# The Downside of Good Peers: How Classroom Composition Differentially Affects Men's and Women's STEM Persistence*

Stefanie Fischer
University of California, Santa Barbara
stefaniejfischer@umail.ucsb.edu

November 7, 2014

**Abstract**

Women are substantially less likely than men to graduate college with a STEM degree. This paper investigates whether class composition can help explain why women are disproportionately more likely to fall out of the STEM "pipeline". Identification comes from a standardized enrollment process at a large public university that randomly assigns freshmen to different mandatory introductory chemistry lectures. Using administrative data, I find that women who are enrolled in a class with higher ability peers are less likely to graduate with a STEM degree, while men's persistence in STEM is unaffected by class composition. I also show that the decline for women is most pronounced for those in the bottom third of the ability distribution. I rule out the possibility that this is driven solely by grades because both men and women receive higher grades in classes with higher ability peers. Overall, these results suggest that class composition as an important factor in determining STEM persistence for women and provide a novel explanation for part of the STEM gender gap in post-secondary education.

*JEL Codes: I20, I230, I240*
Keywords: Higher Education, Gender, STEM, Classroom Composition Effects

*I thank the office of Institutional Research, Planning and Assessment at UC Santa Barbara for access to administrative data. This paper has benefited from many helpful comments from Kelly Bedard, Peter Kuhn, Shelly Lundberg, Dick Startz, Maya Rossin-Slater, Jon Sonstelie, UCSB Labor Lunch Seminar participants, Members of the UCSB Human Capital Working Group, and Brian Duncan and the other participants in the 2014 Graduate Student Workshop affiliated with the WEAI annual conference. All errors are, of course, my own.

# 1 Introduction

Women exit STEM majors at much higher rates than men, which in part contributes to the gender-wage gap because they miss out on the sizable premiums associated with employment in these fields.[1] To date, little is known regarding why women drop out of the STEM pipeline in college. Recent research finds that the instructor-student gender match matters somewhat (Carrell, Page and West (2010) and Hoffmann and Oreopoulos (2009)), and that women's responsiveness to grades explains some of the phenomenon as well (Rask and Tiefenthaler (2008) and Ost (2010)), but much remains unexplained. Gaining a better understanding of the factors that cause women to leave STEM fields during college is important for developing policy aimed at bolstering this group's STEM persistence. This paper examines a novel pathway: *To what extent does the composition of a woman's introductory university STEM course impact STEM persistence?* By and large, these courses are large and competitive, and to the extent that women are more sensitive to these environments, a class composition which magnifies these features may induce marginal women to leave.

I investigate this question using administrative data from a large public research university. A unique feature of this particular setting is the quasi-experimental way in which students are assigned to their first chemistry course, which is a mandatory prerequisite for nearly all STEM majors at most universities, and therefore a significant gateway to success in STEM. Exploiting this quasi-experimental setting, I show that being in a class with higher ability peers reduces the probability that women graduate with a STEM degree and has no effect on men. More specifically, a 15% increase in the number of high ability students in a General Chemistry lecture (one standard deviation) reduces the probability that the average woman graduates with a STEM degree by 3.1 percentage points (6.8%). As one might expect, I further show that the effect is strongest in the bottom third of the math ability distribution. I rule out grades in the same course as the underlying mechanism by showing that there is a positive relationship for both men and women between peer

[1] Paglin and Rufolo (1990); Murnane, Willett and Levy (1995); Grogger and Eide (1995); Brown and Corcoran (1997); Weinberger (1999); Weinberger (2001); Murnane, Willett, Duhaldeborde and Tyler (2000); Rose and Betts (2004)

ability and grades in the STEM gateway course analyzed. These results are informative for at least two reasons. One, they are the first to show that classmate influences are an important factor in determining students' academic success in higher education, at least for women. Two, in contrast to most previous work (discussed below), I focus on STEM major completion rather than grades because STEM completion closely relates to occupation choice, which is an important piece to the gender wage gap story (Murnane et al. (2000) and Rose and Betts (2004)).

Very little is known regarding post-secondary classroom composition effects because isolating the causal impact is difficult. Often students, or more indirectly administrators, influence the student make-up of a classroom. Several studies at the elementary school level, which for the most part rely on data from the large scale randomized experiment Project STAR, find a positive relationship between average classmate ability and achievement (Whitmore (2005), Hanushek, Kain, Markman and Rivkin (2003), Boozer and Cacciola (2001), and Hoxby (2000)). Whether these results extend to a higher education setting, however, is an open question.

To date, the most convincing peer effects study in higher education – which estimates small positive effects on freshman grade point average – relies on the random assignment of students from the United States Air Force Academy to squadrons; which are essentially cohorts (Carrell, Fullerton and West (2009)).[2] This peer group measure is an improvement over previous studies, which define dorm-mates as the peer group, because squadrons capture a more comprehensive set of students' peer interactions.[3] None of these estimates, however, capture the effects that students within a classroom may have on individual outcomes because dorm-mates and squadron members do not necessarily attend the same classes.

There are several ways in which being assigned to a college lecture with relatively higher ability peers could affect students' outcomes. On one hand, this type of environment could be

[2]Lyle (2007) uses a similar military dataset (USMA) and cohort approach and finds no evidence of peer effects. A drawback of both of these studies is that students from military institutions likely are not representative of the general university population, especially women.

[3]The following studies use the random assignment of students to dorms to estimate peer effects and find mixed results. Stinebrickner and Stinebrickner (2006) find small positive peer effects on grades for women. Zimmerman (2003) and Sacerdote (2001) find small positive peer effects on students' grades, grade point average, and the take-up of social networks such as fraternities/sororities. Foster (2006) finds no evidence of peer effects.

performance enhancing. Students may benefit directly from higher ability classmates through knowledge spillovers during class, office hours, or out-of-class group study sessions. Additionally, the average class ability can affect the classroom standard and students may be motivated to work harder to keep up with their high achieving peers, consequently earning higher grades.

On the other hand, a high achieving classroom environment may be harmful in more subtle ways by negatively impacting self-perception. The higher the ability of the peers in a classroom, the harder it is to be ranked highly. This may make the environment more competitive. While in many situations competition can improve performance, contest theory suggests that large gaps in skills between competitors can have the perverse effect of reducing effort incentives. As such, marginal students in this "small fish in a big pond" environment may feel relatively weaker and either reduce their effort resulting in lower grades in STEM courses or exit the STEM pipeline altogether. Brown (2011) provides empirical evidence for this theoretical prediction by showing that the presence of a superstar in a PGA golf tournament is associated with lower performance by the other competitors. In a related vein, Niederle and Vesterlund (2007) and Garratt, Weinberger and Johnson (2013) show that women shy away from competition more than men and that the gender performance gap is exacerbated under competition (Gneezy, Niederle and Rustichini (2003)). These findings suggest that women may be more discouraged by the competitive STEM climate than men.

At least three aspects of the quasi-experimental nature of this setting make it favorable for studying class composition effects. First, the unique no-priority registration process at this institution leaves little room non-random class enrollment (see Section 3.2). Second, the introductory STEM course General Chemistry is a required prerequisite for most STEM majors and students cannot circumvent the course by apply Advanced Placement credits eliminating another avenue for selection. Third, the enrollment of upperclassmen into this introductory course generates substantial exogenous variation in classroom composition (see Section 3.3)

The remainder of the paper is organized as follows. Section 2 describes the data. Section 3 discusses identification and outlines the empirical strategy. Section 4 reports the results. Section 5 concludes.

## 2 Data

All data in this study are drawn from the University of California Santa Barbara (UCSB) administrative data system. UCSB is a research intensive public university with a large undergraduate population. These data include all students who are enrolled in the required introductory university STEM course, General Chemistry (CHEM1A), in a fall quarter during the years 1997 through 2007, and follows each of them through graduation. General Chemistry is a particularly favorable setting to study STEM because at UCSB, and at the majority of other universities, this course is the first prerequisite for most STEM majors[4] and tends to be difficult and competitive.[5] This gateway course is a yearlong sequence consisting of CHEM1A, CHEM1B and CHEM1C.[6] Moreover, STEM majors at UCSB are required to take the course at the university to advance to any follow-on courses. Students are not able to apply Advanced Placement credits to bypass this prerequisite, a strategy that is often used to circumvent introductory math and statistics.

The data contain two types of students, on-track (freshman) and late-track (sophomore or higher). On-track students are defined as taking CHEM1A in the fall quarter of their first year at the university. Late-track students are upperclassmen taking CHEM1A in a fall quarter other than their freshman year. Of this sample, roughly 85% of the observations are on-track students. The other 15% of students are late-track. There are 12,230 on-track students in the sample. Conditional on being a late-track student, 1,291 are upperclassmen who entered the university as freshmen and 621 are transfers. Anecdotally, there are several reasons to believe that late-track CHEM1A students differ on observable and unobservable characteristics from on-track students. First, a portion of the late-track students are transfers. Transfer students typically come from the local city college

[4] Appendix Table A1 lists the STEM majors that require CHEM1A.

[5] One indication of the relative difficulty of the CHEM1A sequence at UCSB, is the large fraction of campus tutoring services allocated to it. In the academic year 2012-2013, about 25% of all the students that used the Campus Learning Assistance Services (CLAS) were General Chemistry students.

[6] CHEM1A is offered in the fall, CHEM1B in the winter and CHEM1C in the spring.

and, on average, have lower high school grades and socioeconomic characteristics. Second, upperclassmen enrolled in CHEM1A, even if they are not transfer students, are behind schedule in their major since CHEM1A is the gateway course for almost all other STEM courses/majors. They are either behind because they switched to a STEM major at some point after their first year or because they needed a year of preparatory courses – remedial math and science – before taking CHEM1A.

