Peer Gender Composition and Undergraduate Achievement and Major Choice

Sidharth Sah*

October 31, 2023

Abstract

Large gender differences exist in the take-up and completion of college majors across academic fields. The degree of gender concentration within fields tends to increase over time spent in college. In this paper, I investigate how the gender composition of peers in first-semester classes impacts women's and men's academic outcomes and major choices. I find that a larger proportion of male peers hurts female academic achievement and decreases female persistence in majors, relative to men in the same classes. The effect of male peers on female non-persistence in major is consistent with the mechanism of poor grades causing negative updates in beliefs about major-specific ability. An ability-based mechanism does not fit for the positive effect of male peers on male persistence in majors. This points to men having some form of taste-based preference for majors associated to classes with more male peers, with increased likelihood of same-major friendship formation being one plausible explanation.

*Contact: sidsah@uchicago.edu. I thank Michael Dinerstein, Christina Brown, and Jack Mountjoy for their invaluable guidance. This paper would not have been possible without the assistance of Robert Dixon, Stephanie Estrada, Bill Hayward, Ben Ost, Gabriela Valencia, and others at the University of Illinois Chicago. I am grateful to Manashi Deshpande, Tom Hierons, Derek Neal, Evan Rose, and many seminar participants for useful comments and feedback.

1 Introduction

While women attend and graduate from college at comparable or higher rates than men, large gender differences persist in the take-up and completion of different college majors. Gender concentration within fields of study has important implications for both equity and efficiency. Given that men dominate many of the majors associated with the highest wages, such as engineering and economics, major choice may be an important contributing factor to the gender pay gap (Brown and Corcoran (1997), Gemici and Wiswall (2014), Patnaik et al. (2020)). Moreover, if there is reason to believe that gendered sorting to majors, and the occupations associated with those majors, does not reflect sorting to comparative advantage, this “friction" may dampen overall economic production (Hsieh et al. (2019)).

A substantial amount of work has gone into understanding the sources of gender differences in college major choice, with proposed factors ranging from high school preparedness (Card and Payne (2017), Aucejo and James (2021)), to preferences over non-pecuniary aspects of major-related occupations (Zafar (2013), Wiswall and Zafar (2017)), to competitiveness (Buser et al. (2014)), to the availability of role models (Carrell et al. (2010)). This paper considers whether peer gender composition within college classes contributes to this gap. In equilibrium, peer composition may reinforce the gendered sorting to fields that stems from other sources if students tend to choose majors related to the classes where they have more same-gender peers.

I investigate this question using administrative data from the University of Illinois Chicago (UIC), a large, public research university. The data provides information on all class registrations, class outcomes, major declarations, and semesters of graduation for two entering cohorts of undergraduate students. I focus on the classes that students enroll in during their first semester at the university based on the idea that incoming students have not yet met their peers and thus cannot coordinate or intentionally select into specific classes based on their gender compositions.

For my primary empirical strategy, I estimate how the gap between female and male outcomes evolves with the class-level gender ratio. I control for course-specific female fixed effects, utilizing variation in peer composition across different lecture times within courses. Doing so accounts for any potential gender differences in course-specific tastes or academic preparation. Given that I am estimating an effect on the *gap* between women and men, I am also able to include fixed effects for each specific lecture time in each semester for each course, in order to control for any gender-neutral sorting or shocks to these specific offerings of the courses. I find that when first-semester classes have more men, women tend to receive worse grades, are less likely to graduate, and are less likely to choose majors associated with those classes, all relative to men attending the same lectures. Placebo tests that use pre-college variables as outcomes show that these results are not driven by observable characteristics of students, such as pre-college academic attainment.

The aforementioned results characterize female outcomes *relative* to male ones. However, it is not immediately clear whether those patterns are driven by the behavior of women, the behavior of men, or both. By omitting the controls for specific iterations of courses, I am able to separately look at the effects of peer composition on male students and on female students. Generally, it seems as though women have worse achievement, in terms of grades and eventual graduation likelihood, in the presence of more men, while men are relatively unaffected by gender ratio for these outcomes. When it comes to major choice, however, it is the case both that women are less likely to declare majors related to male-heavy classes and that men are more likely to declare such majors. If anything, the effects on choice of major are driven mostly by male students.

I next consider what mechanisms might plausibly underlie my results on major choice. Theoretical work on major choice treats the decision as a dynamic problem in which students come into college facing uncertainty regarding their own field-specific abilities and tastes, learn about themselves during early classes, and then decide on a major (Arcidiacono (2004), Arcidiacono et al. (2012), Stinebrickner and Stinebrickner (2014), Arcidiacono et al. (2016)). Under such a framework, peer gender could enter into major choice decisions in two broad ways: students whose grades are affected by peer gender may update their beliefs about their field-specific ability or peer gender may influence beliefs about field-specific tastes. I find both that women receive worse grades in more male-heavy classes and that they are subsequently less likely to opt into majors related to those classes. This suggests that peer composition may affect female major choice indirectly through grades, although it is hard to rule out the possibility of tastes being an alternative or complementary mechanism. On the other hand, it is difficult to explain men's persistence in majors with male-dominated classes via beliefs about ability, as men do not receive better grades in more male-dominated classes.

If grades do not drive the positive effect of male peers on men's major choice, this implies that tastes play a role. I provide suggestive evidence that formation of within-major friendships may provide a plausible “taste-based" mechanism. Specifically, I consider students who enroll in both parts of various two-course sequences that are required for popular majors at UIC. When men have more male peers in their first class in one of these sequences, they tend to have more repeat peers from the first class in their second class in the sequence. It appears that this result reflects a behavioral effect, as the estimates I get for this test are large compared to a distribution of coefficients generated by simulations of random movement into classes. Thus, it may be that case that men form more friendships in male-heavy classes and then choose to study with those friends in the future. This may encourage the choice to declare majors related to those classes.

These results connect to an existing literature on the effects of peer gender in educational contexts. There has been a great deal of work focusing on primary and secondary school contexts. This strand of the literature has generally found that male peers are worse than female peers for the academic performance of all students, although whether this is found to matter more to girls or boys varies across studies and contexts (Hoxby (2000), Whitmore (2005), Lavy and Schlosser (2011), Black et al. (2013), Hu (2015), Gong et al. (2021)).

Some more recent work has considered the impact of peer gender in post-secondary education. Focusing on achievement, De Giorgi et al. (2010) find non-linear effects of class-level gender composition on grades for all students, with a roughly equal gender balance being optimal. Oosterbeek and van Ewijk (2014) find little evidence that the gender composition of workgroups within a class has any impact on outcomes. Hill (2017) uses cross-cohort variation at US universities, in the spirit of Hoxby (2000), to provide evidence that a higher proportion of females in an overall freshman cohort modestly increases male graduation rates. Looking at doctoral programs in STEM fields, Bostwick and Weinberg (2018) find that having more women in a cohort increases degree completion for other women.

The previous papers that are most comparable to this one are Griffith and Main (2019) and Zolitz and Feld (2020). Griffith and Main (2019) exploit random assignment to an introductory class for students entering an undergraduate engineering program at a US university. They find that having more female students in a class increases both grades and persistence beyond the first year of the program for male and female students. Zolitz and Feld (2020) similarly make use of random assignment in the context of compulsory, introductory classes at a Dutch business school. They find that having classes with more female peers increases the likelihood that both men and women select into majors that are more dominated by their own gender.

I contribute to this existing body of work by considering a context where students enter college outside of any particular program and can consider a full array of majors. This allows me to consider heterogeneity across academic fields: I find that both the negative effects of male peers on women and the positive effects of male peers on men are strongly concentrated among male-majority departments. This implies that peer gender effects may be particularly important for many STEM and many highly paid fields. I also explicitly consider whether the the short-term outcome of grades mediates effects on the longer-term outcome of major choice. This prompts my finding that grades do not appear to be the mechanism linking male peers to greater male persistence in majors and that increased friendship formation may be a plausible alternative.

The rest of the paper proceeds as follows. Section 2 describes the administrative data and provides institutional context regarding UIC. Section 3 describes my empirical strategy and presents the results of balance tests used to validate the strategy. Section 4 presents and discusses the results as well as various robustness checks. Section 5 concludes.

2 Data and Institutional Context

My analysis utilizes an administrative dataset from the University of Illinois Chicago, a large, public research university. UIC is one of three universities in the University of Illinois System, and enrolls approximately 21,000 undergraduate students per year. The data covers all undergraduate students who first enrolled at UIC in the Fall of 2015 or Fall of 2016 semesters, including transfers, for a total of 9,797 students.

My identification strategy relies on the assumption that incoming first-semester students do not sort into classes based on the peer composition of those classes, as they have not yet had a formal opportunity to meet their peers. At UIC, incoming students (both freshmen and transfers) are required to attend an on-campus orientation session prior to their first semester of classes. In both summer 2015 and 2016, students attended one of fourteen sessions offered over the course of the summer, registering for their preferred session on a first-come, first-serve basis. At the summer sessions, students met with academic advisors and received course recommendations based on their stated academic interests, prior credits (from either AP/IB examinations or previous college enrollments), and their performance on placement tests taken prior to the orientation. After this meeting, students were free to sign up for classes, subject to capacity constraints.

For the students in the dataset, I observe all class enrollments and outcomes, including grades and class withdrawals, for every semester the student is enrolled at UIC through six or seven years post-matriculation (for students entering in 2015 or 2016, respectively). Here and subsequently I use the term “class" to refer to a specific offering of a “course" that is uniquely identified by a particular semester, lecture time, and instructor. For instance, I would refer to Economics 101 as a “course" and the Fall of 2015, 9 A.M. lecture time for Economics 101 as a “class." I observe the name of the instructor for each class in the set of student-class observations, and, if the class has any associated discussion, laboratory, or experiential sections, the names of the teaching assistants (TAs) that lead those sections. While I do not directly observe characteristics of the instructors or TAs, I infer gender from names.[1] For each student in the sample, the administrative data provides information on several background characteristics, including race, ethnicity, gender, and pre-college academic achievement in the form of high school GPA and a composite ACT score. I also observe all major declarations for all students in the sample. Major information is by semester, meaning that I observe when each student first declares a major and if and when they switch majors.[2] If a student graduates from UIC during the covered period, the data also shows their semester of graduation.

As previously mentioned, my analysis focuses on incoming first-semester students. I thus restrict my main sample to student-class- level observations corresponding to classes

[1] I predict gender using the R package *gender*, which utilizes Social Security data (Mullen (2021)).

[2] At UIC, the large majority of students start out “undeclared" and first declare a major some time after their initial semester.

taken during a student's first semester. My primary estimating equation, equation (1), includes both course-specific female fixed effects and class fixed effects. Estimation of my parameter of interest - the differential effect of class-level male proportion on females compared to males - thus requires observations from courses with multiple classes, each of which has at least one female and one male student. I thereby eliminate observations from all courses and classes that do not meet this requirement from the main sample. This eliminates about 6.6 percent of first-semester observations. The remaining sample consists of 43,351 student-class observations.

Table 1 reports descriptive statistics for the main sample. Panel A reports background characteristics broken down by gender. Entering UIC, men and women look fairly similar, with comparable ethnic compositions and standardized test scores (the differences in means for most of these variables are statistically significant but small in magnitude, compared to the variation within groups). Female students have a statistically significantly advantage in mean high school GPA over their male peers, reflecting the pattern found in the overall population, although, again, the size of the gap is modest (a mean GPA of 3.35 for women compared to 3.22 for men). The overall sample has an average ACT composite score of approximately 24.5, which would place the mean student at approximately the 75th percentile of test-takers, indicating a student body that is academically above average among college-interested students.[3]

Panel B reports individual-level outcomes for the students in the main sample and panel C reports student-class-level outcomes and characteristics. Women tend to academically outperform men at UIC, being more likely to graduate within six years, earn higher grades, and pass their classes. However, despite being more likely to graduate in general, women are substantially less likely to graduate with a STEM major. Among students who graduate within six years, approximately 63% of male students graduate in STEM fields compared to about 49% of women, highlighting the gender differences in choice of field of study.[4]

Panel C also shows that about 40% of student-class observations belong to courses with an associated section, where "section" refers to any additional discussion, laboratory, or practical experience session, typically led by a TA. This subsample, with some additional sample restrictions, is used for a robustness check which relies upon variation in the proportion of male peers in sections within classes, rather than variation in classes within courses. About 20% of student-class observations are from courses for which only one lecture time is offered per year. This subsample is used for a robustness check where the variation in class-level peer composition is based largely on between-cohort variation.

[3] This is based on the 2022-2023 reporting year statistics from ACT, Inc: https://www.act.org/content/dam/act/unsecured/documents/MultipleChoiceStemComposite.pdf

[4] "STEM" is defined according to the 2022 Department of Homeland Security STEM Designated Degree Program List: https://www.ice.gov/doclib/sevis/pdf/stemList2022.pdf. The overall proportion of STEM graduates I observe is somewhat high in part due to the inclusion of certain majors that are not designated as STEM by other definitions, including psychology, economics, and certain pre-health majors.

