value falls to .05 with the addition of course fixed effects. On average, the addition

[36]While this evaluation does not explicitly provide information on the amount of sorting on unobservables, Altonji et al. shows the amount selection on observables to be proportionally informative of the amount of selection on unobservables.

[37]I sample only one individual per section in order to ensure that estimates are unbiased by the fact that each individual would significantly affect average group characteristics if multiple individuals per section were sampled.Bayer et al. (2008) provide a detailed explanation of this bias.

of course fixed effects reduces estimated correlations by 76% for observable characteristics, and the average correlation between individuals and peers across characteristics is .04.

Table 17: Correlation between Individual and Average Characteristics of Classmates

	Unconditional (1)	Conditional on Course (2)
Female	0.283	0.052
Employed	0.323	0.047
Out of State Residency	0.328	0.029
Transfer Student	0.323	0.045
Age 18-21	0.249	0.067
Age 22-26	0.055	0.025
Age 27+	0.095	0.034

Values shown represent correlations between individual characteristics and corresponding mean group characteristics of other individuals in the same section. Column 1 shows unconditional correlations, and the column 2 shows correlations conditional on course fixed effects.

Although correlation estimates in Table 17 are small after conditioning for course fixed effects, these values are not identical to zero. Thus, as a next step, I analyze whether remaining sorting into sections along observables would predict a significant higher probability of an individual getting a job at a firm with an incumbent same-section peer. To do so, I extend Equation 1 to allow for heterogeneity along pair-level covariates:

$$F_{ij\tau} = \rho_{c\tau} + \beta'_{1\tau}X_{ij} + (\gamma_\tau + \beta'_{2\tau}X_{ij})N_{ij} + \epsilon_{ij\tau} \tag{9}$$

where X_{ij} represents a vector of pair-level covariates describing i and j for each characteristic listed in Table 17. In Equation 9, β'_1X measures how the baseline propensity for an individual to start working at a firm with a course peer with whom they share no section varies based on the observable characteristics for each attribute of i and j. I estimate $\hat{\beta}'_1$ with the data and use this estimate to test whether the remaining correlation on observables across sections of a course would lead to a significantly higher predicted probability for a student to work at a firm where a peer with whom they shared more sections works. Specifically, I compare the average estimated value for $\hat{\beta}'_1X$ for pairs who took at least one section together with the observed (and predicted) value of $\hat{\beta}'_1X$ for pairs who shared no sections of their course bundle. This comparison indicates whether differences in block-level correlation by observable characteristics would lead to a higher propensity for an individual to start working at the same firm as a same-section peer, as opposed to a different-section peer.[38]

Table 18 presents predicted propensities for an individual to work in a firm where her peer was incumbent at time of hire τ years after a course, for pairs in no sections together vs. at least one section together. Results indicate that for all years of analysis, $\tau \in \{1, 2, ..., 6\}$, the estimated value of $\hat{\beta}_1'$ predicts that the propensity to work at a firm where a peer is incumbent is actually slightly *lower* for pairs who share one or more sections. For example, three years after the course, the predicted probability that an individual works at a firm where a peer she shared a section with was incumbent at time of hire is .257 percent, which is actually lower than the probability of working at the same firm as a peer from a no shared section, .261 percent. Thus, the small remaining amount of sorting on observables does not explain any of the increased propensity for individuals in the same section to work together in the future. Results from the test provide support for credibility of the identification strategy in showing that positive results cannot be explained by sorting within courses, at least on observable characteristics.

Table 18: Counterfactual Predicted Propensities

τ	All Different Sections	$\geq$ 1 Shared Section
+1 year	0.226	0.223
	(0.105)	(0.106)
+2 years	0.258	0.254
	(0.105)	(0.104)
+3 years	0.261	0.257
	(0.102)	(0.102)
+4 years	0.258	0.253
	(0.100)	(0.099)
+5 years	0.246	0.241
	(0.091)	(0.090)
+6 years	0.235	0.231
	(0.083)	(0.083)

Numbers in table represent the predicted propensity for an individual to work at a firm where the peer is incumbent, τ years after their last course together. Estimates are reported as percentage points.

