The Positive Influence Of Female College Students On Their Male Peers

Andrew J. Hill[1]

JEL Codes: I21; J16

October 2015 (preliminary; comments welcome)

Abstract

Peer gender composition in college affects graduation rates. Using within-college across-cohort variation in freshman enrollment at public four-year colleges in the US indicates that a ten percentage point increase in the proportion of females in a freshman cohort increases cohort male graduation rates by two percentage points. There is no effect on female graduation rates. Effects depend on the college environment. They are more evident in colleges with higher shares of students living on campus, in college housing, and without cars, suggesting that college gender peer effects operate through direct or indirect changes in student behavior rather than being purely compositional. Further analysis using individual level data supports the above findings, and also finds no evidence that within-course gender composition affects course achievement.

[1] University of South Carolina; Email: andrew.hill@moore.ac.edu

1. Introduction

The proportion of females enrolled in institutions of higher education levelled off in the 2000s after increasing through much of the twentieth century. Although there are still marked gender differences in course-taking behavior, females are more likely to choose college majors that were considered the domain of males in the past. Outside of the class environment, universities have responded to evolving student preferences by increasing the number of mixed gender residences, and some universities have even begun to allow opposite gender roommates (see Gordon, 2010, for an account of this in the popular media). The nature and intensity of interactions across genders on college campuses has changed and continues to change both inside and outside the college classroom. This paper takes a first step towards understanding the implications of these changes by estimating the effect of freshman cohort gender composition on academic outcomes.

I exploit within-college across-cohort variation in freshman composition and find that there are statistically and economically significant cohort gender peer effects on graduation rates at public four-year colleges in the US. A 10 percentage point increase in the share of female freshman students increases the subsequent graduation rate of male students in that cohort by about 2 percentage points. There is no clear effect on female graduation rates. Effects are concentrated in colleges where students are more likely to interact outside the classroom: colleges with higher shares of students living in college housing, higher shares of students living on campus, and higher shares of students without cars on campus. This indicates that peer gender composition directly or indirectly affects individual behavior; effects would be evident at all colleges if they were caused by pure compositional changes such as females being admitted at the expense of less able males.

A secondary analysis using course-level data finds no convincing evidence that gender composition effects operate inside the college classroom. This suggests that peer gender composition may affect the general college learning environment rather than the specific class environment, which contrasts with some of the school peer effects literature that finds peer effects are more evident within the classroom than the cohort (Burke and Sass, 2013).

We know that peer gender affects educational outcomes in kindergarten (Whitmore, 2005), elementary school (Hoxby, 2000), middle school (Lee et al, 2014; Lu and Anderson, 2015) and high school (Lavy and Schlosser, 2011; Jackson, 2012). The subset of this literature investigating changes in cohort gender composition in mixed gender school environments typically finds that increasing the share of female students in a cohort within a school improves achievement. The effect found in this paper for college students goes in the same direction. Given that a seventeen year-old in high school and an eighteen year-old in college are not that different, it seems reasonable that cohort peer gender composition may affect the incentives and actions of high school students and college students in similar ways. Furthermore, when students leave their childhood homes to attend college, students not only take classes with their peers, but are likely to reside with them, too.

The potential for college roommates to influence outcomes has long been recognized, and much of the early literature on college peer effects focused on college roommates (Sacerdote, 2001; Zimmerman, 2003). This is partly due to the opportunity for identification presented by institutional random assignment to peer groups in this setting. Peer effects among college roommates are generally estimated to be small in magnitude. Carrell, Fullerton and West (2009), however, find that peer effects are larger when considering college cohorts rather than college roommates. Although they study a very specific population and setting (small cohorts in a military academy), this provides some support for using a broader peer group definition than college roommates. Stinebrickner and Stinebrickner (2006) also argue that college roommates may not be the right place to look for evidence of peer effects in college given students are exposed to a wide set of potentially influential peers at university. More recently, Jain and Kapoor (2015) use institutional random assignment to estimate peer effects in college study groups and roommates simultaneously, finding that roommates were more influential than study groups.

Roommates constitute only one set of peers in college; students also take classes together, play sport together and socialize together. Student behavior in all of these environments is likely to spill over into effects on academic achievement. There are consequently many sets of potentially influential peers in college. I bundle several of these sets together by considering the freshman cohort as the peer group. The benefits and costs of doing so are discussed in the Empirical Strategy section below, after which I include sections that describe the data, report the results, and interpret the findings.

2. Empirical Strategy

2.1. Effects of Freshman Cohort Composition on Graduation Rates Using Aggregate Data

I investigate the effect of college freshman cohort gender composition on students' academic outcomes by estimating the following reduced form equation:

$$y_{ct} = \alpha_c + \beta_t + \gamma_c t + \pi P_{ct} + \varepsilon_{ct} \qquad (1)$$

y_{ct} is a gender-specific aggregate graduation rate for the cohort of freshman students in college c in year t, α_c is a college fixed effect, β_t is a cohort or year fixed effect, and $\gamma_c t$ is a college-specific time trend. P_{ct} is the proportion of female students in the freshman cohort in college c in year t. The parameter π measures the correlation between the share of female freshman students in the college cohort and the aggregate graduation rate for male or female freshman students in that college cohort.

In order for π to have a causal interpretation, the unobserved component of the cohort graduation rate must be uncorrelated with the cohort female share. College fixed effects capture the endogenous sorting of students into colleges. This controls for colleges that consistently have higher shares of female students and higher or lower graduation rates, for example.

There may also be unobserved time-varying factors that are correlated with both changes in the proportion of female students and aggregate graduation rates. In particular, a potential upward trend in the female share of college students due to a widening female-male achievement gap over the past two decades and a potential upward trend in college graduation rates due to grade inflation or higher administrative and student costs associated with student failure over the same period would generate a positive estimate of the parameter π, but would not indicate a causal relationship. The cohort fixed effects and college-specific time trends control confounding factors of this type. Identification therefore relies on college-specific deviations in cohort female shares and cohort graduation rates from their long-term trends. Essentially, we are comparing the graduation rates of cohorts at the same college who are exposed to different cohort female shares by purely random factors. I show in the Data section that there is a fairly considerable amount of idiosyncratic variation in gender shares across cohorts, which is necessary for the identification strategy to work.

My primary specification considers gender peer effects at the college cohort level. I do secondary analysis with semester-specific data at the college department level. The majority of college students primarily interact with other students who entered college in the same year that they did. Freshman students in college are likely to live, attend classes and socialize with other students in their cohorts. The preferences and actions of students in all three settings are likely to affect their academic performance. Defining the peer group as the college cohort captures several dimensions of peer influence that may not be captured by narrower peer groups.

Investigating peer effects at the cohort rather than class level also eliminates concerns related to endogenous sorting into college classes, such as gender differences in course or class selection. This motivation for using college cohorts rather than college classes is similar to the motivation for using within-school grade composition rather than within-school class composition in the school compositional peer effects literature (Hoxby, 2000; Lavy and Schlosser, 2011). Using the freshman cohort rather than cohorts in subsequent years has the added advantage of not being subject to biases introduced by student failure and course repetition (Ciccone and Garcia-Fontes, 2015).

The disadvantage of using this broad definition of the peer group is that cohort gender composition may only be weakly correlated with the gender composition of narrower peer groups in which students primarily interact, such as college residences and classes. Using a noisy proxy for potentially more influential peer groups may attenuate estimated effects. This may be particularly true in larger colleges. Ciccone and Garcia-Fontes (2015), for example, only find cohort compositional peer effects when they restrict their analysis to smaller primary schools and argue that the correlation between grade composition and class composition is too low to find compositional peer effects in larger schools. I explore whether estimated effects differ in magnitude or precision across college size by estimating the primary model on five subsamples based on college size quintiles.

The gender composition of adjacent cohorts is included in the model to test whether estimated effects may be spurious or come about through unexpected channels. Although it is likely that college students interact across cohorts, evidence of peer effects of a similar (or larger) magnitude for adjacent cohorts rather than the actual cohort to which the student belongs would be problematic for the given peer effect interpretation. This can be considered a falsification test. I report results from the following specification.

$$y_{ct} = \alpha_c + \beta_t + \gamma_c t + \sum_{s=t-2}^{t+2} \pi_i P_{cs} + \varepsilon_{ct} \qquad (2)$$

Results from estimating Equations 1 and 2 provide strong evidence of gender-specific peer composition effects using aggregate enrollment and graduation data. I augment these results with a secondary analysis to try and unpack how gender peer effects work in the college environment.

