Classroom Gender Composition and Academic Outcomes: Evidence from Male Cross-Registration at a Women's College

J. Yo-Jud Cheng
Analysis Group

Patrick J. McEwan[1]
Wellesley College

May 2012

Abstract: This paper estimates the effects of male classroom presence on female academic outcomes at a women's college. About 10 percent of courses enroll at least one cross-registered male student. Experimental literature on gender stereotype threat suggests that male presence may exert negative effects on outcomes. Administrative panel data allow controls for student and instructor fixed effects. The paper finds no negative effects of male cross-registrants on academic outcomes, and positive effects for some dependent variables and subsamples. One interpretation consistent with the experimental literature is that the single-sex environment effectively buffers students against gender stereotype threat.

Keywords: gender; peer effects; stereotype threat

[1] Corresponding author: Department of Economics, 106 Central St., Wellesley MA 02481, pmcewan@wellesley.edu.
We thank Larry Baldwin and Ann Hamilton for their guidance on the institutional data. Kristin Butcher, Tom Dee, Adrienne Lucas, Akila Weerapana, and AEFA panel participants provided helpful comments. The paper received partial financial support from a Mellon 23 conference grant to Kristin Butcher and Patrick McEwan, and a Wellesley College Schiff Fellowship to Yo-Jud Cheng. We alone are responsible for any errors or interpretations in this paper.

1. Introduction

A 2006 revision to U.S. Title IX legislation allowed federally-funded public schools to offer single-sex classes.[2] In recent years, several states have taken advantage of this policy change by adopting single-sex education as a cost-effective means of raising student achievement.[3] In contrast, the number of women's colleges in the United States has steadily declined from over 200 in the 1960s to 54 in 2009 (WCC 2009). Despite these diverging paths, empirical research on single-sex schooling—reviewed in section 2—offers little consensus about its likely effects on academic outcomes. Indeed, the decision to attend a single-sex school or college is plausibly correlated with unobserved variables (like student motivation) that affect student outcomes, severely complicating efforts to disentangle causal effects from selection bias.

This paper makes headway by examining a unique natural experiment in which male students cross-register in courses at an all-female college. Wellesley College was founded in 1875 "to provide an excellent liberal arts education for women" at a time when most colleges and universities did not accept women. However, male students have cross-registered in courses at Wellesley College since 1968, when the first of several agreements was signed with the Massachusetts Institute of Technology (MIT). The enrollment of male students remained low until the founding and concurrent establishment of cross-registration with the Olin College of Engineering in 2002. Since then, the proportion of courses each semester with at least one male

[2] The regulations clarify policies in the No Child Left Behind Act (NCLB). Coeducational public schools can offer single-sex classrooms on the condition that they provide a rationale for offering a single-gender class in that subject; provide a coeducational class in the same subject at a "geographically accessible" location (at the same school or at another school); and conduct a review every two years to determine whether single-sex classes are still necessary. Single-sex schools are exempted from providing a rationale for single-sex education and from periodic reviews. However, another school within the same school district must offer "substantially equal" courses, services, and facilities. This school can be either coed or single-sex. Charter schools are exempt from these requirements (NASSPE 2008; DOE 2006).

[3] As of February 2009, at least 540 public schools in the U.S. offered single-sex education opportunities (through single-sex classrooms or single-sex schools) (NASSPE 2009). The number of single-sex schools has increased from two in 1995 to 49 in 2008 (Weil 2008). South Carolina has implemented single-sex programs in over 200 schools (K-12) since the policy revision (SCDOE 2009).

student has increased from approximately 2 percent in the years preceding the agreement to around 10 percent in recent years.

Using student-level panel data from four academic years, we assess whether the classroom presence of males in a female supermajority exerts a negative influence on classroom outcomes, including course grades and subsequent course-taking in subjects. A substantial experimental literature suggests it might. The mere presence of male students in testing situations has caused female students to score lower, perhaps because females unconsciously conform to negative stereotypes that associate women with lower academic ability in certain subject areas, especially traditionally "male" subjects (Spencer et al. 1999; Shih et al. 1999; Inzlicht and Ben-Zeev 2000). These "stereotype threat" effects are generally found in controlled lab settings; our study, in contrast, occurs in natural classroom settings over a longer period.

An immediate econometric concern is that the presence of male cross-registrants in a classroom is correlated with unobserved student or instructor variables that affect outcomes. However, cross-registrants are among the last students to register for courses, and neither group has significant advance knowledge about the other. We also show that levels of course "attrition" (or pre-deadline course dropping among students) are similar across classrooms with and without cross-registered students. Most importantly, the detailed panel data allow us to control for student and instructor heterogeneity by the inclusion of fixed effects in some regression specifications.

Overall, the results show no compelling evidence that the presence of a male student in the classroom lowers the achievement of other (female) students. There are similar findings in subsamples defined by academic department (humanities, social sciences, or natural sciences) and levels of baseline quantitative skills of students. In fact, results in some subsamples suggest

that male presence may *lower* the probability that students obtain a particularly low grade in a course. We explore several explanations for these results. One possibility, also consistent with experimental literature, is that several aspects of the Wellesley College environment act as a "buffer" that mitigates the negative effects of stereotype threat. These factors include a greater recognition of positive characteristics associated with women and the possession of alternative identities associated with positive stereotypes, such as being students at an elite private college (Shih 2004).

In section 2, we frame the empirical analysis by reviewing empirical literature on the effects of single-sex schools, classroom gender composition, and gender stereotype threat. Section 3 provides background on the history of and participation in cross-registration at Wellesley College. Sections 4 and 5 outline the empirical strategy and student panel data. Section 6 present and interpret the regressions, while section 7 concludes.

2. Prior Research

The Effects of Single-Sex Schools

A 1992 report by the American Association of University Women (AAUW) claimed that girls were being "short-changed" by the American coeducational system due to persistent gender gaps in achievement. Several years later, a controversial follow-up report reviewed the research on K-12 single-sex education, concluding that while single-sex education may offer non-academic benefits, these advantages do not consistently translate into "measured improvements in achievement" for all students (AAUW 1998). Other surveys of the research on single-sex education have examined outcomes including academic achievement, labor market outcomes, classroom participation, major choice, instances of sexism, levels of self-esteem, and satisfaction

with educational experiences. These studies have found positive non-academic benefits, but academic effects have been mixed (Riordan 2002; Billger 2006; Campbell and Sanders 2002; Mael 1998). All papers confront well-known challenges of causal inference that similarly afflict the nonexperimental literature on U.S. private school effectiveness (see, e.g., McEwan 2004).

Given legal constraints on single-sex education prior to 2006, the U.S. literature focuses on Catholic, single-sex schools.[4] Using High School and Beyond (HSB) data, Lee and Bryk (1986) found that girls in single-sex schools achieved greater gains in reading and science achievement as compared to their counterparts enrolled in coeducational schools. A follow-up study found that higher education aspirations and interest in graduate work persisted among college-age women that attended single-sex schools (Lee and Marks 1990). A similar longitudinal study using the National Educational Longitudinal Study of 1988 (NELS:88) found no differences in the academic achievement of students in single-sex and coeducational schools (LePore and Warren 1997).