[38] Over 70% of pairs only take one course together, so this test essentially captures a binary difference of sharing vs. not sharing a section.

Specification with Individual Fixed Effects

To further analyze concerns that students sort into sections within a course based on unobservable characteristics, I extend Equation 1 to include individual fixed effects for each member of the pair:

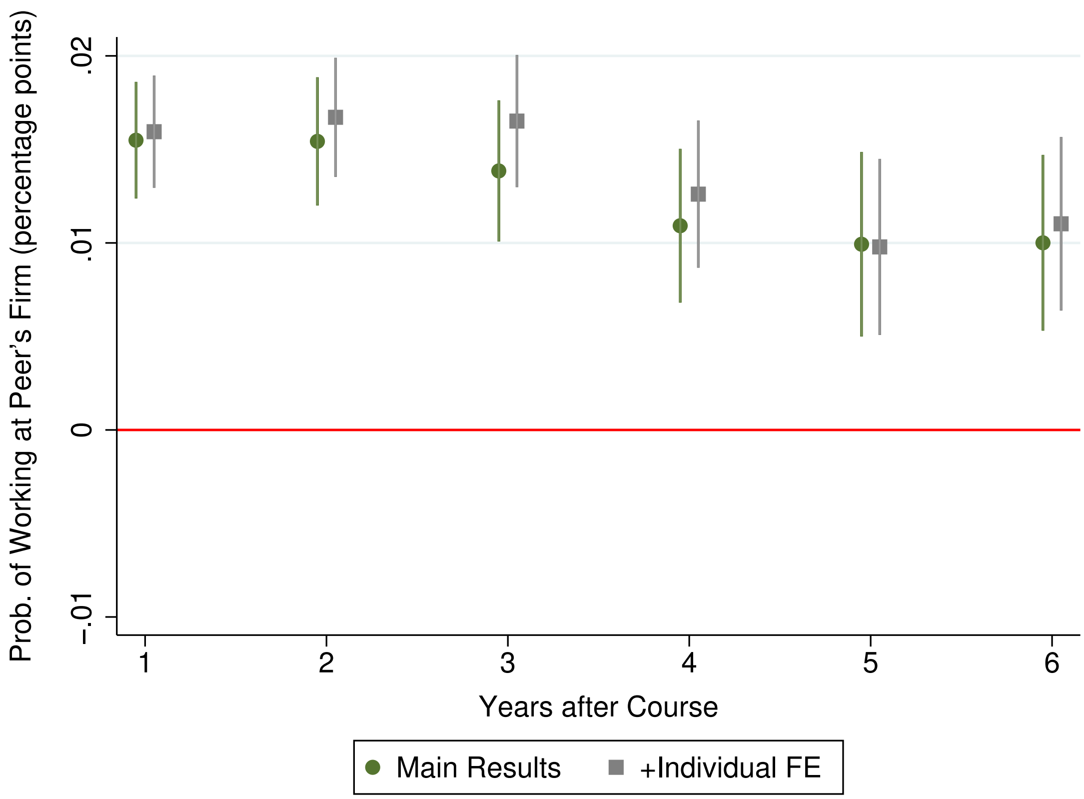
$$F_{ij\tau} = \rho_{c\tau} + \gamma_\tau N_{ij} + \lambda_i + \lambda_j + \epsilon_{ij\tau} \tag{10}$$

where λ_i and λ_j represent fixed effects for i and j, respectively. Since each individual appear multiple times in the data, the inclusion of individual fixed effects tests for one type of sorting on unobservables. Specifically, this analysis accounts for sorting in situations where (i) certain types of workers are more likely to work with those in the same section for unobserved reasons, and (ii) these workers sort themselves into similar sections within a course. For example, suppose some workers are significantly more likely to work at a firm where their peers work because they prefer to work at large firms. The chances of a peer working at the firm is statistically larger for these people, absent any networking effects. If these people who prefer large firms are also more likely to sign up for sections that take place in the morning, then the baseline analysis would incorrectly ascribe this increased propensity of working together to section interaction effects. Equation 10 controls for this concern by capturing any inherent differences across individuals to work with peers with fixed effects.