2.2. Mechanisms of Compositional Peer Effects at College

College gender peer effects on academic achievement could be enacted through compositional or behavioral channels. A change in the graduation rate caused by a change in the proportion of female students may come about because the change in gender composition simply reflects a change in mean ability composition. For example, consider a marginally admitted male student being (randomly) replaced by a female student in a particular freshman cohort. This increases the female share of that cohort. If the marginally admitted male student was less likely to graduate than other admitted males in the cohort, the removal of that male student from the male subsample would increase the male graduation rate, and we would observe a positive correlation between the cohort female share and cohort male graduation rate. This would be a purely compositional effect and not necessarily reflect any actual peer influence. Although this finding would be interesting in its own right for understanding the factors that affect graduation rates, the availability of policy instruments to act on such peer effects may be limited.

The potential behavioral responses to freshman cohort gender composition could come about directly through changes in individual behavior or indirectly through students responding to changes in the college environment. These behavioral mechanisms can be broadly grouped into effects operating inside and outside the college classroom. Inside the classroom, the class or course learning environment and interactions between students and professors could be affected by class gender composition. Outside the classroom, peer gender composition may affect students' study efforts, attitudes towards class, and socializing behavior, which, in turn, are likely to affect academic achievement.

I consider various approaches to identify the mechanisms of freshman cohort peer effects. First, I explore whether effects operate throughout the distribution of female cohort shares. Evidence that effects are nonmonotonic or nonlinear in the female share would rule out effects being driven by purely compositional changes. This is because if the mechanism is purely compositional, marginal effects would be necessarily similar in cohorts with both high and low female shares. This is done by regressing the gender-specific graduation rate on indicators for female share deciles P^i.

$$y_{ct} = \alpha_c + \beta_t + \gamma_c t + \sum_{i=1}^{10} \pi_i P_{ct}^i + \varepsilon_{ct} \qquad (3)$$

The parameter π_i then measures how graduation rates in cohorts with female shares falling into decile P^i differ from graduation rates with female shares in the omitted decile(s) in the same college. This test can only rule out purely compositional effects; it cannot speak to the presence or absence of effects operating through behavioral channels.

The potential for behavioral effects is investigated in the aggregate data by exploring whether there is heterogeneity in the effect by college type. This is done by interacting the proportion of female students in the freshman cohort with fixed college characteristics. The specification is given below with X_c^j denoting college characteristic j.

$$y_{ct} = \alpha_c + \beta_t + \gamma_c t + \pi P_{ct} + \pi_j (X_c^j \times P_{ct}) + \varepsilon_{ct} \qquad (4)$$

If the college characteristic is binary, for example, the coefficient on the interaction term between the college characteristic X_c^j and the female share π_j reveals whether the effect of the proportion of female students in the freshman cohort is more evident in colleges that display this characteristic than colleges that do not. Evidence that effects depend on the college environment indicates that student behavior changes in response to peer composition. This is because purely compositional effects would not vary across colleges; if students are affected by peer composition only in certain environments, then effects must be driven by behavioral responses (direct or indirect) to freshman cohort gender composition. Note that the effect of the actual college characteristic on graduation rates cannot be identified given the inclusion of college fixed effects.

The college characteristics I consider are from the US News Academic Insights dataset. I look at college characteristics that describe how peers may interact outside the classroom - the share of students in college housing, the share of students who live off campus, and the share of students with cars on campus – and college characteristics that describe the extent to which peers may interact inside the classroom – the share of small classes and large classes in the college. Separating characteristics into these two categories helps determine whether peer gender composition effects are likely to occur inside or outside the college classroom (or both). This is important when thinking about what university policies may reinforce positive peer effects and mitigate negative peer effects. This is a necessarily limited and selected set of college characteristics, but is sufficient for arguing that the mechanism is more than purely compositional.

2.3. Effects of Freshman Cohort Composition on Graduation Rates Using Individual Data

The advantage of the above empirical models using college-specific aggregate data over a relatively wide time frame is that estimated effects are representative of the US. These models have limitations, however, particularly if concerned about other dimensions of peer group composition that may change along with gender composition. I complement the above specifications with an analysis using individual level data. Specifically, I use transcript data aggregated to the semester level from the Texas Higher Education Opportunity Project. (Ideally, data would be at the course level with no aggregation, but such data were not available.)

The findings from the primary specification are replicated in the individual level data. This is done by constructing a repeated cross-section with one observation per student where the relevant outcome variables are an indicator for whether the student graduated (and another outcome of total credit hours obtained), and the relevant explanatory variable is the student's freshman cohort female share. The model is otherwise similar, and is estimated separately for males and females.

$$y_{ict} = \alpha_{ic} + \beta_{it} + \gamma_{ic} t + \pi P_{ict} + g(A_{ict}) + \varepsilon_{ict} \qquad (5)$$

y_{ict} is the outcome variable for student i in college c in year t, α_{ic} and β_{it} are indicators for student i's college and cohort, $\gamma_{ic}t$ is a college-specific time trend for student i, and P_{ict} is the proportion of female students in student i's freshman cohort.

As evident in the above equation, I include an additional control for student i's college admissions score (SAT or ACT) A_{ict} in some specifications. The inclusion of a flexible admissions score control means that the parameter π identifies the effect of freshman cohort gender composition for students of equal ability. This speaks to whether effects arise from compositional or behavioral changes. If effects were purely coming about because changes in gender composition reflect changes in ability composition, there would be no effect of freshman cohort gender composition on the probability of graduation for students of the same ability.

I also report results from models in which the effect of the proportion of female peers on the probability of graduation may vary across students of different abilities. This is implemented by considering the effects on students in different within-college admission score deciles A_{ic}^{j}. The effect of cohort gender composition on the probability of graduation for students in the j^{th} decile of admission scores is given by π_j in the below specification. If effects operate through the pure composition channel, there would only be an effect on the probability of graduation for marginally admitted students who would be in the first or second deciles of the within-college admissions score distribution.

$$y_{ict} = \alpha_{ic} + \beta_{it} + \gamma_{ic}t + \sum_{j=1}^{10} \pi_j P_{ict} A_{ic}^{j} + g(A_{ict}) + \varepsilon_{ict} \qquad (6)$$

Using the individual level data, I also explore directly whether the gender difference in the propensity to graduate for marginally admitted students is the same as the gender difference in the propensity to graduate for other (inframarginally) admitted students. This investigates whether replacing a marginally admitted male student with a marginally admitted female student has the potential to affect male and female graduation rates in different ways. If the gender gap in the probability of graduation does not differ for marginally admitted students and other students, and effects are arising purely through the compositional channel, then estimated effects on graduation rates would necessarily be symmetric for males and females. This test is performed by defining indicators for admissions score deciles A^k within colleges and estimating the following model separately for the set of students in each decile $k = 1, \ldots, 10$.

$$y_{ict} = \alpha_{ic} + \beta_{it} + \gamma_{ic}t + \delta^k F_{ict} + \varepsilon_{ict} \qquad (7)$$

Under the hypothesis of effects being compositional rather than behavioral, the extent to which δ^k (the coefficient on the female indicator representing the gender gap in the propensity to graduate) differs for marginally ($k = 1$) and inframarginally ($k > 1$) admitted students determines the extent to which the effects of freshman cohort gender composition may be asymmetric for males and females.

My investigation of within-classroom mechanisms involves investigating whether gender peer effects are evident when the peer group is defined as the set of students likely in the same set of classes together rather than the freshman cohort. Other compositional peer effects papers have looked at how actual within-classroom behavior responds to peer composition (Lavy and Schlosser, 2011; Hill, 2015), but no such data is available in this context. I model the semester-specific average GPA (there is no course-specific GPA) for student i in college c, department d, semester (time period) s and cohort t as a function of the department-specific female share in that semester P_{cdst}.

$$y_{icdst} = \alpha_{cd} + \beta_s + \theta_t + \gamma_{cd}t + \pi P_{cdst} + \varepsilon_{icst} \qquad (8)$$

College-department fixed effects α_{cd}, semester fixed effects β_s, cohort fixed effects θ_t and a college-department linear trend $\gamma_{cd}t$ are included in the model. The variation for identifying the parameter of interest π comes from semester deviations in department female shares and department average GPA scores from their long-term trends.

Students at the same college, entering in the same freshman cohort, and declaring the same major are not necessarily in the same academic peer group. I restrict the sample in various ways to ensure that the set of peers included in the definition are more likely to be taking the same courses. First, only juniors and seniors are considered. Even though freshman and sophomores may have declared a major department, it is likely that their classroom peers do not yet reflect this declaration. Second, the analysis is restricted to students who are deemed to be on-schedule, which is defined as completing at least 12 credit hours per a semester. This is because a junior who has completed only 16 credit hours is unlikely to be taking the same courses as a junior who has completed 24 credit hours even if they entered in the same freshman cohort and declared the same major.