Table 1 presents evidence that on-track students are on average higher achieving than late-track students on observable characteristics. Each column is a separate regression where the outcome is a different predetermined student characteristics. The variable of interest is a 0-1 indicator for being on-track. Year fixed effects are also included in the specification. As reported in Columns 2 and 4, late-track students appear to have lower high school grade point averages and SAT math scores; two characteristics that are predictors of STEM success**.** In fact, late-track students on average have two-thirds of a letter grade lower high school grade point average (see Figure 1).

One desirable feature of this dataset is that on-track students (freshmen) face a no-priority registration policy. On the other hand, late-track students have the ability to selectively enroll in CHEM1A. To alleviate concerns of possible selection issues on this margin, I analyze the outcomes of on-track students only. The only way in which late-track students enter the analysis is through the composition variable (the share of on-track students per lecture), which is the regressors of interest.[7]

Table 2 presents summary statistics for the main sample, which only includes on-track students. From 1997-2007 there are 46 CHEM1A lectures taught by 13 different instructors. The average lecture size is 329 students. On average, on-track students make up 85% of each lecture; the minimum is 71% and the maximum is 96%.[8] The main outcome of interest is STEM completion, defined as graduating with a STEM major from UCSB within five years. Among entering freshmen who take CHEM1A, the average STEM completion rate is 53% for men and 45% for women. Other outcomes used in this analysis are a student's grade in CHEM1A, whether a student takes

[7] While late-track students have the ability to selectively enroll, Table 4 shows that the observable characteristics for this group are orthogonal to class composition. See Section 3.3 for more details.

[8] If the sample is restricted to exclude CHEM1A lectures with 100 or fewer students, the percent freshmen per lecture ranges from 75-96%.

the direct follow-on course (CHEM1B), and a student's grade in CHEM1B. The average CHEM1A grade for males is a 2.65 GPA (on a 4 point scale) and a 2.49 for women. Of all freshmen who take CHEM1A, 85% of men and 80% of women continue on to CHEM1B where the average grade earned in CHEM1B for men and women is 2.59 and 2.58 respectively.

UCSB administrative data also include several socioeconomic measures: race, sex, high school grade point average, SAT math and verbal scores, type of high school (public or private), parents' highest education level, English proficiency and age. Limited information is also available regarding instructors and course times. These include an instructor's sex and a unique instructor identification number, as well as the year, day, and time of the lecture. These data are linked to students.

# 3 Empirical Strategy

## 3.1 Econometric Specification

The primary specification is the following linear probability model:

$$G_{itnd} = \alpha_1 + \beta_1 F_i + \beta_2 lnO_{tnd} + \beta_3 F_i * lnO_{tnd} + \alpha_2 C_{tnd} + \alpha_3 X_i + \alpha_3 M_i + \phi_t + \rho_n + \varepsilon_{itnd} \quad (1)$$

The variable $G_{itnd}$ denotes STEM major completion for an on-track student $i$ who takes CHEM1A in year $t$ with instructor $n$ at time of day $d$; $tnd$ uniquely identifies an individual lecture in a specific year. $F$ is a female indicator variable and $O_{tnd}$ is the number of on-track students in a specific lecture. The log transformation allows one to interpret the on-track estimate as a percent change and takes into account that a one student change is proportionately larger from a small base.[9] The coefficient $\beta_2$ captures the effect of the number of on-track students per lecture on the outcome for men. The coefficient on the interaction term $F_i * lnO_{tnd}$ is the differential effect of the number of on-track students per lecture for women. Thus, for women the percentage point change in STEM

[9] In alternative specifications, I also use percent of on-track students in a lecture as the measure of class composition and obtain quantitatively similar results. This measure is less desirable because it assumes that the marginal effect is constant regardless of the base.

graduation associated with a percent increase in the number of on-track students is $\beta_2 + \beta_3$. $C_{tnd}$ controls for several class level characteristics: the total number of students enrolled in a given lecture and the percent female. The lecture size variable includes the log of the total number of on-track and late-track students enrolled in a given lecture. $X_i$ is a vector of student background characteristics including: race, if a student went to public high school, if English is the only language spoken in the home, the highest level of education attainment of the parent with the highest level of education, high school grade point average, SAT math and verbal scores and age. $M$ indicates that the lecture took place in the morning (starting at 8 a.m. or 9 a.m.). Year fixed effects ($\phi_t$) are included to control flexibly for time trends in STEM completion. Since many instructors appear repeatedly, I include instructor fixed effects ($\rho_n$) to control for time-invariant instructor differences. All standard errors are clustered at the lecture level (instructor/year/time of day).

## 3.2 The Assignment of Students to Lectures

The aim of this study is to understand the differential impact of the number of on-track students in a general chemistry lecture relative to late-track students on STEM completion for men and women. In order to interpret $\hat{\beta}_2$ and $\hat{\beta}_3$ as causal, the variation in number and type of on-track students per lecture must be uncorrelated with unobservable student and instructor characteristics. For example, student preference for lectures with a certain concentration of on-track students would draw a causal interpretation into question. For my purposes, a primary concern is that $\hat{\beta}_2$ and $\hat{\beta}_3$ will be biased if on-track students with similar characteristics systematically enroll in lectures that have a higher (or lower) share of on-track students. To mitigate possible selection issues, only students in the fall quarter of their first year are included in the sample; Hoffmann and Oreopoulos (2009) take a similar approach.[10] This sample of students combined with UCSB's no-priority freshman enrollment process and the Chemistry Department's strictly enforced "no switching" rule leaves little room for selective enrollment for freshmen in CHEM1A.

[10] Freshmen in their first quarter of school have virtually no information about the instructors or classroom composition since they enroll in their first quarter courses prior to the start of school and they have very little flexibility in their class schedules to allow them to be strategic in course registration.

During my sample period, 95% of all first year students attend a fee-based two-day summer orientation either in June, July or August where they register for their fall quarter courses.[11] Importantly, there is no priority based registration during or before summer orientation. In each orientation session, a certain percent of the total seats available in a given "first year" course are made available to that particular orientation session. This equalizes the probability of enrolling in a certain lecture/course across all orientation sessions and eliminates the issue of students who attend earlier orientation dates getting all the "good" classes.

At orientation each student is assigned to a group of 15 students.[12] With this group, students attend seminars about university life and register for their first quarter classes under the guidance of a trained orientation leader. Only one student in an orientation group is able to register at a time, as there is only one laptop per group. Within each group a registration queue is formed by random draw (i.e. creating registration position 1- position 15). The student who draws position 1 registers first. Registration opens at the same time for all groups within an orientation session. According to the Director of Orientation Programs and Parent Services at UCSB, the high demand for CHEM1A, the limited seats, the random registration order, and student's general lack of information about instructors makes strategic CHEM1A registration essentially impossible. The one avenue for selection that one might be concerned with is time of day preferences. While there is no evidence that this is happening in a differential fashion, all specifications include controls for time of day. Making selection even harder, the Chemistry Department strictly enforces an no switching policy. A student can only switch lectures during the first week of the quarter and he must find a student in his desired lecture to replace him in his original one: a one for one switch.[13]

Balance tests for the main sample (entering freshmen cohorts from 1997-2007) are reported in Table 3. Each column in this table corresponds to a separate regression where a different predetermined student characteristic is regressed on $lnO_{itnd}$ and $F_i * lnO_{itnd}$. Year fixed effects are

[11] Each summer 12 freshman orientation dates are offered and students can attend the date of their choice.

[12] Students are placed into orientation groups by declared major, but groups within major are formed randomly.

[13] All information regarding freshman orientation and registration comes from an interview with Kim Equinoa (kim.equinoa@sa.ucsb.edu) who was the director of Orientation Programs and Parent Services at UCSB during the years in which the data for this analysis are from. Information on the Chemistry Department's no switching rule comes from the administrative office within the Chemistry Department.

also included. If selection is present, the coefficient on $lnO_{itnd}$ will attain statistical significance. Furthermore, if gender based selection exists, i.e. all high achieving women enroll in lectures with a low share of on-track students while high achieving men enroll in lectures with a high share of on-track students, the interaction term $F_i * lnO_{itnd}$ will be significant. Although the coefficient on $lnO_{itnd}$ is statistically significant in some cases, the magnitudes are nearly zero. More importantly for my purposes, there is even less evidence of a gender differential in selection across class composition. $F_i * lnO_{itnd}$ is statistically significant in a third of cases but again the magnitudes are minuscule. For instance, Column 2 shows that SAT math and the interaction term are positively related, however, the estimated coefficient indicates that an increase in the log of on-track students in a class of 15% has an additional positive effect for women of 1.7 SAT math points where the SAT math standard deviation is approximately 80 points. Not only are these estimates economically insignificant, the direction works against the main findings of this paper (women's STEM persistence and the average class ability are negatively related) as they are a downward bias. Based on the class composition, there is little evidence that students are selecting into lectures.

The one concern related to student selection that I cannot directly rule out is the possibility that students who attend summer orientation differ from those who do not. Students who do not attend summer orientation register for their fall classes in mid-September prior to the start of the quarter but after all orientation goers have registered. On-track student estimates will be biased if students who do not attend summer orientation are non-representative and systematically register for lectures based on the ratio of on-track to late-track students. If I could observe CHEM1A registration dates I could do a balance test on observables between students who attend summer orientation and those who do not. Since these data are unavailable for my sample period, I have instead obtained registration data for all freshmen enrolled in CHEM1A in fall 2013. Although these students are not in my main sample, the registration behavior should be similar, as the general structure of freshman registration is similar.

