When Sarah Meets Lawrence: The Effects of Coeducation on Women's College Major Choices

Avery Calkins, Ariel J. Binder, Dana Shaat, Brenden Timpe *

October 25, 2021

Abstract

We leverage variation in the timing of U.S. women's colleges' transitions to coeducation to study how exposure to male peers affects women's human capital investments. Our event-study analyses of newly collected historical data find a 3.0 percentage-point (30%) decline in the share of women majoring in STEM after ten years of coeducation. We find little evidence that coeducation altered course offerings, class sizes, competitiveness of STEM classrooms, or the composition of the female student body. Extrapolation of our main estimate suggests that exposure to male peers in college can explain 36% of the contemporary gender gap in STEM.

JEL Codes: I21, I23, I24, J16, J24

*Calkins: RAND Corporation, acalkins@rand.org; Binder: U.S. Census Bureau, ariel.j.binder@census.gov; Shaat: American Institutes for Research, dshaat@air.org; Timpe: University of Nebraska - Lincoln, btimpe@unl.edu. The authors thank Martha Bailey, Dan Bernhardt, John Bound, Charlie Brown, Angela Cools, Manasi Deshpande, Chloe East, Rita Ginja, Donna Ginther, Sara Heller, Brian McCall, Stephanie Owen, Arpita Patnaik, Kevin Stange, Mel Stephens, Sarah Turner, Russell Weinstein, Basit Zafar, several anonymous Reviewers, and various conference and seminar participants for helpful feedback. We are grateful to Brantly Callaway and Pedro Sant'Anna for insightful correspondence about their estimator. Ellen Stolzenberg provided invaluable assistance with the CIRP Freshman Survey. Riley Knust provided excellent research assistance. Calkins and Binder acknowledge financial support from the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health under Award Number T32HD007339 and use of the services and facilities of the Population Studies Center at the University of Michigan, funded by the Eunice Kennedy Shriver National Institute Of Child Health and Human Development of the National Institutes of Health under Award Number P2CHD041028. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. A Rackham Graduate Research Grant from the University of Michigan provided additional support.

In 2016, women earned 57% of all baccalaureate degrees awarded in the United States, but only 37% of degrees awarded in STEM fields.[1] Gender gaps in the choice of major and occupation account for much of the pay gap between college-educated men and women (Brown and Corcoran, 1997; Bertrand, 2017; Blau and Kahn, 2017; Sloane, Hurst and Black, 2019). Beyond inhibiting gender equality, the lack of gender diversity in high-paying fields may also curtail economic productivity and innovation (Hunt, 2016; Hsieh et al., 2019). Designing policies to address these issues requires a complete understanding of the causes of the gender gap in choice of field.

Canonical models of college major choice emphasize heterogeneity in skills and earning potential (Willis and Rosen, 1979; Altonji, Blom and Meghir, 2012; Altonji, Arcidiacono and Maurel, 2016), but recent work attributes gender gaps largely to differences in preferences and subjective beliefs (Zafar, 2013; Wiswall and Zafar, 2015, 2018; Patnaik, Wiswall and Zafar, 2020). An important hypothesis is that these differences are shaped by gendered features of the collegiate environment, such as interactions with male students and faculty that may discourage women from entering male-dominated fields (Ceci, Williams and Barnett, 2009; Hill et al., 2010; Shapiro and Sax, 2011). Efforts to explore this hypothesis, however, have been limited by a lack of plausibly exogenous variation in women's exposure to men in typical collegiate settings.

In this paper, we analyze the impact of gendered collegiate environments on women's major choices by leveraging an important real-world setting: the decline of women's colleges in the United States. While women's colleges numbered in the hundreds in the early 1960s, most have since transitioned to coeducation. These transitions occurred at varying times and were driven by a number of factors, such as an increasingly competitive environment in higher education and the gradual liberalization of Catholic institutions (Goldin and Katz, 2011). They impacted schools at a variety of positions in the American college market—from elite colleges like Sarah Lawrence, to regional schools

[1]These statistics are from the Integrated Postsecondary Education Data System (IPEDS). Throughout this paper we define STEM to include biological sciences, physical sciences, science technology, mathematics and statistics, engineering and engineering technology, and computer science.

like Elms College of Massachusetts, to larger public institutions like Radford University of Virginia. We leverage this variation in an event-study research design that compares the evolution of majoring behavior at newly coeducational colleges to behavior at colleges that did not alter the gender mix of the student body. Under weak assumptions, our methodology provides consistent estimates of the average treatment effect on women at treated schools and is robust to heterogeneity across colleges (Callaway and Sant'Anna, 2021).

We first show that coeducation exposed women to large inflows of male peers and substantially altered their college major choices. Our event-study models leverage hand-collected information on the timing of U.S. colleges' transitions to coeducation, which we merge to data on the near-universe of U.S. baccalaureate institutions.We find that the male share of the student body increased by about 24 percentage points in the ten years following the arrival of coeducational classes. We also estimate a decrease of 3.0 percentage points (30%) in the share of women majoring in any STEM field. This response was driven by decreases in the shares of women majoring in biology, physical sciences, and math. We also find substantial decreases in the shares of women majoring in economics and in business. Correspondingly, we find increases in the shares of women choosing health, psychology and social sciences other than economics.[2] These estimates are robust across multiple comparison groups, including a synthetic control approach.