Table 1 – Descriptive Statistics by Gender - Main Sample

	Women	Men	P-Value of Difference
Panel A: Student characteristics			
Ethnicities			
White	0.30	0.32	0.01
	(0.46)	(0.47)	
Asian	0.20	0.22	0.05
	(0.40)	(0.41)	
Hispanic	0.34	0.33	0.32
	(0.47)	(0.47)	
African American	0.10	0.07	0.00
	(0.30)	(0.25)	
High school GPA	3.35	3.22	0.00
	(0.37)	(0.39)	
ACT composite score	24.02	24.74	0.00
	(4.02)	(3.93)	
Panel B: Student outcomes			
Graduate within 6 Years	0.70	0.64	0.00
	(0.46)	(0.48)	
Graduate with a STEM major within 6 Years	0.34	0.40	0.00
	(0.47)	(0.49)	
Observations	4960	4594	
Panel C: Student-class outcomes and characteristics			
Grade (GPA value)	3.07	2.90	0.00
	(1.05)	(1.14)	
Grade of B or higher	0.81	0.78	0.00
	(0.39)	(0.41)	
Passed class	0.96	0.94	0.00
	(0.20)	(0.24)	
Dropped class	0.05	0.05	0.47
	(0.22)	(0.22)	
Course only offers one class per year	0.18	0.21	0.00
	(0.39)	(0.41)	
Class has an associated section	0.41	0.42	0.00
	(0.49)	(0.49)	
Observations	22162	21189	

Notes: This table reports the means and standard deviations of several variables, by gender, as well as the p-values of t-tests of differences in means across genders for each variable. All statistics are estimated on the observations within the identifying set of the main estimating equation (as described in Section 2) for which the relevant variable is observed. The variables used in Panels A and B are student-level while the variables in Panel C are student-class-level. The numbers of observations listed below Panel B also apply to Panel A. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.

More details on each of these alternate specifications is provided in Section 4.2.

3 Empirical Strategy

My aim is to estimate the differential effect of peer composition on the outcomes of women compared to men. I do so by exploiting variation in peer composition across classes within a course, where a "course" is defined by a department and course number, such as Economics 101, while a "class" is a specific offering of a course with a unique semester, lecture time, and instructor combination. I estimate how the difference in female and male outcomes evolves with the gender ratio across different offerings of the course. For each course, this may then involve variation between lecture times within a semester, say 9 AM compared to 11 AM, and may involve variation between students taking the course in the Fall of 2015 compared to the Fall of 2016.[5]

I focus solely on first-semester students. By doing so, I ensure that the peer composition of classes was not observable to students when they initially enrolled. Students register for first-semester classes in the summer prior to matriculation, before they have had a formal opportunity to meet their peers (other than the relatively small subset who attend their same orientation session). Thus, the students in my sample were not directly selecting into peer groups when they chose classes.

By comparing within courses, I allow men and women to differ in course-specific aptitudes or preferences, accounting for the fact that men and women may have received different kinds of education or may have formed dissimilar interests prior to entering college. Moreover, because my primary empirical strategy estimates how the *gap* between women and men changes with peer composition, I am also able to account for class-level fixed effects. The class fixed effects allow for any kind of gender-neutral sorting to classes, within courses, or class-specific shocks. For instance, if more motivated students tended to take morning classes, as opposed to afternoon ones, this would be picked up by the class fixed effects, so long as the sorting behavior was similar between the male and female populations.

The remaining threat to identification stems from the possibility of differential sorting between men and women to classes. For instance, if women were both more likely to register for morning classes *and* the difference between morning-class and afternoon-class women was larger than the difference between morning-class and afternoon-class men, this would bias my estimates. I argue that this concern is minimal using balance tests,

[5]For the results presented in the main body of the paper, I pool across both within-semester and across-semester variation in classes to maximize power. However, each of these two sources of variation introduces separate concerns for identification. I thus also present results using only within- and only across-semester variation. The version that is only within-semester is presented in Appendix A. The version that is only across-semesters is discussed in Section 4.2 and presented in Appendix B. Both sets of results are similar to my main results, although naturally less precise. This suggests that neither form of variation is solely driving the results.

which I describe later in this section. I also perform multiple robustness checks intended to minimize the extent to which differential sorting may drive my results. I describe these alternative approaches in greater detail in Section 4.2.

I operationalize my identification strategy via the following econometric model of student-class-level outcomes, $Y_{i,r,c}$:

$$Y_{i,r,c} = \alpha_0 + \alpha_1 \times Fem_i \times MP_c + \gamma_r \times Fem_i \times I_r + \delta_c \times I_c + X'_{i,r,c}\beta + u_{i,r,c} \tag{1}$$

where students are indexed by i, courses by r, and classes by c.[6] Fem_i is an indicator variable that takes a value of one if student i is female, while I_r and I_c are indicator variables that take on values of one if outcome $Y_{i,r,c}$ is associated with course r or class c, respectively. MP_c denotes the proportion of male students in class c, $X_{i,r,c}$ contains a vector of observable student and student-class observables, and $u_{i,r,c}$ is an unobservable error term.[7]

The parameter of interest, α_1, measures how the gap between female and male outcomes evolves with the class-level male proportion.[8] Given the focus on an interaction term, I am able to include both course-specific female fixed effects, γ_r, and general class fixed effects, δ_c.[9] The course-specific female fixed effects restrict the identifying variation to be within-course. The class fixed effects pick up any gender-neutral sorting or shocks associated with specific classes, within a course.

In order to identify my parameter of interest, I assume that the residual variation in $Fem_i \times MP_c$, conditional on the fixed effects and controls, is orthogonal to the residual variation in the error. Essentially, I assume that, within a course, students sort to classes in such a way that certain types of male or female students are not more likely to end up in more male-dominated classes.[10] The residual, identifying variation may come from a variety of sources, including capacity constraints on classes, student scheduling constraints, between-cohort variation in the numbers of men and women interested in each course, and idiosyncratic preferences for time slots. These sources may create noise

[6]Classes are nested within courses. Thus, any variable with a c subscript could alternatively be denoted with a double r,c subscript. I omit the course subscripts in these cases for readability.

[7]For regressions taking the form of equation (1), $X_{i,r,c}$ is generally composed of student underrepresented minority status (a dummy for Black, Hispanic, and native students) and instructor-student gender match. The specific set of controls used in each specification is described in the notes for the table reporting the corresponding results.

[8]Specifically, this parameter measures how the gap between female and male outcomes evolves *linearly* with class-level male proportion. Appendix C explores potential non-linearities in the effects. In general, it seems that both the absolute effects of male peers and the differences in effects of male peers on women compared to men are concentrated among the most male-dominated classes.

[9]The focus on a gap between groups of students and use of fixed effects bears some resemblance to Fairlie et al. (2014), although that paper makes use of a combination of individual fixed effects and class fixed effects.

[10]Here, "sorting" refers both to within-semester selections of a specific class time slot and to between-cohort variation. For the between-cohort variation, I still need to assume that students are not deciding whether or not to register for a *course* based on knowledge of the gender composition of that course in that year.

Table 2 – Associations Between Class Male Proportion and Pre-College Characteristics

	Standardized ACT score	High school GPA	Underrepresented minority student	Predicted grade (GPA value)
Female student X	0.026	-0.073	-0.045	-0.079
class male %	(0.110)	(0.057)	(0.058)	(0.070)
Outcome Mean	0.007	3.306	0.523	2.969
Outcome SD	1.000	0.392	0.499	0.410
Observations	31674	31702	43351	23018

Notes: This table reports the results of regressions of pre-college characteristics, at a student-class observation level, on an interaction between a female-student dummy and the class-level male proportion. Each column corresponds to a separate regression. Each regression includes course-specific female fixed effects, class fixed effects, and a control for instructor-student gender match. The predicted grade outcome is based on another (unshown) regression of the GPA value of grades on ACT score, high school GPA, and minority status among students in the main estimation sample. The predicted grade regression only includes the observations used to form the predicted grades: student-class observations from graded classes that had information on both ACT scores and high school GPA. Standard errors are clustered at the class level. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.

in the class-level peer composition, which I argue is uncorrelated with the differences in male and female characteristics.

My fixed effect strategy addresses many forms of potential endogeneity. As previously mentioned, the remaining threat to identification stems from the possibility of differential sorting between men and women to classes. While I cannot fully rule out the possibility of such differential sorting, I do test for it by estimating equation (1) for pre-college attributes that may be predictive of college outcomes: standardized ACT score, high school GPA, and underrepresented minority status.[11] I present the results of this exercise in Table 2.

The first three columns show that there is little systematic association between class male proportion and the differences in male and female pre-college academic aptitude or ethnicity, conditional on the full set of fixed effects. As it may be difficult to interpret the magnitudes of these estimates, I also form predicted course grades by regressing the GPA value of grades on the three pre-college attributes that I tested. I then use these background characteristic-predicted grades as an additional outcome, reported in the fourth column of Table 2. The point estimate suggests that a woman in a 100 percent male class would only be expected to receive a grade that is worth 0.08 fewer GPA points than a women in a 0 percent male class, based on observable characteristics. Along with not being statistically significant, this estimate is absolutely small, less than one tenth of the difference between an A and a B, and relatively small compared to the estimated effects of peer gender on grades that I report in Section 4.1. I interpret this as evidence in favor of

[11]Although I use high school GPA and ACT scores as outcomes in the placebo tests, I do not use them as controls in the specifications reported in the main body of the paper, as these variables are missing for a substantial portion of the sample. In Appendix D, I report results excluding any individual controls and results that include high school GPA and ACT scores as controls. The former set of results are nearly always very similar to those reported in the main text. The latter results generally have similar point estimates to those of my preferred specification, but are less precise due to the curtailed sample sizes.

Table 3 – Sorting by Students to Female Instructors (Within Course)

	Female student	Male student
Female instructor	0.000*	-0.000
	(0.000)	(0.001)
Outcome Mean	0.511	0.486
Observations	43351	43351

Notes: This table reports the results of regressions of dummies for student gender, defined at a student-class observation level, on a dummy for the class being taught by a female instructor. Each column corresponds to a separate regression. Each regression includes course fixed effects. Standard errors are clustered at the class level. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.

the necessary assumption of no differential sorting.

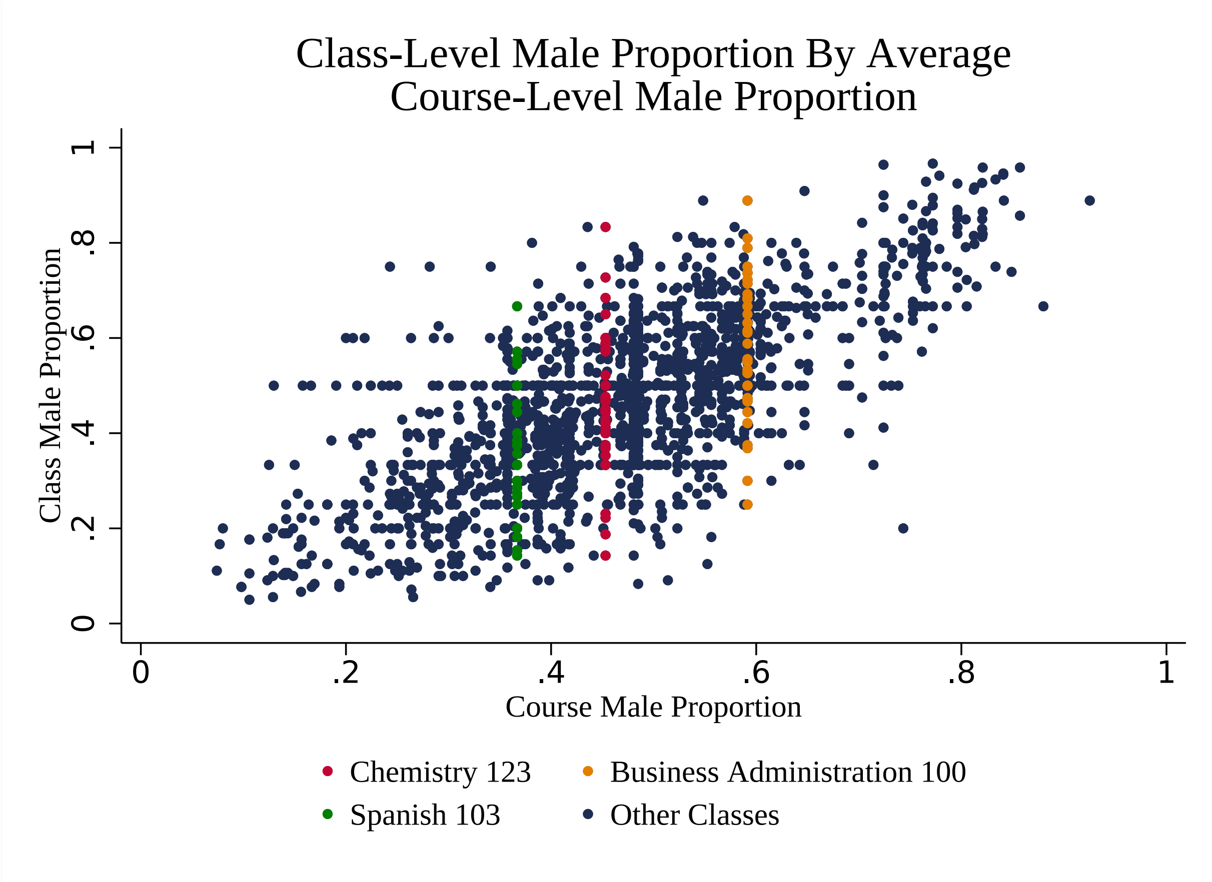
When selecting classes, the notable characteristics that students observe are the time period and, for most classes, the instructor. Thus, one particularly salient potential source of differential sorting is instructor gender. If students were more likely to sort into classes with same-gender instructors and students either performed better in the presence of same-gender instructors or only a certain type of student sorted into classes with same-gender instructors, this would bias my results.[12] I investigate this specific threat directly in Table 3, showing that, within courses, there is no meaningful degree of sorting into or out of female-taught classes by either female or male students. Nonetheless, I include a control for instructor-student gender match in all specifications taking the form of equation (1).