The cohort effects are not specific to college departments, so only capture cohort shocks across college departments. The inclusion of college department-specific cohort fixed effects would result in identifying variation coming from fluctuations in peer gender composition across semesters for students in the same college department-cohort. This would be a concern because variation in gender composition across semesters for students in the same department and cohort is endogenous to the model given that it is likely generated by student failure or students changing majors, which may happen in response to either peer gender composition or GPA.

Equation 8 is still prone to bias if any endogenous changes to female shares and GPA scores that arise over the four junior and senior semesters do not average out across cohorts. I therefore also consider a model in which I restrict the sample to on-schedule juniors in their first junior semester. This specification necessarily excludes either the semester or cohort fixed effects. Using this specification, we no longer need to be concerned about endogenous changes to academic peer group composition that occur in the final two years of college – the period when students are likely taking major-specific courses with other students in their cohort pursuing the same major.

[APPENDIX?] In order to explore whether effects are heterogeneous across the types of college major, college-departments are divided into female share quintiles B^k. I include an additional interaction term between the college-department female share and an indicator for each quintile to explore whether peer gender effects in mostly male departments (such as engineering) differ from peer gender effects in mostly female departments (such as nursing).

$$y_{icdst} = \alpha_{cd} + \beta_s + \theta_t + \gamma_{cd}t + \pi P_{cdst} + \sum_{k=1}^{5} \pi_k \{P_{cdst} \times I(B^k)\} + \varepsilon_{icst} \quad (9)$$

3. Data

Three datasets are used in this paper. The primary analysis uses publicly-available data from the Integrated Postsecondary Education Data System (IPEDS). IPEDS collects information from every college that participates in federal student financial aid programs. I focus on public four-year colleges for which data is available over all the years in the study sample, a sample of 525 colleges. Freshman cohort gender composition is computed from gender-specific aggregate enrollment counts for each institution. Graduation information is available for cohorts that entered between 1996 and 2006. I consider six-year graduation rates as the primary outcome variable; results are similar using other definitions (see Appendix Table 1).

Figure 1 displays freshman cohort gender composition and graduation rates over time. The upper panel plots the average cohort female share both in the cohort's freshman year – the composition measure used in this paper – and upon the cohort's graduation. The average freshman female share is constant at 54 percent through the study sample, indicating that the female-male gender gap in college attainment was already present and stable by the mid-1990s. This is broadly consistent with existing literature (Fortin, Oreopoulos and Phipps, 2015). The average cohort female share at graduation is higher, indicating that admitted females are more likely to graduate than their male counterparts. This is also evident in the lower panel of Figure 1, which shows that female graduation rates are consistently above male graduation over time. Both are trending upward, possibly due to grade inflation or higher administrative and student costs associated with student failure.

The median freshman cohort consists of about 1000 students. We may be concerned that random shocks to composition across cohorts in the same college may generate insufficient variation for the empirical strategy to be effective. The left panel in Figure 2 shows considerable variation in cohort female shares across colleges and cohorts; the interquartile range is about 20 percent. Identifying variation, however, comes from within-college across-cohort variation in gender composition. The distribution of this is shown in the right panel. The standard deviation of the within-college female share is 2.5 percentage points. We will consider gender composition shocks of this magnitude when interpreting whether estimated effects on outcomes are of a meaningful magnitude.

The distributions of college graduation rates for males and females are plotted in Figure 3. Both the distributions across colleges and cohorts (left panel) and within colleges (right panel) are shown. The left panel shows that there are very large differences in college graduation rates across colleges. This is expected given large differences in college and student quality across the US. The right panel shows much smaller differences in cohort graduation rates when considering within-college variation, although these still have a standard deviation of just less than 5 percentage points for both males and females.

The second dataset I use is the Academic Insights product. These data are used to construct the popular US News and World Report college rankings, and are available to institutional subscribers. I use variables from this dataset to explore heterogeneity in gender peer effects by college type. The college characteristics considered describe how peers may interact outside the college classroom - the share of students in college housing, the share of students who live off campus, and the share of students with cars on campus – and college characteristics that describe the extent to which peers may interact inside the college classroom – the share of small classes and large classes in the college. The first set of variables contains most of the descriptions of the college living environment in the US News Academic Insights dataset, and the second set of variables contains most of the descriptions of the class environment in the dataset. I construct time invariant measures of these characteristics for each college by computing means for years in which the data is available.

Colleges are manually matched across the two datasets, resulting in a reduced sample of 165 public four-year colleges. The number of colleges is smaller than in the original sample because satellite campuses are often excluded from the Academic Insights product, which focuses on flagship schools, as well as the relevant college characteristics being missing for a small subset of schools. The original model is estimated on this smaller sample of colleges to confirm that the sample restriction does not affect the primary results. Means and distributions of the college characteristics are plotted in Appendix Figure 1.

The individual level data comes from the Texas Higher Education Opportunity Project (THEOP), a dataset containing information on the universe of students who applied to eight Texas four-year colleges during the 1990s and early 2000s, a period of affirmative action policy changes in Texas. Longitudinal transcript data are available at the student-semester level rather than the student-course level. I observe students' aggregate GPAs for all semesters in which the student is enrolled, the course credits obtained in these semesters, as well as their current departments and fields of study. There is no course-specific information.

The sample is restricted to students who intended to be admitted and were admitted during a fall semester (regular admission). This is to ensure that freshman cohorts are well-defined. Observations from summer terms are excluded from the analysis. Indicators for graduation and time-to-graduation are constructed from the recorded graduation year. I restrict the sample to cohorts who entered college between 1995 and 1999 when looking at gender peer effects on graduation. This is because graduation rates for earlier and later cohorts appear to be based on selected samples (observations in later years, for example, seem to be only the subset of students taking many years to complete their degrees rather than all the enrolled students). A student's total credit hours obtained is computed by summing credit hours obtained across all semesters in which the student is observed. Within-classroom mechanisms are explored using semester-level data as described in the empirical methodology section: course gender composition is approximated by the gender composition of on-schedule juniors and seniors in the same cohort majoring in the same department.

The THEOP data used in the paper contains information on over 80,000 students from eight Texas colleges. Graduation rates are slightly higher than in those reported in the IPEDS data: 59 percent for males and 66 percent for females. The number of departments per institution in the sample ranges from 29 at Rice University to 53 at Texas Tech University. Average department female shares range from around 20 percent in Computer Science and Engineering to above 90 percent in Nursing and Social

Work, but note that the empirical strategy relies on variation across cohorts within the same department to estimate gender peer effects.

4. Results

4.1. Effects of Freshman Cohort Composition on Graduation Rates Using Aggregate Data

The effect of freshman cohort gender composition on male and female graduation rates is reported in Table 1. This table establishes the primary result of the paper. It shows estimates of the parameter of interest π from Equation 1.

Using variation in cohort composition across colleges – the specifications without college fixed effects in columns 1 and 4 – an increase in the female share is associated with a reduction in both male and female graduation rates. The coefficients of -0.33 and -0.25 in columns 1 and 4 indicate that a 10 percentage point increase in the female cohort share is associated with a 3.3 and 2.5 percentage point reduction in male and female graduation rates, respectively. Colleges with higher shares of female freshman students have lower graduation rates on average. However, given that colleges with consistently different gender shares are likely different in several other dimensions, these estimates cannot be given a causal interpretation.

The remaining columns in Table 1 report results from models that include college fixed effects. These parameters are identified from within-college across-cohort variation in gender composition and provide the causal effect of cohort gender composition on graduation rates. Using either linear or quadratic college-specific time trends (columns 2 or 3 and columns 5 or 6), there is a positive effect on male graduation rates and no effect on female graduation rates. The coefficient of 0.2 means that a 10 percentage point increase in the share of female freshman students increases the male graduation rate by 2 percentage points.

Table 2 shows that the positive effect on male graduation rates is evident in colleges of all sizes. The concern in large colleges that cohort gender composition may be too weakly correlated with the gender composition of narrower peer groups (such as college residences) in which students directly interact and may have more influential peers is unfounded; in fact, the only colleges for which the effect is not statistically significant are colleges in the smallest quintile. Students in even the largest colleges experience cohort composition effects.

The estimates in the first column of Table 3 show that gender peer composition effects for male students are only statistically evident within a student's own cohort; the gender composition of adjacent cohorts has no effect on graduation rates. This is Equation 2 in the Empirical Strategy, and provides a falsification test to support the interpretation of effects. The second column shows no effect on female students, which is consistent with Table 1.

The magnitude of the estimated effect of the female share on male graduation rates in Table 1 to 3 is consistently around 0.2. The economic significance of this effect is interpreted by considering a one standard deviation composition shock to a representative college freshman cohort with 1000 male students and 1000 female students. The within-college standard deviation in the female share – the relevant variation for the empirical strategy – is 2.5 percentage points (Figure 2). Replacing 50 male students with 50 female students would increase the female share by 2.5 percentage points. We assume that 50 male students randomly substitute away from the given college (choose to attend another college) and 50 female students randomly substitute towards the given college.