For this group of students – all freshmen enrolled in CHEM1A in fall of 2013 – I observe their CHEM1A registration date and time as well as demographics and CHEM1A instructor characteristics. Ninety percent of this sample attended a summer orientation/registration (slightly lower than the main sample). Comparing the observable characteristics of students who attend orientation and those who do not, the only group who is underrepresented in orientation attendance is underrepresented minorities; 38% of the orientation attending group are URMs compared to 49% of the non-orientation attending group. There appears to be no selection into orientation attendance by gender, parent's education level, whether one has a high school GPA in the top half of the distribution for the sample, whether one scores in the top half of the SAT math or SAT verbal distribution for the sample, type of high school one attended, and English language learner status. Most importantly, the data indicate that there is no statistically significant difference in the share of on-track students in a lecture or begin time of the lecture for those who attend summer orientation and those who do not. Appendix Table A2 reports these results.

## 3.3 Source of Variation in Classmate Ability

The approach used to identify classroom composition effects in this study is similar to that of Hoxby (2000). She estimates classroom peer effects among elementary age students by exploiting plausibly random variation in the gender and racial composition within a given grade and school over time. I exploit exogenous variation in the ability composition across CHEM1A lectures within a year and over time. Variation in CHEM1A classroom composition is driven by late-track student enrollment patters. Late-track students register for fall classes the previous spring, before on-track enrollment. Based on estimated freshmen fall enrollment, the university holds a fraction of the CHEM1A seats in each lecture for incoming freshmen; in some cases late-track students fill all of the seats allotted to them in a given lecture and in other cases they do not.

For instance, suppose for simplicity that there are 4 lectures and each has a maximum enrollment of 100. Further suppose that 30 percent of the seats (30 seats in this case) in each lecture in a given year are made available to late-track students. If in one of the lectures all 30 seats are filled, in the second only 25 are filled, and in the third and fourth only 20 and 15 respectively are taken, then variation will arise in the number of late-track students. The second stage of the process is that on-track students are assigned (virtually at random) to the remaining seats. Although in theory late-track students have the ability to select into specific lectures which is why I don't analyze their outcomes, the balance test reported in Table 4 show no sign this group is selecting into lectures based on the composition of the class. There is some evidence that better late-track students select morning lectures but, importantly for this paper's empirical approach, time of day is uncorrelated with the share of on-track students.

# 4 Results

Results from the main specification (Equation 1) – which estimates the differential impact of the number of on-track students in a class for men and women – are reported in Table 5. Column 1 of Table 5 reports results for the full sample and shows that increasing the number of on-track students in a class by 15% reduces the probability that a woman graduates with a STEM major by 3.1 percentage points (see Column 1, panel B). Increasing the number of on-track students by 15% in the average class is equivalent to adding 44 more on-track students to a class with 281 on-track students, which is about one standard deviation. To give context to the magnitude of the results, the average STEM graduation rate for women is 45% and 53% for men. Thus, a 15% increase in the number of on-tracks student in a lecture decreases the STEM graduation rate for an average woman from 45% to about 42%, which is a decrease of 6.6%. For men, Column 1 suggests that there is no statistically significant relationship between the ability of the students in his CHEM1A lecture and the rate at which he persists in STEM.[14]

While the aim of this analysis is to understand the total effect of a student's classmates on that student's STEM outcomes, it is worth noting that in theory the total estimated composition effect can embody three distinct effects; exogenous peer effects (also known as contextual effects), endogenous peer effects, and correlated effects. Correlated effects are present when individuals

[14] I re-run Equation 1 using a probit model rather than a linear probability model and obtain similar results. I also use percent on-track students in the class as well as number of on-track students in the class rather than natural log of the number of on-track students and obtain similar results.

in the same group behave similarly because they have similar individual characteristics (Moffitt et al. (2001)). This is often caused by students self-selecting into a group. The random assignment of students to lectures ensures that the composition estimates are free of correlated effects. Exogenous peer effects arise when a student's classmates' predetermined observable characteristics (high school grade point average, SAT scores etc.) affect her outcomes. Endogenous peer effects are present when a student's classmates' outcomes affect her outcome. For instance, although in practice this seems rather unlikely, if a student's classmates' decision to remain in STEM affects her decision to stay, then the composition estimates in this study will capture not only the the exogenous effect but also this endogenous effect.[15] That said, although in the peer effects literature it is often the goal of the empiricists to solve the reflection problem which involves isolating the exogenous effect net of the other two effects, in this this study I am interested in learning how classmates' affect a student's STEM completion. As such, my composition estimates are capturing this total effect and the presence of an endogenous effect does not undermine the empirical findings.

## 4.1 Sensitivity and Heterogeneity Analysis

One factor that threatens the causal interpretation of these results is student assignment to lectures with a particular composition based on characteristics that correlate with student STEM persistence. Although a balance tests reveal that students are statistically similar on observables across classes (Table 3), lectures added at the last minute to meet a larger than expected CHEM1A demand are a particular concern.[16] The data do not allow one to specifically identify whether a class is added at the last minute, however, the very small lectures (i.e. those with fewer than 100 students) are likely to be the added course if one exists. Importantly, these very small lectures also

[15]If one believes that endogenous effects are present and assuming that both the exogenous and endogenous effects are negative, my estimate of the "total effect" will overestimate the exogenous effect. That is, the estimate will be inflated by a social multiplier, which is commonly known as the reflection problem, and be more negative than the true exogenous effect.

[16]Although many years the Chemistry Department accurately estimates the demand for CHEM1A, there are cases where they add an additional lecture the week before the fall term begins.

appear to be correlated with the percent of on-track students in the class. The average percent of on-track students for these lectures is 72 compared to 86% for the entire sample. It is also the case that on-track students assigned to these smaller add-on lectures have the greatest potential to be non-representative. For instance, the small percent of on-track students who do not attend a summer orientation/registration session and also enroll in CHEM1A (which is on average 5% percent of an incoming freshman class) are most likely assigned to an add-on lecture during the first week of school. One would expect this non-summer orientation attending group of students to be less advantaged, thereby dampening the estimated on-track student effect found in the main specification.[17] Column 2 of Table 5 reports the estimates for the subsample which excludes lectures with fewer than 100 students; variation in percent on-track student per lecture ranges from 75 to 96%.[18] Results for this subsample indicate that potential add-on lectures are not driving the main findings. In fact, the magnitude of the estimated on-track student effects for women and men are not statistically different from the estimated effects using the whole sample.

Additionally, understanding which group of students is driving the main result is important for developing and implementing interventions. Columns 3-5 of Table 6 report results disaggregated by SAT math score. The effect is strongest for women in the bottom third of the SAT math distribution. A 15% increase in the number of on-track students in a class reduces the probability by 5.7 percentage points that a women in this SAT math group completes college with a degree in STEM, and this effect is statistically different from the estimated effects in the other two SAT categories (Columns 4 and 5). Consistent with the main finding, the men in all subsamples appear to be unaffected by the classroom composition. While it is intuitive that women in the lower part of the math ability distribution are the group most affected by classroom composition since they are the group most at risk of dropping out of STEM, these results oppose the findings in Carrel et al. (2010). They find that the group influenced by STEM interventions are women at the top of the SAT math distribution. In particular, they document that women in the top 25% of the SAT math

[17] Non-orientation attending students are likely less advantaged because summer orientation is an additional cost. According to UCSB office of Orientation Programs and Parent Services, the most common reason students do not attend orientation is due to summer employment.

[18] A balancing test for the subsample is statistically the same as the balancing test for the main sample.

distribution with female STEM instructors are more likely to graduate with a STEM major.

One might wonder if the results truly are a gender effect. It is possible that I am capturing an underrepresented minority effect or merely picking up the fact that all students at the bottom end of the SAT math distribution are less likely to graduate with a degree in STEM. Columns 1 and 2 of Table 6 report results from the specifications outlined in Equations 2 and 3 respectively. These models, which are extensions of equation (1), include a triple interaction term allowing one to disentangle differences in the on-track student effect across gender and race (Column 1), as well as gender and position in the SAT math distribution (Column 2).

$$G_{itnd} = \alpha_1 + \beta_1 F_i + \beta_2 lnO_{tnd} + \beta_3 URM_i + \beta_4 F_i * URM_i + \beta_5 lnO_{tnd} * URM_i + \beta_6 F_i * lnO_{tnd} + \beta_7 F_i * URM_i * lnO_i + \alpha_2 C_{tnd} + \alpha_3 X_i + \alpha_3 M_i + \phi_t + \rho_n + \varepsilon_{itnd} \quad (2)$$

$$G_{itnd} = \alpha_1 + \beta_1 F_i + \beta_2 lnO_{tnd} + \beta_3 Low_i + \beta_4 F_i * Low_i + \beta_5 lnO_{tnd} * Low_i + \beta_6 F_i * lnO_{tnd} + \beta_7 F_i * Low_i * lnO_i + \alpha_2 C_{tnd} + \alpha_3 X_i + \alpha_3 M_i + \phi_t + \rho_n + \varepsilon_{itnd} \quad (3)$$

$URM_i$ denotes whether a student is an underrepresented minority (black, Hispanic, American Indian or Filipino) and $Low_i$ indicates whether a student falls in the bottom third of the SAT math distribution for the sample in a given year. All other variables are as defined in Section 3. Column 1 shows that all women, regardless of race, have STEM persistence rates that are negatively affected by the number of on-track students in her CHEM1A class. As reported in Column 1 Panel B, URM and non-URM women experience a 3.0 percentage point decline in STEM persistence as a result of an increased number of on-track students. Again, there is no detectable class composition effect for men, URM or non-URM.