Given that identification is only possible via variation in class-level male proportion within courses, it is of interest how much such variation exists in the data. Figure 1 plots class-level male proportion against average course-level male proportion. While there is naturally a high degree of correlation between the two, there is meaningful variation between classes within course, particularly for courses that are closer to the center of the distribution and have many different observed classes. I plotted the three non-general requirement courses that have the most classes in separate colors, revealing that all three have classes spanning much of the range of possible gender ratios. This is the case in spite of the on-average male domination of Business Administration 100 and on-average female domination of Spanish 103.

The exact variation that I exploit is plotted in Figure 2. The blue line gives the raw distribution of class-level male proportion in the main estimation sample while the maroon line displays the residual variation conditional on course fixed effects. As would be expected, taking out course-level variation substantially condenses the distribution.

[12]There is recent literature suggesting that female instructors may improve the achievement of female college students (see for example Hoffman and Oreopoulos (2009) and Carrell et al. (2010)), although there are mixed findings regarding how female instructors affect the future course and major selections of female students (see for example Bettinger and Long (2005) and Price (2010)).

Figure 1



The remaining variation is concentrated within a span of about twenty percentage points.

3.1 Within-Gender Empirical Strategy

Estimating equation (1) can establish whether or not peer gender composition creates a separation in the outcomes of male and female students. However, it does not reveal whether this is driven by the outcomes of students of a particular gender or by the simultaneous behavior of both genders. In order to estimate how the female and male populations each separately respond to peer gender composition, I estimate several within-gender regressions, of the forms

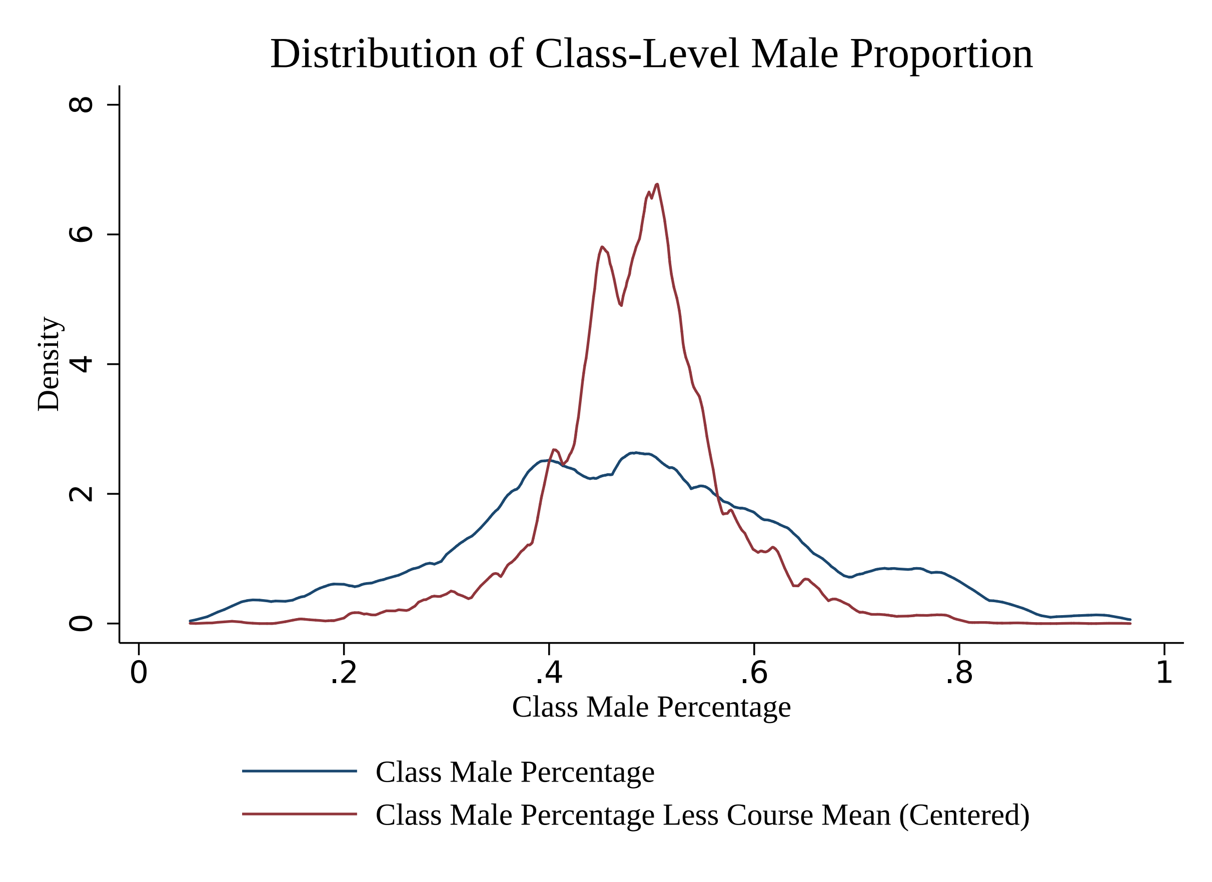
$$Y^F_{i,r,c} = \alpha^F_0 + \alpha^F_1 \times MP_c + \gamma^F_r \times I_r + (X^F)'_{i,r,c}\beta^F + u^F_{i,r,c} \tag{2}$$

$$Y^M_{i,r,c} = \alpha^M_0 + \alpha^M_1 \times MP_c + \gamma^M_r \times I_r + (X^M)'_{i,r,c}\beta^M + u^M_{i,r,c} \tag{3}$$

where equation (2) is estimated only on the set of female students and equation (3) is estimated only on the set of males.

By necessity, these specifications exclude class fixed effects. Thus, if there is any kind of absolute sorting to classes that have more male students among students of either gender, this will threaten the validity of my results. While this provides a weaker argument for

Figure 2



identification, I again perform the placebo test of putting pre-college characteristics (and the grades predicted by those pre-college characteristics) on the left-hand side of equations (2) and (3). The results of this test are presented in Table 4.

The placebo test shows no strong evidence of students of either gender sorting to more male-dominated classes, within course, in terms of observable background characteristics. Focusing on the predicted grade outcome, which has the most interpretable magnitude, the results indicate that moving from an all-female to an all-male class would be expected to shift a female student's grade by 0.035 GPA points and a male student's grade by 0.006 GPA points, based on the average association between class male percentage and student observables. These predictions are again small not only in an absolute sense but also relative to the estimated effect of male peers on female grades that I report and discuss in Section 4.3.

4 Results

4.1 Main Results

I now turn to the estimation of how the gap between female and male outcomes evolves with peer gender composition. Regression coefficients for the interaction between being a

Table 4 – Associations Between Class Male Proportion and Pre-College Characteristics, By Gender

	Standardized ACT score	High school GPA	Underrepresented minority student	Predicted grade (GPA value)
Only women:				
Class male %	0.007	-0.052	-0.063	-0.035
	(0.074)	(0.036)	(0.041)	(0.044)
Outcome Mean	-0.080	3.371	0.506	3.005
Outcome SD	1.005	0.379	0.500	0.411
Observations	16383	16393	22162	12166
Only men:				
Class male %	-0.116	-0.033	0.002	0.006
	(0.087)	(0.041)	(0.042)	(0.047)
Outcome Mean	0.102	3.235	0.542	2.929
Outcome SD	0.985	0.394	0.498	0.406
Observations	15291	15309	21189	10852

Notes: This table reports results of regressions of pre-college characteristics, at a student-class observation level, on class-level male proportion. Each cell corresponds to a separate regression with the outcome given by the column header. Top row results are estimated only on female students and bottom row results only on male students. Each regression includes class fixed effects and a control for instructor gender. The predicted grade outcome is based on another (unshown) regression of the GPA value of grades on ACT score, high school GPA, and minority status among students in the main estimation sample. The predicted grade regression only includes the observations used to form the predicted grades: student-class observations from graded classes that had information on both ACT scores and high school GPA. Standard errors are clustered at the class level. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.

female student and class-level male proportion on class and college outcomes are reported in Table 5. I report results from a variety of specifications with differing fixed effects. Results from my preferred specification, characterized by equation (1), are displayed in column 4. In addition, the table presents results when including no fixed effects (column 1), including course and course-specific female fixed effects but no class fixed effects (column 2), and class fixed effects but no course-specific female fixed effects (column 3). Controls that are made redundant by fixed effects are excluded from the relevant specifications. Standard errors are clustered at the class level.

I estimate each model for six outcomes, exploring a range of short- and long-term effects. As an immediate outcome, I consider the GPA point value of the grade received in the class. The intermediate outcomes include dummy variables for whether or not a student is observed to take any future classes in the same academic department as the given class, whether the student switches major to another department (conditional on having declared a major in the department of the given class in the first semester), and whether the student goes on to declare a major in the same department.[13] Because

[13]"Departments" are defined according to the UIC academic catalogue: https://catalog.uic.edu/ucat/degree-programs/degree-minors/. These departments can span multiple majors, although the number of majors is usually small. For instance, the English Department houses

Table 5 – Estimated Effect of Class Male Proportion on Student Outcomes

	(1)	(2)	(3)	(4)
Grade (GPA value) [Mean = 2.99]				
Female student X class male %	-0.539***	-0.125	-0.280***	-0.309**
	(0.110)	(0.143)	(0.090)	(0.151)
Observations	33071			
Take a future course in same department [Mean = .64]				
Female student X class male %	-0.013	-0.024	-0.118***	-0.067
	(0.068)	(0.044)	(0.038)	(0.047)
Observations	43351			
Switch major to another department [Mean = .14]				
Female student X class male %	0.372***	-0.034	0.054	0.031
	(0.056)	(0.111)	(0.064)	(0.146)
Observations	4750			
Declare a major in same department [Mean = .05]				
Female student X class male %	-0.018	-0.053**	-0.024	-0.047*
	(0.025)	(0.021)	(0.016)	(0.025)
Observations	37153			
Graduate within six years [Mean = .68]				
Female student X class male %	-0.208***	-0.092*	-0.022	-0.106*
	(0.038)	(0.052)	(0.033)	(0.056)
Observations	43351			
Graduate with a major in same department [Mean = .13]				
Female student X class male %	-0.417***	-0.047*	-0.003	-0.046*
	(0.075)	(0.025)	(0.024)	(0.028)
Observations	43351			
Fixed effects				
Course and course-female	No	Yes	No	Yes
Class	No	No	Yes	Yes
Controls				
Student gender	Yes	No	Yes	No
Instructor gender	Yes	Yes	No	No
Class male %	Yes	Yes	No	No

Notes: This table reports results of regressions of student-class outcomes on an interaction between a female-student dummy and class male proportion. Each cell corresponds to a separate regression, with outcome given by the row header and fixed effects and controls given by the column foot. Each regression includes controls for student underrepresented minority status and instructor-student gender match. All regressions are estimated on the observations within the identifying set of column (4) (as described in Section 2) for which the relevant outcome is observed. Standard errors are clustered at the class level. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.

only a small minority of students have declared majors in their first semester at UIC, the switching major outcome is estimated on only a small subset of students, and is consequently imprecise. The longest term outcomes are dummy variables for whether or not a student graduates within six years of enrollment and whether or not a student graduates with a declared major in the same academic department as the given class, again within six years of enrollment.

Column 1 of Table 5 reveals a general pattern of women having relatively worse academic performance and lower likelihood of persistence in more male-dominated academic departments. Introducing the full set of fixed effects in column 4 reveals the extent to which this pattern is explained by the causal effect of class-level peer gender composition. Starting with the immediate impact of male peers, the best estimate suggests that going from a class with no males to a class with all males would drive down the average female grade by a third of a GPA point, relative to men in the same class. This decrease is approximately one third of the difference between an A and a B.[14]

Considering other outcomes reveals that peer gender influences outcomes throughout the college career. My preferred specification suggests that having a first-year class with more men decreases the relative likelihood that women will go on to declare a major related to that class, graduate from college, and graduate with a major related to that class, all relative to male peers in the same class. All of these estimates are statistically significant and large relative to the mean likelihoods of these outcomes. These results suggest that peer gender composition influences the decision making of college students in ways that go beyond the impact within a specific class. Whether the longer-term outcomes are a direct function of the short-term grade outcome or not is a question that I will turn to in Section 4.5.

The results are generally similar, in terms of sign, across different specifications. Under the assumption that the results of the specification used in column 4 are the "truth," the high degree of similarity between those results and the column 2 results, which exclude class fixed effects, might be taken as evidence that the exclusion of class fixed effects is not a major threat to identification. This is reassuring for the interpretation of the within-gender results, which are estimated without class fixed effects and are presented in Section 4.3.

4.2 Robustness Checks

The primary threat to the validity of the results presented in the prior section comes from the possibility of differential sorting, whereby the difference between men and

both the English major and the Teaching of English major.

[14]Greater context on the magnitudes of the estimates is provided in Section 4.3, discussing the implied effect sizes given the range of actually observed class-level male proportions. This discussion is postponed as it is easier to think about the magnitudes of effects on students of each gender than the magnitude of the effect on the gap between genders.

women in a class is systematically related to the class gender ratio, within a course. In order to assuage these concerns, I report results from two sets of alternative specifications that may offer stronger arguments against differential sorting. The first set, reported in Appendix B, restricts to the set of courses for which only one class is offered per year. That is, if Economics 101 only had one lecture time in each of Fall of 2015 and Fall of 2016, Economics 101 would be included in this subset. For these courses, students in a given cohort do not have the ability to select between class offerings in ways that may be correlated with class gender makeup. The only way students could react to class characteristics so as to create differential sorting would be on the extensive margin of whether or not to take the course at all (during their first semester). Under the assumption that year-specific class characteristics do not have a large impact on course take-up, this specification largely relies on cohort-level variation. The results using this subsample of courses are generally similar to my main results, although less precise.