Figure 12 shows estimated coefficients for Equation 10, compared to estimated coefficients for the primary specification in Equation 1. Results indicate that estimates including fixed effects produce very similar estimates to the original specification on the effect of taking an additional section together on the propensity for a student to work at a firm with an incumbent peer for any of the years of analysis. All results are reported with 95% confidence bars and confidence intervals overlap between the original and fixed effects specifications for all estimates.[39] The findings provide further support that results are not driven by sorting on unobservable characteristics into sections within a course.

[39] Overlapping confidence intervals do not necessarily imply values are not statistically different. In the next steps of the paper, I plan to formally compare regression coefficients across models with high-dimensional fixed effects.

Figure 12: Effects of Taking Additional Section on Probability of Working at Peer's Firm: Including Individual Fixed Effects



Note: Coefficient estimates represent the effect of taking an additional section together and have been multiplied by 100 to reflect percentage changes. Standard errors are clustered by course bundle. Estimates are displayed with 95% confidence intervals.

Appendix B.2 Testing for Sorting on Pre-Existing Relationships

Next, I assess whether students tend to sign up for sections with pre-existing friends in a way that significantly affects the outcome of interest. This behavior would bias estimates if friends have correlated characteristics that make it more likely they will end up at the same firm, independent of classroom interactions.[40] If certain pairs of students are more likely to sign up in sections with friends, this effect would not necessarily be picked up by the sorting tests in Section B.1 if sets of friends sort together, but not systematically, across sections.

Since I do not observe directly whether a pair of students share a pre-existing relationship, I proxy for this information by looking at geographic proximity between a pair, as well as

[40]Students who are pre-existing friends may also be more likely to help out their classmates, but I worry less about this scenario since I still pick up a "networking" effect.

whether they worked together previously. The intuition is that pairs who live closer together or who were prior coworkers are more likely to have established social ties before taking a course together.[41] These metrics are motivated by prior research showing both residential proximity and former coworkers serve as significant channels for networking.[42]

To test whether sorting on pre-existing relationships drives outcomes, I first analyze whether students do in fact exhibit a higher propensity to enroll in the same section with peers who they worked with previously peers from a closer geographic proximity:

$$N_{ij} = \rho_{c\tau} + \phi_1 Prior\_Cowork_{ij} + \phi_2 Dist_{ij} + \epsilon_{ij\tau} \tag{11}$$

where $Prior\_Cowork$ represents an indicator for whether i and j worked at the same firm prior to taking courses together, and $Dist$ measures the travel distance in miles between the pair's high schools. The outcome of interest, N_{ij}, served as the an independent variable in Equation 1 and represents the number of sections i and j share, and ρ_c denotes a course bundle fixed effect. The coefficients ϕ_1 and ϕ_2 measure whether pairs who worked together previously or live closer enroll in more sections together, conditional on the bundle of courses shared. If students do display an increased propensity to sign up with these peers, I next look at whether the observed sorting is enough to significantly affect outcomes of interest in Equation 1. If students do not display an increased propensity to sign up with peers with whom they share pre-existing relationships, as measured by distance and prior work histories, this kind of sorting is likely not a big concern in this analysis.

Table 19 shows estimation results for Equation 11. Column (1) displays unconditional results, and column (2) includes course bundle fixed effects. Results from column (2) indicate pairs who were prior coworkers take approximately .04 more sections together, compared to pairs who did not work together before taking courses. An increase in the travel distance between pairs by one standard deviation decreases the number of sections individuals share by .01 sections. These findings indicate students display an increased propensity to enroll in sections with individuals with whom they were connected to previously, as measured by working together prior to the course and geographic proximity.

[41] I measure residential proximity of a pair by calculating the travel distance between their high schools of attendance, and I measure prior coworker status using an indicator for whether the pair worked at the same place for at least one quarter in the three years prior to their first course together.

[42] See Topa (2001), Bayer et al. (2008), Hellerstein et al. (2011), Schmutte (2015), Cingano and Rosolia (2012), Hensvik and Skans (2016), and Glitz (2017).