In higher education, nonexperimental studies have found that women's college graduates attain higher levels of education and obtain significant labor market benefits in terms of salary and occupational attainment (Tidball and Kistiakowsky 1976; Tidball 1985, 1986; Ledman et al. 1995; Riordan 1994). Other studies (Kinzie et al. 2007; Smith 1990; Smith et al. 1995) have found that women's college graduates report higher levels of academic engagement and overall satisfaction with college, though Smith (1995) notes that the findings do not have a strong causal interpretation. A related literature focuses on the transition of women's colleges to coeducation. Following one such transition, Bilger (2002) found that female students were less likely to

[4] An international literature finds mixed effects. Two studies of 8th grade students in Nigeria and Thailand found that girls enrolled in single-sex schools earned higher scores in math than their counterparts enrolled in coeducational schools (Lee and Lockheed 1990; Jimenez and Lockheed 1989), while two studies on affective, non-achievement outcomes found no significant differences between girls in single-sex and coeducational Belgian elementary and Irish secondary schools (Brutsaert and Bracke 1994; Cairns 1990).

pursue male-dominated majors and occupations (but were more likely to earn a professional degree), relative to a comparison group of all other public and private colleges and universities in the United States. In one of the few studies of transitioning classrooms, Canada and Pringle (1995) observed Wheaton College's shift from a women's college to coeducation, finding that the number of follow-up interactions with teachers initiated by female students decreased as the proportion of males in a class increased.

The Effects of Classroom Gender Composition

More recent studies have side-stepped the empirical challenge of self-selection into single-sex schools by analyzing random or plausibly random variation in *classroom* gender composition. Using experimental data from Tennessee STAR, Whitmore (2005) found that the average percentile rank of standardized math and reading scores improved for both boys and girls in kindergarten and second grade classes that were over 50 percent female, but not in other grades. Hoxby (2000) employed year-to-year fluctuations in gender composition, finding gains in writing, but not in reading or math, as the proportion of girls increased in a class.[5] This research has greater internal validity than studies of exclusively single-sex schools, but is best generalized to a relatively narrow range of credibly random, "noisy" variation in classroom gender composition in coeducational settings.

The Effects of Gender Stereotype Threat

This paper examines the effects of classroom gender composition in the context of a female supermajority. A growing literature on stereotype threat suggests a mechanism through

[5] An Israeli study, using a related empirical approach, identified positive effects on academic achievement for girls as the proportion of girls in a class increased, especially when the proportion exceeded 55 to 60 percent (Lavy and Schlosser 2007).

which the presence of even one male student in the classroom could influence academic performance. These studies randomly assign gender-based "primes"—or cues that make subjects aware of a specific identity—to subjects and evaluate their subsequent test performance. Several studies have shown that women, as a group, are susceptible to stereotype threat in areas that are typically associated with men, such as natural sciences and math, but are not susceptible to these effects in other subject areas (Spencer et al. 1999; Shih et al. 1999; Inzlicht and Ben-Zeev 2000). For example, Inzlicht and Ben-Zeev (2000), found that the physical presence of a male in a testing environment with a small group of people could effectively act as a prime that elicited stereotype threat among women and lowered performance.[6] Other studies have found similar results with larger group sizes and different levels of interaction between group members (Beaton et al. 2007; Sekaquaptewa and Thompson 2003).

A more recent literature studies evaluates "buffers" that could minimize the negative effects of stereotype threat on performance. McGlone and Aronson (2007) found that that by priming female participants' identities as students at an elite private college, the gender gap between the achievement of male and female participants decreased significantly on a standardized math test. Other stereotype threat buffers have focused on disseminating information on its effects. In one intervention, female participants were informed about the theory of stereotype threat before taking a standardized math test (Johns et al. 2005). These

[6] High-achieving female undergraduates at Brown University were randomly assigned to groups with one of three gender compositions (three females; two females and a male; or a female and two males) and given math and verbal Graduate Record Exam (GRE) questions. The math performance of females decreased as the relative number of males in the group increased. There were no statistically significant differences on the verbal portion of the exam, and the men's test scores were consistent across all groups. The performance deficits on math assessments persisted regardless of whether subjects were told that their scores would be anonymized (Inzlicht and Ben-Zeev 2003).

women performed equally well as male participants, while females that were not exposed to this intervention performed significantly worse than men in their treatment group.[7]

Most stereotype threat studies occur within controlled experimental settings, which may limit their generalizability (Levitt and List 2007a, 2007b). One study conducted in the natural setting of regular high school classrooms found that stereotype effects were evident among women taking mathematical assessments (Keller and Dauenheimer 2003), but another found no apparent stereotype effects under any conditions (Rivardo et al. 2008). The latter study suggested that stereotype threat may be less salient in small liberal arts colleges where women are in the majority—even in coeducational settings—and the learning environment is more nurturing and personal.

3. Cross-Registration at Wellesley College

History and Background

Wellesley College was founded in 1875 as a women's college. Cross-registration programs that allowed men to enroll in Wellesley classes started in 1968 with the Massachusetts Institute of Technology (MIT), and were expanded with the founding of the Twelve College Exchange[8] (1969), and agreements with Brandeis University (1984), Babson College (1984) and Olin College of Engineering (2002), in addition to a series of much smaller exchange programs (OIR 2008).

Most of these cross-registration agreements were negotiated as a cost-effective means of expanding course offerings without having to create new departments or hire new faculty.

[7] Other interventions that have significantly boosted standardized test scores, grades, and academic engagement for females have taught students to view intelligence as malleable (rather than fixed) and to attribute academic difficulties to external sources (Good et al. 2003; Aronson et al. 2001).

[8] For the 2008-2009 academic year, participating institutions were: Amherst College, Bowdoin College, Connecticut College, Dartmouth University, Mt. Holyoke College, Smith College, Trinity College, Vassar College, Wellesley College, Wesleyan College, and Wheaton College.

However, the exchange program with MIT may have been used partially as a solution to the "pressure at MIT for more women" that was "cheaper than building a dormitory for women" on MIT's campus (Fleming 1969). A bus service between Wellesley and MIT was instituted at the start of the cross-registration program in 1968, and a shuttle between Wellesley, Babson, and Olin was established in 2007 to facilitate cross-registration. In the recent years included in this paper's data, over two-thirds of cross-registering students are Olin students. Olin students take a distribution requirement in Arts/Humanities/Social Sciences (AHS) before graduation. Many fulfill this requirement at Wellesley due to the more limited selection of AHS courses at their home institution.

Both male and female students at partner institutions are eligible to participate in Wellesley's cross-registration program. Cross-registrants cannot register for classes until Wellesley's initial registration period—in the semester prior to enrollment—has passed. In practice, most cross-registrants do not enroll until the new semester has begun, given the administrative burden of pre-registering. Anecdotal evidence suggests that Wellesley students pre-register for courses with no prior knowledge of the future enrollment of male students, and that cross-registrants select classes with minimal information on the Wellesley students that have registered.

Cross-registrants are allowed to enroll in more than one course, but courses typically cannot exceed one-half of a student's total course hours each semester. They must abide by Wellesley's course add/drop deadlines, and therefore cannot add Wellesley courses after the second week of classes. However, any student may drop a course without repercussions during a three-week drop period. Course withdrawals after this period are denoted WDR on the official transcript, and are interpreted by students and faculty as a signal of lower performance.

Participation

In 2007-2008, the most recent year of our dataset, 178 cross-registrants enrolled in 260 Fall or Spring semester courses at Wellesley College (along with 2,336 degree-seeking students). In that academic year, 55 percent of cross-registrants were male, and 71 percent were Olin students. The general growth in cross-registration, particularly among males in the wake of the Olin agreement, is illustrated in Figure 1.

In 2007-2008, 21 percent of classes at Wellesley College enrolled at least one cross-registrant, and 11 percent enrolled a male student. Figure 2 charts the steady growth in the proportions of courses with cross-registrants after the Olin agreement. Among members of the Class of 2008, 98 percent took at least one course during their undergraduate career with a cross-registered student, and 85 percent took a course with a male student. Eighty-five percent took at least 1/8 of their courses (a college semester) with a cross-registered student, and 27 percent took at least 1/8 of their courses with a male student.