The other alternative specification focuses on peer gender composition in laboratory, discussion, or practical experience sections that are associated with classes. Because many classes have multiple associated sections, using section-level variation allows for the inclusion of class-specific female fixed effects, which would account for any kind of differential sorting to classes between men and women. The sorting to sections within classes might be considered “more idiosyncratic" than the higher level sorting to classes, as sections have tighter capacity constraints, may have less observable information than classes because some sections do not provide information on the TA in charge (and TAs may have less information available about them than faculty), and students may prioritize class selections over section selections, resulting in sections being subject to greater scheduling constraints (if class times are chosen “first" by students). The section-level analysis provides some supporting evidence for the results reported in the main paper, although the results are generally not statistically significant in the overall sample. Interestingly, it seems that section-level peer effects may be strongly concentrated in sections for female-majority classes. Appendix E further discusses the section-level analysis and reports the results.

4.3 Results Within Gender

I now consider whether the effects I find in Section 4.1 are driven more by the behavior of female or male students. Estimates of how class-level male proportion affects outcomes within gender, using regressions of the forms of equations (2) and (3), are presented in Table 6. It appears that having more males in a class may be harmful to both the grades received and the likelihood of taking a future course in the same department for female students. For male students, the estimated effects are negative, but small in magnitude and not statistically significant. The divergence between the two groups thus stems from the fact that women are more negatively affected by male peers than men are, rather than

Table 6 – Estimated Effect of Class Male Proportion on Student Outcomes, By Gender

	Grade (GPA value)	Future course in dept.	Switch major out of dept.	Declare major in dept.	Graduate within 6 years	Graduate in dept.
Only women:						
Class male %	-0.171	-0.071**	-0.064	-0.021	-0.070**	-0.005
	(0.106)	(0.029)	(0.065)	(0.016)	(0.034)	(0.018)
Outcome Mean	3.068	0.663	0.141	0.060	0.707	0.135
Outcome SD	1.051	0.473	0.348	0.237	0.455	0.341
Observations	17284	22162	2537	18977	22162	22162
Only men:						
Class male %	-0.046	-0.047	-0.027	0.031**	0.022	0.042**
	(0.129)	(0.036)	(0.095)	(0.015)	(0.039)	(0.019)
Outcome Mean	2.901	0.625	0.135	0.049	0.647	0.119
Outcome SD	1.143	0.484	0.341	0.216	0.478	0.324
Observations	15787	21189	2213	18176	21189	21189

Notes: This table reports results of regressions of student-class outcomes on class-level male proportion. Each cell corresponds to a separate regression, with outcome given by the column header. Top row results are estimated only on female students and bottom row results only on male students. Each regression includes course fixed effects and controls for student underrepresented minority status and instructor gender. All regressions are estimated on the observations within the identifying set of the main specification (as described in Section 2) for which the relevant outcome is observed. Standard errors are clustered at the class level. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.

male students providing an academic benefit to one another. This finding is consistent with studies from both primary and secondary school contexts, which have generally found that more male-heavy academic environments are worse for the academic outcomes of all students, compared to more female-heavy ones (Hoxby (2000); Lavy and Schlosser (2011); Gong et al. (2021)). A variety of mechanisms have been proposed for this pattern, including disruptive behavior (Lavy and Schlosser (2011)) and teacher responses to class composition (Gong et al. (2021)). In the context of UIC, it is the case that that the male population has a lower average high school GPA than the female population. Thus, it might be suspected that the observed effect of male peers is really a function of low-ability peers. However, controlling for peer ability, in the form of average high school GPA, only increases the estimated negative effect of males on grades, for both men and women.[15] To the extent that ability is well captured by this measure, it appears that the effect of male students on achievement is not driven by ability and likely reflects other attributes of males or male-dominated environments.

There is a divergence in the signs of the effects of male peers on men versus women for longer term outcomes. In spite of the result on grades, it appears that, if anything, having more male peers makes male students more likely to graduate within six years. Female students, on the other hand, are significantly less likely to graduate when they

[15] Within-gender results including the peer ability control are reported in Appendix F.

have more male peers in their first-semester classes. As with grades, another achievement outcome, the effect on the gap between men and women appears to be driven primarily by a negative effect of male peers on female students.

Results on choice of major, however, appear to be driven more by men. While female students may be somewhat less likely to declare majors corresponding to their more male-dominated classes, there is a larger, positive effect of male peers on the likelihood of male students choosing a given major. Similarly, I estimate a null effect of male peers on female likelihood of graduating with a related major, contrasting with a significant, positive effect on the same outcome for male students. Taking these results together suggests that male students harm the academic achievement of female students while having more ambiguous effects on the achievement of other males. However, it seems that the presence of male students encourages other male students to persist in majors while having a more moderate impact on the choices of female students.

The estimates in Table 6 are linear in class-level male proportion. Thus, the reported numbers compare outcomes between classes with no men and classes with all men, which is an extreme comparison given the distribution of classes students are likely to take. Due to the inclusion of course fixed effects, these results are estimated using only within-course variation. In the estimation sample, going from a class at the 10th percentile of class male proportion to the 90th percentile, within a course, would correspond to a shift in the class-level male proportion of about 0.2.[16] The estimates in Table 6 would thus imply that going from a 10th percentile class to a 90th percentile class, within a course, would decrease female grades by 0.03 GPA points, reduce female likelihood of graduation by 1.4%, and increase the likelihood of men graduating in a related department by 0.8%, on average.

Of course, when considering the full range of classes both within *and* across courses, students are exposed to more extreme variation in peer composition. Across all classes in the estimation sample, going from a a class at the 10th percentile of class male proportion to the 90th percentile would imply a change in class male proportion of 0.46. If my estimates are externally valid to how peer composition affects outcomes when comparing any two classes, rather than just two classes within a course, then going from a 10th percentile class to a 90th percentile class would be expected to decrease female grades by 0.08 GPA points and increase the likelihood of men graduating in a related department by 1.9%. However, it is difficult to gauge the extent to which my estimates may be valid for these kinds of comparisons.

The results presented in Tables 5 and 6 are also aggregated across courses from all academic departments at UIC. It may be of interest how the effects vary across different departments. Appendix G presents results broken down by whether a class is in a male-majority or female-majority department. The results indicate that both the negative effects of male peers on women and the positive effects of male peers on men are strongly

[16]The within-course variation in class male proportion is plotted in Figure 2.

concentrated among male-majority departments. Indeed, it appears that there is little if any effect of peer gender in classes within female-majority departments. Given that most STEM majors and most highly paid majors are male-dominated, this suggests that peer gender effects may be particularly relevant to policy-makers who care about increasing female representation in STEM or about the role of major choice in perpetuating the gender pay gap.

4.4 Relationship Between Short- and Long-Term Outcomes

My analysis has thus far considered a range of outcomes that span a student's college career, ranging from the immediate outcome of grade in a first-semester class to later outcomes like choice of major and graduation. It is interesting to consider how the short- and long-term outcomes interact. In a dynamic model of major choice, as in Arcidiacono (2004), students enter college considering multiple majors while facing uncertainty about both their major-specific ability levels and their major-specific tastes, learn through experimentation, and eventually make a final major choice. First-semester classes may provide important information about both ability and tastes in a way that could be influenced by peer composition. The preceding results suggest that male peers affect grades. If a student, naive to the influence of peer gender on grades, receives a poor grade in a male-heavy class for a given major, they may negatively update their belief about their ability in that major. Peer gender may also influence beliefs about tastes in multiple ways. If students care directly about major-level gender composition, for instance if they dislike being in a gender minority, they may update their beliefs about the overall gender composition of a major based on the gender composition of the first class for that major. Even if a student does not have explicitly think about gender composition, peer gender may influence her beliefs about taste for a major if class gender composition influences classroom environment or how many friends she makes in a class. Thus, peers may influence major choice both indirectly through the impact of grades on beliefs about ability or directly through beliefs about taste.

Prior work finds that women are more likely to opt out of a major in response to a poor grade than men, which might suggest that women would be more susceptible to an indirect effect of male peers on major choice through grades (Rask and Tiefenthaler (2008), Ahn et al. (2019)). The within-gender results, presented in Table 6, could be seen as broadly concordant with these findings. Women appear to be more likely to receive a bad grade in a class with more male peers and are subsequently less likely to take future courses in the same department and may be modestly less likely to declare a major within that department. Men, on the other hand, experience little to no effect of male peers on grades, but are more likely to persist in a major when exposed to more male peers. For women it thus seems plausible that any peer effects on major choice flow through grades and beliefs about abilities, although it is not possible to rule out beliefs about tastes as a complementary or alternative mechanism. For men it seems as though effects on major choice must come from a channel other than grades and their impact on beliefs about ability. In order to further tease apart how peer gender affects major choice, I consider the effects of peer composition in finer subsets of the overall sample.

I first consider effect heterogeneity by ability level, as measured by high school GPA, with the results reported in Table 7. Looking at the effects on grades, it appears as though the negative effect of male students on grades is concentrated among the lower half of the ability distribution, for both women and men. For women, effects on major choice and graduation are also concentrated in the lower half of the ability distribution. Women who face stronger effects on grades also having stronger effects on major choice is consistent with the notion that the latter effect is a function of the former. For men, however, the positive effect of male peers on major choice is concentrated in the lower half of the ability distribution. Thus, the male students who are harmed more by male peers in terms of grades are also more likely to persist in majors where they have more male peers. This pattern would be surprising if grades are the mechanism driving the effect of male peers on male major choice.

I also directly consider how longer-term outcomes are mediated by the intermediate grade outcome. Specifically, I compare effects between the subset of students who received an A and those who received a worse grade. Given that nearly half of first-semester grades in the sample are As, it might be reasonable to think that UIC students would consider an A to be a "good" grade and anything else to be "bad."[17] As would be predicted if beliefs about ability was an important mechanism, women who receive As exhibit little if any effect of male peers on major choice, while women who receive worse grades in a class appear to avoid the major associated with that class. For men, on the other hand, the positive effect of male peers on major choice appears to be concentrated among men who received "bad" grades. This further suggests that male peers do not influence male major choice decisions indirectly through beliefs about ability, as men receiving inferior signals about their major-specific ability sort into the majors associated with their male-heavy classes at a higher rate.

4.5 Evidence on the Effect of Class Gender Composition on Friendship Formation

For women, it is plausible that if peer gender influences major choice, it does so indirectly through grades, although it is not possible to rule out the existence of other mechanisms. For men it does not seem that grades are an intermediary connecting peer gender to major choice. It thereby remains to consider other ways in which peer gender may influence male major choice. Prior work has found that having at least one peer choose a given major may increase the likelihood of take-up of that major, with one potential

[17] Approximately 41% of grades in the main estimation sample are As.

Table 7 – Estimated Effect of Class Male Proportion on Student Outcomes, By Gender and Pre-College Academic Ability

	Grade (GPA value)	Future course in dept.	Switch major out of dept.	Declare major in dept.	Graduate within 6 years	Graduate in dept.
Women with above average HS GPAs:						
Class male %	-0.182	-0.122***	-0.195	0.006	-0.037	-0.004
	(0.132)	(0.040)	(0.259)	(0.018)	(0.046)	(0.021)
Outcome Mean	3.298	0.668	0.394	0.053	0.759	0.060
Outcome SD	0.927	0.471	0.489	0.224	0.428	0.238
Observations	6897	9260	386	8750	9260	9260
Men with above average HS GPAs:						
Class male %	-0.110	0.075	0.276	0.018	-0.054	0.042*
	(0.178)	(0.056)	(0.556)	(0.023)	(0.062)	(0.024)
Outcome Mean	3.237	0.650	0.309	0.042	0.736	0.059
Outcome SD	0.961	0.477	0.463	0.200	0.441	0.235
Observations	4612	6529	314	6101	6529	6529
Women with below average HS GPAs:						
Class male %	-0.278	0.009	-0.400	-0.030	-0.131**	0.023
	(0.193)	(0.052)	(0.401)	(0.026)	(0.065)	(0.027)
Outcome Mean	2.748	0.600	0.349	0.052	0.523	0.051
Outcome SD	1.175	0.490	0.478	0.222	0.500	0.220
Observations	5283	7133	232	6701	7133	7133
Men with below average HS GPAs:						
Class male %	-0.188	-0.035	-0.677	0.030	0.054	0.051**
	(0.214)	(0.055)	(0.439)	(0.020)	(0.066)	(0.020)
Outcome Mean	2.593	0.589	0.282	0.043	0.495	0.050
Outcome SD	1.218	0.492	0.451	0.203	0.500	0.218
Observations	6269	8780	301	8210	8780	8780

Notes: This table reports results of regressions of student-class outcomes on class-level male proportion. Each cell corresponds to a separate regression, with outcome given by the column header and the subset of students used for estimation given by the row header. Each regression includes course fixed effects and controls for student underrepresented minority status and instructor gender. All regressions are estimated on the observations within the identifying set of the main specification (as described in Section 2) for which the relevant outcome is observed. Standard errors are clustered at the class level. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.