Table 19: Effects of Prior Coworkers and Travel Distance on Number of Sections Together

	(1)	(2)
Previous Coworker	0.135***	0.0418***
	(0.00797)	(0.00202)
Travel Distance	-0.00828***	-0.0110***
	(0.00264)	(0.00135)
Course FE		X
N	25,731,304	25,718,668

*** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered by course bundle. Travel distance is measured using z-scores.

Since students display an increased propensity to sign up with peers they worked with or lived closer to, one concern is that unobservable traits that lead people to sign up for the same section also affect their propensity to work at the same firm, independent of network effects. In this situation, the coefficient γ in Equation 1 would overestimate the true effect of taking an additional section with a peer on the propensity of getting a job at a firm where the peer is incumbent. To test for this, I extend Equation 1 to include controls for prior coworkers and travel distance:

$$F_{ij\tau} = \bar{\rho}_{c\tau} + \bar{\gamma}_\tau N_{ij} + \phi_1 Prior\_Cowork_{ij} + \phi_2 Dist_{ij} + \bar{\epsilon}_{ij\tau} \tag{12}$$

where as before, $F_{ij\tau}$ denotes an indicator for whether i works at a firm where j was incumbent at time of hire, τ years after their last course together. I control for sorting into courses with a course bundle fixed effect, $\bar{\rho}_c$, and N_{ij} measures the number of sections i and j took together out of their shared courses. Additionally, I add variables $Prior\_Cowork$ and $Dist$ to account for the role sorting into sections along pre-existing relationships plays in affecting subsequent job-finding outcomes. If estimates of $\bar{\gamma}$ do not differ significantly from estimates of γ in Equation 1, this implies sorting into sections based on pre-existing relationships as captured by former coworkers and peers in closer geographical proximity does not drive results.[43]

[43] It is possible that students systematically form pre-existing relationships in ways that are not captured by residential proximity or prior work history. Since students typically enroll in college either out of high school or after working for some time, geographic proximity based on high school and former workplace contacts should capture a significant portion of an individual's relationships. Furthermore, many other places where people form relationships will also be geographically correlated (e.g. family, places of worship, social groups).

Figure 13 plots estimation results of Equation 12 for up to six years of analysis. Additionally, the graph includes original estimation results from Equation 4. All estimates are reported with 95% confidence bars. Results indicate travel distance and prior work relationships do not drive the effect of taking an additional section on the propensity for an individual to work at a firm where a peer is incumbent for any of the years of analysis. For the first four years after taking a course together, including controls for pre-existing relationships actually increases the point estimates of taking a section together on the propensity to work at a peer's firm, although these results are not statistically different. Estimation results for Equation 12 six years after the last course together are no longer statistically different from zero, due in large part to declining sample sizes across years. Additionally, for all years of analysis, coefficient estimates from the original analysis and estimates with the addition of controls for sorting on pre-existing relationships have overlapping confidence intervals.[44]