The increase in the female share increases the male graduation rate by 0.5 percentage points (0.025*0.2). If the initial male graduation rate were 50 percent (the mean male graduation rate in the sample), 475 of the remaining 950 male students would graduate. The 0.5 percentage point increase in the graduation rate causes 480 male students to graduate: 5 more male students This is interpreted as a small, meaningful effect, especially given the labor market differences for individuals with four-year degrees and those with some college (Oreopoulos and Petronijevic, 2013).

Alternatively, the effect on a representative college of 1000 male students and 1000 female students can be reinterpreted to highlight the potential for effects to operate through the pure composition channel. Rather than assuming that the 50 male and 50 female students who substitute away and towards the given college are randomly changing their college choice, we could assume that they are marginally admitted students. With the graduation rate as before, 500 out of the original 1000 male students would graduate and 560 out of the original 1000 females would graduate (assuming the mean graduation rates of 0.50 and 0.56). After the shock to the cohort female share – replacing 50 males with 50 females – 480 out of 950 male students graduate. Given there is no effect on female graduation rates, the graduation rate for females remains at 0.56, so 588 out of 1050 female graduate.

If only 20 out of the 50 replaced male students would have graduated (a graduation rate of 40 percent for this subset of male students), the effect could be driven by a simple change in ability composition. Under this hypothesis, in order for there to be no effect on female graduation rates, the marginally admitted females must graduate at the same rate as inframarginally admitted females, so 28 out of the 50 added female students graduate (a graduation rate of 0.56 for this subset of female students). The observed pattern of effects is only consistent with the pure composition channel if shocks to cohort gender composition are driven by marginally admitted males with lower probabilities of graduation than their inframarginal male peers and marginally admitted females with the same probability of graduation as other admitted females. This possibility is explored using individual level data, but we first explore potential mechanisms using aggregate data.

4.2. Mechanisms of Compositional Peer Effects at College

The Empirical Strategy explains how freshman cohort gender composition effects may come about through both pure compositional and behavioral channels. The first test of the pure composition channel is provided by Equation 3, which explores the effects of freshman cohort gender composition throughout the distribution of cohort female shares rather than assuming the female share enters linearly. As discussed, effects on graduation are necessarily monotonic and linear in the cohort female share if changes in gender peer composition simply reflect changes in ability composition; nonmonotonicity or nonlinearity necessarily indicates the presence of behavioral effects.

We cannot rule out pure composition effects driving the effect for males when looking at estimates from Equation 3 plotted in Figure 4. Male graduation rates are lowest in cohorts with female shares in the first decile and increase consistently to the tenth decile (relative to the omitted fifth and sixth deciles). Figure 5 shows that there is no effect on female graduation rates, indicating that the absence of effects on female graduation rates in the primary specification is not driven by considering a linear-in-means model rather than a more flexible functional form. Although the estimates plotted in Figure 4 allow the presence of pure composition effects by providing no evidence of nonmonotonicity or nonlinearity, they cannot otherwise inform our understanding of the channel through which peer gender composition affects college outcomes.

I estimate models taking the form of Equation 4 to explore whether gender peer effects are more evident in colleges with certain characteristics. Heterogeneity in the effect by college type would provide evidence of behavioral effects as pure composition effects could not depend on the college environment. Results in Table 4 indicate that effects depend on college type. Freshman peer gender composition has more of an effect on male graduation rates in colleges with higher shares of students in college housing, colleges with lower shares of students living off campus, and colleges with lower shares of students with cars on campus (columns 2 to 4). For example, in colleges with 10 percent of students living in college housing, the effect of the proportion of female peers on the male graduation rate is only 0.09 (-0.04 + 0.96*0.1), while it is 0.28 and 0.46 in colleges with 30 and 50 percent of students living in college housing, respectively. Column 1 confirms that the overall effect in this smaller sample of colleges is similar to the overall effect for all colleges in Table 1.

Interestingly, and contrasting with previous estimates, the cohort female share also exerts a positive effect on female graduation rates in colleges with these characteristics (columns 8 to 10). Although the magnitudes of the estimates are smaller for females, which explains the absence of an average treatment effect when considering all colleges combined (column 7), this provides further support for the gender of college peers causing a behavioral response.

Results in columns 5, 6, 11 and 12 show that gender peer effects do not depend on the size of college classes in a statistically significant way, although the standard errors of these estimates are large. The absence of heterogeneity in the class size dimension suggests effects may not be operating inside the college classroom, although subsequent tests would be required to confirm this.

Taking these results as a whole, gender peer effects operate in colleges in which students are more likely to interact with their college peers: colleges with students living in college housing, on campus, and without cars. Idiosyncratic positive shocks to the freshman female share in these colleges appear to positively affect the general learning environment for both males and (possibly) females. Lavy and Schlosser (2011) argue that an increase in the proportion of girls within a grade improves academic outcomes through changes in the classroom environment. I cannot make this claim as strongly – they directly test for changes in the reported classroom environment – but results appear to be consistent with their findings that females exert positive externalities in learning environments.

4.3. Effects of Freshman Cohort Composition on Graduation Rates Using Individual Data

Table 5 shows that the effect on male graduation found in Table 1-3 can be replicated using individual level data from THEOP (Equation 5). The coefficient of 0.92 in column 1 means that a 10 percentage point increase in the freshman cohort female share increases the probability of graduation by 9.2 percentage points for an individual male student admitted to a four-year college in Texas during the study period. Column 2 adds flexible controls for admissions scores. Comparing students with the same ability (upon college admission), peer gender composition still affects the likelihood of graduation for males and is very similar in magnitude to the previous estimate without ability controls. There is also the expected positive effect on total credit hours obtained for males reported in column 3. Columns 4 to 6 show that there is no statistically significant effect of cohort gender composition on females, as was found in the primary specification using aggregate data.

The individual level data can also be used to further probe the mechanisms. Estimates plotted in Figures 6 show that the effect of the cohort female share on the likelihood of graduation is similar for male students of different abilities (Equation 6). The effect is not statistically distinguishable from the overall average treatment effect (the horizontal red line) for students in all but the second admissions score decile, although there is suggestive evidence that students of higher ability are less prone to gender peer effects. Figure 7 plots the corresponding estimates for females. These results suggest that very low ability female students may benefit from female peers and very high ability females may be harmed by female peers, although the null hypothesis of no cohort female share effects cannot be rejected for female students in all but the first two deciles. The absence of heterogeneity by ability provides evidence that effects are not driven by the gender and ability composition of marginally admitted students only.

As discussed in the last paragraph of the first Results subsection, for effects to operate purely through the composition channel given the pattern of estimated effects, it must be the case that marginally admitted males graduate at lower rates than inframarginally admitted males while marginally admitted females graduate at the same rate as inframarginally admitted females. This is directly tested by investigating whether the female-male graduation gap is larger for marginally admitted students (in the first and second admissions score deciles) than inframarginal students (in the remaining deciles) (Equation 7). Results are plotted in Figure 8, showing no difference in the gender gap across students in the different admission score deciles. This provides further evidence that the effect is not purely compositional.

The final table attempts to look at whether course gender composition affects course performance (Equation 8). If so, this would indicate that college peer effects operate inside the classroom, which would have different policy implications than if effects operate through cohort peers more generally. Estimates in Table 6 show a very small positive effect on male course achievement from course female share, and no effect on females. A 10 percentage point increase in the share of females majoring in the same department causes a 0.02 increase in average semester GPA per semester for male junior or senior students (column 1). The estimated effect is less precise when the sample is restricted to students in their first junior semester (column 2), and the inclusion of controls for lagged GPA (column 3) removes the effect for males. Overall, I interpret these results as indicating that within classroom effects are at most responsible for only a small component of the overall freshman cohort gender composition effect.

Conclusion

This paper provides the first evidence that female students improve the college learning environment for their male peers. An increase in the freshman cohort female share increases the cohort graduation rate for males at public four-year colleges in the US, while there is no overall effect for females.

The effect does not appear to be driven by females replacing marginally admitted males (which would increase the mean ability of admitted males holding all else equal). This interpretation is supported by a set of secondary findings. First, there is heterogeneity in the effect by college type. Generally, freshman cohort effects are more evident in colleges in which students are more likely to interact outside class. This could not be the case if effects operated through a purely compositional channel. Second, estimates using individual level data show that controlling for individual student ability does not affect the pattern of results, as well as indicate that effects on males are evident through most of the ability distribution. A male student of median ability is more likely to graduate if he is exposed to a higher share of females in his freshman cohort, which is not consistent with a purely compositional mechanism. Taken as a whole, these findings suggest that the peer effect mechanism is operating through direct or indirect effects on student behavior.