Table 8 – Estimated Effect of Class Male Proportion on Student Outcomes, By Gender and Realized Course Grade

	Future course in dept.	Switch major out of dept.	Declare major in dept.	Graduate within 6 years	Graduate in dept.
Women who received As:					
Class male %	-0.043	-0.045	-0.019	0.010	-0.015
	(0.041)	(0.072)	(0.027)	(0.048)	(0.031)
Outcome Mean	0.714	0.102	0.082	0.833	0.182
Outcome SD	0.452	0.302	0.274	0.373	0.386
Observations	7454	1073	6231	7454	7454
Men who received As:					
Class male %	0.047	0.096	-0.000	0.001	0.023
	(0.059)	(0.103)	(0.025)	(0.054)	(0.030)
Outcome Mean	0.715	0.083	0.068	0.817	0.190
Outcome SD	0.451	0.276	0.253	0.387	0.392
Observations	5874	949	4731	5874	5874
Women who did not receive As:					
Class male %	-0.060	0.071	-0.059**	-0.017	-0.052*
	(0.049)	(0.116)	(0.027)	(0.059)	(0.029)
Outcome Mean	0.649	0.132	0.051	0.666	0.135
Outcome SD	0.477	0.339	0.220	0.472	0.342
Observations	9830	1180	8273	9830	9830
Men who did not receive As:					
Class male %	-0.036	0.145	0.051**	0.072	0.031
	(0.054)	(0.144)	(0.024)	(0.062)	(0.028)
Outcome Mean	0.634	0.145	0.049	0.594	0.111
Outcome SD	0.482	0.352	0.216	0.491	0.314
Observations	9913	973	8499	9913	9913

Notes: This table reports results of regressions of student-class outcomes on class-level male proportion. Each cell corresponds to a separate regression, with outcome given by the column header and the subset of students used for estimation given by the row header. Each regression includes course fixed effects and controls for student underrepresented minority status and instructor gender. All regressions are estimated on the observations within the identifying set of the main specification (as described in Section 2) for which the relevant outcome is observed. Standard errors are clustered at the class level. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.

explanation being the direct utility value of studying with a friend (De Giorgi et al. (2010)). If men are more likely to form friendships with other men, then having a more male-dominated first class in a major may increase the expected number of friends who are interested in the same major. This could then translate into an increased willingness to declare and persist in that major.

While I cannot directly observe friendships in the data, I provide suggestive evidence using taking classes together as a proxy. Specifically, I consider all two-course sequences that are required for at least one of the twenty most popular majors at UIC and are commonly taken by first year students. A two-course sequence is defined as a set of two courses with one being a pre-requisite for the other. For instance, General Chemistry I and General Chemistry II is a two-course sequence that is required for the Biology and Chemistry majors, among others, and is recommended as first-year coursework for students interested in those majors. For students who take the first course in one of these sequences during their first semester and take the second course in a later semester, I define as outcomes the number of peers, of either gender, who were in the same class for both the first and second course and the number of peers who were in the same laboratory or discussion section for both the first and second course.[18]

I estimate the association between class-level male proportion and the number of same-subsequent-class peers using equations (2) and (3). Similarly, I look at the link between section-level male proportion and same-subsequent-section peers using the estimating equations

$$Y^F_{i,r,c,s} = \eta^F_0 + \eta^F_1 \times MP_s + \theta^F_c \times I_c + \varepsilon^F_{i,r,c,s} \tag{4}$$

$$Y^M_{i,r,c,s} = \eta^M_0 + \eta^M_1 \times MP_s + \theta^M_c \times I_c + \varepsilon^M_{i,r,c,s} \tag{5}$$

where sections are indexed by s, MP_s denotes the proportion of male students in section s, η^g_1 measures how the number of repeated peers evolves with MP_s for students of gender g, θ^g_c are class fixed effects, and $\varepsilon^g_{i,r,c,s}$ is the error term. Equation (4) is estimated only on the sample of female students and equation (5) only on male students. I utilize eleven total sequences between the class and section analysis.[19]

[18]This necessarily means that students who only take the first course in a sequence and never take the second are excluded. It also means that each peer pair is "double-counted" in the sense that any pair of students taking the same class for both courses in a sequence will be reflected in the outcome variable of both students.

[19]For class-level results, I use all sequences for which there are at least two classes for the first course in the sequence in both Fall 2015 and Fall 2016: General Chemistry I and General Chemistry II; Introduction to Psychology and Introduction to Research in Psychology; Introduction to UIC and Professional Development and Business Professional Development II; Calculus I and Calculus II; Calculus II and Calculus III; General Physics I and General Physics II; Principles of Microeconomics and Microeconomics: Theory and Applications; Principles of Macroeconomics and Macroeconomics in the World Economy: Theory and Applications; and Biology of Cells and Organisms and Biology of Populations and Communities. For the section-level results, I use the same set of sequences, excluding Introduction to UIC and Professional Development and Business Professional Development II as they have no associated sections and including Program Design

It is almost certainly the case that some students will enroll in the same class twice in a row by chance. However, male students being systematically more likely to take future classes with the same peers when their initial class has more men may represent an increased likelihood of forming friends and coordinating future enrollment with them.

The results of this exercise are presented in Table 9. While the estimated effects of peer composition on the number of same-subsequent-class peers are imprecise, the results suggest that more male-dominated classes tend to result in more continued peers in the next class for both men *and* women. Similarly, having more men in the section of a first class increases the likelihood of having any same-subsequent-section peers for both men and women, even within a given class. The observed results may reflect homophily in friendship formation among men with some complementary explanation for women, some greater overall degree of friendship formation in more male-dominated environments, or other explanations. In any case, if men are more likely to form same-major friends in more male-dominated environments, this offers one plausible mechanism for greater male persistence in majors associated to male-dominated classes, in spite of male peers seemingly not providing academic benefits to male students.

Table 9 reports statistical significance based on the null hypotheses of coefficients equalling zero. However, it is not clear that the absence of intentional sorting to classes with prior classmates implies a zero coefficient. For instance, the point estimates in Table 6 would suggest that male peers encourage male persistence in major to a greater extent than they discourage female persistence. Taking these coefficients literally would imply that more male-heavy classes would have greater net persistence into related majors, which could mechanically generate the patterns seen in Table 9 (if students are more likely to take the second course in a sequence if they are also persisting in a related major). To address this, I benchmark the estimates I get from the data against simulated coefficients.

For the simulations, I use the same sample of students used for estimation in Table 9. I leave fixed their first-course enrollments and simulate fully random movement into classes for the second course in each sequence, preserving the observed sizes of the second classes. That is, if a student is observed taking the Fall 2015, 9 AM General Chemistry I and the Spring 2016, 9 AM General Chemistry II, in a simulation, I “leave" them in the Fall 2015, 9 AM General Chemistry I but randomly “place them" in a class for General Chemistry II, such that each simulated General Chemistry II class has the same number of students as is observed in the data. I then estimate the same regression models reported in Table 9 using the simulated data. This exercise fully conditions on which students choose to take both courses in a sequence. The comparison between my estimates from the data and the estimates from the simulations is thus based only on selection of class, conditional on choosing to complete the sequence.

I and Program Design II and Introduction to Criminology, Law, and Justice and Foundations of Law and Justice as, although there is not sufficient variation in these sequences for class-level analysis, there is enough for section-level analysis.

Table 9 – Estimated Effect of Class and Section Male Proportion on the Number of Peers Taking Future Classes Together

	Number of peers in the same subsequent class	Number of peers in the same subsequent section
Only women:		
Class male %	4.738	-
	(3.679)	-
Section male %	-	0.362*
	-	(0.218)
Outcome Mean	9.278	0.348
Outcome SD	10.665	0.754
Observations	1027	880
Only men:		
Class male %	4.394	-
	(2.654)	-
Section male %	-	0.306
	-	(0.207)
Outcome Mean	8.459	0.550
Outcome SD	9.442	0.970
Observations	1319	1179

Notes: The first column of this table reports results of regressions of number of peers from a student's first class in a two-course sequence who take the same second class on class male proportion, conditional on course fixed effects. The second column reports results of regressions of number of peers from a student's first section in a sequence who take the same second section on section male proportion, conditional on class fixed effects. Each cell corresponds to a separate regression. All regressions are estimated on the set of students who take the first course in one of the listed two-courses sequences during their first semester and take the second course in a subsequent semester. Top panel regressions are only estimated on female students and bottom panel regressions only on male students. Standard errors are clustered at the first-class level. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.

The results are presented in Figure 3, with the distributions of simulated coefficients (from 1,000 simulations) appearing in blue and the estimates from the original data appearing in red. It appears that, for both women and men, random enrollments for second courses would not yield coefficients of zero on class-level male proportion. However, it is clear that the estimate magnitudes I observe in the data are unlikely to occur based on purely random enrollment. Focusing on the results for men, the class-level coefficient lies at the 98th percentile of simulated coefficients and the section-level coefficient lies at the 96th percentile. It is thus reasonable to conclude that men (and women) are more likely to take classes and sections with former classmates when their initial classes are more male-heavy, even fully accounting for the extensive margin of course selection. This may reflect a higher degree of friendship formation in more male-dominated classes, which would provide a plausible explanation for the positive effect of male peers on male persistence in majors.

Figure 3

5 Conclusion

Gender differences in college major take-up and completion are of interest to policy-makers, largely due to the substantial labor market implications. I investigate the role of peer composition in driving gender differences in college student achievement and major choice using a large administrative dataset that provides information on two cohorts of students at the University of Illinois Chicago over the courses of their college careers. I focus on incoming first-semester students, who enroll for courses without knowledge of peer enrollments, in order to minimize the risk of intentional sorting into peer groups. A rich set of fixed effects accounts for gender differences in performance at a course level, utilizing variation in lecture times within a course, and accounts for gender-neutral sorting or shocks associated with specific class times. I find that when a class has more male students, women receive worse grades, are less likely to declare majors associated with that class, and are less likely to graduate, all relative to men attending the same lectures. My identification strategy is supported by balance tests showing that I am comparing across classes that have very similar students in terms of observable characteristics. The findings are further validated by two robustness checks: one only uses cross-cohort variation to avoid issues of sorting to different classes within a semester and the other allows for arbitrary gendered sorting to specific classes and instead focuses on variation across TA-led sections within a class.

The aforementioned results concern how men and women diverge based on peer composition. I further estimate student responses by gender in order to determine whether men or women are responding more to peer composition. I find that women receive worse grades and are less likely to eventually graduate when they have more male peers, while male achievement is not substantially affected by peer gender. On the other hand, when a first-semester class has more men, men are more likely to declare a related major, while the effect on women's major choice is negative but more modest in magnitude than the effect on men.

In a simple model of college major choice, as in Arcidiacono (2004), we can think of students as entering college with uncertainty about their major-specific abilities and tastes, learning about themselves during classes taken early in college, and making major decisions based on what they learned. The fact that I observe women receiving worse grades in male-dominated classes and subsequently being less likely to pursue related majors implies that male peers may affect women's major choice via beliefs about ability, although it is impossible to rule out taste as a mechanism. However, the lack of a positive effect of male peers on men's grades makes it unlikely that the positive effect of male peers on men's major choice is explained by beliefs about ability. One alternative explanation could be that men form more friends in more male-dominated classes, which encourages them to continue taking courses in the same field in order to continue studying with those friends. I find that men in more male-dominated classes take future classes with more repeated peers than men in less male-dominated classes, suggesting that friendship formation is a plausible mechanism for the effect of male peers on male major choice. Future work would do well to further consider mechanisms of peer gender on student outcomes, potentially leveraging different kinds of data to directly study friendship networks.

My results suggest that increasing the proportion of students of a given gender in a class yields positive results for the other students of the same gender. One policy implication could be that any policy that increases the representation of gender-minority students in an academic field will have a greater than anticipated impact. The presence of more gender-minority students should encourage other students of that gender to perform well and persist in the field, creating a total effect larger than the direct impact of the policy. While my estimates are necessarily based only on the relatively modest amount of variation in gender composition across classes within a course, future work on the impact of larger differences in gender composition would be useful. This would be informative about how much peer effects may matter for the efficacy of policies that change class gender ratios by large amounts.

References

Ahn, T., Arcidiacono, P., Hopson, A., and Thomas, J. R. (2019). Equilibrium grade inflation with implications for female interest in stem majors. *NBER Working Papers*.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics*, 121(1-2):343–375.

Arcidiacono, P., Aucejo, E., Maurel, A., and Ransom, T. (2016). College attrition and the dynamics of information revelation. *NBER Working Papers*.

Arcidiacono, P., Hotz, V. J., and Kang, S. (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, 166:3–16.

Aucejo, E. and James, J. (2021). The path to college education: The role of math and verbal skills. *Journal of Political Economy*, 129(10):2721–2993.

Bettinger, E. P. and Long, B. T. (2005). Do faculty serve as role models? the impact of instructor gender on female students. *American Economic Review*, 95(2):152–157.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2013). Under pressure? the effect of peers on outcomes of young adults. *Journal of Labor Economics*, 31(1):119–153.

Bostwick, V. K. and Weinberg, B. A. (2018). Nevertheless she persisted? gender peer effects in doctoral stem programs. *NBER Working Papers*.

Brown, C. and Corcoran, M. (1997). Sex-based differences in school content and the male-female wage gap. *Journal of Labor Economics*, 15(3):431–465.

Buser, T., Niederle, M., and Oosterbeek, H. (2014). Gender, competitiveness, and career choices. *The Quarterly Journal of Economics*, 129(3):1409–1447.

Card, D. and Payne, A. A. (2017). High school choices and the gender gap in stem. *NBER Working Papers*.

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

De Giorgi, G., Pellizzari, M., and Redaelli, S. (2010). Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics*, 2(2):241–275.