Appendix C Alternative Network Mechanisms

In Progress

Appendix D Measuring Effects of Network Jobs

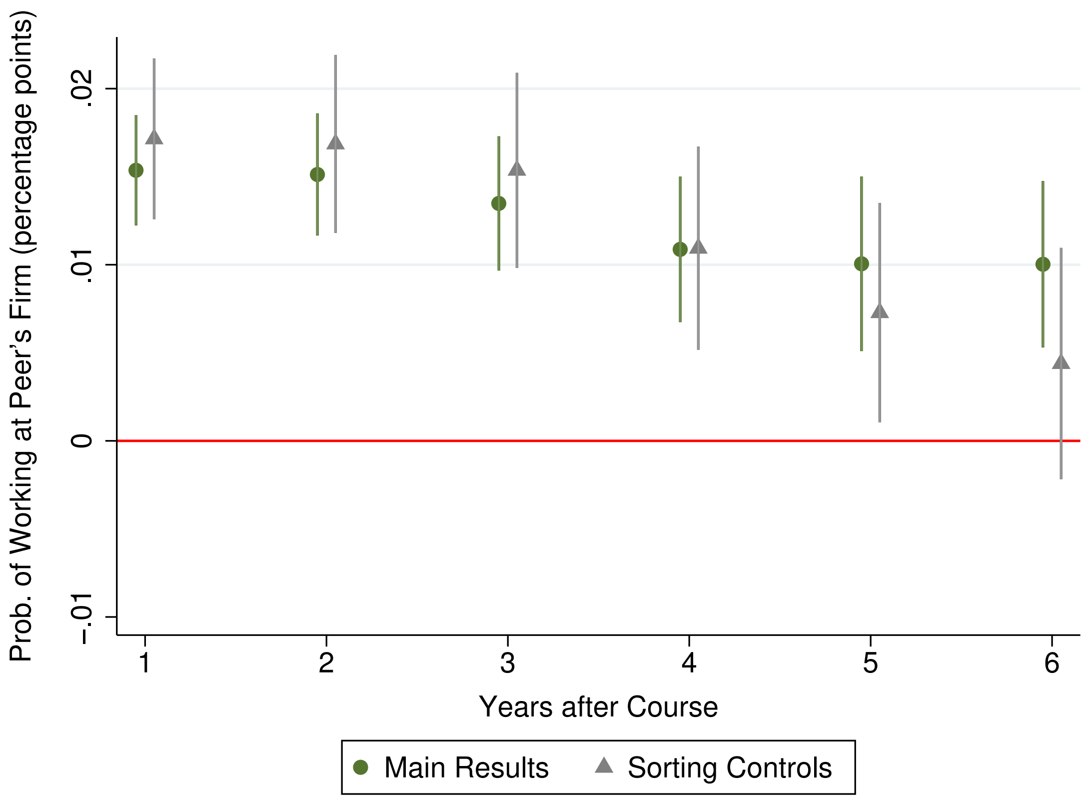
This section provides appendix notes for Section 5 of the paper.

Appendix D.1 Specification with Gender Heterogeneity

One concern with the estimation of $Prob(Ref_{imt} = 1)$ arises in the analysis of referral effects by gender. Table 5 indicates the effects of taking an additional section together on the propensity to work at a peer's firm does not differ by gender. However, women have a higher baseline propensity than men to work at a firm with an incumbent peer, independent of classroom interaction effects. Since baseline propensity to work with a peer factors into the calculation of $Prob(Ref_{imt} = 1)$, failure to account for gender heterogeneity in baseline values will bias estimates of gender differences in the effects of obtaining a job through a

[44] Overlapping confidence intervals are suggestive of closeness but do not necessarily imply values are not statistically different. In the next steps of the paper, I plan to formally compare regression coefficients across models with high-dimensional fixed effects.

Figure 13: Effects of Taking Additional Section on Probability of Working at Peer's Firm: Including on Pre-Existing Relationship Sorting Controls



Coefficient estimates have been multiplied by 100 to reflect percentage changes, and standard errors are clustered by course bundle. Estimates are displayed with 95% confidence intervals.

classmate. To address this, I extend Equation 2 to allow for heterogeneous effects of taking an additional section together and in baseline propensity to work with a peer:

$$F_{ijt} = \mathbb{1}_{\rho_{ct}+\phi_{1\tau}Fem_i+(\gamma_t+\phi_{2\tau}Fem_i)N_{ij}+\epsilon_{ij\tau}>0}$$

$$\downarrow$$

$$Prob(Ref_{i_jt} = 1|N = n) = \begin{cases} 1 - \frac{\frac{1}{1+exp(-(\rho_{ct}+\delta'_{1t}Fem_i))}}{\frac{1}{1+exp(-(\rho_{ct}+\delta'_{1t}Fem_i+(\gamma_t+\delta'_{2t}Fem_i)n_{ij}))}} & F_{ijt} = 1 \\ 0 & F_{ijt} \neq 1 \end{cases} \tag{13}$$

I use $Prob(Ref_{i_jt} = 1)$ in Equation 13 as the preferred pair-level referral specification both for analyzing the effects of getting a job through a classmate by gender, as well as aggregate effects of getting a job through a classmate. Apart from allowing probabilities to vary by student gender, the methodology behind the derivation of Equation 13 remains the same as that of Equation 5.