It is not feasible to significantly alter freshman peer gender composition in college, and it is clear that doing so would generate many benefits and costs for both individuals and society not considered in this paper. Results do, however, indicate that there may have been an unexpected benefit for admitted males from gender affirmative action policies favoring females in the past and a cost to potential affirmative action policies favoring males that may be implemented to reduce the female-male college attainment gap. Furthermore, to the extent that increasing the proportion of females improves the general learning environment, males may actually benefit from mixed gender residences rather than the conventional wisdom that opposite gender dorm mates provide a distraction.

One hypothesis is that the presence of female peers may reduce excessive partying on campus, which could be interpreted as consistent with Kremer and Levy (2008) who find negative spillovers from alcohol consumption among male college roommates. More evidence would be needed to substantiate such an argument, but it is one pointer for further research in the presence of intensifying debates about the college hook-up culture.

REFERENCES

1. Burke, Mary A., and Tim R. Sass. "Classroom peer effects and student achievement." *Journal of Labor Economics* 31, no. 1 (2013): 51-82.
2. Carrell, S., R. Fullerton, and J. West, “Does Your Cohort Matter? Measuring Peer Effects in College Achievement,” *Journal of Labor Economics* 27 (2009), 439–464.
3. Ciccone, A., and W. Garcia-Fontes, “Peer Effects When Boys And Girls Learn Together: A Birth Cohort Approach,” unpublished manuscript (2015).
4. Fortin, N. M., P. Oreopoulos, and S. Phipps, “Leaving Boys Behind,” *Journal of Human Resources* 50(3) (2015).
5. Garlick, R., “Academic Peer Effects with Different Group Assignment Policies: Residential Tracking Versus Random Assignment,” unpublished manuscript (2014). **[NOT CITED?]**
6. Gordon, Larry. "Mixed-gender dorm rooms are gaining acceptance." LA Times (2010).
7. Hill, Andrew J. 2015. "The Girl Next Door: The Effect of Opposite Gender Friends on High School Achievement." *American Economic Journal: Applied Economics*, 7(3): 147-77.
8. Hoxby, Caroline. Peer effects in the classroom: Learning from gender and race variation. No. w7867. National Bureau of Economic Research, 2000.
9. Jackson, C. Kirabo. "Single-sex schools, student achievement, and course selection: Evidence from rule-based student assignments in Trinidad and Tobago." *Journal of Public Economics* 96, no. 1 (2012): 173-187.
10. Jain, Tarun, and Mudit Kapoor. "The impact of study groups and roommates on academic performance." *Review of Economics and Statistics* 97, no. 1 (2015): 44-54.
11. Kremer, Michael, and Dan Levy. "Peer effects and alcohol use among college students." *The Journal of Economic Perspectives* 22, no. 3 (2008): 189-189.
12. Lavy, V., and A. Schlosser, “Mechanisms and Impacts of Gender Peer Effects at School,” *American Economic Journal: Applied Economics* 3 (2011), 1–33.
13. Lee, Soohyung, Lesley J. Turner, Seokjin Woo, and Kyunghee Kim. All or Nothing? The Impact of School and Classroom Gender Composition on Effort and Academic Achievement. No. w20722. National Bureau of Economic Research, 2014.
14. Lu, Fangwen, and Michael L. Anderson. "Peer Effects in Microenvironments: The Benefits of Homogeneous Classroom Groups." *Journal of Labor Economics* 33, no. 1 (2015): 91-122.
15. Oreopoulos, Philip, and Uros Petronijevic. Making college worth it: A review of research on the returns to higher education. No. w19053. National Bureau of Economic Research, 2013.
16. Sacerdote, B., “Peer Effects with Random Assignment: Results for Dartmouth Roommates,” *Quarterly Journal of Economics* 116 (2001), 681–704.
17. Stinebrickner, R., and T. Stinebrickner, “What Can Be Learned about Peer Effects Using College Roommates? Evidence from New Survey Data and Students from Disadvantaged Backgrounds,” *Journal of Public Economics* 90 (2006), 1435–1454.
18. Whitmore, Diane. "Resource and peer impacts on girls' academic achievement: Evidence from a randomized experiment." *American Economic Review* (2005): 199-203.
19. Zimmerman, D., “Peer Effects in Academic Outcomes: Evidence from a Natural Experiment,” *Review of Economics and Statistics* 85 (2003), 9–23.