The effects of coeducation on women's college major choices could operate through two primary channels. First, coeducation could affect STEM-inclined women's propensities to enroll at former women's colleges (a *composition effect*). Second, the transition to a coeducational campus environment could influence a woman's choice of major, conditional on her enrollment choice (an *environmental effect*). We investigate these mechanisms using a novel linkage between our hand-collected data on the expansion of coeducation and survey data on the career aspirations and academic records of college freshmen. We find no evidence of a change in the *intended* major choices, career interests, or

[2]Most health majors are in nursing or allied health fields, rather than pre-professional degree programs.

family-related aspirations of entering freshman women at newly coeducational colleges. Female students' high school grade point averages rose after schools began admitting men, suggesting the changes in ultimate major choices cannot be explained by a decline in academic ability. We also find no evidence that the switch to coeducation was accompanied by changes in a range of gender-neutral determinants of major choice, such as changes in class sizes, binding capacity constraints in quantitative courses, changes in the the menu of degrees offered, increased academic competition, or labor market shocks that could alter the return to STEM degrees. We do find a statistically significant increase in the male share of faculty, suggesting that students at former women's colleges encountered more men not only as classmates, but also as instructors. However, this effect is only one-fifth as large as the effect on the male share of peers.

These results suggest that the composition effect is small relative to the environmental effect, and that the environmental effect is primarily driven by exposure to male peers rather than other channels. We conclude that our estimated effects of coeducation on women's college major choices are driven by *gendered peer effects*—i.e. greater salience and enforcement of gender roles, gender stereotypes, and marriage market considerations that accompany a shift to a coeducational environment. In the concluding discussion, we present a simple calculation that suggests that exposure to male peers explains 36% of the contemporary gender gap in STEM major choice. Though we do not have the ability to distinguish among the various gender-normative channels that may determine this effect, our findings suggest that social interventions on college campuses that target these channels may be important in increasing the representation of women in science.

This paper contributes to several strands of literature. While gender differences in student characteristics such as math ability (Turner and Bowen, 1999; Speer, 2021), STEM grades (Calkins, 2020; Goldin, 2015; Astorne-Figari and Speer, 2019), and high school course-taking (Ceci et al., 2014; Card and Payne, 2017) all contribute to gender differences in college major choice, they only account for a small portion of the gap (Patnaik, Wiswall and Zafar, 2020). We

build on literature suggesting that the gender gap is driven by heterogeneous preferences over fields of study (e.g., Zafar, 2013; Wiswall and Zafar, 2015, 2018; Patnaik, Wiswall and Zafar, 2020), finding that exposure to male peers in college may be an important determinant of these preferences.

Our study of real-world exposure to male peers furthers literature that uses laboratory and field experiments to analyze the role of the social environment on gender gaps in economic behavior (see Bertrand, 2011 for a review). Women's willingness to compete and aptitude in competition have been shown to be lower in mixed-gender relative to single-sex environments (Gneezy, Niederle and Rustichini, 2003; Niederle and Vesterlund, 2007, 2008; Kamas and Preston, 2012). In addition, these differences appear to depend on prevailing social norms (Gneezy, Leonard and List, 2009), exposure to male peers in elementary school (Booth and Nolen, 2012*a*,*b*), and whether actions will be observed by male peers (Bursztyn, Fujiwara and Pallais, 2017). Our paper helps link this literature to real-world gender gaps in educational outcomes.

Our analysis also contributes to a growing literature on gendered peer effects in post-secondary educational environments. Prior work tends to leverage highly specialized settings—such as STEM doctoral programs in Ohio (Bostwick and Weinberg, 2018), the first coeducational class at West Point (Huntington-Klein and Rose, 2018), or introductory classes at a single Dutch or Swiss university (Zölitz and Feld, 2018; Shan, 2021). In addition to analyzing a wide range of U.S. colleges, one feature of our context is that we study exposure to men in a mostly female environment, rather than exposure to women in a mostly male environment. This is closer to representative of the American higher education landscape today, where women outnumber men on most campuses. Our results may also have broader implications for the labor market—for instance, they may shed light on the expected effects of inflows of men to traditionally female-dominated occupations like nursing.

Finally, our paper contributes to literature on the educational roles of women's colleges. Early studies found that graduates of women's colleges earned higher income and had higher occupational prestige than graduates

of coeducational colleges (Riordan, 1994; Tidball, 1980, 1989). Women at women's colleges also tend to report greater satisfaction with educational aspects of the college experience and greater support in their educational endeavors (Smith, 1990; Miller-Bernal, 1993; Kinzie et al., 2007).[3] Related to our paper is a case study by Billger (2002) that found a decrease in the number of women choosing traditionally male majors and occupations after a single women's college transitioned to coeducation. We expand on this literature by exploiting variation in the date of transition to coeducation in a causal framework and by considering the near-universe of historical women's colleges in the United States.[4]

1 Women's Colleges' Transitions to Coeducation

1.1 Historical context

Women's colleges have been a part of higher education in the United States since 1836. Most early women's colleges were located in the Northeast and were progressive institutions designed to expand educational opportunities for women (Chamberlain, 1988). Their footprint grew as Protestant- and Catholic-affiliated schools opened, mostly in the South and Midwest (Harwarth, Maline and DeBra, 1997). Non-denominational public universities, such as Texas Women's University, and private women's colleges, such as Sarah Lawrence, added to the ranks in the 20th century. While the precise number is a subject of debate, by the 1960s, it was estimated that between 233 and

[3] In addition, Dasgupta and Asgari (2004) found that students at a women's college were less likely to form negative stereotypes about women's STEM abilities than woman students at a coeducational college.

[4] In an important study of the historical drivers of the rise and decline of single-sex education in the United States, Goldin and Katz (2011) also assembled data on institutions' dates of conversion to from single-sex environments to coeducation. Although our data was collected independently, the time periods of the two datasets partially overlap, and we verify that our transition dates largely coincide during the era where it is possible to compare them. We thank the authors for generously providing their data.

315 U.S. colleges served a strictly female student body (Harwarth, Maline and DeBra, 1997).

The modern decline of women's colleges began in earnest in the late 1960s. While much of this trend was driven by changes in demand for single-sex education, the timing of and approach to the transition to coeducation was characterized by substantial heterogeneity (Miller-Bernal and Poulson, 2004; Thomas, 2008). The liberalization of Catholic education following the Vatican II council opened the door to coeducation in traditional Catholic schools (Goldin and Katz, 2011). Some institutions also worried about the implications of proposed equal rights legislation (Thomas, 2008). These forces led to a wave of changes, with roughly one-half of women's colleges disappearing or converting to coeducation by the early 1970s. Other schools resisted coeducation but eventually followed suit, with more colleges ending single-sex education nearly every year since. Still more chose to embrace their single-sex mission for good: Thirty-four women's colleges, and only three men's colleges, were in operation in 2016.