4. Empirical Strategy

To identify the impact of male classroom presence on academic outcomes, we estimate variants of the following regression by ordinary least squares:

$$(1)\quad O_{ijkt} = \beta_0 + \beta_1 X_{jkt} + \beta_2 X_{jkt}^{male} + \beta_3 M_{jkt} + \beta_4 \cdot M_{jkt} \cdot X_{jkt} + \beta_5 \cdot M_{jkt} \cdot X_{jkt}^{male} + I_{ijkt}\lambda + C_{jkt}\delta + \mu_t + \varepsilon_{ijkt}$$

where O_{ijkt} is the academic outcome of female student *i* (excluding cross-registrants) obtained in course *j*, taken with instructor *k* in term *t*. (In section 5 we discuss the academic outcomes and

the dataset in detail.) X_{jkt} is a dummy variable, measured at the level of course-by-instructor-by-term cells, indicating whether *any* cross-registered student is enrolled (gauging a combined "stranger" effect on academic outcomes of Wellesley students). X_{jkt}^{male} is a dummy variable indicating presence of at least one male cross-registrant, the coefficient on which we cautiously interpret as a "male presence" effect on academic outcomes. Standard errors are adjusted for multi-way clustering at the levels of students and course-by-professor-by-term cells, following the procedure in Cameron, Gelbach, and Miller (forthcoming).

Our initial estimates control for a vector of student variables (I_{ijkt}), including pre-college test scores, race and ethnicity, and class year (section 5 provides more details). Course variables (C_{jkt}) include department fixed effects, dummy variables indicating class size, and the level of a course (whether an intermediate 200-level or an introductory 100-level).[9] Finally, we control for an indicator of whether an *instructor* is male (M_{jkt}), and interactions of this variable with X_{jkt} and X_{jkt}^{male}; the coefficient on the second interaction indicates whether a "male presence" effect is exacerbated or ameliorated in classrooms with male instructors. The μ_t are fixed effects indicating each academic term.

Estimates of β_2 do not have a causal interpretation if $\text{cov}(X_{jkt}^{male}, \varepsilon_{ijkt}) \neq 0$. For example, students may add or drop a course based on male presence, and cross-registered students may choose courses on the basis of classroom peers. In either case, student sorting may lead to correlations between X_{jkt}^{male} and unobserved student attributes that affect outcomes. While such sorting is unlikely during pre-registration, any registered student can drop a course penalty-free during the first three weeks of the semester. Our dataset does not record the identity or academic

[9] The department fixed effects are primarily included to capture variation in grading standards. However, inter-departmental variation substantially diminished in the wake of a "grade deflation" policy, instituted in Fall 2004. As described in section 5, we limited our sample to academic terms following this policy.

outcomes of such students (who may be replaced by other students during a 2-week add period if a course falls below maximum capacity). To partly assess the likelihood of sample selection bias, section 5 reports evidence that course "attrition" is not associated with the presence of cross-registered students, whether male or female.

However, to further guard against bias, we report estimates of equation (1) that include fixed effects for students and instructors (facilitated by within-student and within-instructor variation in the presence of cross-registrants over eight academic terms). The inclusion of instructor fixed effects prevents the inclusion of a separate male instructor variable that, in any case, is plausibly correlated with unexplained variation in outcomes.[10] However, the interaction effects between instructor gender and presence of cross-registered students can still be estimated.

5. Data

Sample

The panel dataset includes 37,829 grade-by-student-by-term observations on 3,197 degree-seeking students, during four academic years that include substantial cross-registration (2004-2005 to 2007-2008).[11] The sample reflects two exclusions. First, it excludes grades in advanced 300-level courses because such courses usually exhibit substantially less grade variation than introductory or intermediate courses (see footnote 10). Second, it focuses on first-year and sophomore students, to facilitate identification of effects of early male presence on

[10] Unlike Hoffmann and Oreopoulos (2009a), our primary goal is not to determine whether female students fare worse (or better) with male professors. In contrast to the university context considered by those authors, there is clearly much scope at Wellesley for sorting on the basis of instructors' gender which is well-known even during pre-registration.

[11] The data do not include prior terms because of a sharp change in grading policy. In Fall 2004, Wellesley College began requiring that mean course grades not exceed a B+ (or 3.33 grade point average). The policy applies to introductory 100-level and intermediate 200-level courses, but not advanced 300-level courses; it also does not apply to courses with less than 10 students enrolled. The most immediate effect of the policy was to lower mean grades in a handful of departments, principally in the humanities and social sciences, relative to "tougher-grading" departments (Butcher, McEwan, and Weerapana 2010).

subsequent course selection (in any case, juniors and especially seniors mainly enroll in 300-level courses).

Dependent Variables

We estimate equation (1) using three dependent variables. The first is the letter grade, recoded numerically according to College standards.[12] The modal grade is a B+ (3.33), with a mean of 3.29 (see Table 1). This dependent variable has two notable limitations. First, 11% of the sample did not receive a standard letter grade because they chose to take the course "credit/no-credit", because they withdrew after the no-penalty drop deadline, or because they received a "permanent incomplete" after failing to complete a course requirement. Of these, credit/no-credit elections are the most important (8% of observations). Non-letter grades do not contribute to the official grade point average, raising the possibility that students "game" the policy to raise the cumulative grade point average (for evidence, see Butcher et al. 2010). The immediate concern is that non-random credit/no-credit elections introduce sample selection bias into estimates of equation (1). A second concern is that professors may respond to a negative shock in course performance—perhaps due to the presence of a cross-registered student—by shifting a portion of the grade distribution rightward (albeit constrained by a mean grade point average of 3.33); this type of re-scaling would mask the impact of male presence.

We employ a second dependent variable—a dummy variable indicating "low" achievement—that addresses these concerns. Low achievement, in this case, is defined as students who received a letter grade of C- or below, who received "no-credit" (defined officially as a C- or below), who withdrew from the course after the course drop deadline, or who failed to complete its requirements. The variable is missing for less than 200 grade observations,

[12] See http://www.wellesley.edu/Registrar/gradingsystem.html.

substantially reducing sample selection. Low course performance is rare—5% of grades—and course performance in the left tail of the class distribution is less likely to be charitably adjusted upward. In addition, students in the left tail are presumably the most challenged by the course material and are therefore more vulnerable to the effects of stereotype threat (Spencer et al. 1999).

The third dependent variable is the number of subsequent courses taken by the student in an academic department (thus, a student in Economics 101 who only takes Economics 102 in later semesters is coded as 1 for that grade-by-student-by-term observation). Following Hoffmann and Oreopoulos (2009b), we use it as simple measure of ongoing interest in the subject, which would presumably be compromised by an unhappy or difficult experience in an introductory or intermediate course.

Independent Variables

Twenty-two percent of course grades in the sample are awarded in classes with at least one cross-registered student, and the great majority of these classes have only one such student (Table 1). Likewise, 10 percent of course grades are awarded in classes with at least one male student present. Overall, 43 percent of course grades are given by a male instructor.

Additional student control variables summarized in Table 1 include admissions test scores, as well as students' scores on a Quantitative Reasoning placement exam taken by first-year students during an orientation week (Butcher, McEwan, and Taylor 2010). The last two columns of Table 1 confirm that the means of student variables, including race and ethnicity, are similar across courses with no cross-registrants and courses with at least one male cross-registrant.

Relative to grades received in classes without cross-registrants, grades received in classes with male cross-registrants are more likely to be in the social sciences and Economics, and less likely to be in the natural sciences or Mathematics (consistent with the curricular strengths of the main sending institution, as described in section 3). This, in turn, contributes to differences in the relative proportions of grades awarded by male instructors, and in smaller or larger classes. To ensure that course-related differences do not bias results, regressions always include academic department fixed effects and the course variables in Table 1.

6. Results

Full Sample Results

Table 2 reports the full-sample results for four variants of equation (1), estimated with each of the three dependent variables. Columns (1), (5), and (9) report a baseline specification that applies minimal controls for term fixed effects. When instructors are male, there is some evidence of slightly lower grades, on average, and a higher probability of low achievement. However, the presence of a male cross-registrant apparently *lowers* the probability of low achievement by a substantial 1.4 percentage points.