FIGURES

Figure 1. Freshman cohort gender composition and graduation rates over time

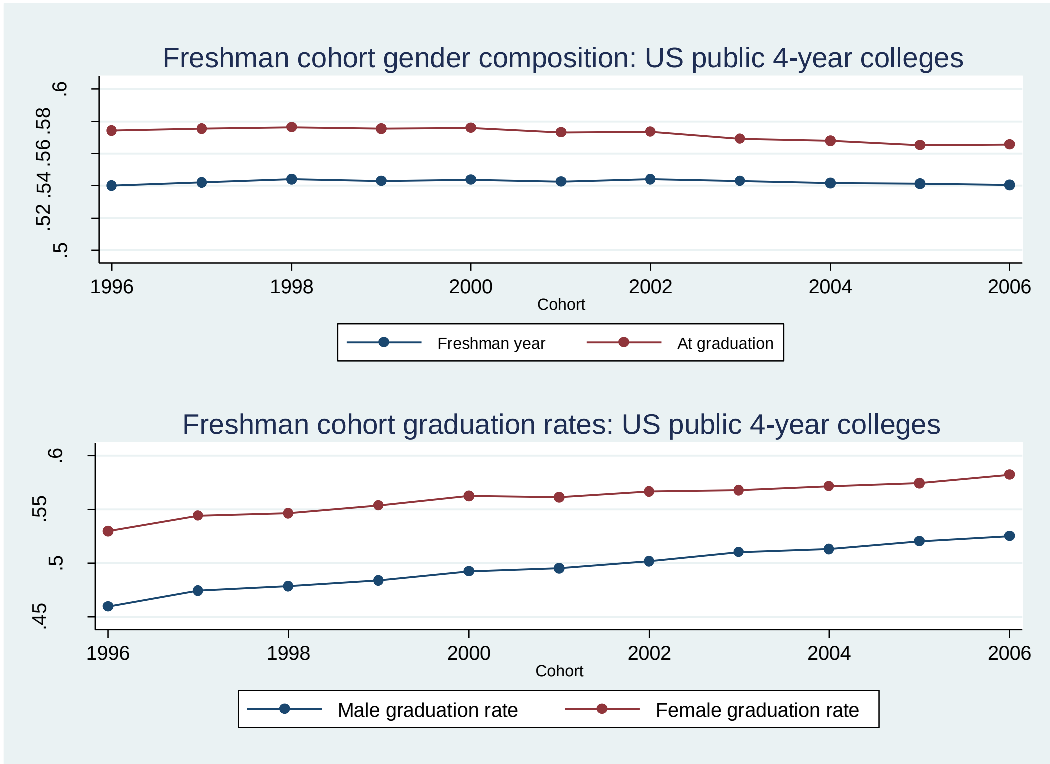


Figure 2. Distribution of freshman cohort gender composition

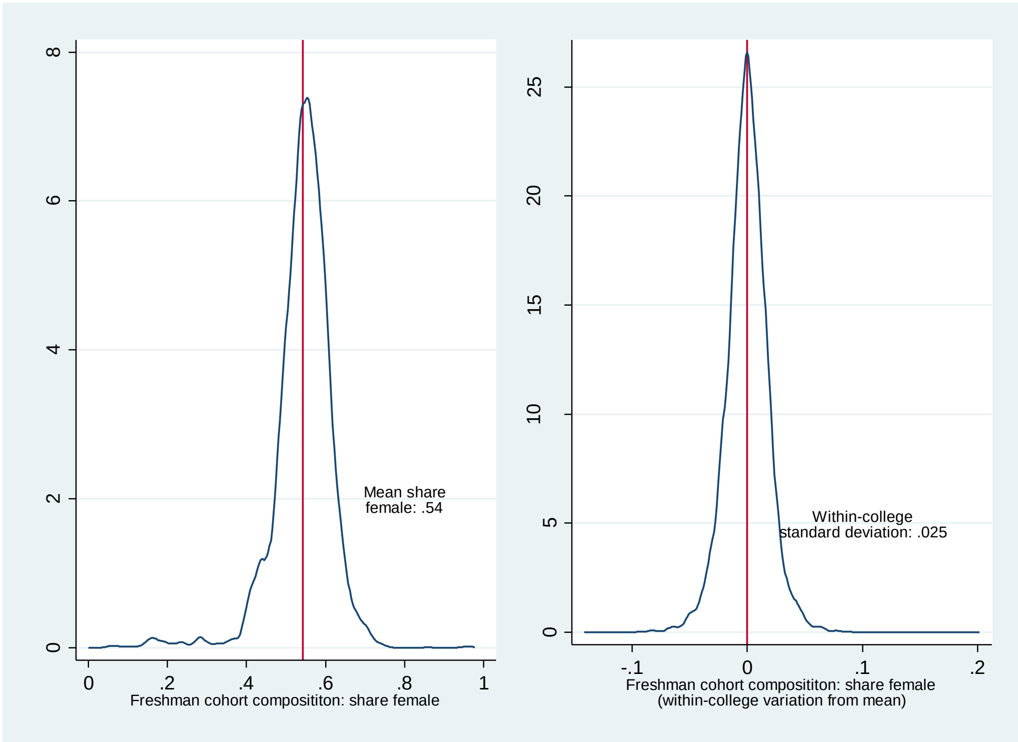


Figure 3. Distribution of college cohort graduation rates

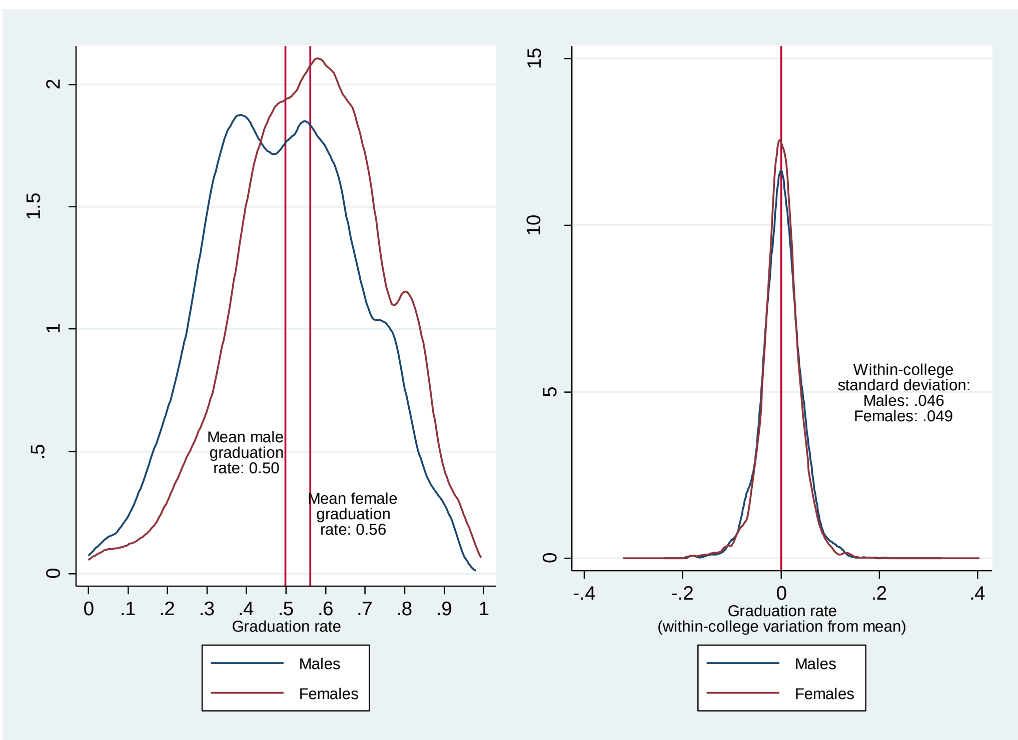


Figure 4. Estimated effect of freshman female share on males by share female deciles

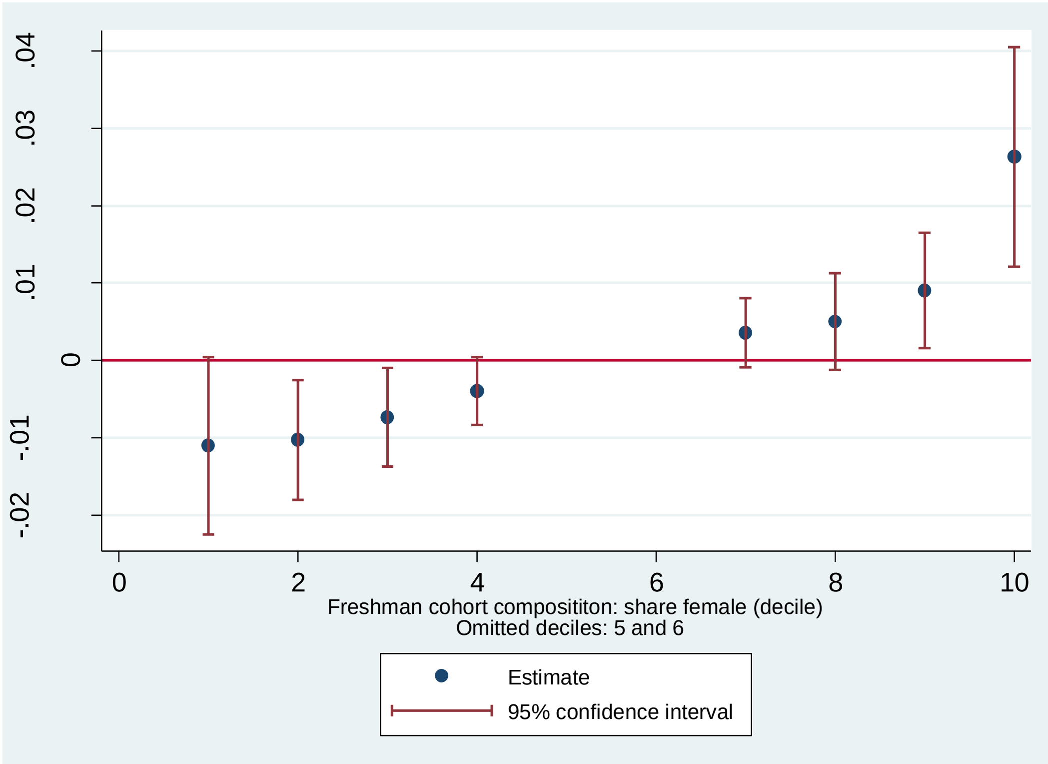


Figure 5. Estimated effect of freshman female share on females by share female deciles

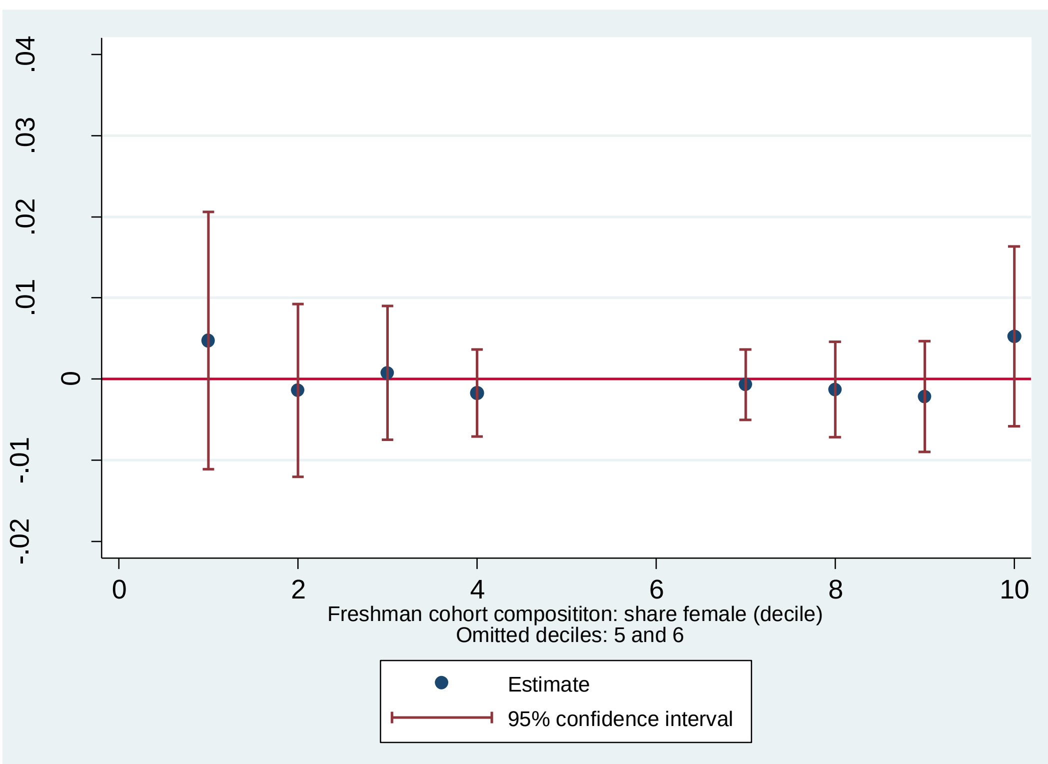


Figure 6. Estimated effect of freshman female share on males by admission score deciles