Results presented in Table 6 Column 2 further support a gender story. These results show that only women (and not men) in the bottom third of the SAT math distribution for the sample have STEM persistence rates that are affected. In fact, women in this group are 4.0 percentage points less likely to graduate in STEM as a result of a 15% increase in the number of on-track students in a class. The results for the women are statistically different from zero and statistically different from men in this same SAT math group.

$$G_{itnd} = \alpha_1 + \beta_1 F_i + \beta_2 lnO_{tnd} + \beta_3 B_i + \beta_4 F_i * B_i + \beta_5 lnO_{tnd} * B_i + \beta_6 F_i * lnO_{tnd} + \tag{4}$$

$$\beta_7 F_i * Low_i * lnO_i + \alpha_2 C_{tnd} + \alpha_3 X_i + \alpha_3 M_i + \phi_t + \rho_i + \varepsilon_{itnd}$$

Finally, socioeconomic status may also play a role in one's willingness to leave STEM when placed in a lecture with a higher share of on-track students. I use a similar triple difference specification – as outlined in Equation 4 – and examine the differential effect of on-track concentration on STEM persistence by gender and by parent's level of education. $B_i$ indicates if a student has at least one parent with a bachelor's degree. Results presented in Column 3 of Table 6 show that all women, regardless of whether her parent is a college graduate, have an increased probability of exiting STEM. Consistent with all other specifications, the persistence rate for men in all subgroups seems to be statistically unrelated to the composition of the class. Together, these findings provide strong evidence that *women* in the bottom third of the math ability distribution are the group most affected by the ability of their classmates. There is no evidence to support the conjecture that it is merely reflecting minority status, being in the bottom of the ability distribution, or socioeconomic status.

## 4.2 Possible Mechanisms

These reported findings raise the question: *Why are women in the bottom third of the math ability distribution less likely to graduate with a STEM degree if their first experience with STEM is in a setting with higher ability classmates, and why are men unaffected by this factor?* Because little is known regarding post-secondary classroom composition effects and student outcomes in general, less is known about the mechanisms at work, and in particular why composition matters more for women. One relevant study – a project funded by the National Science Foundation and conducted by the Goodman Research Group – surveyed roughly 25,000 undergraduate women enrolled in engineering programs across 53 institutions between 1999-2001 with the goal of identifying ".. aspects of women's educational experiences that are critical to their retention in engineering."

Although engineering students comprise only a small part of the greater group of women in STEM, lessons learned from the engineering study likely have some relevance for other STEM areas since many early prerequisites are common across STEM majors. This study finds that women who leave engineering are most likely to leave in their freshman or sophomore year. The top two reasons for leaving are: (1) they are dissatisfied with their grades and/or the heavy workload, and (2) they dislike the overall climate of the major, including the competitive nature and discouraging faculty and peers. Both of these factors may be directly influenced by the class composition.

### 4.2.1 Grades

First, I investigate grades in the initial course as a possible mechanism. In addition to the study by the Goodman Research Group, several studies in economics find that women are more responsive to grades than men, and as a result exit STEM majors.[19] If all CHEM1A classes are graded on a similar curve (i.e. 10 percent of the class earns an A grade, 35 percent earns a B grade etc.), then students in lectures with more high ability classmates will receive lower grades relative to their counterpart (those in lectures with fewer on-track students). For instance, suppose that there are two CHEM1A lectures with equal enrollment in a given fall quarter and one has more on-track students than the other. Relying on the fact that lectures with more on-track students are overall higher ability (Table 1), then a student receiving a score of 77% in the lecture with more on-track students will be assigned a lower final letter grade than if she was in a class with fewer on-track students.

To explore this possibility, I use a specification similar to Equation 1 with a student's CHEM1A grade as the outcome. CHEM1A grade is a variable taking on values from 0 to 4. If grades are the underlying mechanism, students in classes with more on-track peers should receive lower final grades. The results presented in column 1 of Table 7, however, show no sign of this. In fact, both men and women experience a bump in CHEM1A grade as a result of an increase in the

[19] The average grade in STEM courses is much lower than humanities, social sciences, arts and interdisciplinary courses. Rask and Tiefenthaler (2008) and Ost (2010) show that women in STEM are more responsive to grades than men.

number of on-track students in a class. Columns 2-7 further show that this result is robust to a variety of subsamples. It appears that all students – male, female, URM, non-URM, the low SAT math scoring group and high SAT math scoring group – experience a marginal increase in their grade. These results are also robust to the subsample which excludes lectures with fewer than 100 students. Quantile regression results (available upon request) further reveal that a positive effect is detectable at all places in the distribution for both men and women which suggests that the overall grade distribution for classes with a higher share of on-track students is shifted to the right. Moreover, if grades are driving the result, controlling for CHEM1A grade in the main specification (Equation 1) should diminish the composition estimates. Table 8 reports such estimates and shows that the effect remains despite controlling for initial course grade, and the differential effect persists as well.

It is possible that the grade findings are a result of positive peer effects as it has been shown that higher ability peer groups elicit higher individual grades (Stinebrickner and Stinebrickner (2006); Han and Li (2009); Zimmerman (2003); Sacerdote (2001); Carrell et al. (2009)). Given the structure of the data, however, I cannot rule out that the positive grade effect is merely an artifact of instructors adjusting their grade scales based on the overall ability of the students. For instance, positive effects will emerge if instructors increase everyone's grade in a class because on-whole they are high achieving; a channel that instructor fixed effects will not capture because this behavior is time-varying.

### 4.2.2 Course Climate

General Chemistry is *competitive* and women, when given the option, are more likely to select out of competitive environments (Niederle and Vesterlund (2007), Garratt et al. (2013)). The competitive environment in the introductory chemistry course comes from at least three source, (1) the course is required for most STEM majors and students are required to keep a "C" average in the introductory sequence to advance to upper division courses, (2) because many STEM majors are graduate school and medical school prerequisites, students are motivated to maintain a high grade point average, and (3) grades are assigned based on a curve. Increasing the share of high ability students therefore increases competition in an already competitive environment.

Additionally, the composition of the class could affect students' *self-perception* about their immediate and future success in the major. Presumably, all students enter the initial course with an expectation about how they will do. Throughout the course they learn about their relative standing and update beliefs about themselves accordingly. Individuals in lectures with relativity higher ability classmates may adjust these believes differently than those who are not. For example, Pop-Eleches and Urquiola (2011) show that students who just make it into better high schools receive a bump in exam scores but also report feeling marginalized and relatively weaker compared to students who are placed in classes with lower ability classmates. To the extent that women's self-perception about their future success is more negatively affected by the ability of those around them, it could explain their much lower retention rate. Along these lines, if women are more *risk averse*, then the marginal women may switch to majors where they perceive having a higher chance of "making it" while marginal men gamble by staying in STEM. Consistent with this idea, Kuziemko, Buell, Reich and Norton (2014) show that men are more likely to gamble to avoid low rank whereas women accept it.[20]

There are many reasons to believe that women might be turned off by the climate in STEM while men are not. As a result, marginal women may become discouraged and either exert lower effort or quit STEM altogether. Some evidence supporting the idea that women become discouraged is in the take-up of the direct follow-on course, CHEM1B. Twenty percent of women exit the general chemistry sequence after the initial course compared to 15% of men and some of this is attributable to the composition of the introductory course. Table 9 reports results for a linear probability model similar to Equation 1 where the outcome is equal to one if a student takes CHEM1B and zero otherwise. This table shows that women's and men's CHEM1B take-up is unrelated to CHEM1A classroom composition for the whole sample (Column 1), but when broken out by SAT math subgroups the results indicate that CHEM1B take-up for women in the bottom third of the

[20] Although data for this study comes from the laboratory and manipulates an individual's rank in the wealth distribution, it is reasonable that the detected behavioral response extends to a classroom ability distribution.

SAT math distribution – which is the group with STEM major completion most affected by classroom composition – is negatively related to the number of on-track students in a class (Column 5). The estimated effect is statistically different from zero and statistically different from the effect for men in this same group, but due to the noisy estimates in the other subsamples, I cannot reject that the effect for women is the same as the estimated effects for women from the other SAT math groups (the middle third and the top third). These results are robust to a specification that controls for CHEM1A grade as well.

Next, I estimate a small negative relationship between number of on-track students and CHEM1B grade for women and find no statistically significant effect for men (Table 10, Column 1). An increase in the number of on-track students of 15% is related to a reduction in a woman's CHEM1B grade of 0.08 grade points which is 27% of a letter grade. Consider this estimate understated (less negative than one would expect) because some students – particularly those near the bottom of the ability distribution, see Table 9 – have already exited the pipeline as a result of the introductory class composition and are no longer in the sample. It is very likely that this follow-on grade finding reflects a student's overall discouraged feeling – i.e. lack of effort – particularly if she feels marginalized in the introductory course. It is also possible that learning or mastering the fundamental skills needed to successfully complete a STEM degree is lower for women in introductory classes with higher quality students. However, this later explanation, is not supported in my data, as I find a positive relationship between the ability composition in a classroom and CHEM1A grade for women.

Finally, I show that women are responding to the composition of their introductory course by switching into majors that are relatively less quantitative.[21] Table 11, Column 2 reports that increasing the number of on-track students in a class by 15% leads to 3.3 percentage point increase in the probability that a woman graduates with a humanities, social sciences, art, or interdisciplinary major.[22] This table shows that women are still graduating (as shown in column 3) but on

[21] Appendix Table A3 outlines by sex the percent of students in each major category at entry and at graduation.

[22] I get similar results when I exclude Economics (Econ, Econ-Math, and Econ-Accounting) and Psychology from this group.

average they are graduating in majors that are, lower paying, less quantitative and arguably less competitive.

Although I can not directly point to the channel by which high ability classmates adversely affect women's STEM retention, I can rule out that the effect is operating through grades. There are, however, various other channels through which the climate may discourage women including competition, self-perception and risk aversion. Consistent with this notion, I show that women both directly after the initial course and at other points in their STEM major pursuit give-up on STEM as a response to the share of high ability classmates, in favor of majors that have a more female friendly environment.