Fairlie, R. W., Hoffman, F., and Oreopoulos, P. (2014). A community college instructor like me: Race and ethnicity interactions in the classroom. *The American Economic Review*, 104(8):2567–2591.

Gemici, A. and Wiswall, M. (2014). Evolution of gender differences in post-secondary human capital inverstments: College majors. *International Economic Review*, 55(1):23–56.

Gong, J., Lu, Y., and Song, H. (2021). Gender peer effects on students' academic and noncognitive outcomes. *Journal of Human Resources*, 56(3):686–710.

Griffith, A. L. and Main, J. B. (2019). First impressions in the classroom: How do class characteristics affect student grades and majors? *Economics of Education Review*, 69:125–137.

Hill, A. J. (2017). The positive influence of female college students on their male peers. *Labour Economics*, 44:151–160.

Hoffman, F. and Oreopoulos, P. (2009). A professor like me: The influence of instructor gender on college achievement. *The Journal of Human Resources*, 44(2):479–494.

Hoxby, C. (2000). Peer effects in the classroom: Learning from gender and race variation. *NBER Working Papers*.

Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2019). The allocation of talent and u.s. economic growth. *Econometrica*, 87(5):1439–1474.

Hu, F. (2015). Do girl peers improve your academic performance? *Economics Letters*, 137:54–58.

Lavy, V. and Schlosser, A. (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, 3(2):1–33.

Mullen, L. (2021). *gender: Predict Gender from Names Using Historical Data*. R package version 0.6.0.

Oosterbeek, H. and van Ewijk, R. (2014). Gender peer effects in university: Evidence from a randomized experiment. *Economics of Education Review*, 38:51–63.

Patnaik, A., Wiswall, M. J., and Zafar, B. (2020). College majors. *NBER Working Papers*.

Price, J. (2010). The effect of instructor race and gender on student persistence in stem fields. *Economics of Education Review*, 29:901–910.

Rask, K. and Tiefenthaler, J. (2008). The role of grade sensitivity in explaining the gender imbalance in undergraduate economics. *Economics of Education Review*, 27(6):676–687.

Stinebrickner, R. and Stinebrickner, T. R. (2014). A major in science? initial beliefs and final outcomes for college major and drop. *Review of Economic Studies*, 81:426–472.

Whitmore, D. (2005). Resource and peer impacts on girls' academic achievement: Evidence from a randomized experiment. *American Economic Review*, 95(2):199–203.

Wiswall, M. and Zafar, B. (2017). Preference for the workplace, investment in human capital, and gender. *The Quarterly Journal of Economics*, 133(1):457–507.

Zafar, B. (2013). College major choice and the gender gap. *The Journal of Human Resources*, 48(3):545–595.

Zolitz, U. and Feld, J. (2020). The effect of peer gender on major choice in business school. *IZA Institute of Labor Economics Discussion Paper Series*.

Appendix A - Only Utilizing Within-Semester Variation

The results presented throughout the main body of the text utilize variation in classes within courses both between the Fall of 2015 and the Fall of 2016 and across lecture times within each semester. However, each form of variation may introduce separate identification concerns. I present here results that only utilize within-semester variation. Appendix B presents results for only across-semester variation. For the results here, I estimate equation (1), but instead of using course-specific female fixed effects, I use course-by-year-specific female fixed effects. That is, in lieu of a fixed effect for women in Econ 101, I include one fixed effect for women in Econ 101 in Fall of 2015 and another fixed effect for women in Econ 101 in Fall of 2016. The results are presented in Table 10. The point estimates are quite similar to what is seen in Table 5 in the main text, although they are naturally somewhat less precise.

Table 10 – Estimated Effect of Class Male Proportion on Student Outcomes, Using only Within-Semester Variation

	Grade (GPA value)	Future course in dept.	Switch major out of dept.	Declare major in dept.	Graduate within 6 years	Graduate in dept.
Female student X class male %	-0.227	-0.064	-0.083	-0.031	-0.078	-0.032
	(0.169)	(0.051)	(0.231)	(0.024)	(0.061)	(0.027)
Outcome Mean	2.988	0.644	0.138	0.054	0.677	0.127
Outcome SD	1.099	0.479	0.345	0.227	0.467	0.333
Observations	33071	43351	4750	37153	43351	43351

Notes: This table reports the results of regressions of student-class outcomes on an interaction between a female-student dummy and class male proportion. Each regression includes course-specific female fixed effects, class fixed effects, and controls for student underrepresented minority status and instructor-student gender match. All regressions are estimated on the observations within the identifying set of the main specification (as described in Section 2) for which the relevant outcome is observed and the given course only has only class in each of Fall of 2015 and Fall of 2016. Standard errors are clustered at the class level. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.

Appendix B - Estimation Using Courses with Only One Class Per Year

Table 11 – Estimated Effect of Class Male Proportion on Student Outcomes, For Courses that Offer One Class Per Year

	Grade (GPA value)	Future course in dept.	Switch major out of dept.	Declare major in dept.	Graduate within 6 years	Graduate in dept.
Female student X	-0.463	-0.131	0.168	-0.064	-0.322*	-0.133
class male %	(0.291)	(0.150)	(0.181)	(0.078)	(0.183)	(0.101)
Outcome Mean	3.057	0.561	0.110	0.065	0.716	0.239
Outcome SD	1.051	0.496	0.313	0.246	0.451	0.426
Observations	6758	8481	1986	5848	8481	8481

Notes: This table reports the results of regressions of student-class outcomes on an interaction between a female-student dummy and class male proportion. Each regression includes course-specific female fixed effects, class fixed effects, and controls for student underrepresented minority status and instructor-student gender match. All regressions are estimated on the observations within the identifying set of the main specification (as described in Section 2) for which the relevant outcome is observed and the given course only has only class in each of Fall of 2015 and Fall of 2016. Standard errors are clustered at the class level. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.

Appendix C - Exploring Non-Linearity of Effects

Throughout the main body of the paper, I report results from specifications estimating linear effects of class-level male proportion. However, it is reasonable to think that important non-linearities in the effects of peer gender composition may exist. In this appendix, I investigate this possibility. Specifically, I estimate regressions of the forms

$$Y^{F}_{i,r,c} = \alpha_0^{F,NP} + \sum_{j=2}^{6} \alpha_j^{F,NP} I_j + \gamma_r^{F,NP} \times I_r + (X^{F,NP})'_{i,r,c}\beta^{F,NP} + u^{F,NP}_{i,r,c} \tag{6}$$

$$Y^{M}_{i,r,c} = \alpha_0^{M,NP} + \sum_{j=2}^{6} \alpha_j^{M,NP} I_j + \gamma_r^{M,NP} \times I_r + (X^{M,NP})'_{i,r,c}\beta^{M,NP} + u^{M,NP}_{i,r,c} \tag{7}$$

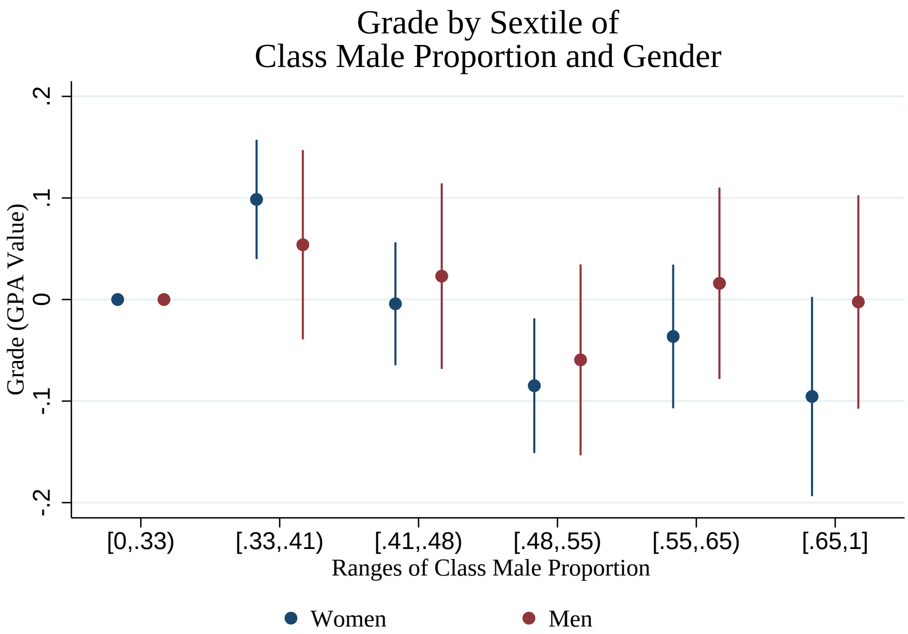
These regressions are nearly identical to equations (2) and (3), except, rather than estimating linear effects of class-level male proportion, they estimate coefficients on indicators of being in the second through sixth sextiles of class-level male proportion, $I_2, ..., I_6$ (with the first sextile being the omitted category). This will flexibly capture the form of non-linearities in the effects of peer gender. These regressions all include course fixed effects and controls for student underrepresented minority status and instructor gender, as with the within-gender specifications described in the main text. The results of this exercise are presented in Figures 4-9, for each of the six main outcomes considered throughout the paper. For each outcome, equations (6) and (7) are estimated separately, but the estimates are plotted on the same graphs for concision and ease of comparison.

The results generally indicate that both the absolute effects of male peers and the differences in effects on women compared to men are concentrated among the most male-dominated classes, conditional on course. Looking at the effects on grade and graduation, there seems to be a generally declining pattern of female achievement with increasing male proportion, with the most negative point estimates being on the top sextile of class-level male proportion, for both outcomes. This is also where the differences in the point estimates of the effects on men and women are largest (although the differences are not significant).

Similarly, there seems to be an increasing likelihood of male declaration of a major with the male proportions of associated first-semester classes. The largest positive point estimate on this outcome for men, and the biggest difference relative to women, is for the top sextile of class male proportion. The same holds for the top sextile of class male proportion and graduation within a department associated to a class as an outcome.

Many of the most male-dominated majors, both nationally and at UIC, are STEM majors, among the most highly paid majors on average, or both. If policy makers have particular interest in representation in STEM and the gender wage gap, a concentration of the effect of peer gender in the most heavily male-dominated environments underscores the importance of peer gender to real-world outcomes of interest. It is precisely in many

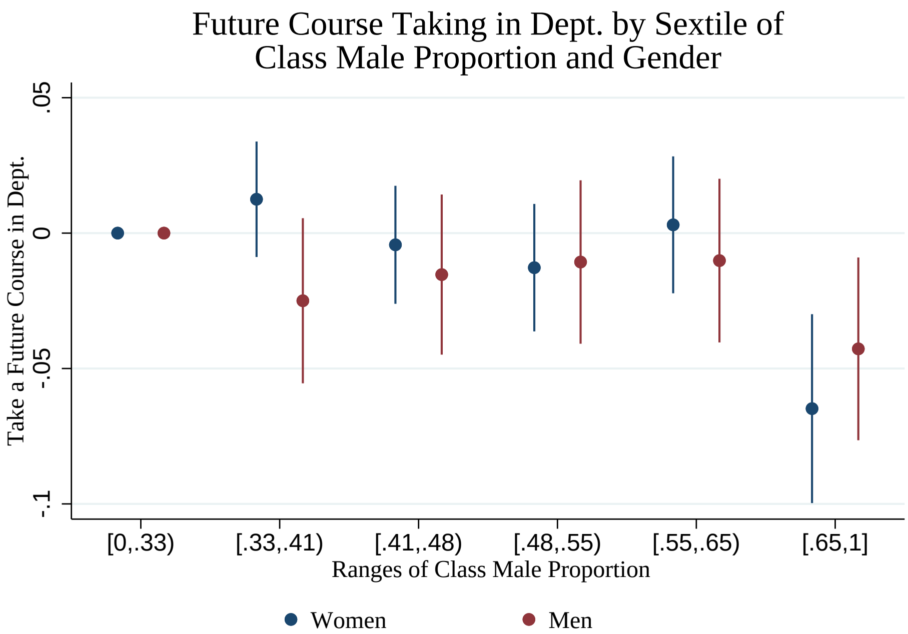
Figure 4



Notes: This figure presents the coefficients and 95% confidence intervals from regressions of the GPA value of grades on indicators for being in each sextile of class-level male proportion, conditional on course fixed effects and controls for instructor gender and student minority status. Coefficients for men and women are estimated separately. The first sextile is the omitted category.

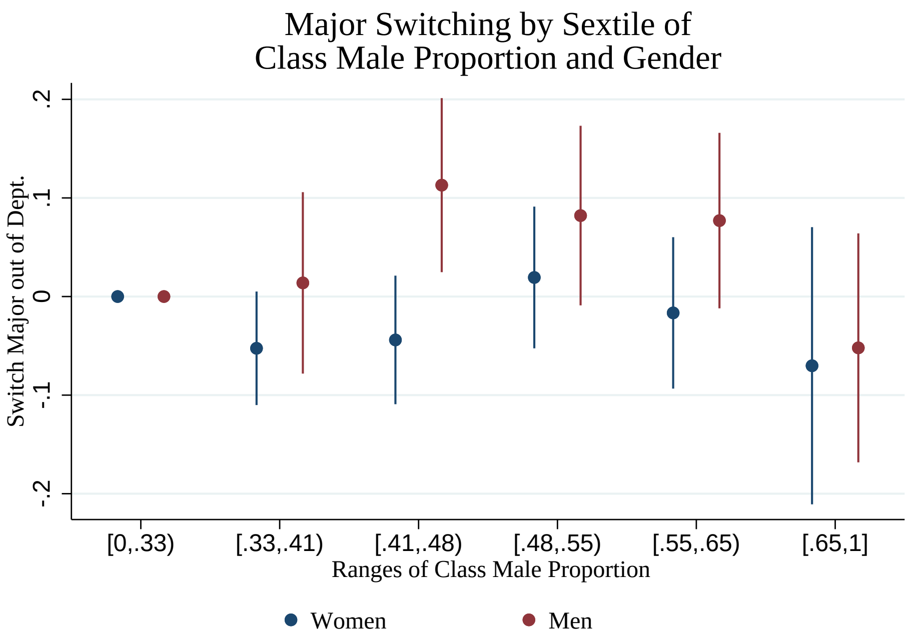
STEM majors and many highly paid majors where we would expect peer gender to matter most, based on these results.