Our research design leverages variation in the existence and timing of coeducation events across comparable groups of institutions to estimate causal effects of coeducation on women's choices of major. This design is internally valid under the assumption that the timing of coeducation at a particular school is not correlated with idiosyncratic trends in women's major choices at that school. We provide evidence supportive of this assumption.

1.2 Expected effects of coeducation

Guided by the prior literature and a simple Roy model (presented in Appendix A), we discuss a number of mechanisms by which women's field choices at former women's colleges may be affected by the transition to coeducation.

First, women's college j's transition to coeducation may alter the enrollment decisions made by women high school seniors. Women who prefer a single-sex environment may substitute away from j and toward other women's colleges, while women who prefer a co-educational environment may substitute

toward j. If women's pre-enrollment propensities to major in STEM are correlated with preferences for single-sex educational environments, we may see changes in majoring behavior that are driven by changes in the *composition* of students.

For women who continue to enroll at college j, the transition to coeducation may impact major choices by changing the (perceived or realized) payoffs or feasibility of majoring in STEM relative to non-STEM fields. This effect, which we refer to as the *environmental effect*, may itself be the product of several different factors associated with the transition to coeducation. For example, coeducation could introduce psychological costs related to competition with men for course grades and professors' attention (Niederle and Vesterlund, 2007). This could steer women toward non-STEM majors, to the extent that men disproportionately entered STEM classrooms (which we show occurred in our setting; see also Kahn and Ginther, 2017). The presence of men on campus could also increase the salience of gender norms and stereotypes (Akerlof and Kranton, 2000; Schmader, 2002; Steele, 1997), leading women to choose non-STEM fields out of a sense of social conformity. Gender norms could also play a role through the marriage market if majoring in a high-paying, STEM-related field is seen as a negative signal to potential future spouses (Bertrand, Kamenica and Pan, 2015; Bursztyn, Fujiwara and Pallais, 2017). We refer to the collection of mechanisms related to changes in the gender composition of women's classmates as *gendered peer effects*.

Other mechanisms, such as changes in the educational inputs provided by newly coeducational colleges, may also factor into the environmental effect. For instance, if colleges hire more male faculty to prepare for the arrival of male students, one byproduct could be a weakening of role model effects that have been shown to draw women into quantitative fields (Carrell, Page and West, 2010; Kofoed and McGovney, 2019; Bottia et al., 2015). In addition, even holding the gender mix of the student body constant, coeducation could influence women's major choices through gender-neutral peer effects—i.e., by changing the average ability of peers and, potentially, the competitiveness of STEM courses. Moreover, it is possible that coeducation accompanied cur-

ricular changes made in the face of budget shortfalls that created capacity constraints in STEM courses (e.g., a de-funding of lab spaces).

2 Data and Research Design

Because little reliable information exists on the history of women's colleges in the United States, we start by manually collecting a list of women's colleges and their dates of transition to coeducation, covering the period of 1966-2016. We view this as one important contribution of our study. Our collection effort, which we detail in Appendix B, identified 219 institutions that were women's colleges at some point in their history. 155 of these institutions transitioned to coeducation during our sample period.

Figure 1 documents the distribution of the years of transition to coeducation at women's colleges in our sample. The modal transition date is between 1969 and 1971, before the passage of Title IX in 1972.[5] Although Title IX affected the educational environment and outcomes of women at public institutions (Rim, 2020), it is unlikely to have caused meaningful change in the women's colleges in our sample—the majority of which were private and received little federal funding.[6,7]

[5]Our sample has switch dates that are more concentrated in the late 1960s and early 1970s than our full data on former women's colleges. A number of institutions that changed to coeducation in the 1980s and 1990s either closed shortly thereafter, did not have STEM programs in 1966, or had large numbers of male graduates before their official transition to coeducation.

[6]At private colleges, Title IX initially only applied to the particular program receiving federal money until 1987, rather than *all* programs at that college (see footnote 15 of Rim (2020)). Aside from student aid grants, federal money was probably not a large driver of decisions at the (mostly small and private) colleges in our sample: direct federal appropriations were unlikely, and the median expenditure on research per year was $0, according to the IPEDS finance data.

[7]Appendix Table A2 presents linear probability model estimates of ever transitioning to coeducation on the full sample, and transitioning before 1972 on our sample of switchers. We included controls for region, religiosity, size (i.e. log enrollment in the first year of observation) and college selectivity rating. While most of these variables significantly predict the existence and timing of a transition to coeducation, they jointly explain less than 22% of the variation in each outcome, suggesting the availability of substantial institutional heterogeneity to analyze in an event-study framework. (This, of course, does not prove that coeducation was randomly assigned conditional on these variables. We test this identifying

2.1 Data on student outcomes

Our main analyses rely on data from the Higher Education General Information Survey (HEGIS) and the Integrated Postsecondary Education Data System (IPEDS). HEGIS and IPEDS provide information from 1966–1986 and 1987–2016, respectively, on the number of degrees awarded by year, institution, major, and gender. HEGIS collected data from all institutions of higher education subject to Title IV of the Higher Education Act of 1965. IPEDS collects data from all post-secondary institutions in the United States who participate in federal financial assistance programs—a larger sample than HEGIS. Many schools participate in IPEDS voluntarily even if they do not accept federal financial assistance, so the coverage of U.S. baccalaureate-degree-granting institutions in IPEDS is nearly universal. For the remainder of the paper we will refer to the HEGIS and IPEDS data as IPEDS for concision.[8]

We use these data to construct a measure of the share of graduates of gender G that earn a degree in field f at institution i in academic year t:

$$s^G_{f,it} = \frac{\sum_\mu D^G_{\mu,it} \times \mathbb{1}(\mu \in f)}{\sum_\mu D^G_{\mu,it}}. \tag{1}$$

$D^G_{\mu,it}$ is the number of degrees in major μ earned by graduates from institution i in year t of gender G. The data only provide a measure of degrees awarded: we do not observe individuals who matriculate but do not finish their degrees, or the time to completion for those who finish their degrees. We therefore cannot investigate intermediate channels of degree progress.