Appendix D.2 Testing the Pair-Level Independence Assumption

To test the plausibility of the pair-level probability independence assumption, I estimate upper and lower bounds for the value of $Prob(Ref_{it} = 1)$. I then estimate the effects of referrals on labor market outcomes to test sensitivity of results. If the direction of results do not change and the the magnitude does not vary much between bounds, this provides support for the plausibility of the independence assumption. If a student gets hired at a firm with one or zero incumbent classmates, there is no need to calculate bounds since I do not need to aggregate over pair-level probabilities to calculate the probability an individual works in a job obtained through an incumbent classmate.

In situations where students get hired at a firm with more than one incumbent classmate, I start by estimating bounds for the value of $Prob(Ref_{it} = 1)$ if the independence assumption were relaxed. To do so, without loss of generality, suppose again that i starts working at firm with three incumbent classmates, j, k, and l.:

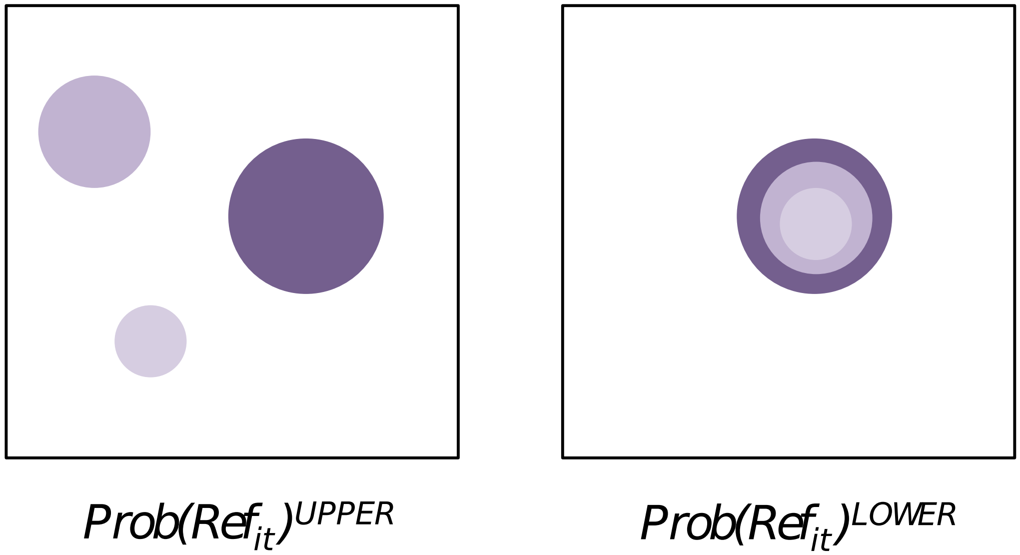
$$Prob(Ref_{it})^{UPPER} = \max\{Prob(Ref_{i_j t}) + Prob(Ref_{i_k t}) + Prob(Ref_{i_l t}), 1\} \tag{14}$$

$$Prob(Ref_{it})^{LOWER} = \max\{Prob(Ref_{i_j t}), Prob(Ref_{i_k t}), Prob(Ref_{i_l t})\} \tag{15}$$

where $Prob(Ref_{it})^{UPPER}$ and $Prob(Ref_{it})^{LOWER}$ represent the upper and lower bounds of the probability of referral, respectively. Figure 14 depicts this intuition graphically. The highest value of the probability i works in a referral job occurs if pair probabilities were mutually exclusive, in which case the probability of referral is denoted as the sum of the pair probabilities, with the value capped at one. On the other hand, the minimum value of the probability i works in a referral job occurs if pair probabilities were all nested within one another. In this case, the probability of referral would be equal to the value of the largest pair probability.

Next, I use $Prob(Ref_{it})^{UPPER}$ and $Prob(Ref_{it})^{LOWER}$ in place of $Prob(Ref_{it})$ to analyze lower and upper bounds of the effects of referrals on earnings and turnover rate. If estimates of referral effects at the bounds still produce results consistent with a referral-based job search

Figure 14: Probability Bounds



Note: Different circles represent pair-level referral probabilities.

model, this provides support for the independence pair-level probabilities as a reasonable assumption in calculating $Prob(Ref_{it})$.[45] Additionally, I estimate effects of referrals using a restricted sample of individuals who have at most one incumbent peer at a firm at time of hire. This creates a selected sample size, but one benefit of this estimation is that I do not need to aggregate pair-level observations to obtain an estimate of $Prob(Ref_{it})$. Due to the large number of employers in the economy, less than four percent of jobs involve students getting hired at a firm with more than one incumbent peer.