Figure 5



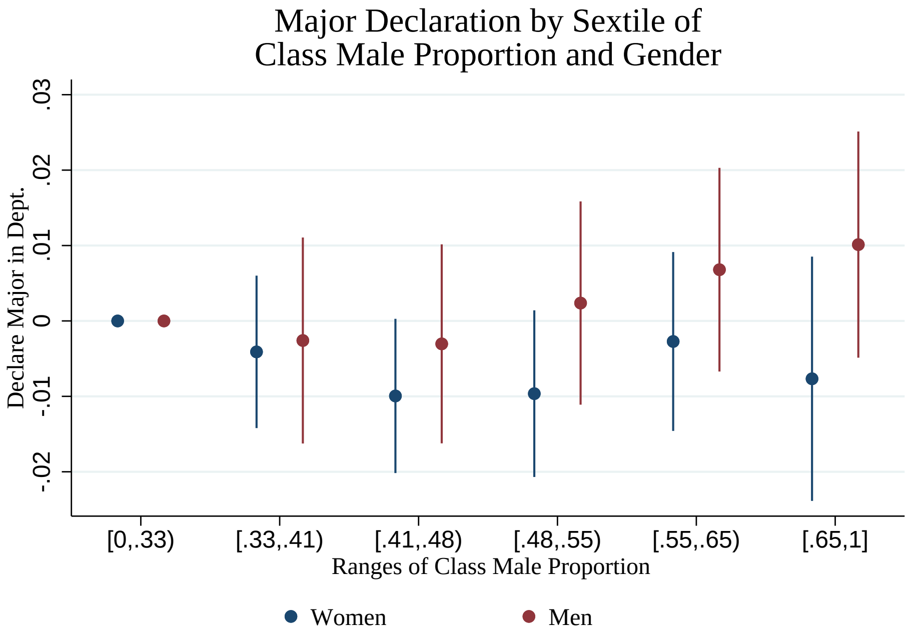
Notes: This figure presents the coefficients and 95% confidence intervals from regressions of an indicator for taking any future class in the same department on indicators for being in each sextile of class-level male proportion, conditional on course fixed effects and controls for instructor gender and student minority status. Coefficients for men and women are estimated separately. The first sextile is the omitted category.

Figure 6



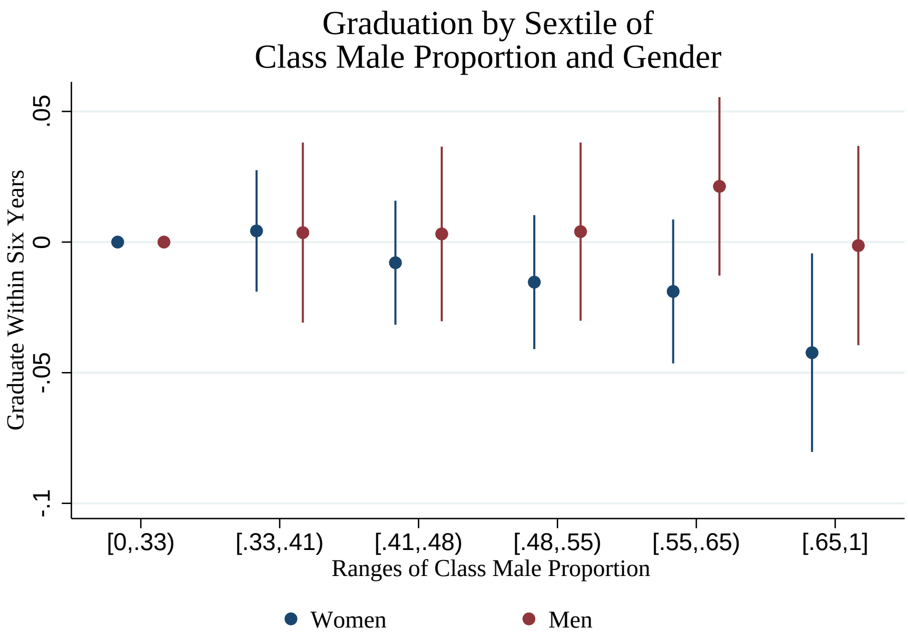
Notes: This figure presents the coefficients and 95% confidence intervals from regressions of an indicator of switching out of a major on indicators for being in each sextile of class-level male proportion, conditional on course fixed effects and controls for instructor gender and student minority status. Coefficients for men and women are estimated separately. The first sextile is the omitted category.

Figure 7



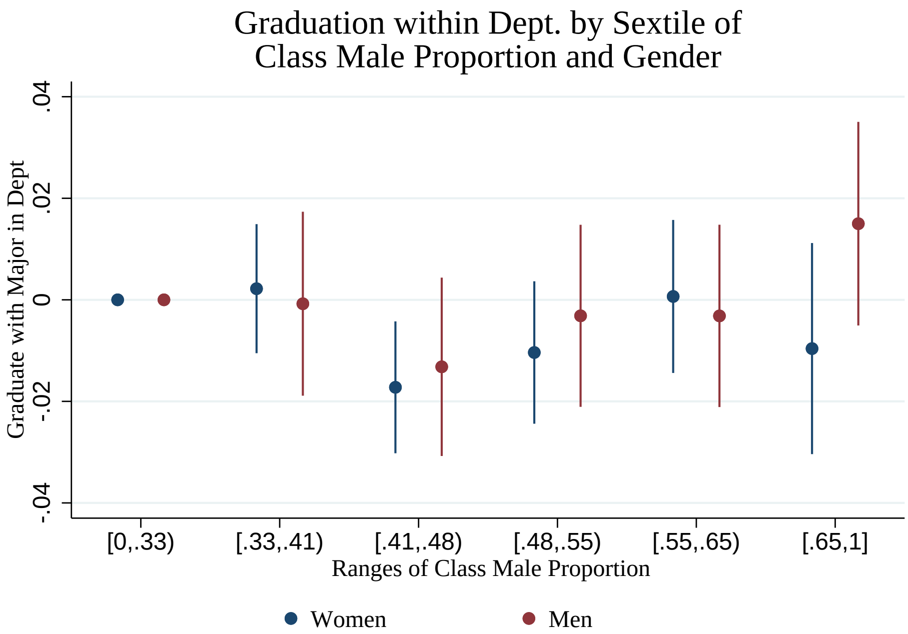
Notes: This figure presents the coefficients and 95% confidence intervals from regressions of an indicator for ever declaring a major in the same department on indicators for being in each sextile of class-level male proportion, conditional on course fixed effects and controls for instructor gender and student minority status. Coefficients for men and women are estimated separately. The first sextile is the omitted category.

Figure 8



Notes: This figure presents the coefficients and 95% confidence intervals from regressions of an indicator for graduating within six years on indicators for being in each sextile of class-level male proportion, conditional on course fixed effects and controls for instructor gender and student minority status. Coefficients for men and women are estimated separately. The first sextile is the omitted category.

Figure 9



Notes: This figure presents the coefficients and 95% confidence intervals from regressions of an indicator for graduating within six years with a declared major in the same department on indicators for being in each sextile of class-level male proportion, conditional on course fixed effects and controls for instructor gender and student minority status. Coefficients for men and women are estimated separately. The first sextile is the omitted category.

Appendix D - Alternative Control Schemes

Table 12 reports the results of estimating equation (1) while including both course-specific female fixed effects and class fixed effects but varying the sets of additional student and class-student controls. The specifications in column 1 include no controls beyond the fixed effects, column 2 specifications include controls for student minority status and student-instructor gender match (just as in column 4 of Table 5 in the main body of the paper), and column 3 includes the same controls as column 2 in additional to controls for high school GPA and ACT composite score. Table 12 reveals that the set of controls used has very little impact on the estimated effects of class male proportion on student outcomes. Importantly, including controls for individual pre-college academic aptitude has little impact on the qualitative interpretation of the results. It does however curtail the precision of the results, due to the required exclusion of the portion of the sample that is missing information on either high school GPA, ACT score, or both.

Table 12 – Estimated Effect of Class Male Proportion on Student Outcomes Under Varying Sets of Controls

	(1)	(2)	(3)
Grade (GPA value) [Mean = 2.99]			
Female student X class male %	-0.319**	-0.309**	-0.227
	(0.150)	(0.151)	(0.175)
Observations	33071	33071	23018
Take a future course in same department [Mean = .64]			
Female student X class male %	-0.067	-0.067	-0.079
	(0.047)	(0.047)	(0.055)
Observations	43351	43351	31645
Switch major to another department [Mean = .14]			
Female student X class male %	0.018	0.031	0.255
	(0.149)	(0.146)	(0.649)
Observations	4750	4750	1232
Declare a major in same department [Mean = .05]			
Female student X class male %	-0.047*	-0.047*	-0.049*
	(0.025)	(0.025)	(0.028)
Observations	37153	37153	29710
Graduate within six years [Mean = .68]			
Female student X class male %	-0.111*	-0.106*	-0.104
	(0.057)	(0.056)	(0.066)
Observations	43351	43351	31645
Graduate with a major in same department [Mean = .13]			
Female student X class male %	-0.047*	-0.046*	-0.052*
	(0.028)	(0.028)	(0.029)
Observations	43351	43351	31645
Controls			
Student underrepresented minority status	No	Yes	Yes
Student-instructor gender match	No	Yes	Yes
High school GPA	No	No	Yes
ACT composite score	No	No	Yes

Notes: This table reports the results of regressions of student-class outcomes on an interaction between a female-student dummy and class male proportion. Each cell corresponds to a separate regression, with outcome given by the row header and controls given by the column foot. Each regression includes course-specific female fixed effects and class fixed effects. All regressions are estimated on the observations within the identifying set of the main specification (as described in Section 2) for which the relevant outcome is observed. Standard errors are clustered at the class level. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.

Appendix E - Section-Level Analysis

The results reported in the main body of the paper exploit variation in peer composition across classes within a course. This strategy assumes that there is no systematic, gendered sorting to the classes, within courses, that have more male students. This assumption is supported by the fact that first-semester students enroll in classes with little or no access to information about peer enrollments and by the highly comparable observable characteristics of students across classes with differing gender ratios, as reported in Table 2. However, there could still be reasonable concern that sorting to class characteristics, such as time slot or instructor traits, may induce differential sorting on unobservables between men and women, which would bias my results.

For classes that have associated laboratory, discussion, or practical experience sections, there is an additional level of variation to exploit: peer composition across sections, within classes. Utilizing this variation lets me to allow for any arbitrary pattern of sorting to classes by students. I then assume that, within a class, there is not meaningful sorting to sections that have more men. There are reasons why sorting to sections within a class may be “more" exogenous than the initial sorting to classes: sections have tighter capacity constraints, differences between sections may be less visible or salient than differences between classes, and section choices may be more constrained by schedules if choice of preferred classes is prioritized by students over choice of preferred sections. However, focusing on sections, rather than classes, introduces the notable drawback of reducing the available portion of the sample, as not all classes have sections, and reducing useful variation, as no classes have as many different sections as there are different classes within certain courses. Moreover, estimation using section-level variation implies a different parameter than estimation using class-level variation. It may be that class peers and section peers matter differently, depending on what the mechanisms are via which peers influence outcomes.

For my analysis of the effects of section peers on student outcomes, I emphasize two estimating equations. Analogous to my primary estimating equation for class-level analysis, equation (1), I estimate

$$Y_{i,r,c,s} = \pi_0 + \pi_1 \times Fem_i \times MP_s + \rho_c \times Fem_i \times I_c + \upsilon_s \times I_s + X'_{i,r,c,s}\tau + \varepsilon_{i,r,c,s} \quad (8)$$

where sections are indexed by s, MP_s denotes the proportion of male students in section s, I_s is an indicator variable that take on a value of one if outcome $Y_{i,r,c,s}$ is associated with section s, and $\varepsilon_{i,r,c,s}$ is the error term. π_1 captures how the gap between female and male outcomes evolves with MP_s. This specification includes both class-specific female fixed effects, ρ_c, section fixed effects, υ_s, and controls for student and student-section characteristics, τ. This allows for both any kind of sorting to classes and gender-neutral sorting to sections, meaning that differential sorting to sections within class by men compared to women is the only threat to identification.

In practice, the results of estimating equation (8) are too imprecise to be interpretable. I therefore focus on a specification without section fixed effects,

$$Y_{i,r,c,s} = \eta_0 + \eta_1 \times Fem_i \times MP_s + \theta_c \times Fem_i \times I_c + \iota_c \times I_c + X'_{i,r,c,s}\kappa + \varepsilon_{i,r,c,s} \tag{9}$$

This specification still includes class-specific female fixed effects, here denoted θ_c, but foregoes section fixed effects for class fixed effects, ι_c. Identification with this model relies upon an assumption of no sorting, absolute or differential, to sections within a class. This is similar in spirit to the within-gender, class-level analysis presented in the main text.