Our primary focus is on the share of women completing degrees in quantitative fields such as math, biology, and the physical sciences. We pool these three majors, as well as engineering, engineering technology, and computer science, into a comprehensive STEM field. We also examine economics. Finally, we aggregate all other individual majors into nine fields, and examine

assumption extensively below.)

[8]The data from the HEGIS 1970 issue on degrees awarded has historically not been available in digital form. Because this is an important time period for our study, we digitized the 1970 issue of the HEGIS data on degrees awarded.

the effects of coeducation on the full distribution of degrees awarded across these eleven fields (see Section 3.2).

After merging the hand-collected data on schools' dates of transition to the IPEDS data, we made a number of sample restrictions, which we detail in Appendix B.3. These restrictions exclude cases in which the transition to coeducation was a gradual process rather than a single event, some likely errors in the IPEDS data, institutions that did not regularly offer STEM programs, and institutions that closed shortly after transitioning. Finally, we drop schools that we do not observe for at least four years prior to their transition to coeducation. We are left with a clean sample of 77 institutions that switched to coeducation. As shown in Figure 1 (blue bars), the distribution of years of transition for our analysis sample is similar to that of the full sample of switching schools – most institutions began admitting men in the late 1960s and early 1970s, but a substantial number remained female-only until much later. Our sample also includes 927 schools that form the basis of our comparison groups, including 29 women's colleges that did not open their doors to men during our sample period.[9]

To examine some of the mechanisms by which coeducation may alter women's decisions to major in STEM, we require panel information on the underlying abilities and curricular preferences of women and their peers. We collect these data from The Freshman Survey (TFS), which is produced by the Higher Education Research Institute (HERI) at UCLA. Our sample includes nearly 13 million college freshmen who entered baccalaureate institutions between 1966 and 2006. The survey elicits information from entering freshmen at these participating institutions on their college and career plans, academic preparedness, and other characteristics. Staff at HERI linked these data to our data on dates of transition to mixed-sex classes, allowing us to study the evolution of these outcomes before and after schools adopted coeducation.

One drawback of The Freshman Survey is that it is not collected at all

[9]Institutions that did not initially offer STEM programs or that closed shortly after the transition make up the bulk of coeducation events dropped from the sample. Returning them to the sample, which gives us a total of 119 coeducation events, does not substantially change our main results (see Appendix Figure A1).

schools, nor is it collected regularly at all participating colleges. Because our research design requires observing schools both before and after the transition to coeducation, we limit our sample of switching colleges to institutions that appear at least once in the five years before and once in the 10 years after the reform. This leaves us with 30 treated institutions. As shown in Appendix Figure A2, most but not all of these schools are present in any given year: we observe the average treated school 9.5 times in the 15-year window around the switch to coeducation. Our pool of comparison schools is constructed analogously to our analysis of degree completions: our preferred approach relies on 31 colleges that admit virtually no men during our sample, but as a robustness check, we also construct a larger comparison group of 127 female-only and coeducational schools that did not switch during our sample period. We restrict our comparison group to schools that provide data in the same set of academic years as our treated schools.

2.2 Summary statistics

Table 1 presents average institutional characteristics, as measured in the IPEDS data, for our main samples. Column 1 reports sample means for our preferred sample of switching women's colleges in the year prior to each school's switch to coeducation. Columns 2 and 3 report means for our samples of comparison colleges – non-switching women's colleges and all non-switching colleges, respectively – observed in the same set of academic years as the set analyzed in Column 1.

As shown in Columns 1 through 3, the comparison groups are similar in the share of women choosing STEM. The switchers are also similar to the always-women's colleges in initial enrollment and the number of graduate degrees awarded. However, some important differences exist as well. For example, the switching sample is 80% privately controlled, compared to 100% of always-women's colleges and 56% of all non-switchers. Half of our switchers were at some point affiliated with the Catholic church, compared to 31% of the always-women's colleges and no more than 6% of the schools in the other

comparison group. Switchers are also less selective on average than always-women's colleges, and more selective than non-switchers in general.[10] Our empirical strategy adjusts for differences in these observable characteristics, and is valid if switching and non-switching schools exhibit *conditional* parallel trends in major choice after such an adjustment.

Finally, Column 4 records average institutional characteristics in the year prior to the transition to coeducation for our sub-sample of 30 switching colleges that contain sufficient HERI data. Although this sample is slightly more private and selective than the main switching sample (Column 1), it is generally quite similar. Indeed, we show that the two samples yield very similar estimates of the effects of coeducation on women's STEM major choices (see Section 4.1).

As men joined the student body at former women's colleges, their presence was felt disproportionately in traditionally "male" disciplines. Figure 2 reports the distribution of majors chosen by students at our sample of former women's colleges and a comparison group of coeducational schools. Relative to men at schools that were coeducational throughout our sample period (light blue), the first cohorts of men who entered former women's colleges were less likely to choose heavily quantitative majors (medium blue). They were less than half as likely to major in STEM and more likely to specialize in health or social sciences other than economics. However, relative to women at these schools (dark blue), they were still more likely to choose relatively quantitative, high-return fields such as STEM, economics, and business. This suggests that the gender mix of the classroom in these fields changed substantially, as did classrooms in some other fields.