Table 20 shows bounds estimates on the effects of referrals on log earnings. All specifications include both individual and firm fixed effects. Column (1) of Table 20 shows estimation results using the original specification, $Prob(Ref_{it})$, from Table 7. Column (2) displays lower bound estimates of referral effects using $Prob(Ref_{it})^{UPPER}$ and column (3) displays upper bound estimates of referral effects using $Prob(Ref_{it})^{LOWER}$.[46] Bounds estimates indicate referrals increase earnings initially by 11.3-18.1%, and these effects fade by 1.2-2.5 percentage points per quarter at the firm. All estimates are statistically significant and indicate

[45] Note I only require the pair-level independence assumption in situations in which a student gets a job at a firm where multiple peers are incumbent. Due to the large number of employers in the economy, students have zero or one incumbent peers in 96% of cases.

[46] $Prob(Ref_{it})^{UPPER}$ denotes the upper bound estimate for whether i works in a referral job, which means it overstates the true probability of a referral, leading to an underestimate of the effect of referrals on earnings. The opposite applies to $Prob(Ref_{it})^{LOWER}$.

referrals increase initial earnings, with effects fading with tenure at firm. Next, in column (4), I estimate effects of referrals on the restricted sample of cases in which an individual has at most one incumbent peer at the firm at time of hire. These estimates are very similar the original specification in column (1), with referrals increasing initial earnings by 14.5% and effects decreasing by 2.2 percentage points per quarter at the firm. Overall, these estimates provide support for the plausibility of the pair-level referral probability assumption to calculate $Prob(Ref_{it})$.

Table 20: Bounds Estimates: Effects of Referrals on Earnings

	Original	Lower Bound	Upper Bound	Restricted
	(1)	(2)	(3)	(4)
Referral	0.136***	0.113***	0.181***	0.145***
	(0.017)	(0.014)	(0.025)	(0.036)
Referral×Tenure	-0.016***	-0.012***	-0.025***	-0.022***
	(0.002)	(0.002)	(0.003)	(0.005)
Individual FE	X	X	X	X
Firm FE	X	X	X	X
N	3,937,551	3,937,551	3,937,551	3,758,499

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The outcome of interest is log earnings, and all results are estimated using a linear probability model. Estimates with firm fixed effects include interactions of firm effects with tenure.

Next, Table 21 shows bounds estimates on the effects of referrals on turnover. As in the estimates for earnings, column (1) shows estimation results using the original specification, $Prob(Ref_{it})$, from Table 7. Column (2) displays lower bound estimates of referral effects using $Prob(Ref_{it})^{UPPER}$ and column (3) displays upper bound estimates of referral effects using $Prob(Ref_{it})^{LOWER}$. Lower bound estimates show referrals decrease turnover initially by 2.8 percentage points, and they subsequently increase by 0.2 percentage points per quarter. These estimates are statistically significant and are close in magnitude to estimation results using the original specification. Results using an upper bound on the calculation for the probability an individual works in a referral job indicate referrals decrease turnover initially by 2.2 percentage points and effects subsequently increases by 0.4 percentage points per quarter. While initial referral effects are not statistically significant, and effects of referrals interacted with tenure are only weakly significant, coefficients have the same sign and similar magnitudes to the estimates in the original specification. Estimates in column (4) using the restricted sample also show coefficients with matching signs to the original specification, although once again estimates are not statistically significant.

Table 21: Bounds Estimates: Effects of Referrals on Turnover

	Original	Lower Bound	Upper Bound	Restricted
	(1)	(2)	(3)	(4)
Referral	-0.031***	-0.028***	-0.022	-0.011
	(0.011)	(0.009)	(0.017)	(0.025)
Referral×Tenure	0.003**	0.002**	0.004*	0.003
	(0.001)	(0.001)	(0.002)	(0.004)
Prior Turnover Controls	X	X	X	X
Firm FE	X	X	X	X
N	3,209,052	3,209,052	3,209,052	3,044,160

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The outcome of interest is turnover, and all results are estimated using a linear probability model. Estimates with firm fixed effects include interactions of firm effects with tenure.