# 5 Discussion and Conclusion

It has been well documented that women are less likely than men to persist in STEM majors and careers. This study targets a unique group of students, those taking General Chemistry in their first quarter of college, to better understand how one's first collegiate experience in STEM explains STEM major graduation rates. Relying on data containing roughly 12,000 first year university students from 11 entering cohorts between 1997-2007, I estimate how the ability of one's classmates, as measured by the share of on-track students per class, in a required STEM major course affects a student's STEM major completion.

In summary, women who are assigned to a STEM lecture with higher ability peers early in their university career are less likely to persist in STEM. I show that this result is driven by women in the bottom third of the SAT math distribution. Men in this same SAT bracket (or any subsample for that matter) do not experience this same negative relationship between classmate ability and STEM persistence.

I have ruled out the possibility that women earn lower grades in classes with higher achieving classmates and as such are less likely to persist. The number of on-track students per class and grades in the initial course are positively related. On the other hand, I cannot rule out the possibility that women's decision to exit is a response to the climate created in classes with higher ability peers. In fact, I find some evidence consistent with a story that marginal women are deterred by the climate, become discouraged, and eventually exit to majors with a more female friendly environment. I show that at least some of these women are leaving the STEM pipeline immediately after the initial General Chemistry course, as a share of them do not persist to the direct follow-on as a result of the ability of their classmates in the initial course.

This study is the first to provide an analysis of the relationship between class composition and STEM degree completion in higher education, and to document the differential response by gender. Although these estimates do not provide direct policy implications, they do reveal (1) an important group for policy to target, and (2) suggest that, for women, a behavioral response may be present. Thus, if the goal of policy is to bolster the participation of women in STEM fields, deepening our understanding of the channels through which classmates affect a woman's STEM behavior is important.

# References

**Boozer, Michael and Stephen E Cacciola**, "Inside the'Black Box'of Project STAR: Estimation of peer effects using experimental data," *Yale Economic Growth Center Discussion Paper*, 2001, (832).

**Brown, Charles and Mary Corcoran**, "Sex-Based Differences in School Content and the Male-Female Wage Gap," *Journal of Labor Economics*, 1997, *15* (3), 431–65.

**Brown, Jennifer**, "Quitters never win: The (adverse) incentive effects of competing with superstars," *Journal of Political Economy*, 2011, *119* (5), 982–1013.

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

**_ , Richard L Fullerton, and James E West**, "Does Your Cohort Matter? Measuring Peer Effects in College Achievement," *Journal of Labor Economics*, 2009, *27* (3), 439–464.

**Foster, Gigi**, "It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism," *Journal of public Economics*, 2006, *90* (8), 1455–1475.

**Garratt, Rodney J, Catherine Weinberger, and Nick Johnson**, "The state street mile: Age and gender differences in competition aversion in the field," *Economic Inquiry*, 2013, *51* (1), 806–815.

**Gneezy, Uri, Muriel Niederle, and Aldo Rustichini**, "Performance in competitive environments: Gender differences," *QUARTERLY JOURNAL OF ECONOMICS-CAMBRIDGE MASSACHUSETTS-*, 2003, *118* (3), 1049–1074.

**Grogger, Jeff and Eric Eide**, "Changes in college skills and the rise in the college wage premium," *Journal of Human Resources*, 1995, pp. 280–310.

**Han, Li and Tao Li**, "The gender difference of peer influence in higher education," *Economics of Education review*, 2009, *28* (1), 129–134.

**Hanushek, Eric A, John F Kain, Jacob M Markman, and Steven G Rivkin**, "Does peer ability affect student achievement?," *Journal of applied econometrics*, 2003, *18* (5), 527–544.

**Hoffmann, Florian and Philip Oreopoulos**, "A professor like me the influence of instructor gender on college achievement," *Journal of Human Resources*, 2009, *44* (2), 479–494.

**Hoxby, Caroline**, "Peer effects in the classroom: Learning from gender and race variation," Technical Report, National Bureau of Economic Research 2000.

**Kuziemko, Ilyana, Ryan W Buell, Taly Reich, and Michael I Norton**, "Last-place Aversion: Evidence and Redistributive Implications," *The Quarterly Journal of Economics*, 2014, *129* (1), 105–149.

**Lyle, David S**, "Estimating and interpreting peer and role model effects from randomly assigned social groups at West Point," *The Review of Economics and Statistics*, 2007, *89* (2), 289–299.

**Moffitt, Robert A et al.**, "Policy interventions, low-level equilibria, and social interactions," *Social dynamics*, 2001, *4*, 45–82.

**Murnane, R.J., J.B. Willett, and F. Levy**, "The growing importance of cognitive skills in wage determination," Technical Report, National Bureau of Economic Research 1995.

**_ , _ , Y. Duhaldeborde, and J.H. Tyler**, "How important are the cognitive skills of teenagers in predicting subsequent earnings," *Journal of Policy Analysis and Management*, 2000, *19* (4), 547–568.

**Niederle, Muriel and Lise Vesterlund**, "Do women shy away from competition? Do men compete too much?," *Quarterly Journal of Economics*, 2007, *122*, 1067–1101.

**Ost, Ben**, "The role of peers and grades in determining major persistence in the sciences," *Economics of Education Review*, 2010, *29* (6), 923–934.

**Paglin, M. and A.M. Rufolo**, "Heterogeneous human capital, occupational choice, and male-female earnings differences," *Journal of Labor Economics*, 1990, pp. 123–144.

**Pop-Eleches, Cristian and Miguel Urquiola**, "Going to a better school: Effects and behavioral responses," Technical Report, National Bureau of Economic Research 2011.

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

**Rose, H. and J.R. Betts**, "The effect of high school courses on earnings," *Review of Economics and Statistics*, 2004, *86* (2), 497–513.

**Sacerdote, Bruce I**, "Peer Effects with Random Assignment: Results for Dartmouth Roommates," *Quarterly Journal of Economics*, 2001, *116* (2), 681–704.

**Stinebrickner, Ralph and Todd R 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*, 2006, *90* (8), 1435–1454.

**Weinberger, Catherine J**, "Mathematical college majors and the gender gap in wages," *Industrial Relations: A Journal of Economy and Society*, 1999, *38* (3), 407–413.

**Weinberger, C.J.**, "Is Teaching More Girls More Math the Key to Higher Wages," *Squaring Up: Policy Strategies to Raise Womens Income in the US*, 2001.

**Whitmore, Diane**, "Resource and peer impacts on girls' academic achievement: Evidence from a randomized experiment," *American Economic Review*, 2005, pp. 199–203.

**Zimmerman, David J**, "Peer effects in academic outcomes: Evidence from a natural experiment," *Review of Economics and Statistics*, 2003, *85* (1), 9–23.