2.3 Empirical strategy

We exploit variation in the timing (and existence) of women's colleges' transitions to coeducation to estimate effects of coeducation on women's major choices and the determinants of those choices. Because we expect these

[10] "Selective" refers to a 1972 Barron's rating of 1-3. Some institutions have missing data—we code these institutions as non-selective.

outcomes to evolve dynamically, we report event-study estimates that trace out changes in majoring behavior at newly coeducational colleges relative to colleges that did not change the gender mix of the student body.

Our empirical strategy relies on our ability to use our pool of comparison colleges to construct the counterfactual distribution of majors that women would have chosen in the absence of male classmates. We do so by implementing a slightly modified version of the doubly-robust estimator of Callaway and Sant'Anna (2021). A detailed description of our implementation of the estimator can be found in Appendix C. The doubly-robust estimator relies on two complementary strategies: an inverse propensity weighting approach that pares our comparison group to schools with similar observed characteristics, and an outcome regression that uses the resultant comparison group to impute a counterfactual trend in the outcome of interest, conditional on observed characteristics. We use this approach to estimate $\alpha_{g,t}$, the average treatment effect on the treated (ATT), separately for each year t and "cohort" of switching schools g (e.g., $g = 1969$ for all schools that adopted coeducation in the 1969-1970 academic year).

This approach is robust to heterogeneity in treatment effects across schools and to concerns about weighting and inappropriate comparisons that may bias estimates of the ATT that rely on conventional regression-based techniques (Callaway and Sant'Anna, 2021; Sun and Abraham, 2021; Borusyak, Jaravel and Spiess, 2021; Goodman-Bacon, 2021; see Appendix C). In addition, this approach provides flexibility to aggregate the individual estimates $\hat{\alpha}_{g,t}$ based on our choice of target parameter. In our main results, we report event-study estimates by aggregating over time relative to the switch to coeducation:

$$\hat{\beta}_k = \sum_g \sum_t \frac{\omega_{g,t}\hat{\alpha}_{g,t} 1\{g - t = k\}}{\sum_g \sum_t \omega_{g,t} 1\{g - t = k\}} \tag{2}$$

where $\omega_{g,t}$ is the number of female students in the first year we observe schools in cohort g. This effectively weights our event-study estimates by school size. The coefficients β_k flexibly trace out the effect of coeducation at each year k

relative to the transition.

In addition, to provide a summary measure and increase statistical power, we report overall difference-in-difference estimates of the effects of coeducation on field choices and their underlying determinants. Because we expect the effect of coeducation to manifest gradually – as men filter into the school gradually, and as successive cohorts of women are exposed to more years of coeducational courses and campus life – we estimate "short-run" and "long-run" effects:

$$\hat{\beta}_{SR} = \sum_g \sum_t \frac{\omega_{g,t}\hat{\alpha}_{g,t} 1\{0 \leq g - t \leq 4\}}{\sum_g \sum_t \omega_{g,t} 1\{0 \leq g - t \leq 4\}} \tag{3}$$

$$\hat{\beta}_{LR} = \sum_g \sum_t \frac{\omega_{g,t}\hat{\alpha}_{g,t} 1\{5 \leq g - t \leq 9\}}{\sum_g \sum_t \omega_{g,t} 1\{5 \leq g - t \leq 9\}} \tag{4}$$

Our preferred estimates construct the counterfactual trend for a given group g using a pool of comparison schools that consists of women's colleges that never switched to coeducation, as well as switching schools that did not transition for at least 10 years after g.[11] Given that we found imbalances on selectivity and Catholic affiliation between switchers and non-switching women's colleges in Table 1, we specify these variables as controls in the propensity score model and the regression-based prediction of trends in the outcome. As a robustness check, we repeat the exercise with a comparison group of "never-treated" women's colleges only (dropping those that eventually switch), and then again with a comparison group composed of all non-switching schools, including coeducational institutions. Because coeducational institutions also differ from switching institutions in terms of size and baseline shares of students majoring in STEM (Table 1), when using this set of schools as a com-

[11] This approach corresponds approximately to the "not-yet-treated" version of the estimator of Callaway and Sant'Anna (2021). We use only schools that switch at least 10 years later because this allows us to maintain a stable comparison group for a decade after cohort g's transition.

parison group we expand our vector of controls to include school size (proxied by number of degrees awarded in pre-reform years) and a linear trend in the share of students graduating in STEM in the five years before treatment year g.

For all estimates, we conduct inference using a block bootstrap clustered at the school level with 1,000 replications. We report 95% pointwise confidence intervals in our event-study figures.[12]

Our estimates rest on the identifying assumption common to all difference-in-difference designs: in the absence of the switch to coeducation and conditional on covariates X, women's choices of major at our treated schools would have trended in parallel to those in the comparison group. While this assumption is fundamentally unverifiable, our estimation strategy allows us to conduct several placebo tests. First, our event-study approach enables us to estimate β_k before the reform. A series of coefficients that depart significantly from 0 would be suggestive of a violation of "parallel trends." Second, as described above, we test the sensitivity of our results to the choice of comparison group. Third, in Appendix D, we conduct a synthetic control analysis on the main sample and verify that the results match those generated by the event studies.

2.4 Tests for confounding labor market shocks

In addition to these indirect tests, we directly test whether the switch to coeducation was correlated with certain labor market conditions that might determine women's STEM major choices. We use the March CPS to construct four measures of labor market conditions at the state-year level (Ruggles et al., 2020) and then test for changes in these indicators that coincide with the switch to coeducation (Pei, Pischke and Schwandt, 2018). If these estimates depart substantially from 0, they would cast doubt on the validity of the parallel trends assumption.