These effects are perhaps driven by selection on observables or unobservables. Indeed, additional specifications that control for student fixed effects (but still condition on *M*), show no evidence that students perform worse in courses with male instructors. However, the apparently beneficial (and counterintuitive) effects of attending class with a male cross-registrant are more robust. After controlling for students fixed effects (column (6)), the presence of a male cross-registrant still lowers the probability of low achievement by 1.6 percentage points; further inclusion of instructor fixed effects renders the effect statistically insignificant, but it remains

around 1 percentage point. Other coefficient estimates on male cross-registrants (and its interaction with male instructor) are smaller and statistically insignificant in all specifications.

In addition to the specifications in Table 2, we re-estimated all regressions after replacing X_{jkt} and X_{jkt}^{male} with variables indicating the proportion of class enrollments that are cross-registrants and male cross-registrants, respectively. Prior stereotype threat experiments found that the proportion of males in a group mediated the size of effects (Inzlicht and Ben-Zeev 2000). In our sample, variation in these proportions is partly confounded with class sizes, since the great majority of classes with a cross-registrant enroll only one (Table 1). (Our specifications continue to control for broad categories of class size, because College-imposed ceilings for mean grade points are enforced less stringently in classes with less than 10 students.) None of these results, available from the authors, suggest that higher proportions of males in the classroom negatively affect outcomes, either alone or combined with a male instructor.

Although not reported in Table 2, coefficients on student control variables in columns (2) and (6) are generally large, of the predicted sign, and statistically significant.[13] For example, an increase of 1 standard deviation in the Quantitative Reasoning test score increases grade point by 0.08 points (15 percent of a standard deviation), on average; the same increase lowers the probability of low achievement by 1 percentage point, relative to a sample mean of 5 percent. It bears emphasis that these effects are conditional on a range of other admissions test scores such as the ACT and SAT scores, which are also correlated with students' grades.

Attrition and Cross-Registration

[13] Full results are available from the authors. For similar results from other academic years, see McEwan and Soderberg (2006) and Butcher, McEwan, and Taylor (2010).

The data include a grade-level observation for each student that did not drop the course by the official deadline (but regrettably excludes observations for "attriting" students who registered but dropped by the official deadline). If attrition responds to the presence or absence of cross-registrants, then sample selection could introduce correlations between X_{jkt}, X_{jkt}^{male}, and unobserved attributes of remaining students correlated with outcomes. To partly assess the likelihood of such bias, we assess whether a course-level proxy of attrition is similar across courses with and without cross-registrants.

Students register for classes using an online system that records the order in which they enroll in a course. Thus, students with higher registration sequence numbers were among the last to register before the add deadline. We calculate course-level attrition as $1-\frac{N}{T}$, where N is the final number of grade-level observations in a course and T is highest registration sequence number in the dataset (a proxy for the largest number of students ever enrolled).[14] In 2,933 course-level observations over eight academic terms, the mean attrition is a relatively high 0.34, consistent with the high levels of student "churn" observed during the initial weeks of class.

Table 2 regresses this measure on variables from equation (1) that indicate presence of cross-registrants. The simplest specification in column (1), controlling only for term fixed effects, suggests slightly higher attrition (2.6 percentage points) when any cross-registrant is present, and a similarly-sized decline when the cross-registrant is male. However, both coefficients move closer to zero, turning statistically insignificant, when course controls (including department fixed effects) and instructor fixed effects are included in the final columns. Thus, while mean course attrition is high, there is no evidence of differential attrition across

[14] In fact, a student with an even higher sequence number could have dropped before the deadline, suggesting that our measure of attrition may be slightly under-stated.

courses with and without cross-registered students, lending support to the causal interpretation of our results.

Subsample Results

In Table 4, we re-estimate the regressions from Table 2 (focusing on the preferred fixed effects specifications) in three subsamples: humanities courses, social sciences courses (including economics), and natural sciences and mathematics. Given the experimental literature on stereotype threat reviewed in section 2, we are especially interested in whether negative effects are more strongly observed in the third subsample.

In humanities courses (panel A), the presence of both a male cross-registrant and a male instructor actually appears to increase the number of subsequent courses taken by .4-.5 (in contrast to the pure "stranger" interaction effect with male instructors, which lowers course-taking, on average, by about the same amount).

In both panels B and C, the presence of male cross-registrants has pronounced effects on *lowering* the probability of low course achievement (as in the full sample), especially in science and math courses. Consistent with this, panel C further suggests that male presence has large positive effects on grade points—at least 0.14, or more than one-quarter standard deviation. Few other coefficients are statistically significant, although the point estimates in panel C curiously suggest that these positive effects of male cross-registrant presence are essentially nullified by the additional presence of a male instructor. In summary, there is no compelling evidence that the presence of male students in classrooms harms academic outcomes, and in some subsamples it actually appears to raise average outcomes.

Finally, Table 5 re-estimates fixed effects specifications in two samples: students with test scores in the bottom half of the Quantitative Reasoning assessment taken during orientation, and students in the top half. We use this assessment instead of SAT or ACT scores because it is missing fewer observations than other admissions tests, and it is highly correlated with other assessments as well as academic outcome measures. Overall, these results do not overturn prior conclusions. Among students with lower baseline quantitative skills (panel A), male classroom presence does not affect outcomes or may slightly improve it. As in the humanities subsample, male presence interacted with male instructors slightly increases subsequent course-taking, and male presence has small effects on lowering the probability of low achievement.

Discussion

Overall, we find no evidence that the presence of male students exerts a negative impact on the academic achievement or course selection of students at an all-female college. The lack of negative effects is generally inconsistent with the broader experimental literature on gender stereotype threat that often finds substantial effects among female students exposed to seemingly-innocuous gender primes, such as the presence of a male in a testing situation.

The results could be attributed to several factors. First, it is plausible that negative effects exist, but are only triggered when the proportion of males in classroom exceeds an (unknown) threshold. Because the vast majority of classes with male students have just one cross-registrant, we are unable observe effects that would only emerge with greater variation in classroom gender composition. Second, grade scaling could prevent the identification of shifts in academic achievement caused by the registration of male students. However, it seems likely

that difficult classroom experiences due to male presence would at least be reflected in subsequent course selection, and this was not the case.

Third, the college environment could act as a buffer that effectively mitigates the negative consequences of stereotype threat. Shih (2004) suggests that individuals who are highly-identified with a stereotyped group (as students at a women's college are likely to be) and frequently interact with other members of their group are more aware of the group's positive aspects and can thereby overcome its negative associations. Other mitigating factors that may serve as buffers against stereotype threat include the greater influence of alternative identities, such as being a student at an elite private college, and the nurturing and personal learning environment found at liberal arts colleges (McGlone and Aronson 2007; Rivardo et al. 2008). Indeed, the positive academic effects found in some sub-samples could be explained by *additional* effort female students employ in "threatened" situations, such as a class with a male cross-registrant. Some literature suggests that subjects do so in order to disprove negative stereotypes and to compensate for their supposed disadvantages (Shih 2004).

Finally, it may simply be that effects of gender stereotype threat are close to zero, and less salient in non-experimental classroom settings in general (Levitt and List 2007a).

7. Conclusion

This paper examined the effects of cross-registered male students on academic outcomes of female students at Wellesley College, a liberal arts college that admits only women. Since a cross-registration agreement with the Olin College of Engineering in 2002, male cross-registration has increased. Now, around 10 percent of courses have at least one male enrolled. We use four academic years of panel data, including dependent variables that measure course

grades and subsequent course selection in subjects. To control for potential correlation of male presence with unobserved variables that affect achievement, we include student and instructor fixed effects, among other control variables.