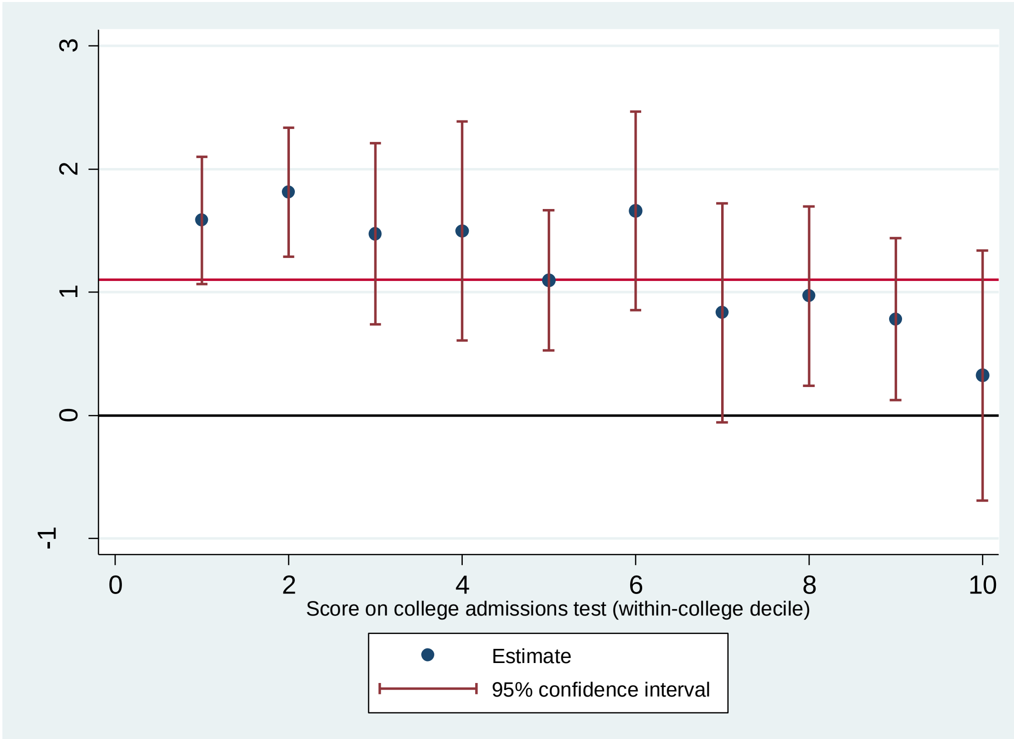


Figure 7. Estimated effect of freshman female share on females by admission score deciles

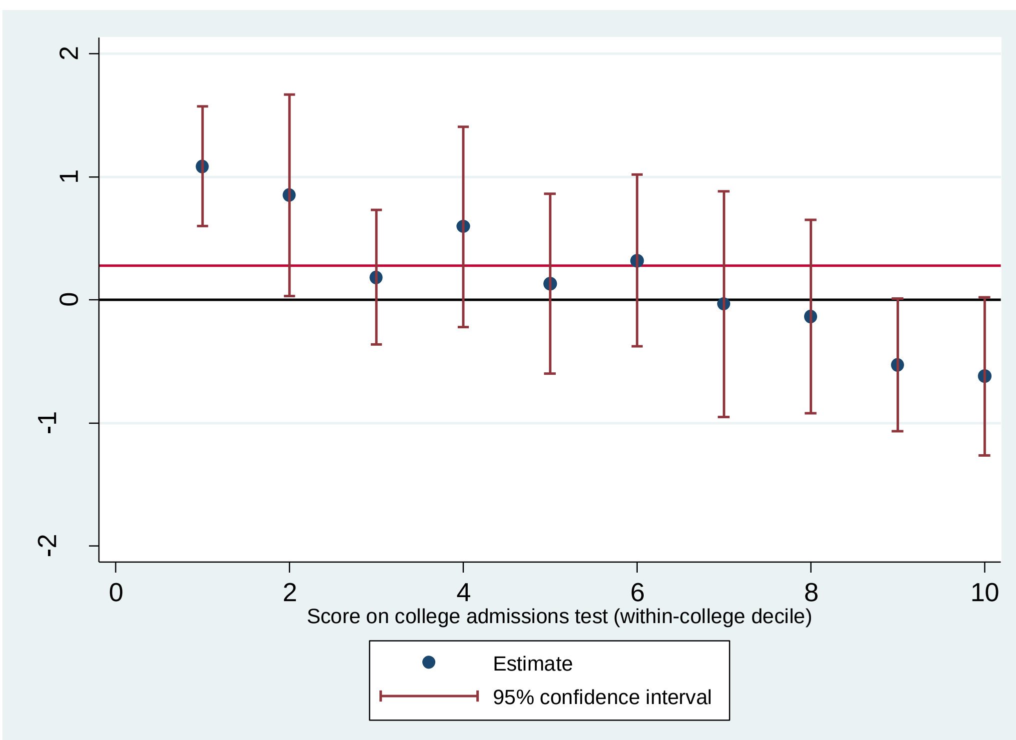
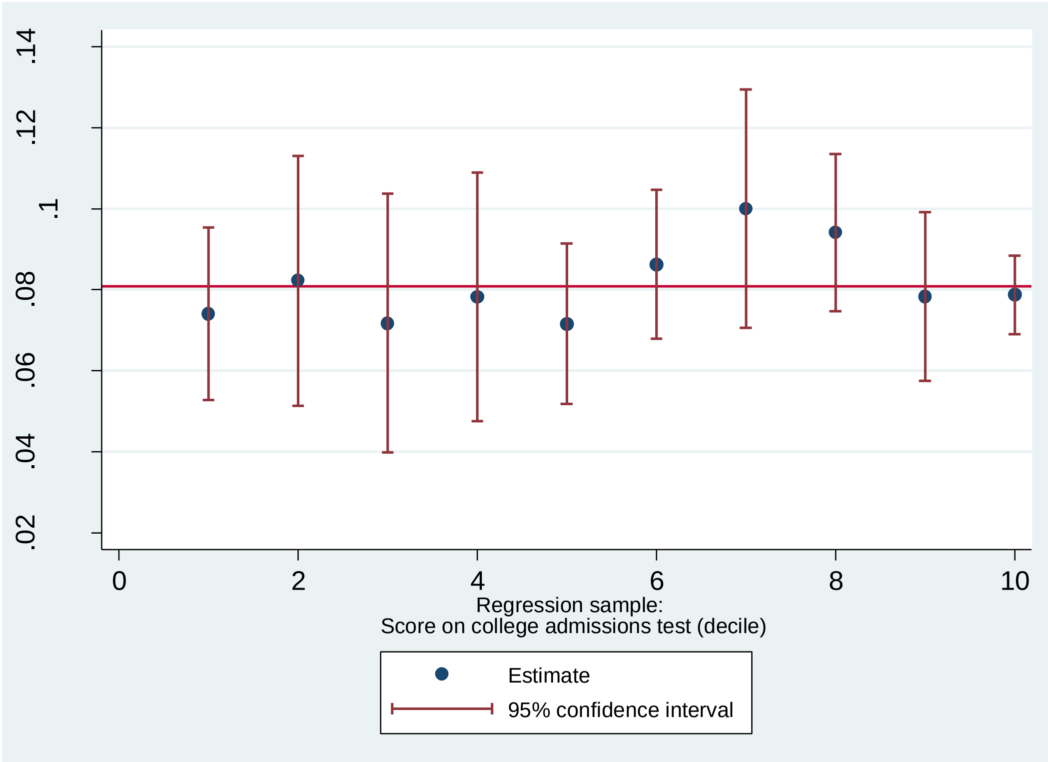
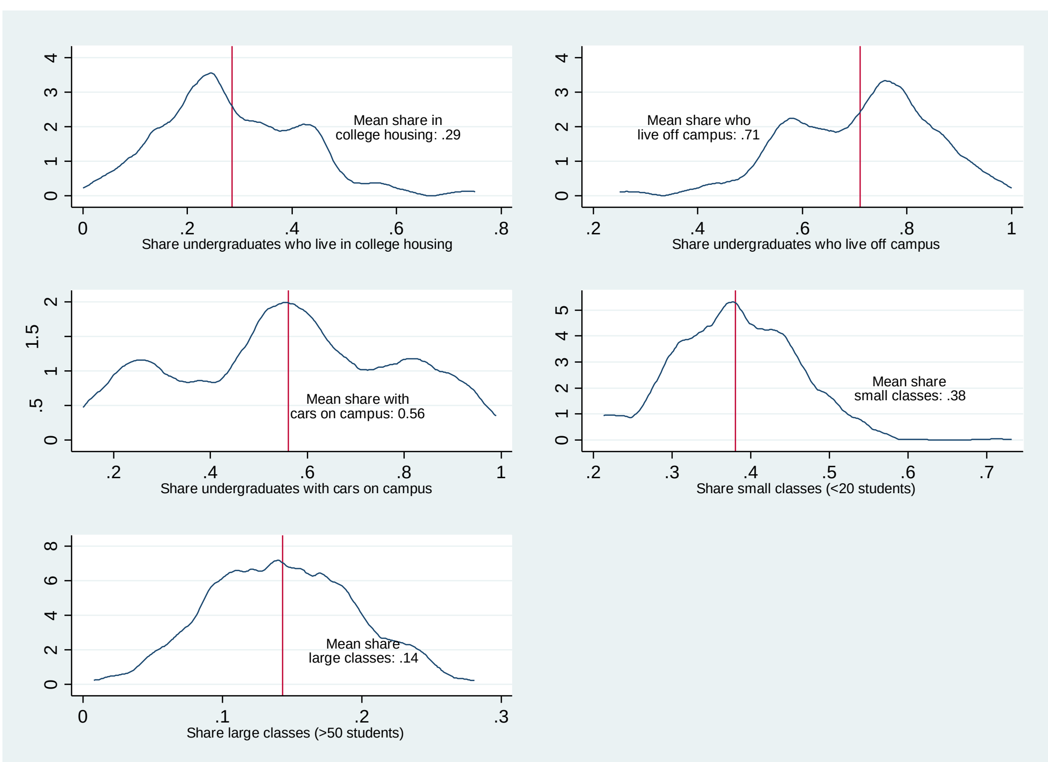