The results of these exercises are shown in Figure A3. We find no effect

[12] In Appendix C, we also report event-study results with simultaneous confidence intervals that account for concerns about multiple testing across event-study periods (Callaway and Sant'Anna, 2021).

on the unemployment rate, a common measure of the overall health of the labor market. The event-study is quite flat, with the 95% confidence interval generally ruling out any effects larger than one-half of 1 percent in either direction. We also find no correlation between coeducation and the ratio of employment in STEM-related occupations to employment in non-STEM occupations among college-educated workers (our estimated long-run effect is -0.0013 with a standard error or 0.0035). We do find a short-run increase in the earnings of *male* workers in STEM occupations (relative to male workers in non-STEM occupations) following coeducation (panel C). However, to the extent that this estimate indicates an increase in demand for STEM workers, we would expect it to *increase* the share of students majoring in STEM fields. Furthermore, this effect dissipates after a few years, and the estimated long-run effect is indistinguishable from 0 (0.043, s.e. 0.034). Finally, the relative earnings among *women* in STEM occupations exhibits a slight *decrease*, but just as is the case for men, it quickly dissipates and is indistinguishable from 0 in the long run (-0.014, s.e. 0.023).

These results suggest that relevant labor market conditions were orthogonal to the timing of coeducation. They are consistent with the findings of Goldin and Katz (2011), in their partially overlapping sample, that business cycle and other macroeconomic conditions did not appear to drive coeducation.

3 Results

3.1 The effects of coeducation on exposure to male students and quantitative major choices

We start by examining how the change to coeducation increased women's exposure to male peers. We use two outcome variables to measure women's exposure to men: (a) the male share of all bachelor's degrees awarded at each institution and (b) the male share of total undergraduate enrollment at each institution. Panels A and B of Figure 3 report event study estimates for each of these outcomes. Coeducation induced a gradual increase in the male share of

graduates beginning the same year as the transition to coeducation. The long-run estimate $\hat{\beta}_{LR}$ in our preferred sample is 0.192, indicating that the male share of graduates increased by 19.2 percentage points by years 5 to 9 post-transition (Column 1 of Table A2). Coeducation induced a more immediate, and larger, increase in the male share of undergraduate enrollment, with an estimated long-run effect of 0.216 in our preferred sample. Combined with the distribution of men's major choices shown in Figure 2, these results suggest that women, particularly in STEM-related classrooms, experienced substantial inflows of male peers.

Next, we examine the effects of coeducation on women's choices of quantitative majors. Panels C through F of Figure 3 report event study estimates for the shares of women graduating with degrees in biology, math, physical sciences, and economics. Estimated long-run effects for these outcomes are presented in Columns 3 through 6 of Table 2. The event-study paths are all relatively flat until event time 1 or 2, then trend downward sharply until event time 6 or 7, and finally level off at a negative long-run value. The individual yearly point estimates for years 6 through 10 approach statistical significance for math and economics, and are statistically significant for biology and physical sciences. The sharp drop observed after event time 1 coincides with the class of female students who were sophomores when the first coeducational class enrolled at the college (assuming that students graduate in four years). The preferred long-run estimates indicate a 1.7 percentage-point decline in the share of women majoring in biology, a 0.6 percentage-point decline for physical sciences, a 0.6 percentage-point decline for math, and a 0.6 percentage-point decline for economics, in years 5 to 9 post-transition. These declines are relative to contemporaneous trends in women's major choices at similar colleges that did not transition to coeducation. The estimates are all statistically significant at the 10% level or lower. Panels B and C of Table 2 report estimates based on our alternative comparison groups, and show slightly larger and more statistically significant declines.

In Figure 4, we present similar estimates for our comprehensive measure of STEM. The estimated long-run effects are also presented in Table 2. Once

again, the event-study path is relatively flat until event time 1 and then drops sharply, indicating a drop beginning with students who were most likely sophomores when the college transitioned to coeducation. In the long run, we find a 3.0 percentage-point decline in the share of women majoring in STEM, relative to contemporaneous trends in women's major choices at similar colleges that did not transition to coeducation (Panel A, Column 1). To provide context for this estimate, the second row records the share of female students majoring in STEM during event times 5 to 9, under the counterfactual scenario in which the college did not transition to coeducation. We compute these counterfactual shares by adding the contemporaneous average trend in STEM majoring at comparison schools (adjusted for controls) onto the pre-treatment average share of women students majoring in STEM at switching schools. Table 2 reports a counterfactual share of 9.9%—relative to this baseline, coeducation caused a $3.0/9.9 = 30\%$ decline in the share of women majoring in STEM.

3.2 The effect of coeducation on the full distribution of major choices

What majors were chosen by women at formerly-women's colleges in lieu of STEM and economics? To answer this question, we classify all individual majors into eleven major concentrations: STEM, art, business, economics, education, health, home economics, humanities, psychology, social sciences other than economics and psychology, and all other fields. (For further detail, see Appendix E.1.2.)[13] Panel A of Figure 5a presents estimates of $\hat{\beta}_{LR,f}$ for each field f, ranked in increasing order.[14] STEM experienced the largest outflows of women, with business and economics experiencing smaller outflows. In contrast, we estimate that health, psychology and social sciences other than economics experienced inflows of women, although none of these estimates are

[13] "Other" contains a set of small majors, many of which were not likely offered by small private women's colleges (e.g. agriculture and forestry). The ones that likely were offered include interdisciplinary majors and social services.

[14] Appendix Figure A4 and Appendix Tables A2 and A3 present the event study and long-run estimates for other fields.

statistically significant.

Figure 5b presents semi-elasticity responses that account for the relative sizes of each field. For a given field f, the semi-elasticity is given by $\epsilon_{LR,f} = \beta_{LR,f}/c^W_{LR,f}$, where $c^W_{LR,f}$ is the long-run share of women majoring in f in the counterfactual scenario in which the school did not transition to coeducation. With this specification, the observed gendered pattern of substitution strengthens: Economics, STEM, and business experience the greatest (proportional) outflows of women, while social sciences other than economics, health, and psychology experience the greatest inflows.