Overall, there are no systematic effects of male students on the academic outcomes of Wellesley students. Indeed, the presence of male students has a slightly positive effect on academic achievement (by lowering the probability of obtaining very low grades), especially in classes in the social and natural sciences. We explored several possible explanations for these findings. The most plausible is that Wellesley College itself "buffers" students against gender stereotype threat, perhaps by making students aware of the positive aspects of the stereotyped gender, encouraging alternative identities (e.g. students at an elite private college), or providing a nurturing learning environment (Shih 2004; McGlone and Aronson 2007; Rivardo et al. 2008). The positive effects are consistent with a small experimental literature suggesting that subjects might devote extra effort to disprove negative stereotypes (Shih 2004).

The paper is limited to gender variation within a single institution, which provides considerable advantages from the perspective of internal validity. Assignment of males to classrooms is more plausibly exogenous to female students' outcomes than in many research settings, and high-quality administrative data facilitate careful controls for remaining student and instructor heterogeneity. However, it also introduces limitations to the external validity of findings. The results may be attributable to characteristics shared by Wellesley College students that differ from the general population, or institutional factors that are specific to the college.

References

American Association of University Women Educational Foundation (AAUW). 1992. *How Schools Shortchange Girls: The AAUW Report*. Ed. Susan McGee Bailey. Washington DC: National Education Association.

American Association of University Women Educational Foundation (AAUW). 1998. *Separated by Sex: A Critical Look at Single-Sex Education for Girls*. Washington DC: AAUW.

Aronson, Joshua, Carrie B. Fried, and Catherine Good. 2001. Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology* 38: 113-25.

Beaton, Ann, Francine Tougas, Natalie Rinfret, Nathalie Huard, and Marie-Noelle Delisle. 2007. Strength in numbers? Women and mathematics. *European Journal of Psychology of Education* 22(3): 291-306.

Billger, Sherrilyn M. 2002. Admitting men into a women's college: A natural experiment. *Applied Economics Letters* 9: 479-83.

Billger, Sherrilyn M. 2006. Reconstructing school segregation: On the efficacy and equity of single-sex schooling. IZA Discussion Paper No. 2037.

Brutsaert, Herman, and Piet Bracke. 1994. Gender context of the elementary school: Sex differences in affective outcomes. *Educational Studies* 20(1): 3-11.

Butcher, Kristin F., Patrick J. McEwan, and Corrine H. Taylor. 2010. The effects of quantitative skills training on college outcomes and peers. *Economics of Education Review* 29: 187-99.

Butcher, Kristin F., Patrick J. McEwan, and Akila Weerapana. 2010. Gaming responses to a grade deflation. Working paper.

Cairns, Ed. 1990. The relationship between adolescent perceived self-competence and attendance at single-sex secondary school. *British Journal of Educational Psychology* 60: 207-11.

Cameron, A. Colin, Jonah Gelbach, and Douglas Miller. Forthcoming. Robust inference with multi-way clustering. *Journal of Business and Economic Statistics.*

Campbell, Patricia B., and Jo Sanders. 2002. Challenging the system: Assumptions and data behind the push for single-sex schooling. In Amanda Datnow and Lea Hubbard, Eds., *Gender in Policy and Practice*. New York: RoutledgeFalmer.

Canada, Katherine, and Richard Pringle. 1995. The role of gender in college classroom interactions: A social context approach. *Sociology of Education* 68(3): 161-86.

Department of Education (DOE). 2006. Nondiscrimination on the Basis of Sex in Education Programs or Activities Receiving Federal Financial Assistance; Final Rule. 34 CFR Part 106. Federal Register: October 25, 2006, Volume 71, Number 206.

Fleming, Phyllis J. 1969. *Wellesley-MIT Exchange Program*. Letter to Ms. Adams, Mr. Phibbs, and Mr. Schneider. Archives. Clapp Library, Wellesley College, Wellesley, MA.

Good, Catherine, Joshua Aronson, and Michael Inzlicht. 2003. Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Applied Developmental Psychology* 24: 645-62.

Hoffmann, Florian, and Philip Oreopoulos. 2009a. A professor like me: The influence of instructor gender on college achievement. *Journal of Human Resources* 44(2): 479-94.

Hoffmann, Florian and Philip Oreopoulos. 2009b. Professor qualities and student achievement. *The Review of Economics and Statistics* 91(1): 83-92.

Hoxby, Caroline. 2000. Peer effects in the classroom: Learning from gender and race variation. NBER Working Paper No. 7867. Cambridge, MA: NBER.

Inzlicht, Michael, and Talia Ben-Zeev. 2000. A threatening intellectual environment: Why females Are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science* 11(5): 365-71.

Inzlicht, Michael, and Talia Ben-Zeev. 2003. Do high-achieving female students underperform in private? The implications of threatening environments on intellectual processing. *Journal of Educational Psychology* 95(4): 796-805.

Jimenez, Emmanuel, and Marlaine E. Lockheed. 1989. Enhancing girls' learning through single-sex education: Evidence and a policy conundrum. *Educational Evaluation and Policy Analysis* 11(2): 117-42.

Johns, Michael, Toni Schmader, and Andy Martens. 2005. Teaching stereotype threat as a means of improving women's math performance. *Psychological Science* 16(3): 175-79.

Keller, Johannes, and Dirk Dauenheimer. 2003. Stereotype threat in the classroom: Dejection mediates the disrupting threat effect on women's math performance. *Personality and Social Psychology Bulletin* 29: 371-81.

Kinzie, Jillian, Auden D. Thomas, Megan M. Palmer, Paul D. Umbach, and George D. Kuh 2007. Women students at coeducational and women's colleges: How do their experiences compare? *Journal of College Student Development* 48(2): 145-65.

Lavy, Victor, and Analia Schlosser. 2007. Mechanisms and impacts of gender peer effects at school. NBER Working Paper No. 13292. Cambridge, MA: NBER.

Ledman, Robert E., Marianne Miller, and Darrel R. Brown. 1995. Successful women and women's Colleges: Is there an intervening variable in the reported relationship? *Sex Roles* 33(7/8): 489-97.

Lee, Valerie E., and Anthony S. Bryk. 1986. Effects of single-sex secondary schools on student achievement and attitudes. *Journal of Educational Psychology* 78(5): 381-95.

Lee, Valerie E., and Marlaine E. Lockheed. 1990. The effects of single-sex schooling on achievement and attitudes in Nigeria. *Comparative Education Review* 34(2): 209-31.

Lee, Valerie E., and Helen M. Marks. 1990. Sustained effects of the single-sex secondary school experience on attitudes, behaviors, and values in college. *Journal of Educational Psychology* 82(3): 578-92.

Lee, Valerie E., and Helen M. Marks. 1992. Who goes where? Choice of single-sex and coeducational independent secondary schools. *Sociology of Education* 65(3): 226-53.

LePore, Paul C., and John Robert Warren. 1997. A comparison of single-sex and coeducational Catholic secondary schooling: Evidence from the National Educational Longitudinal Study of 1988. *American Educational Research Journal* 34(3): 485-511.

Levitt, Steven D., and John A. List. 2007a. Viewpoint: On the generalizability of lab behaviour to the field. *Canadian Journal of Economics* 40(2): 347-70.

Levitt, Steven D., and John A. List. 2007b. What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives* 21(2): 153-74.

Mael, Fred A. 1998. Single-sex and coeducational schooling: Relationships to socioemotional and academic development. *Review of Educational Research* 68(2): 101-29.

McEwan, Patrick J. 2004. The potential impact of vouchers. *Peabody Journal of Education* 79(3): 57-80.

McEwan, Patrick J., and Soderberg, Kristen. 2006. Roommate effects on grades: Evidence from first-year housing assignments. *Research in Higher Education* 47(3): 347-70.

McGlone, Matthew S., and Joshua Aronson. 2007. Forewarning and forearming stereotype-threatened students. *Communication Education* 56(2): 119-33.

National Association for Single Sex Public Education (NASSPE). 2008. The legal status of single-sex public education. Downloaded May 2009 from http://www.singlesexschools.org/legal.html.