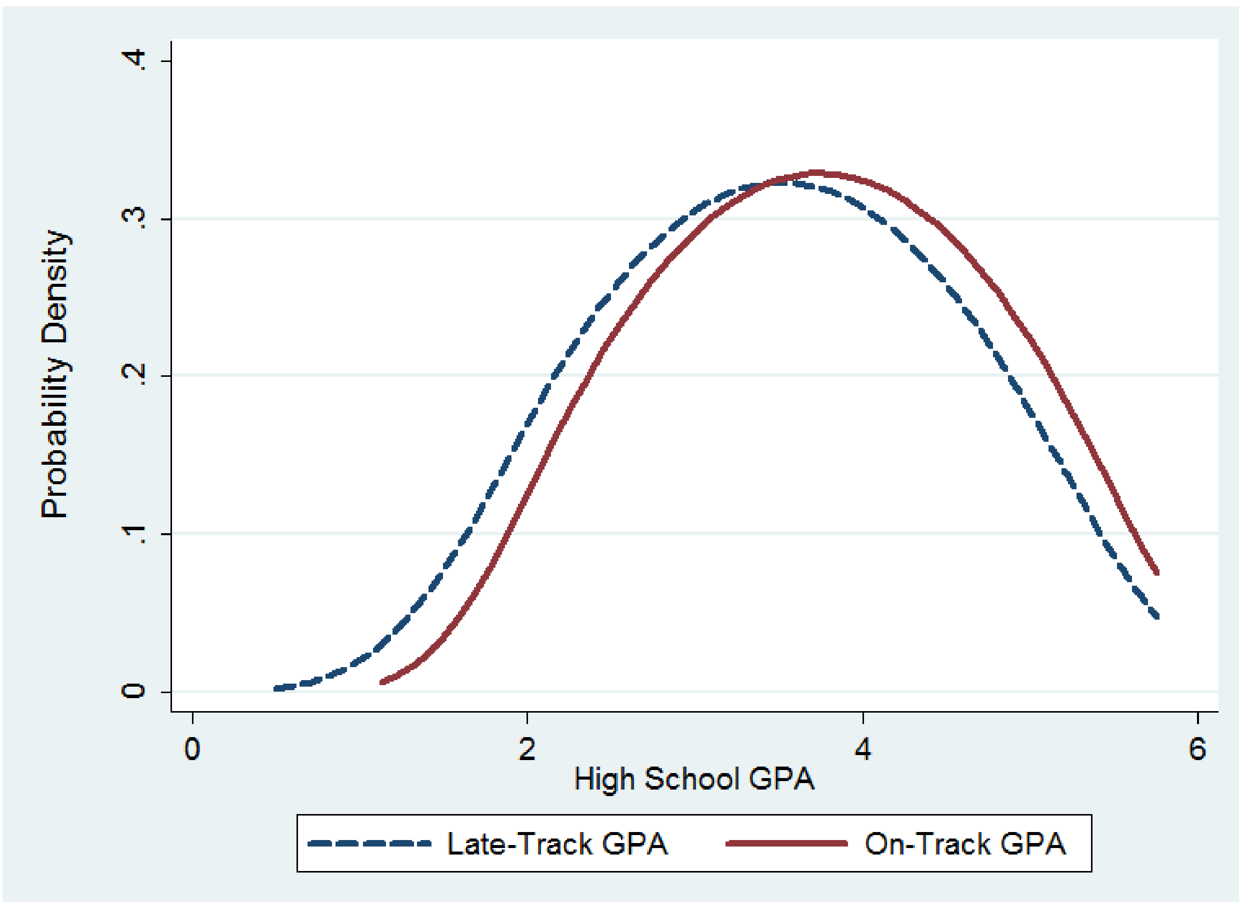


Figure 1: Differences in High School GPA Between On-Track and Late-Track

Table 1: On-Tracks are Relatively Higher Achieving than Late-Tracks

| | On-Track | Late-Track | Diff. (1) - (2) |
|---|---|---|---|
| | (1) | (2) | (3) |
| **Student Background Characteristics** | | | |
| Women | 0.49 | 0.52 | -0.03*** |
| | (0.01) | (0.01) | (0.01) |
| URM (underrepresented minority) | 0.32 | 0.32 | 0.01 |
| | (0.47) | (0.47) | (0.01) |
| High school grade point average | 3.75 | 3.52 | 0.23*** |
| | (0.32) | (0.44) | (0.01) |
| SAT math score | 612.74 | 604.34 | 8.40*** |
| | (80.84) | (80.95) | (2.25) |
| SAT verbal score | 569.58 | 575.91 | -6.34*** |
| | (84.59) | (83.66) | (2.33) |
| Attended public high school | 0.85 | 0.86 | -0.01 |
| | (0.36) | (0.35) | (0.01) |
| English is only language spoken in home | 0.67 | 0.69 | -0.02* |
| | (0.47) | (0.46) | (0.01) |
| No parent graduated from college | 0.33 | 0.30 | 0.03*** |
| | (0.47) | (0.46) | (0.01) |
| **Outcomes** | | | |
| Graduate with STEM major | 0.49 | 0.44 | 0.06*** |
| | (0.50) | (0.50) | (0.01) |
| Graduate | 0.81 | 0.82 | -0.014 |
| | (0.40) | (0.38) | (0.01) |
| CHEM1A grade | 2.57 | 2.26 | 0.31*** |
| | (0.93) | (1.11) | (0.03) |
| Took follow-on course (CHEM1B) | 0.83 | 0.63 | 0.20*** |
| | (0.38) | (0.48) | (0.01) |
| Grade in follow-on course (CHEM1B) | 2.58 | 2.53 | 0.05* |
| | (0.87) | (0.97) | (0.03) |
| **Observations** | 12,230 | 1,935 | 14,165 |

Notes: On-track students are enrolled in the first quarter of General Chemistry (CHEM1A) in the fall quarter of their freshman year at UCSB from 1997 to 2007. Late-track students are enrolled in CHEM1A during this time frame but are taking the course as an upperclassman or transfer student. URM stands for underrepresented minorities and includes all race categories except white, Asian and Indian. Level of significance is indicated as follows: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard deviations are in parentheses.

Table 2: Summary Statistics (sample includes on-tracks only)

| | Women (1) | Men (2) |
|---|---|---|
| **Classroom Characteristics** | | |
| % on-track in a lecture | 0.85 | 0.85 |
| | (0.05) | (0.05) |
| CHEM1A lecture size | 329.37 | 328.53 |
| | (46.73) | (47.37) |
| **Student Background Characteristics** | | |
| URM (underrepresented minority) | 0.34 | 0.31 |
| | (0.47) | (0.46) |
| High school grade point average | 3.80 | 3.70 |
| | (0.31) | (0.33) |
| SAT math score | 587.59 | 636.41 |
| | (78.16) | (75.49) |
| SAT verbal score | 564.88 | 573.97 |
| | (82.63) | (85.60) |
| Attended public high school | 0.86 | 0.84 |
| | (0.35) | (0.37) |
| English is only language spoken in home | 0.69 | 0.65 |
| | (0.46) | (0.48) |
| No parent graduated from college | 0.36 | 0.29 |
| | (0.48) | (0.46) |
| **Outcomes** | | |
| Graduate with STEM major | 0.45 | 0.53 |
| | (0.50) | (0.50) |
| Graduate | 0.82 | 0.80 |
| | (0.39) | (0.40) |
| CHEM1A grade | 2.49 | 2.65 |
| | (0.95) | (0.91) |
| Took follow-on course (CHEM1B) | 0.80 | 0.85 |
| | (0.40) | (0.36) |
| Grade in follow-on course (CHEM1B) | 2.58 | 2.59 |
| | (0.87) | (0.87) |
| **Observations** | 5,942 | 6,288 |

Notes: The sample includes only on-track students, those enrolled in the first quarter of General Chemistry (CHEM1A) in the fall quarter of their freshman year at UCSB between the years 1997 and 2007. URM stands for underrepresented minorities and includes all race categories except white, Asian and Indian.

Table 3: Balance Tests – Are On-Track Students Selectively Enrolling?

| | URM | SAT Math | SAT Verbal | H.S. GPA | Parent is College Grad | Public H.S. |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Panel A: Point Estimates** | | | | | | |
| Ln(no. of on-track) | -0.076*** | -0.043 | 5.032 | 0.043** | 0.064** | 0.001 |
| | (0.028) | (4.696) | (5.040) | (0.020) | (0.028) | (0.021) |
| Ln(no. of on-track) X Fem. | 0.025 | 12.66** | 7.110 | -0.011 | 0.030 | -0.053* |
| | (0.039) | (6.325) | (6.906) | (0.027) | (0.039) | (0.027) |
| **Panel B: Estimated effects in %-pts. associated with a 15% increase in no. of on-track** | | | | | | |
| The differential effect | 0.30 | 1.80** | 1.00 | 0.00 | 0.40 | -1.00* |
| Women | -0.70* | 1.80*** | 1.70** | 0.00 | 1.00*** | -1.00*** |
| Men | -1.00*** | -0.01 | 0.70 | 0.01** | 1.00*** | 0.00 |
| **Observations** | 12,230 | 12,142 | 12,142 | 12,160 | 12,230 | 12,230 |

Notes: Each column is a separate regression and also includes year fixed effects. On-track students are those enrolled in the first quarter of General Chemistry (CHEM1A) in the fall quarter of their freshman year at UCSB between the years 1997 and 2007. URM stands for underrepresented minorities and includes all race categories except white, Asian and Indian. Level of significance is indicated as follows: *** p<0.01, ** p<0.05, * p<0.1. Standard deviations are in parentheses.

Table 4: Balance Tests – Are Late-Track Students Selectively Enrolling?

| | URM | SAT Math | SAT Verbal | H.S. GPA | Parent is College Grad | Public H.S. |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Panel A: Point Estimates** | | | | | | |
| Ln(no. of on-track) | -0.003 | -4.549 | -1.5072 | 0.041 | 0.006 | -0.045 |
| | (0.051) | (11.6906) | (9.730) | (0.043) | (0.055) | (0.039) |
| Ln(no. of on-track) X Fem. | -0.056 | -5.082 | -6.653 | 0.062 | 0.020 | 0.068 |
| | (0.078) | (14.890) | (14.680) | (0.065) | (0.081) | (0.056) |
| **Panel B: Estimated effects in %-pts. associated with a 15% increase in no. of on-track** | | | | | | |
| The differential effect | -1.00 | -0.70 | -0.90 | 0.01 | 0.00 | 1.00 |
| Women | -1.00 | -1.40 | -1.10 | 0.01* | 0.00 | 0.00 |
| Men | 0.00 | -0.60 | -0.20 | 0.01 | 0.00 | -1.00 |
| **Observations** | 1,935 | 1,443 | 1,443 | 1,896 | 1,935 | 1,935 |

Notes: Each column is a separate regression and also includes year fixed effects. On-track students are those enrolled in the first quarter of General Chemistry (CHEM1A) in the fall quarter of their freshman year at UCSB between the years 1997 and 2007. URM stands for underrepresented minorities and includes all race categories except white, Asian and Indian. Level of significance is indicated as follows: *** p<0.01, ** p<0.05, * p<0.1. Standard deviations are in parentheses.

Table 5: The Effect of the Number of On-Track Students on STEM Major Completion

| | Full Sample | Lectures > 100 | Bottom 1/3 of SAT Math | Middle 1/3 of SAT Math | Top 1/3 of SAT Math |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| **Panel A: Point Estimates** | | | | | |
| Ln(no. of on-track) | -0.100 | -0.037 | -0.232 | -0.71 | -0.057 |
| | (0.105) | (0.103) | (0.171) | (0.299) | (0.138) |
| Ln(no. of on-track) X Fem. | -0.112*** | -0.137*** | -0.166** | 0.010 | -0.055 |
| | (0.030) | (0.045) | (0.067) | (0.061) | (0.077) |
| Instructor, Year, Time of day FE | X | X | X | X | X |
| Student Characteristics | X | X | X | X | X |
| **Panel B: Estimated effects in %-pts. associated with a 15% increase in no. of on-track** | | | | | |
| Women | -3.10** | -2.40* | -5.70** | -0.90 | -1.60 |
| Men | -1.50 | -0.50 | -3.20 | -1.00 | -0.80 |
| **Observations** | 12,230 | 12,122 | 4,206 | 3,438 | 4,586 |

Note: Each column is a separate specification. Controls include percent female in a class, log class size, year and instructor fixed effects, whether the lecture was held in the morning, a vector of student background characteristics, and a student's declared major at entry. Student background characteristics include: gender, race (black, Hispanic, Asian, American Indian, Filipino, Indian and white is the omitted group), if a student went to public high school, if English is the only language spoken in the home, the highest level of education attainment of the parent with the highest level of education, high school grade point average, SAT math and verbal scores and age. Clustered standard errors are in parentheses , *** p<0.01, ** p<0.05, * p<0.1. Clusters are by CHEM1A lecture (class). A 15% increase in number of on-track students in a class is the equivalent of increasing the number of on-track students by about 1 standard deviation (44 students). URM stands for underrepresented minorities and includes all race categories except white, Asian and Indian.