4 Mechanisms

What drove these responses of women's field choices to coeducation? As discussed in Section 1.2, if the switch to coeducation affected women's enrollment decisions, these responses could be attributed to a *composition effect* that altered the preferences or abilities of the female student body. Alternatively, these responses could be due to *environmental effects* that influence women's field choices conditional on having enrolled at a transitioning college. Environmental effects could be attributed either to gender-related factors stemming directly from the exposure to male peers, or non-gendered influences stemming from changes in the educational inputs provided by the college (e.g. student/faculty ratios, course offerings, availability of lab spaces, ability of peers, and other factors). We explore these potential mechanisms below. While it is impossible to precisely quantify all of these mechanisms with the higher education data that is available during our sample period, the evidence we present suggests a prominent role for *gendered peer effects*, in which women's field choices were directly affected by exposure to male classmates.

4.1 Effects on student preferences

We leverage the HERI data to estimate the effects of coeducation on various proxies for women's field preferences. To verify that inferences drawn from the

HERI sub-sample are externally valid with respect to the full IPEDS sample (recall Section 2.1), we first re-estimate the STEM event study and long-run estimates using only the 30 switchers identified in the HERI sub-sample. Results based on the preferred specification of the comparison group are shown in Appendix Figure A1. We find very similar effects in the two samples.

The first four panels of Figure 6 show event-study estimates of the effect of coeducation on the share of newly-enrolled freshman women *intending* to major in STEM, biology, physical sciences, and math.[15] The event studies are mostly flat, absent a few statistically insignificant and transitory changes. In all cases, the event study figures bear little resemblance to the sharp decreases in *actual* choice of major that we observe in Figures 3 and A1.[16]

Table 3 summarizes these estimates by presenting long-run effect estimates for each major, using each of our three comparison groups. Across all outcomes and comparison groups, our estimates are never statistically distinguishable from 0.

These estimates suggest the effects on STEM degree completions cannot be explained by coeducation-driven changes in intended choice of major among entering freshman women. Appendix Figure A9 reports results from a parametric bootstrap that provides insight into the statistical robustness of the difference between the estimates from these two samples (Johnston and DiNardo, 1997; Goodman-Bacon, 2018).[17]. Under a range of plausible assumptions, we can rule out effects on STEM *intentions* of the magnitude that would be necessary to explain our estimated effects on STEM degree *completions*. For example, under the assumption that the two estimates are uncorrelated, the 90% confidence interval for their difference excludes 0. Furthermore, our confidence interval tightens considerably if the two effects are positively correlated.

The final two panels of Figure 6 present event-study estimates for career

[15]The HERI data does not report intention to major in economics separately from other social sciences, so we do not include it in this analysis.

[16]Note that this exercise is *not* a pure test of the composition effect: students could rationally anticipate an environmental effect from the arrival of male peers and change their intended coursework, even before matriculating. Nonetheless, the small effects on intended major that we document would suggest little scope for composition effects.

[17]This procedure is described in detail in Appendix C.1

and family ambitions. We find no evidence of a decline in the share of freshman women who report a STEM field as their probable career (Panel E). If anything, there is a slight increase in event years 6 and 7 that then reverses. Panel F tests for changes in the share of freshman women who say that raising a family is either a "very important" or "essential" personal goal. We estimate a remarkably flat pattern, suggesting that coeducation did not increase the family-related goals of the subsequent female student body (perhaps at the cost of quantitative coursework and career goals).

Is it plausible that coeducation would induce such a small change in coursework, career and family preferences among freshman women yet such a large change in their eventual field choices? Previous work suggests that students' career plans are not set in stone when they matriculate; rather, their decisions are shaped substantially by the campus environment and other factors over the course of their collegiate careers (Zafar, 2011; Stange, 2012; Gong, Stinebricker and Stinebrickner, 2019; Owen, 2020). Even among students that declare a STEM major, nearly half end up switching to another field (Altonji, Arcidiacono and Maurel, 2016; Astorne-Figari and Speer, 2019). These findings suggest that women may not anticipate the effects of coeducation on their field choices at the time of enrollment. On the other hand, if preferences regarding the gender mix of the collegiate environment matter more than—or are uncorrelated with—preferences for specific fields, women could rationally expect that coeducation will affect their choice of field yet still select into newly coeducational schools in an unbiased manner. We view our results as consistent with either of these classes of educational choice models.

4.2 Effects on student ability

Panels A through C of Figure 7 test for changes in freshman women's reported high school grade point averages (GPAs). GPA is the most consistently-observed measure of student ability in the HERI data.[18] Following the switch

[18]The HERI Freshman Survey consistently provides GPA only in letter-grade categories, so it is not possible to reconstruct a student's specific GPA. We instead code GPA as the share attaining a GPA that is equivalent to a particular letter grade or higher.

to coeducation, there was no decrease in the share of female freshman earning at least an A or at least an A-, in high school (effects at lower points in the distribution are similar and shown in Appendix Figure A10). If anything, our preferred estimates show slight long-run increases that are statistically insignificant.

As a final indirect test for changes in women students' STEM-related preferences and abilities, we use the IPEDS data to examine whether coeducation caused STEM-inclined women to enroll at comparable institutions that remained women-only. Specifically, we examine whether the share of students majoring in STEM at a given women's college i in year t is positively correlated with the enrollment-weighted share of (formerly) women's colleges, in the same Census region and with the same Barron's selectivity rating as i, that transitioned to coeducation before t. Appendix F provides further details on the methodology. Our results indicate that transitions to coeducation exerted trivial effects on STEM degree completions among women at comparable women's colleges. In sum, our various tests indicate that the *composition effect* of coeducation on women's college major choices, if it exists, is quite small.