National Association for Single Sex Public Education (NASSPE). 2009. Single-Sex Schools / Schools with single-sex classrooms / what's the difference? Downloaded May 2009 from http://www.singlesexschools.org/schools-classrooms.htm.

Office of Institutional Research (OIR). 2008. *Factbook 2008-2009.* Wellesley, MA: Wellesley College.

Riordan, Cornelius. 1994. The value of attending a women's college: Education, occupation, and income benefits. *The Journal of Higher Education* 65(4): 486-510.

Riordan, Cornelius. 2002. What do we know about the effects of single-sex schools?: Implications for public schools. In Amanda Datnow and Lea Hubbard, Eds., *Gender in Policy and Practice*. New York: RoutledgeFalmer.

Rivardo, Mark G., Michael E. Rhodes, and Brandi Klein. 2008. Lack of stereotype threat at a liberal arts college. *College Student Journal* 42(3): 832-41.

Sekaquaptewa, Denise, and Mischa Thompson. 2003. Solo status, stereotype threat, and performance expectancies: Their effects on women's performance. *Journal of Experimental Social Psychology* 39: 68-74.

Shih, Margaret, Todd L. Pittinsky, and Nalini Ambady. 1999. Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science* 10(1): 80-3.

Shih, Margaret. 2004. Positive stigma: Examining resilience and empowerment in overcoming stigma. *Annals of the American Academy of Political and Social Science* 591: 175-185.

Smith, Daryl G. 1990. Women's colleges and coed colleges: Is there a difference for women? *The Journal of Higher Education* 61(2): 181-97.

Smith, Daryl G., Lisa E. Wolf, and Diane E. Morrison. 1995. Paths to success: Factors related to the impact of women's colleges. *The Journal of Higher Education* 66(3): 245-66.

South Caroline Department of Education (SCDOE). 2009. *Office of Public School Choice – Single-Gender Education: Single-Gender Initiatives.* Downloaded May 2009 from http://ed.sc.gov/agency/Innovation-and-Support/Public-School-Choice/SingleGender/.

Spencer, Steven J., Claude M. Steele, and Diane M. Quinn. 1999. Stereotype threat and women's math performance. *Journal of Experimental Social Psychology* 35: 4-28.

Steele, Claude M., and Joshua Aronson. 1995. Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology* 69(5): 797-811.

Tidball, M. Elizabeth, and Vera Kistiakowsky. 1976. Baccalaureate origins of American scientists and scholars. *Science* 193(4254): 646-52.

Tidball, M. Elizabeth. 1985. Baccalaureate origins of entrants into American medical schools. *The Journal of Higher Education* 56(4): 385-402.

Tidball, M. Elizabeth. 1986. Baccalaureate origins of recent natural science doctorates. *The Journal of Higher Education* 57(6): 606-20.

Weil, Elizabeth. 2008, March 2. Teaching boys and girls separately. *New York Times*.

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

Women's College Coalition (WCC). 2009. Colleges and universities by name. Downloaded May 2009 from http://www.womenscolleges.org/colleges/byname.

Table 1: Summary statistics

	Full sample					Subsamples	
						X=0	$X^{male}=1$
Variable	Mean	S.D.	Obs.	Min.	Max.	Mean	Mean
Grade point	3.279	(0.567)	33507	0	4	3.276	3.286
Low achievement	0.049	(0.215)	37663	0	1	0.049	0.041
Number of subsequent courses in department	1.377	(2.131)	37829	0	18	1.359	1.379
X: Any cross-registrant in class	0.221	(0.415)	37829	0	1	0.000	1.000
1 cross-registrant in class	0.174	(0.379)	37829	0	1	0.000	0.711
>1 cross-registrant in class	0.047	(0.211)	37829	0	1	0.000	0.289
X^{male}*: Male cross-registrant in class*	0.098	(0.297)	37829	0	1	0.000	1.000
1 male cross-registrant in class	0.086	(0.280)	37829	0	1	0.000	0.881
>1 male cross-registrant in class	0.012	(0.107)	37829	0	1	0.000	0.119
M: Male instructor	0.430	(0.495)	37829	0	1	0.443	0.333
Quantitative Reasoning score	0.000	(1.000)	37311	-3.97	1.83	0.000	0.053
ACT score	29.397	(2.776)	10774	19	36	29.403	29.480
SAT writing / 100	6.927	(0.719)	22007	3.2	8	6.930	6.946
SAT math / 100	6.812	(0.643)	35174	4.3	8	6.812	6.848
SAT verbal / 100	6.941	(0.669)	35174	3.7	8	6.939	6.978
African-American student	0.050	(0.218)	37829	0	1	0.050	0.046
Asian/Asian-American student	0.246	(0.431)	37829	0	1	0.246	0.265
Latina student	0.058	(0.233)	37829	0	1	0.057	0.052
Other race/ethnicity	0.146	(0.353)	37829	0	1	0.147	0.146
Unknown race/ethnicity	0.084	(0.277)	37829	0	1	0.081	0.087
White student	0.416	(0.493)	37829	0	1	0.418	0.404
200-level class (vs. 100-level)	0.496	(0.500)	37829	0	1	0.477	0.511
Class size <10	0.044	(0.206)	37829	0	1	0.050	0.037
Class size ≥10 and <25	0.547	(0.498)	37829	0	1	0.568	0.480
Class size ≥25	0.409	(0.492)	37829	0	1	0.383	0.483
Humanities	0.366	(0.482)	37829	0	1	0.375	0.364
Social Science/Economics	0.346	(0.476)	37829	0	1	0.329	0.401
Natural Sciences/Math	0.249	(0.433)	37829	0	1	0.264	0.170

Notes: See text for details of sample and variable definitions.

Table 2: Effects of cross-registrant presence on academic outcomes, full sample

	Dependent variable: *Grade point*				Dependent variable: *Low achievement*				Dependent variable: *Number of subsequent courses in department*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
X	0.0151	-0.0014	-0.0129	-0.0057	0.0064	0.0099*	0.0101*	0.0043	0.2046**	0.0553	0.0591	0.0584
	(0.018)	(0.016)	(0.018)	(0.016)	(0.006)	(0.005)	(0.006)	(0.006)	(0.095)	(0.075)	(0.090)	(0.082)
X^{male}	-0.0045	-0.0052	0.0047	0.0163	-0.0141**	-0.0139**	-0.0157**	-0.0095	-0.0635	-0.0792	-0.0854	0.0061
	(0.023)	(0.020)	(0.024)	(0.020)	(0.007)	(0.007)	(0.008)	(0.008)	(0.130)	(0.100)	(0.116)	(0.108)
M	-0.0298***	0.0000	0.0023	--	0.0155***	0.0074**	0.0054	--	0.0192	0.0392	0.0405	--
	(0.010)	(0.009)	(0.012)		(0.003)	(0.003)	(0.004)		(0.059)	(0.042)	(0.054)	
*M *X*	0.0111	0.0104	0.0245	0.0101	-0.0037	-0.0060	-0.0052	-0.0042	-0.3277**	-0.1682*	-0.1656	-0.1878
	(0.029)	(0.025)	(0.028)	(0.024)	(0.009)	(0.008)	(0.010)	(0.010)	(0.142)	(0.100)	(0.120)	(0.117)
$M * X^{male}$	-0.0333	-0.0308	-0.0442	-0.0527	0.0071	0.0079	0.0075	0.0067	0.2536	0.1230	0.1104	0.0520
	(0.038)	(0.032)	(0.037)	(0.033)	(0.013)	(0.012)	(0.013)	(0.013)	(0.185)	(0.137)	(0.161)	(0.166)
Observations	33507	33507	33507	33507	37663	37663	37663	37663	37829	37829	37829	37829
Adjusted R^2	0.002	0.112	0.484	0.526	0.002	0.034	0.248	0.271	0.122	0.260	0.357	0.396
Student controls	N	Y	N	N	N	Y	N	N	N	Y	N	N
Course controls	N	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y
Student f.e.	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y
Instructor f.e.	N	N	N	Y	N	N	N	Y	N	N	N	Y
Term f.e.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Notes: *** indicates statistical significance at 1%; ** at 5%; * at 10%. Standard errors in parentheses are adjusted for multi-way clustering at two levels: students and course-by-instructor-by-term cells. Student controls include the ACT, SAT writing, SAT math, SAT verbal, race and ethnicity dummy variables, class-year dummy variables, and dummy variables indicating missing values of prior controls. Course controls include academic department fixed effects, and dummy variables indicating class size and 200-level courses.