Table 6: Heterogeneity Analysis – STEM Major Completion for Various Subgroups

| | URM Effect? | Low Ability Effect? | Low SES Effect? |
|---|---|---|---|
| | (1) | (2) | (3) |
| **Panel A: Point Estimates** | | | |
| Ln(no. of on-track) | -0.096 | -0.116 | -0.143 |
| | (0.114) | (0.106) | (0.119) |
| Ln(no. of on-track) X Fem. | -0.125** | 0.802* | -0.087 |
| | (0.060) | (0.472) | (0.070) |
| Ln(no. of on-track) X URM | -0.050 | | |
| | (0.070) | | |
| Ln(no. of on-track) X Fem. X URM | 0.033 | | |
| | (0.101) | | |
| Ln(no. of on-track)X bottom 1/3 | | 0.013 | |
| | | (0.064) | |
| Ln(no. of on-track) X Fem. X bottom 1/3 | | -0.144* | |
| | | (0.084) | |
| Ln(no. of on-track) X Parent col. grad | | | 0.054 |
| | | | (0.063) |
| Ln(no. of on-track) X Fem. X Parent col. grad | | | -0.042 |
| | | | (0.087) |
| **Panel B: Estimated effects in %-pts. associated with a 15% increase in no. of on-track** | | | |
| Non-URM – women | -3.10** | | |
| Non-URM – men | -1.30 | | |
| URM – women | -3.30** | | |
| URM – men | -2.00 | | |
| Bottom 1/3 – women | | -4.00** | |
| Bottom 1/3 – men | | -1.40 | |
| Top 2/3 – women | | -2.00 | |
| Top 2/3 – men | | -1.60 | |
| Col. grad parent – women | | | -3.00** |
| Col. grad parent – men | | | -1.20 |
| No col. grad parent – women | | | -3.20** |
| No col. grad parent – men | | | -2.10 |
| **Observations** | 12,230 | 12,230 | 12,230 |

Note: Each column is a separate specification. The Column 1 regression also includes a dummy for URM and an interaction term between URM and woman. URM stands for underrepresented minorities and includes all race categories except white, Asian and Indian. The Column 2 regression also includes a dummy for being in the bottom 1/3 of the SAT math distribution and the interaction between being in the bottom and a woman. The Column 3 regression also includes a dummy for having at least one parent with a college degree and the interaction between being that dummy and woman. Additionally, all three regressions include controls for percent female in a class, class size, year and instructor fixed effects, whether the lecture was held in the morning, a vector of student background characteristics, and a student's declared major at entry. Student background characteristics include: gender, race (black, Hispanic, Asian, American Indian, Filipino, Indian and white is the omitted group), if a student went to public high school, if English is the only language spoken in the home, the highest level of education attainment of the parent with the highest level of education, high school grade point average, SAT math and verbal scores and age. Clustered standard errors are in parentheses , *** p<0.01, ** p<0.05, * p<0.1. Clusters are by CHEM1A lecture (class). A 15% increase in number of on-track students in a class is the equivalent of increasing the number of on-track students by about 1 standard deviation (44 students).

Table 7: The Effect of the Number of On-Track Students on CHEM1A Grade

| | Full Sample | Lectures > 100 | Non-URMs | URMs | Bottom 1/3 | Middle 1/3 | Top 1/3 |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| **Panel A: Point Estimates** | | | | | | | |
| Ln(no. of on-track) | 0.831* | 0.856* | 0.799* | 1.005* | 1.376** | 0.893** | 0.766* |
| | (0.441) | (0.447) | (0.431) | (0.546) | (0.610) | (0.416) | (0.453) |
| Ln(no. of on-track) X Fem. | 0.029 | 0.005 | 0.036 | 0.010 | -0.079 | -0.092 | 0.121 |
| | (0.068) | (0.125) | (0.087) | (0.111) | (0.092) | (0.119) | (0.148) |
| Instructor, Year, Time of day FE | X | X | X | X | X | X | X |
| Student Characteristics | X | X | X | X | X | X | X |
| **Panel B: Estimated effects in %-pts. associated with a 15% increase in no. of on-track (0.3 grade points = 1 letter grade)** | | | | | | | |
| Women | 0.12* | 0.12* | 0.11* | 0.14* | 0.18** | 0.11** | 0.12** |
| Men | 0.12* | 0.12* | 0.11* | 0.14* | 0.19** | 0.13** | 0.11* |
| **Observations** | 12,230 | 12,122 | 8,281 | 3,949 | 4,206 | 3,438 | 4,586 |

Note: Each column is a separate specification. "Bottom 1/3" refers to the bottom 1/3 of the SAT math distribution. "Middle 1/3" and "Top 1/3" refer to the middle 1/3 and top 1/3 of the SAT math distribution for the class respectively. Controls include percent female in a class, log class size, year and instructor fixed effects, whether the lecture was held in the morning, a vector of student background characteristics, and a student's declared major at entry. Student background characteristics include: gender, race (black, Hispanic, Asian, American Indian, Filipino, Indian and white is the omitted group), if a student went to public high school, if English is the only language spoken in the home, the highest level of education attainment of the parent with the highest level of education, high school grade point average, SAT math and verbal scores and age. Clustered standard errors are in parentheses , *** p<0.01, ** p<0.05, * p<0.1. Clusters are by CHEM1A lecture (class). A 15% increase in number of on-track students in a class is the equivalent of increasing the number of on-track students by about 1 standard deviation (44 students). URM stands for underrepresented minorities and includes all race categories except white, Asian and Indian.

Table 8: STEM Major Completion – Controls for CHEM1A Grade

| | Full Sample | Lectures > 100 |
|---|---|---|
| | (1) | (2) |
| **Panel A: Point Estimates** | | |
| Ln(no. of on-track) | -0.256* | -0.185 |
| | (0.140) | (0.139) |
| Ln(no. of on-track) X Fem. | -0.117*** | -0.138*** |
| | (0.034) | (0.053) |
| CHEM1A Grade | 0.169*** | 0.169*** |
| | (0.005) | (0.005) |
| Instructor, Year, Time of day FE | X | X |
| Student Characteristics | X | X |
| **Panel B: Estimated effects in %-pts. associated with a 15% increase in no. of on-track** | | |
| Women | -5.20*** | -4.50** |
| Men | -3.60* | -2.60 |
| **Observations** | 12,230 | 12,122 |

Note: Each column is a separate specification. Controls include percent female in a class, log class size, year and instructor fixed effects, whether the lecture was held in the morning, a vector of student background characteristics, and a student's declared major at entry. Student background characteristics include: gender, race (black, Hispanic, Asian, American Indian, Filipino, Indian and white is the omitted group), if a student went to public high school, if English is the only language spoken in the home, the highest level of education attainment of the parent with the highest level of education, high school grade point average, SAT math and verbal scores and age. Clustered standard errors are in parentheses , *** p<0.01, ** p<0.05, * p<0.1. Clusters are by CHEM1A lecture (class). A 15% increase in number of on-track students in a class is the equivalent of increasing the number of on-track students by about 1 standard deviation (44 students). URM stands for underrepresented minorities and includes all race categories except white, Asian and Indian.

Table 9: The Effect of the Number of On-Track Students on CHEM1B Enrollment

| | Full Sample | Lectures > 100 | Non-URMs | URMs | Bottom 1/3 | Middle 1/3 | Top 1/3 |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| **Panel A: Point Estimates** | | | | | | | |
| Ln(no. of on-track) | 0.000 | 0.015 | -0.039 | 0.110 | -0.221 | 0.141 | 0.125 |
| | (0.086) | (0.093) | (0.065) | (0.191) | (0.167) | (0.160) | (0.110) |
| Ln(no. of on-track) X Fem. | -0.045** | -0.058* | -0.049 | -0.025 | -0.098* | 0.040 | -0.035 |
| | (0.019) | (0.031) | (0.037) | (0.051) | (0.050) | (0.044) | (0.047) |
| Instructor, Year, Time of day FE | X | X | X | X | X | X | X |
| Student Characteristics | X | X | X | X | X | X | X |
| **Panel B: Estimated effects in %-pts. associated with a 15% increase in no. of on-track (0.3 grade points = 1 letter grade)** | | | | | | | |
| Women | -0.60 | -0.60 | -1.10 | 1.20 | -4.50** | 2.50 | 1.30 |
| Men | 0.01 | 0.20 | -0.40 | 1.50 | -3.10 | 2.00 | 1.60 |
| **Observations** | 12,230 | 12,122 | 8,281 | 3,949 | 4,206 | 3,438 | 4,586 |

Note: Each column is a separate specification. "Bottom 1/3" refers to the bottom 1/3 of the SAT math distribution. "Middle 1/3" and "Top 1/3" refer to the middle 1/3 and top 1/3 of the SAT math distribution for the class respectively. Controls include percent female in a class, log class size, year and instructor fixed effects, whether the lecture was held in the morning, a vector of student background characteristics, and a student's declared major at entry. Student background characteristics include: gender, race (black, Hispanic, Asian, American Indian, Filipino, Indian and white is the omitted group), if a student went to public high school, if English is the only language spoken in the home, the highest level of education attainment of the parent with the highest level of education, high school grade point average, SAT math and verbal scores and age. Clustered standard errors are in parentheses , *** p<0.01, ** p<0.05, * p<0.1. Clusters are by CHEM1A lecture (class). A 15% increase in number of on-track students in a class is the equivalent of increasing the number of on-track students by about 1 standard deviation (44 students). URM stands for underrepresented minorities and includes all race categories except white, Asian and Indian.