Figure 8. Estimated gender gap in graduation rate by performance on college admissions tests



Appendix Figure 1. Distributions of college characteristics



TABLES

Table 1. Effect of freshman cohort composition on overall college graduation rates

	Male graduation rates			Female graduation rates		
Mean	*0.50*			*0.56*		
	1	2	3	4	5	6
Share females	-0.33 **	0.19 **	0.17 **	-0.25 *	0.00	-0.01
	(0.11)	(0.04)	(0.04)	(0.12)	(0.04)	(0.04)
College fixed effects	No	Yes	Yes	No	Yes	Yes
College-specific time trend	Linear	Linear	Quadratic	Linear	Linear	Quadratic
Number of observations	5775					
Number of colleges	525					

Notes. Regressions weighted by gender-specific college cohort size and include cohort fixed effects and college-specific time trends. Robust standard errors clustered at the college level.
*** Significant at 1% level. * Significant at 5% level. + Significant at 10% level.*

Table 2. Effect of freshman cohort composition on overall college graduation rates by college size

	Male graduation rates									
College size quintile	*1*		*2*		*3*		*4*		*5*	
Mean freshman cohort	*343*		*679*		*1147*		*1871*		*4075*	
Mean share male	*0.34*		*0.36*		*0.43*		*0.39*		*0.61*	
	1		2		3		4		5	
Share females	0.06		0.16	*	0.20	*	0.21	+	0.19	*
	(0.06)		(0.07)		(0.08)		(0.12)		(0.08)	
Number of observations	1155		1155		1155		1155		1155	
Number of colleges	105		105		105		105		105	

Notes. Regressions weighted by gender-specific college cohort size and include college fixed effects, cohort fixed effects and quadratic college-specific time trends. Robust standard errors clustered at the college level.
*** Significant at 1% level. * Significant at 5% level. + Significant at 10% level.*

Table 3. Effect of lagged and leading freshman cohort composition on overall college graduation rates

	Male graduation rates		Female graduation rates
Mean	*0.50*		*0.56*
	1		2
Share females (t-2)	-0.03		-0.06
	(0.07)		(0.07)
Share females (t-1)	0.08		0.00
	(0.09)		(0.08)
Share females (t)	0.17	+	-0.09
	(0.09)		(0.09)
Share females (t+1)	0.08		-0.06
	(0.09)		(0.09)
Share females (t+2)	0.05		-0.09
	(0.07)		(0.08)
Number of observations	3675		3675
Number of colleges	525		525

Notes. Regressions weighted by gender-specific college cohort size and include college fixed effects, cohort fixed effects and quadratic college-specific time trends. Robust standard errors clustered at the college level.

*** Significant at 1% level. * Significant at 5% level. + Significant at 10% level.*

Table 4. Effect of freshman cohort gender composition interacted with college characteristics

College characteristic			Share of students in college housing		Share of students who live off campus		Share of students with cars on campus		Share of classes that are small (<20)	Share of classes that are large (>50)
Mean			*0.29*		*0.71*		*0.56*		*0.38*	*0.14*
	Male graduation rates *(Mean: 0.57)*									
	1		2		3		4		5	6
Share females	0.22	**	-0.04		0.89	**	0.90	**	-0.03	0.26
(standard error)	(0.07)		(0.12)		(0.33)		(0.25)		(0.34)	(0.18)
Share females *										
College characteristic			0.96	*	-0.93	*	-0.80	*	0.65	-0.29
(standard error)			(0.43)		(0.43)		(0.32)		(0.81)	(1.12)
	Female graduation rates *(Mean: 0.64)*									
	7		8		9		10		11	12
Share females	-0.03		-0.25	*	0.63	*	0.46	*	-0.15	-0.06
(standard error)	(0.07)		(0.13)		(0.30)		(0.21)		(0.40)	(0.16)
Share females *										
College characteristic			0.90	*	-0.88	*	-0.43		0.33	0.32
(standard error)			(0.39)		(0.40)		(0.28)		(0.96)	(1.00)
Number of observations	1815		1782		1749		935		1815	1815
Number of colleges	165		162		159		85		165	165

Notes. Regressions weighted by gender-specific college cohort size and include college fixed effects, cohort fixed effects and linear college-specific time trends. Robust standard errors clustered at the college level.
*** Significant at 1% level. * Significant at 5% level. + Significant at 10% level.*

Table 5. Effect of freshman cohort composition on individual student college graduation

	Males					Females				
	Graduated				Total credit hours	Graduated				Total credit hours
Mean	*0.59*				*94.79*	*0.66*				*96.26*
	1		2		3		4	5		6
Share females	0.92	*	1.10	*	38.87	*	0.02	0.29		-0.74
	(0.34)		(0.31)		(12.01)		(0.38)	(0.28)		(18.82)
Admissions score fixed effects	No		Yes		Yes		No	Yes		Yes
Number of students			43067					44391		
Number of colleges			8					8		

Notes. All regressions include college fixed effects, cohort fixed effects and quadratic college-specific time trends. Robust standard errors clustered at the college level.
*** Significant at 1% level. * Significant at 5% level. + Significant at 10% level.*

Table 6. Effect of college-cohort-term-department (academic peer group) gender composition on individual student term GPA

	Males			Females		
	All junior and senior terms	First junior term only		All junior and senior terms	First junior term only	
Mean	*3.09*	*3.05*		*3.26*	*3.22*	
	1	2	3	4	5	6
Share females	0.17 **	0.10 +	0.03	-0.01	0.02	0.05
(standard error)	(0.03)	(0.05)	(0.03)	(0.03)	(0.05)	(0.03)
[p-value]	0.00	0.07	0.41	0.66	0.76	0.12
Semester fixed effects	Yes	No	No	Yes	No	No
Lagged cumulative GPA fixed effects	No	No	Yes	No	No	Yes
Number of students	56695	46278	45602	65328	55149	54311
Mean semesters per student	2.8	1	1	2.8	1	1
Number of college-departments	311	302	302	323	315	315
Number of colleges	8	8	8	8	8	8

Notes. All regressions include college-department fixed effects, cohort fixed effects and linear college-department-specific time trends. Robust standard errors clustered at the college level.
*** Significant at 1% level. * Significant at 5% level. + Significant at 10% level.*

Appendix Table 1. Effect of freshman cohort composition on 4-year, 5-year and 6-year graduation rates

	Male graduation rates					
	4-year		5-year		6-year	
Mean	*0.23*		*0.20*		*0.06*	
	1		2		3	
Share females	0.09	**	0.07	**	0.01	
	(0.03)		(0.03)		(0.02)	
Number of observations			5775			
Number of colleges			525			

Notes. Regressions weighted by gender-specific college cohort size and include college fixed effects, cohort fixed effects and quadratic college-specific time trends. Robust standard errors clustered at the college level.
*** Significant at 1% level. * Significant at 5% level. + Significant at 10% level.*