Table 3: Effects of cross-registrant presence on course attrition

	Dependent variable: *Course attrition*		
	(1)	(2)	(3)
X	0.026***	0.009	-0.008
	(0.009)	(0.009)	(0.010)
X^{male}	-0.028**	-0.020	-0.007
	(0.013)	(0.012)	(0.014)
M	0.003	-0.001	--
	(0.006)	(0.006)	
$M * X$	-0.012	-0.015	0.010
	(0.015)	(0.014)	(0.016)
$M * X^{male}$	0.020	0.016	-0.010
	(0.021)	(0.019)	(0.021)
Observations	2933	2933	2933
Adjusted R^2	0.009	0.184	0.247
Course controls	N	Y	Y
Instructor f.e.	N	N	Y
Term f.e.	Y	Y	Y

Notes: *** indicates statistical significance at 1%; ** at 5%; * at 10%. Robust standard errors are in parentheses. Course attrition is calculated as 1-[final course enrollment ÷ number of students ever registered in course prior to drop deadline]; the sample mean is 0.341.

Table 4: Effects of cross-registrant presence on academic outcomes, by categories of department

	Dependent variable: *Grade point*		Dependent variable: *Low achievement*		Dependent variable: *Number of subsequent courses in department*	
	(1)	(2)	(3)	(4)	(5)	(6)
A. Humanities courses						
X	-0.0154	-0.0235	0.0028	0.0018	0.0919	0.1076
	(0.024)	(0.024)	(0.009)	(0.008)	(0.107)	(0.103)
X^{male}	0.0122	0.0171	0.0046	0.0106	-0.0951	-0.0799
	(0.032)	(0.031)	(0.011)	(0.011)	(0.137)	(0.141)
M	0.0191	--	0.0018	--	-0.0476	--
	(0.016)		(0.005)		(0.062)	
$M * X$	0.0171	-0.0253	0.0070	0.0074	-0.3505**	-0.4690***
	(0.044)	(0.045)	(0.016)	(0.016)	(0.161)	(0.177)
$M * X^{male}$	-0.0412	-0.0167	-0.0273	-0.0335	0.4277*	0.4497*
	(0.058)	(0.058)	(0.020)	(0.022)	(0.243)	(0.264)
Observations	11624	11624	13794	13794	13838	13838
Adjusted R^2	0.578	0.621	0.339	0.352	0.554	0.586
B. Social Sciences/Economics courses						
X	0.0204	0.0266	-0.0003	-0.0064	0.0046	0.0211
	(0.024)	(0.020)	(0.008)	(0.009)	(0.119)	(0.120)
X^{male}	-0.0389	0.0009	-0.0225**	-0.0193*	-0.0124	0.0717
	(0.031)	(0.028)	(0.011)	(0.011)	(0.152)	(0.159)
M	0.0201	--	-0.0005	--	0.1391*	--
	(0.018)		(0.006)		(0.083)	
$M * X$	0.0042	-0.0143	0.0049	0.0053	-0.0502	-0.0439
	(0.036)	(0.034)	(0.013)	(0.013)	(0.166)	(0.170)
$M * X^{male}$	-0.0047	-0.0436	0.0100	0.0134	-0.1363	-0.2402
	(0.056)	(0.050)	(0.018)	(0.019)	(0.235)	(0.261)
Observations	12208	12208	13002	13002	13083	13083
Adjusted R^2	0.584	0.626	0.382	0.403	0.554	0.573
C. Natural Sciences/Mathematics courses						
X	-0.0844	-0.0382	0.0215	0.0109	-0.0274	0.0169
	(0.065)	(0.057)	(0.025)	(0.026)	(0.188)	(0.185)
X^{male}	0.2395***	0.1451*	-0.0648**	-0.0490	0.0533	-0.0602
	(0.077)	(0.080)	(0.029)	(0.033)	(0.225)	(0.254)
M	-0.0156	--	0.0192*	--	-0.0648	--
	(0.025)		(0.011)		(0.075)	
$M * X$	0.1046	0.0825	-0.0233	-0.0178	0.2128	0.1125
	(0.083)	(0.079)	(0.033)	(0.033)	(0.243)	(0.252)
$M * X^{male}$	-0.2366**	-0.1655	0.0729	0.0648	-0.2722	0.0277
	(0.108)	(0.106)	(0.045)	(0.045)	(0.299)	(0.334)
Observations	8296	8296	9392	9392	9426	9426
Adjusted R^2	0.713	0.742	0.500	0.525	0.587	0.605
Course controls	Y	Y	Y	Y	Y	Y
Student f.e.	Y	Y	Y	Y	Y	Y
Instructor f.e.	N	Y	N	Y	N	Y
Term f.e.	Y	Y	Y	Y	Y	Y

Notes: *** indicates statistical significance at 1%; ** at 5%; * at 10%. Standard errors in parentheses are adjusted for multi-way clustering at two levels: students and course-by-instructor-by-term cells. Student controls include the ACT, SAT writing, SAT math, SAT verbal, race and ethnicity dummy variables, class-year dummy variables, and dummy variables indicating missing values of prior controls. Course controls include academic department fixed effects, and dummy variables indicating class size and 200-level courses.

Table 5: Effects of cross-registrant presence on academic outcomes, by quantitative skills

	Dependent variable: *Grade point*		Dependent variable: *Low achievement*		Dependent variable: *Number of subsequent courses in department*	
	(1)	(2)	(3)	(4)	(5)	(6)
A. Bottom half of proficiency test distribution						
X	-0.0060	-0.0088	0.0157*	0.0097	0.1207	0.1321
	(0.025)	(0.022)	(0.009)	(0.009)	(0.104)	(0.101)
X^{male}	0.0008	0.0231	-0.0207*	-0.0090	-0.1771	-0.1141
	(0.031)	(0.026)	(0.012)	(0.012)	(0.131)	(0.139)
M	0.0180	--	0.0025	--	0.0495	--
	(0.016)		(0.006)		(0.062)	
$M * X$	0.0061	0.0117	-0.0024	-0.0107	-0.2142	-0.2279
	(0.037)	(0.034)	(0.015)	(0.015)	(0.140)	(0.141)
$M * X^{male}$	-0.0383	-0.0612	-0.0018	-0.0053	0.3235*	0.2686
	(0.050)	(0.044)	(0.020)	(0.021)	(0.195)	(0.221)
Observations	14730	14730	16713	16713	16793	16793
Adjusted R^2	0.495	0.555	0.247	0.292	0.360	0.406
B. Top half of proficiency test distribution						
X	-0.0191	0.0006	0.0078	0.0019	0.0019	0.0040
	(0.019)	(0.019)	(0.006)	(0.007)	(0.099)	(0.093)
X^{male}	0.0070	0.0082	-0.0119	-0.0106	-0.0032	0.0693
	(0.025)	(0.024)	(0.008)	(0.009)	(0.126)	(0.118)
M	-0.0121	--	0.0074*	--	0.0283	--
	(0.014)		(0.004)		(0.061)	
$M * X$	0.0410	0.0081	-0.0115	-0.0087	-0.1168	-0.1542
	(0.029)	(0.028)	(0.010)	(0.011)	(0.138)	(0.139)
$M * X^{male}$	-0.0469	-0.0376	0.0139	0.0169	-0.0533	-0.0775
	(0.040)	(0.040)	(0.014)	(0.015)	(0.181)	(0.190)
Observations	18337	18337	20439	20439	20518	20518
Adjusted R^2	0.465	0.511	0.257	0.283	0.365	0.411
Course controls	Y	Y	Y	Y	Y	Y
Student f.e.	Y	Y	Y	Y	Y	Y
Instructor f.e.	N	Y	N	Y	N	Y
Term f.e.	Y	Y	Y	Y	Y	Y

Notes: *** indicates statistical significance at 1%; ** at 5%; * at 10%. Standard errors in parentheses are adjusted for multi-way clustering at two levels: students and course-by-instructor-by-term cells. Student controls include the ACT, SAT writing, SAT math, SAT verbal, race and ethnicity dummy variables, class-year dummy variables, and dummy variables indicating missing values of prior controls. Course controls include academic department fixed effects, and dummy variables indicating class size and 200-level courses.