I present the results of estimating equations (8) and (9) on pre-college characteristics, and the grades predicted by those characteristics, in Table 13. The top panel displays the results of estimating equation (8) and the bottom the results for equation (9). The estimates show that the differences between observable male and female characteristics do not change systematically with section-level male proportion, whether conditioning on section fixed effects or not. Importantly, observable differences between students predict no change in grade differential between men and women when going from a wholly female to a wholly male section, again whether conditioning on section fixed effects or not. I interpret this as evidence that the identifying assumptions for both specifications are satisfied in this sample.

Table 14 presents the estimated effects of section peer composition on student outcomes, across different specifications. Column 1 includes no fixed effects and documents a strong negative association between female academic performance and major choice with section-male percentage, matching the pattern seen in the class-level analysis. Column 4 reports the results of estimating equation (8), yielding results that, as previously mentioned, are too imprecise to make meaningful inference from. Column 2 reports results from equation (9), which I consider the preferred specification for the section-level analysis. Here, although no results are significant at conventional levels, there is suggestive evidence of a similar pattern to the results of the class-level analysis: worse female academic achievement in the face of more male peers. The strongest evidence of a section peer effect is for grades, with relatively meager evidence of effects on major choice.

I conclude my discussion of the section-level analysis with one interesting note on the distribution of estimated effects of section peer gender. It seems that any negative effect male section peers have on women, relative to men, is concentrated among female-majority classes. Table 15 reports results on the subset of classes that are female-majority while Table 16 does the same for classes that are male-majority. Focusing on the second column of Table 15, the point estimates of the effects of male section peers among female-majority classes are quite similar to the estimated effects of male class peers that are reported in the main body of the paper. The same is not true among male-majority classes, as seen in Table 16. In particular, coefficients on the differential effects of male peers on grades, major

Table 13 – Associations Between Section Male Proportion and Pre-College Characteristics

	Standardized ACT score	High school GPA	Underrepresented minority student	Predicted grade (GPA value)
With section fixed effects:				
Female student X	-0.140	0.063	-0.015	0.093
section male %	(0.185)	(0.076)	(0.082)	(0.099)
Outcome Mean	0.000	3.282	0.517	2.783
Outcome SD	1.000	0.385	0.500	0.452
Observations	13604	13612	19106	11870
With class fixed effects:				
Female student X	-0.000	0.004	0.007	0.024
section male %	(0.110)	(0.042)	(0.047)	(0.055)
Outcome Mean	0.000	3.282	0.517	2.783
Outcome SD	1.000	0.385	0.500	0.452
Observations	13604	13612	19106	11870

Notes: This table reports the results of regressions of pre-college characteristics, at a student-class observation level, on an interaction between a female-student dummy and the section-level male proportion. Each cell corresponds to a separate regression. Top panel regressions include section fixed effects. Bottom panel regressions include class fixed effects. All regressions include class-specific female fixed effects. The predicted grade outcome is based on another (unshown) regression of the GPA value of grades on ACT score, high school GPA, and minority status among students in the main estimation sample. The predicted grade regressions only include the observations used to form the predicted grades: student-class observations from graded classes that had information on both ACT scores and high school GPA. Standard errors are clustered at the class level. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.

declarations, and graduation in a department are all negative in Table 15 and much larger in magnitude in Table 15 than in Table 16. It would be interesting to see if future work, which may have more data on sections, documents a similar pattern. Considering why the effects are concentrated among certain classes could help shed light on the mechanisms by which peers influence outcomes.

Table 14 – Estimated Effect of Section Male Proportion on Student Outcomes

	(1)	(2)	(3)	(4)
Grade (GPA value) [Mean = 2.83]				
Female student X section male %	-0.619***	-0.206	-0.196	-0.083
	(0.126)	(0.178)	(0.134)	(0.215)
Observations	16939			
Take a future course in same department [Mean = .64]				
Female student X section male %	-0.178***	-0.052	-0.190***	0.002
	(0.062)	(0.059)	(0.048)	(0.076)
Observations	19106			
Switch major to another department [Mean = .15]				
Female student X section male %	0.412***	-0.060	0.174	0.002
	(0.088)	(0.191)	(0.131)	(0.394)
Observations	2052			
Declare a major in same department [Mean = .08]				
Female student X section male %	-0.034	-0.015	-0.054*	-0.013
	(0.031)	(0.029)	(0.030)	(0.040)
Observations	16292			
Graduate within six years [Mean = .67]				
Female student X section male %	-0.228***	-0.048	-0.012	0.074
	(0.041)	(0.067)	(0.054)	(0.086)
Observations	19106			
Graduate with a major in same department [Mean = .14]				
Female student X section male %	-0.498***	-0.029	-0.071**	0.015
	(0.076)	(0.038)	(0.033)	(0.044)
Observations	19106			
Fixed effects				
Class and class-female	No	Yes	No	Yes
Section	No	No	Yes	Yes
Controls				
Student gender	Yes	No	Yes	No
TA gender	Yes	Yes	No	No
Section male %	Yes	Yes	No	No

Notes: This table reports the results of regressions of student-class outcomes on an interaction between a female-student dummy and section male proportion. Each cell corresponds to a separate regression, with outcome given by the row header and fixed effects and controls given by the column foot. Each regression includes controls for student underrepresented minority status and instructor-TA gender match. All regressions are estimated on the observations within the identifying set of the main specification (as described in Section 2) for which the relevant outcome is observed. Standard errors are clustered at the class level. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.

Table 15 – Estimated Effect of Section Male Proportion on Student Outcomes, Among Female Majority Classes

	(1)	(2)	(3)	(4)
Grade (GPA value) [Mean = 2.83]				
Female student X section male %	-0.240	-0.330	-0.184	-0.056
	(0.193)	(0.233)	(0.255)	(0.317)
Observations	8236			
Take a future course in same department [Mean = .64]				
Female student X section male %	-0.156	-0.116	-0.148	-0.127
	(0.097)	(0.082)	(0.096)	(0.101)
Observations	9552			
Switch major to another department [Mean = .15]				
Female student X section male %	0.534***	0.416	0.016	0.399
	(0.172)	(0.308)	(0.156)	(0.637)
Observations	998			
Declare a major in same department [Mean = .08]				
Female student X section male %	-0.028	-0.034	-0.112**	-0.054
	(0.049)	(0.043)	(0.051)	(0.057)
Observations	8221			
Graduate within six years [Mean = .67]				
Female student X section male %	-0.130*	0.086	0.058	0.182
	(0.073)	(0.088)	(0.092)	(0.113)
Observations	9552			
Graduate with a major in same department [Mean = .14]				
Female student X section male %	-0.280***	-0.101*	-0.114**	-0.055
	(0.076)	(0.057)	(0.051)	(0.062)
Observations	9552			
Fixed effects				
Class and class-female	No	Yes	No	Yes
Section	No	No	Yes	Yes
Controls				
Student gender	Yes	No	Yes	No
TA gender	Yes	Yes	No	No
Section male %	Yes	Yes	No	No

Notes: This table reports the results of regressions of student-class outcomes on an interaction between a female-student dummy and section male proportion. Each cell corresponds to a separate regression, with outcome given by the row header and fixed effects and controls given by the column foot. Each regression includes controls for student underrepresented minority status and instructor-TA gender match. All regressions are estimated on the observations within the identifying set of the main specification (as described in Section 2) for which the relevant outcome is observed and the given class is female-majority. Standard errors are clustered at the class level. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.

Table 16 – Estimated Effect of Section Male Proportion on Student Outcomes, Among Male Majority Classes

	(1)	(2)	(3)	(4)
Grade (GPA value) [Mean = 2.83]				
Female student X section male %	-0.374*	-0.010	-0.099	-0.069
	(0.192)	(0.264)	(0.209)	(0.284)
Observations	8703			
Take a future course in same department [Mean = .64]				
Female student X section male %	-0.072	0.044	0.026	0.156
	(0.092)	(0.093)	(0.087)	(0.115)
Observations	9554			
Switch major to another department [Mean = .15]				
Female student X section male %	-0.080	-0.911*	0.041	-0.253
	(0.211)	(0.497)	(0.273)	(0.569)
Observations	1054			
Declare a major in same department [Mean = .08]				
Female student X section male %	0.123***	0.006	0.055	0.033
	(0.046)	(0.039)	(0.049)	(0.054)
Observations	8071			
Graduate within six years [Mean = .67]				
Female student X section male %	-0.296***	-0.155*	-0.128	-0.044
	(0.066)	(0.094)	(0.087)	(0.123)
Observations	9554			
Graduate with a major in same department [Mean = .14]				
Female student X section male %	-0.280***	0.046	0.003	0.091
	(0.097)	(0.046)	(0.057)	(0.060)
Observations	9554			
Fixed effects				
Class and class-female	No	Yes	No	Yes
Section	No	No	Yes	Yes
Controls				
Student gender	Yes	No	Yes	No
TA gender	Yes	Yes	No	No
Section male %	Yes	Yes	No	No

Notes: This table reports the results of regressions of student-class outcomes on an interaction between a female-student dummy and section male proportion. Each cell corresponds to a separate regression, with outcome given by the row header and fixed effects and controls given by the column foot. Each regression includes controls for student underrepresented minority status and instructor-TA gender match. All regressions are estimated on the observations within the identifying set of the main specification (as described in Section 2) for which the relevant outcome is observed and the given class is male-majority. Standard errors are clustered at the class level. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.

Appendix F - Within-Gender Peer Effects, When Controlling for Peer Ability

Table 17 – Estimated Effect of Class Male Proportion on Student Outcomes, By Gender And Controlling for Average Peer Ability

	Grade (GPA value)	Future course in dept.	Switch major out of dept.	Declare major in dept.	Graduate within 6 years	Graduate in dept.
Only Women:						
Class male %	-0.235*	-0.078**	-0.146	-0.021	-0.111***	-0.009
	(0.130)	(0.033)	(0.217)	(0.017)	(0.041)	(0.019)
Outcome Mean	3.068	0.663	0.141	0.060	0.707	0.135
Outcome SD	1.051	0.473	0.348	0.237	0.455	0.341
Observations	12180	16393	618	15451	16393	16393
Only Men:						
Class male %	-0.103	-0.002	-0.234	0.031**	0.014	0.055***
	(0.159)	(0.041)	(0.337)	(0.016)	(0.047)	(0.018)
Outcome Mean	2.901	0.625	0.135	0.049	0.647	0.119
Outcome SD	1.143	0.484	0.341	0.216	0.478	0.324
Observations	10881	15309	615	14311	15309	15309

Notes: This table reports the results of regressions of student-class outcomes on class-level male proportion. Each cell corresponds to a separate regression, with outcome given by the column header. Top row results are estimated only on female students and bottom row results only on male students. Each regression includes course fixed effects and controls for student underrepresented minority status, instructor gender, and average peer HS GPA. All regressions are estimated on the observations within the identifying set of the main specification (as described in Section 2) for which the relevant outcome is observed. Standard errors are clustered at the class level. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.

Appendix G - Heterogeneity of Effects Between Female- and Male-Majority Departments

Table 18 – Estimated Effect of Class Male Proportion on Student Outcomes, By Gender and Department-Level Male Proportion

	Grade (GPA value)	Future course in dept.	Switch major out of dept.	Declare major in dept.	Graduate within 6 years	Graduate in dept.
Women in female-majority-department classes:						
Class male %	-0.067	-0.044	-0.016	-0.016	-0.037	0.002
	(0.117)	(0.030)	(0.073)	(0.016)	(0.040)	(0.019)
Outcome Mean	3.148	0.674	0.141	0.058	0.710	0.146
Outcome SD	1.016	0.469	0.348	0.233	0.454	0.353
Observations	12266	15203	2016	12704	15203	15203
Men in female-majority-department classes:						
Class male %	-0.054	-0.063	0.036	0.009	-0.010	0.008
	(0.159)	(0.040)	(0.125)	(0.013)	(0.050)	(0.017)
Outcome Mean	2.950	0.604	0.171	0.039	0.628	0.099
Outcome SD	1.131	0.489	0.377	0.194	0.483	0.299
Observations	8503	10757	1006	9417	10757	10757
Women in male-majority-department classes:						
Class male %	-0.501**	-0.160**	-0.329***	-0.039	-0.176***	-0.029
	(0.233)	(0.072)	(0.125)	(0.042)	(0.066)	(0.045)
Outcome Mean	2.873	0.637	0.138	0.063	0.698	0.111
Outcome SD	1.109	0.481	0.345	0.243	0.459	0.314
Observations	5018	6959	521	6273	6959	6959
Men in male-majority-department classes:						
Class male %	-0.038	-0.015	-0.122	0.076**	0.087	0.105**
	(0.223)	(0.075)	(0.147)	(0.037)	(0.063)	(0.043)
Outcome Mean	2.845	0.647	0.104	0.060	0.666	0.140
Outcome SD	1.154	0.478	0.306	0.237	0.472	0.347
Observations	7284	10432	1207	8759	10432	10432

Notes: This table reports the results of regressions of student-class outcomes on class-level male proportion. Each cell corresponds to a separate regression, with outcome given by the column header and the subset of students used for estimation given by the row header. Each regression includes course fixed effects and controls for student underrepresented minority status and instructor gender. All regressions are estimated on the observations within the identifying set of the main specification (as described in Section 2) for which the relevant outcome is observed. Standard errors are clustered at the class level. * indicates significance at a level of 0.1, ** at a level of 0.05, and *** at a level of 0.01.