Overall, estimation results using upper and lower bounds, as well as the restricted estimator, provide estimates in line with the main results from Table 7. This provides further support for the validity of the independence assumption in calculating the probability an individual works in a job obtained through a classmate.

Appendix D.3 Upper Bound on Measurement Error

In Progress

Appendix D.4 Alternative Estimation Strategy: Mixture Model

To circumvent the problem of measurement error in latent variable analysis, I propose a mixture model analysis as an alternative estimation strategy. The mixture model recovers parameters of interest without identifying Ref_{imt}. Intuitively, I am trying to measure the effects of getting a job at a firm through an incumbent classmate. However, if a student starts working at a firm with an incumbent peer in the data, I do not observe whether or not the job was obtained through the peer. Thus, I need to find a way of measuring the effect of getting a job through a peer without actually observing this variable in the data. The mixture model uses the observed data to model the presence of subpopulations of students obtained their jobs through classmates vs. independently of classmates. I implement this approach in three steps:

1. Recover all parameters from Equation 3 other than α_1, α_2 using the set of employment spells that definitely are not obtained from an incumbent classmate (i.e., jobs in which students get hired at a firm with no incumbent classmates):

$$\begin{aligned} Y_{imt} = &\hat{\delta}' X_{imt} + \hat{\beta}_1 Course_{imt} + \hat{\beta}_2 (Course \times Tenure)_{imt} + \\ &+ \hat{\eta}_i + \hat{f}_m + \hat{\gamma}_t + \epsilon_{imt} \end{aligned}$$

   where Y_{imt} represents log earnings and other variables correspond to definitions from Section 5.

2. Calculate residual values for potentially referred jobs (i.e., jobs in which students get hired at a firm with an incumbent peer they shared a section with):

$$\begin{aligned} \hat{R}_{imt} = &Y_{imt} - \hat{\delta}' X_{imt} - \hat{\beta}_1 Course_{imt} - \hat{\beta}_2 (Course \times Tenure)_{imt} - \\ &- \hat{\eta}_i - \hat{f}_m - \hat{\gamma}_t \end{aligned}$$

   Note that residuals are estimated from the subsample of individuals who have at least one job spell in which they do not work with a former classmate, or else they would not have recovered parameter values in Step 1. Employment spells in which students get hired at a firm where at least one incumbent peer works fall into one of two groups: 1) jobs obtained with the help of an incumbent peer or 2) jobs obtained without the help of any incumbent peers. To the extent that jobs obtained through classmates differ in earnings from jobs obtained without the help of classmates all else equal, the residual value, $\hat{R}_{imt}$, captures these differences.

3. Estimate α_1, α_2 in mixture model using maximum likelihood estimation:

$$\begin{aligned} \ell(\alpha_1, \alpha_2) = \sum_i \sum_r ln \big[&\hat{p}_{imr} \prod_{\tau=1}^{T(r)} f(\hat{R}_{im\tau} - \alpha_1 - \alpha_2 Tenure) + \\ &(1 - \hat{p}_{imr}) \prod_{\tau=1}^{T(r)} f(\hat{R}_{im\tau}) \big] \end{aligned} \tag{16}$$

   In Equation 16, I sum over all potentially referred employment spells r of individual i

for all individuals who work at least one spell in a job without any incumbent classmates. For each employment spell, I assess residual values over each quarter at firm, τ. I assume a normal probability distribution on the distribution of residual values, $f(\cdot)$. The probability weights represent the probability a job was obtained through a classmate, $\hat{p}_{imr}$, and the probability the job was obtained independently of the classmate, $(1-\hat{p}_{imr})$. I use the calculated probability of obtaining a job through a classmate from Section 5, $Prob(Ref_{imt} = 1)$ for $\hat{p}_{imr}$. I can also extend Equation 16 to look at turnover by using a logistic distribution for $f(\cdot)$ since this represents a binary variable. In the next step of the paper, I plan to estimate the parameters of interest, α_1, α_2 using the mixture model. This will provide an alternative estimation strategy to deal with the measurement concern.