Figure 1: Number of students cross-registered in at least one course, by gender and year

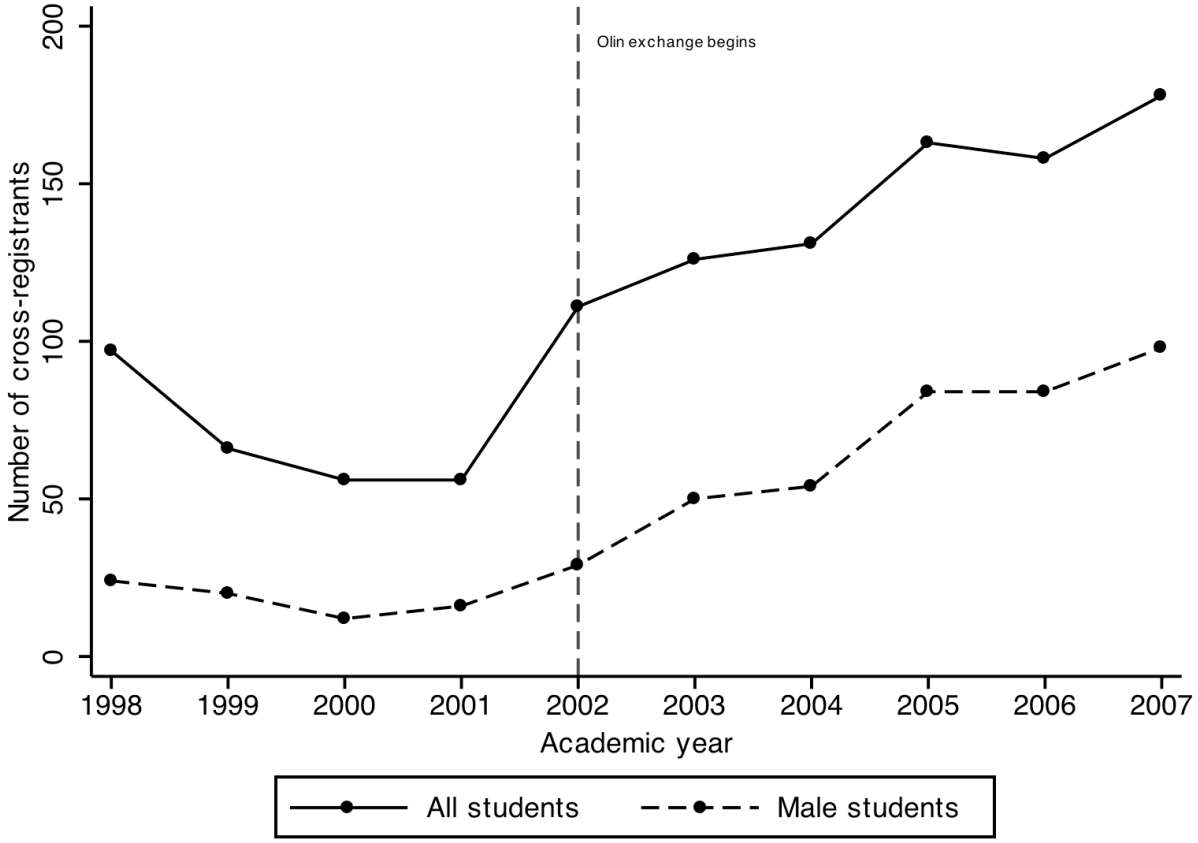
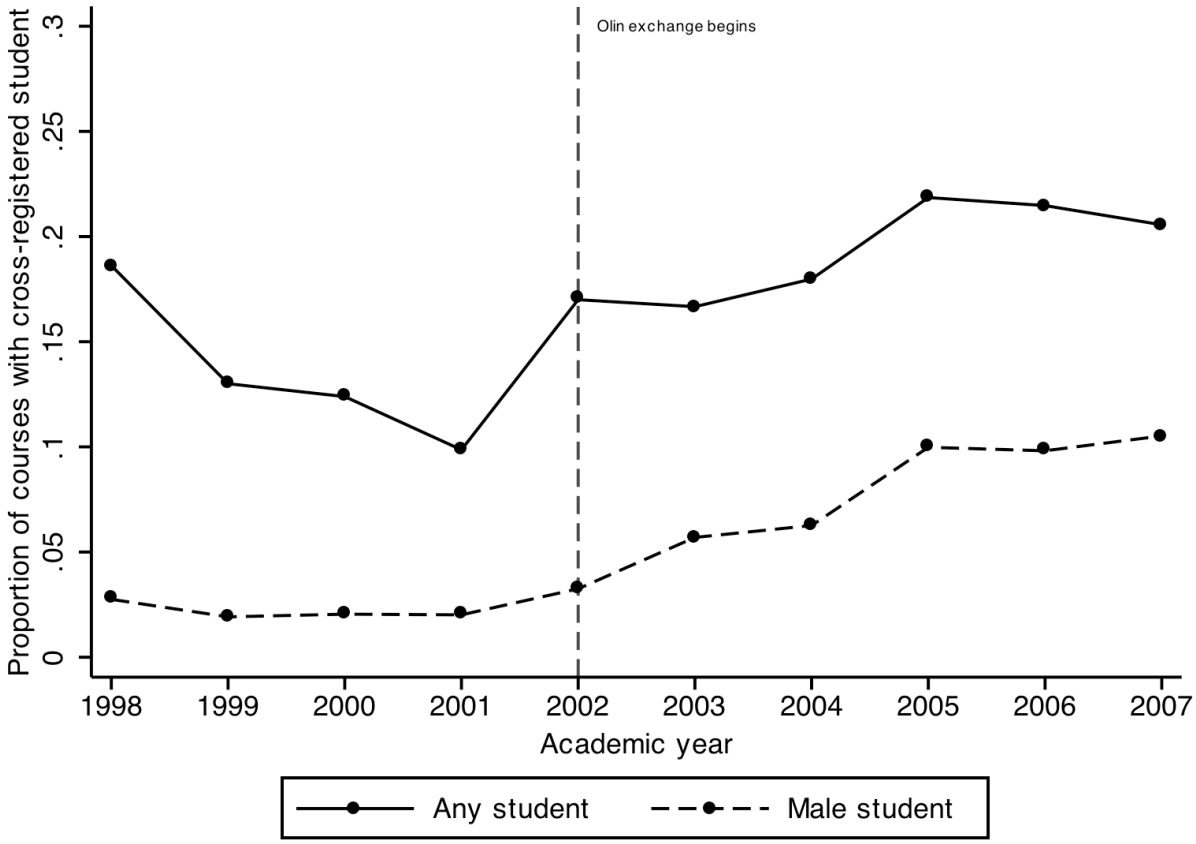


Figure 2: Proportion of courses with at least one cross-registered student, by gender and year



Notes: Each dot indicates the unweighted proportion of all courses taught in Fall and Spring semesters of an academic year that enroll a cross-registered student.