Table 10: The Effect of the Number of On-Track Students on CHEM1B Grade

| | Full Sample | Lectures > 100 | Non-URMs | URMs | Bottom 1/3 | Middle 1/3 | Top 1/3 |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| **Panel A: Point Estimates** | | | | | | | |
| Ln(no. of on-track) | -0.431 | -0.374 | -0.288 | -0.660** | -0.208 | -0.625 | -0.555* |
| | (0.260) | (0.281) | (0.320) | (0.287) | (0.373) | (0.585) | (0.276) |
| Ln(no. of on-track) X Fem. | -0.130* | -0.205** | -0.163* | -0.156 | -0.096 | -0.244 | 0.136 |
| | (0.067) | (0.099) | (0.085) | (0.143) | (0.131) | (0.169) | (0.132) |
| Instructor, Year, Time of day FE | X | X | X | X | X | X | X |
| Student Characteristics | X | X | X | X | X | X | X |
| **Panel B: Estimated effects in %-pts. associated with a 15% increase in no. of on-track (0.3 grade points = 1 letter grade)** | | | | | | | |
| Women | -0.08** | -0.08** | -0.06 | -0.11** | -0.04 | -0.12 | -0.06 |
| Men | -0.06 | -0.05 | -0.04 | -0.09** | -0.03 | -0.09 | -0.08** |
| **Observations** | 10,086 | 10,003 | 6,942 | 3,144 | 3,187 | 2,857 | 4,042 |

Note: Each column is a separate specification. "Bottom 1/3" refers to the bottom 1/3 of the SAT math distribution. "Middle 1/3" and "Top 1/3" refer to the middle 1/3 and top 1/3 of the SAT math distribution for the class respectively. Controls include percent female in a class, log class size, year and instructor fixed effects, whether the lecture was held in the morning, a vector of student background characteristics, and a student's declared major at entry. Student background characteristics include: gender, race (black, Hispanic, Asian, American Indian, Filipino, Indian and white is the omitted group), if a student went to public high school, if English is the only language spoken in the home, the highest level of education attainment of the parent with the highest level of education, high school grade point average, SAT math and verbal scores and age. Clustered standard errors are in parentheses , *** p<0.01, ** p<0.05, * p<0.1. Clusters are by CHEM1A lecture (class). A 15% increase in number of on-track students in a class is the equivalent of increasing the number of on-track students by about 1 standard deviation (44 students). URM stands for underrepresented minorities and includes all race categories except white, Asian and Indian.

Table 11: Where are the Women Going?

| | Graduate in Non-STEM | Graduate |
|---|---|---|
| | (1) | (2) |
| **Panel A: Point Estimates** | | |
| Ln(no. of on-track) | 1.22 | -0.035 |
| | (0.099) | (0.113) |
| Ln(no. of on-track) X Fem. | 0.107*** | -0.074* |
| | (0.030) | (0.038) |
| Instructor, Year, Time of day FE | X | X |
| Student Characteristics | X | X |
| **Panel B: Estimated effects in %-pts. associated with a 15% increase in no. of on-track** | | |
| Women | 3.20** | -0.50 |
| Men | 1.70 | -1.50 |
| **Observations** | 12,230 | 12,230 |

Note: Each column is a separate specification. Controls include percent female in a class, log class size, year and instructor fixed effects, whether the lecture was held in the morning, a vector of student background characteristics, and a student's declared major at entry. Student background characteristics include: gender, race (black, Hispanic, Asian, American Indian, Filipino, Indian and white is the omitted group), if a student went to public high school, if English is the only language spoken in the home, the highest level of education attainment of the parent with the highest level of education, high school grade point average, SAT math and verbal scores and age. Clustered standard errors are in parentheses , *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Clusters are by CHEM1A lecture (class). A 15% increase in number of on-track students in a class is the equivalent of increasing the number of on-track students by about 1 standard deviation (44 students). URM stands for underrepresented minorities and includes all race categories except white, Asian and Indian.

# Appendix

Table A1: STEM Majors

| Major Requires CHEM1A | STEM Majors |
|---|---|
| X | Biology |
| X | Biochemistry |
| X | Biopsychology |
| X | Chemistry |
| X | Engineering |
| X | Computer Engineering |
| | Computer Science |
| X | Earth Science |
| X | Ecology |
| X | Environmental Science |
| | Mathematics |
| | Statistics |
| X | Geophysics |
| X | Hydrology |
| X | Zoology |
| X | Pharmacology |
| X | Physics |
| X | Physiology |
| X | Physical Geography |

Table A2: Balance Test – Orientation Attendees vs. Non-Attendees

| | Summer Orientation Group | Did Not Attend Orientation | Diff. (1) - (2) |
|---|---|---|---|
| | (1) | (2) | (3) |
| **Student Background and Class Characteristics** | | | |
| No. On-track is above median | 0.48 | 0.46 | 0.02 |
| | (0.01) | (0.04) | (0.04) |
| Lecture is at 8 or 9 a.m. | 0.35 | 0.39 | -0.04 |
| | (0.01) | (0.04) | (0.04) |
| Female | 0.49 | 0.52 | -0.03 |
| | (0.01) | (0.04) | (0.04) |
| URM (underrepresented minority) | 0.38 | 0.48 | -0.10** |
| | (0.01) | (0.04) | (0.04) |
| High school GPA is above median | 0.51 | 0.46 | -0.05 |
| | (0.01) | (0.04) | (0.04) |
| SAT math score is above median | 0.51 | 0.51 | 0.00 |
| | (0.01) | (0.04) | (0.04) |
| SAT verbal score is above median | 0.54 | 0.48 | 0.06 |
| | (0.01) | (0.04) | (0.04) |
| Attended public high school | 0.92 | 0.93 | -0.01 |
| | (0.01) | (0.02) | (0.02) |
| English is only language spoken in home | 0.54 | 0.48 | 0.06 |
| | (0.01) | (0.04) | (0.04) |
| No parent graduated from college | 0.37 | 0.43 | -0.05 |
| | (0.01) | (0.04) | (0.04) |
| **Observations** | 1,624 | 178 | 1,802 |

Note: URM stands for underrepresented minorities and includes all race categories except white, Asian and Indian. Column 3 uses an asterisk system to denote whether the differences in means are significant. Level of significance is indicated as follows: *** p<0.01, ** p<0.05, * p<0.1. Standard deviations are in parentheses. The sample includes only those students enrolled in the first quarter of General Chemistry (CHEM1A) in the fall quarter of their freshman year at UCSB for the year 2013.

Table A3: Major Composition by Gender (%)

| | Women | | Men | |
|---|---|---|---|---|
| | Intended Major at Entry | Major at Grad. | Intended Major at Entry | Major at Grad. |
| Hard Science | 15.07 | 14.29 | 49.37 | 33.77 |
| Bio/Environ. Sci. | 51.81 | 29.88 | 24.84 | 17.45 |
| Social Science | 3.48 | 19.27 | 2.79 | 15.79 |
| Human./Arts/Interd. | 3.08 | 36.57 | 2.55 | 32.97 |
| Undeclared | 26.55 | 0.00 | 20.45 | 0.01 |

Note: See Appendix Table A4 for the majors that fall into each major category: hard science, biology/environmental studies, social sciences, and humanities/arts/interdisciplinary. The sample includes only those students enrolled in the first quarter of General Chemistry (CHEM1A) in the fall quarter of their freshman year at UCSB between the years 1997 and 2007.

Table A4: Majors by Category

| **Hard Sci.** | **Bio & Env. Studies** | **Social Sci.** | **Human., Arts, & Interdis.** |
|---|---|---|---|
| Biochemistry | Biochemistry-Molecular Biology | Anthropology | Art History |
| Chemistry | Biological Sciences | Geography | Art Studio |
| Chemical Engineering | Biopsychology | Geophysics | Asian & American Studies |
| Computer Engineering | Physiology | Physical Geography | Asian Studies |
| Electrical Engineering | Biology | Economics-Accounting | Black Studies |
| Earth Science | Cell & Develp. Biology | Economics-Mathematics | Chicana and Chicano Studies |
| Hydrological Sciences | Microbiology | Economics | Chinese |
| Mechanical Engineering | Environmental Studies | Political Science | Classics |
| Pharmacology | | Psychology | Communication Studies |
| Physics | | Sociology | Comparative Literature |
| Computer Science | | Business Economics | Creative Studies |
| Mathematics | | | Dance |
| Financial Math & Stats | | | Dramatic Art |
| Statistics | | | English |
| Zoology | | | Feminist Studies |
| Aquatic Biology | | | Film & Media Studies |
| Ecology and Evolution | | | Financial Mathematics & Statistics |
| Computer Science | | | French |
| | | | Germanic Languages |
| | | | Global Studies |
| | | | History or History of Public Policy |
| | | | Interdisciplinary Studies |
| | | | Italian Cultural Studies |
| | | | Japanese |
| | | | Language, Culture & Society |
| | | | Latin Am/Iberian Studies |
| | | | Law & Society |
| | | | Linguistics |
| | | | Medieval Studies |
| | | | Middle Eastern Studies |
| | | | Music & Music Composition |
| | | | Philosophy |
| | | | Portuguese |
| | | | Religious Studies |
| | | | Slavic Languages & Literatures |
| | | | Spanish |