4.3 Other changes in the school environment

One important remaining possibility is that the coeducation-induced inflow of new students affected the level of competition (e.g. for grades or other career-advancing opportunities like internships) in STEM-related classrooms. We explore this possibility by analyzing how coeducation affected women's *relative* positions in the ability distribution. Given that men disproportionately choose quantitative majors—including men enrolling at newly coeducational colleges (Figure 2)—we would expect a decline in women's positions in the ability distribution to result in increased competition particularly in STEM-related classrooms. We measure freshman women's average ability position at school i in year t by computing the share of *all* entering freshman with a lower high school letter grade GPA than a given freshman woman, and then averaging this measure across all freshman women. Panel C of Figure 7 reports event-study

estimates of the effect of coeducation on this relative ability measure. We document a statistically significant long-run *increase* in the average woman's ability position of about 4-5 percentiles. This finding casts doubt on the hypothesis that women at newly coeducational colleges were competed out of STEM classrooms by higher-ability male peers.

Alternatively, the switch to coeducation may have altered women's field choices because certain classrooms became disproportionately crowded — or even so full that women could not enroll in them. While we do not observe detailed data on classroom size, we can evaluate the possibility that average classroom size changed by analyzing student-faculty ratios. The results in Panel D of Figure 7 suggests that schools quickly increased their faculty ranks to keep up with the influx of male students. After a small and statistically insignificant bump in years 2 and 3, the student-faculty ratio kept pace or even fell at former women's colleges relative to our preferred comparison group.

A related possibility is that transitioning colleges changed course offerings in a bid to attract additional students or due to budgetary shortfalls. Although we exclude colleges that closed shortly after the transition to coeducation (and verified in Figure A1 that including these schools does not affect our STEM estimates), it is possible that operational schools de-funded lab spaces or otherwise changed course offerings in a manner that discouraged quantitative field choices. While we do not observe course catalogs, we construct a proxy by counting the number of individual majors in which a degree was awarded at school i in year t. The resulting estimates in panel E of Figure 7 are very small and statistically insignificant, suggesting little change in the menu of degree options available to students.

In Appendix section G, we perform a complementary test of the hypothesis that coeducation created capacity constraints that limited the production of STEM graduates. The logic behind this test is that in the face of *binding* resource constraints, a women's college might choose to open its doors to men (to expand tuition revenue) and curtail its most expensive programs (to cut costs). To assess this hypothesis, we compare overall growth in degrees earned by field following coeducation to field-specific estimates of the marginal

cost of instruction from Hemelt et al. (2018). Contrary to the hypothesis, we find a slight *positive* correlation between marginal cost and growth in degrees earned, both within STEM fields and overall. These findings suggest that the estimated effect of coeducation on women's STEM major choices was not driven by binding resource constraints.

Finally, we test for *gendered role model effects* by examining the gender composition of the faculty. In fall 1971, 51% of faculty were female at women's colleges, compared to only 23% at coeducational institutions. Although information on faculty in the IPEDS data is spotty, we observe faculty numbers by sex in select fall semesters between 1971 and 1989, after which it is reported consistently. Event-study estimates of the female share of faculty are presented in Panel F of Figure 7. We do observe a small decrease in the female share of faculty, with a statistically significant long-run estimate of 4.5 percentage points.[19] However, this effect is about one-fifth the size of the observed 21.6 percentage point increase in the male share of undergraduates. Therefore, while we cannot completely rule out gendered role model effects, the results suggest that *gendered peer effects* were the primary mechanism through which coeducation influenced women's college major choices.

5 Discussion

Our paper takes advantage of a unique natural experiment in the history of American higher education—the transitions of hundreds of women's colleges to coeducation at varying times during the 1960s-2000s—to isolate the contributions of gendered collegiate environments to gender disparities in field choice. This analysis expands a literature that has emphasized the role of non-pecuniary factors, such as subjective beliefs and preferences, on major choice, but has yet to explore how major choices actually respond to changes in non-pecuniary features of the environment.

Drawing on a newly assembled historical dataset, we estimate event study

[19]This finding contrasts slightly with Kaplan (1978), who notes that the faculty of Vassar College became substantially more male after the transition to coeducation.

specifications that compare the evolution of women's major choices at newly coeducational colleges to those at comparable colleges that transitioned at different times (or did not transition at all). In the long run, we find that the share of women majoring in STEM fell by around 3.0 percentage points (30%) relative to control colleges. We also estimate negative effects of coeducation on women's economics and business major choices. Our analysis suggests that these reductions are largely driven by greater exposure to male peers, with some evidence that coeducation caused a small decrease in opportunities for women to interact with same-gender faculty role models. We find no evidence that coeducation changed the composition of matriculating female freshman, created stricter competition for grades and opportunities in STEM-related classrooms, changed the menu of courses offered, or created capacity constraints in high-cost fields.

What do our estimates imply about the contribution of gendered peer effects to the overall gender gap in STEM observed in recent years? We can gauge this with a simple exercise. According to our 2016 data, 28 percent of baccalaureate degrees awarded to men were in STEM fields, but only 11.5 percent of degrees awarded to women were in STEM fields. In addition, 57 percent of total degrees were awarded to women, so the average woman's potential peer group was 43 percent male. What would be the effect of reducing this number to zero? As discussed in Section 3, coeducation caused a 21.6 percentage-point increase in the male share of undergraduates and a 3.0 percentage-point decline in the share of women graduating with STEM degrees. Thus, a 43-percentage-point reduction in the male share of undergraduates would be associated with a $3.0 \cdot 43/21.6 = 5.97$-percentage-point increase in the share of women majoring in STEM. This amounts to 36 percent of the 16.5-percentage-point gap.

Of course, this counterfactual exercise must be qualified in several respects. Our sample of switchers and women's colleges primarily contain students who live on campus, which is less often true of students at 4-year colleges today. In addition, most coeducation events occurred in the 1960s-80s, when gender roles may have been more salient than they are today. These considerations suggest that our exercise overstates the aggregate role of gendered peer effects.

On the other hand, the men who entered newly coeducational colleges were less likely than the average man to enter traditionally male fields (Figure 2). This suggests that women at newly coeducational colleges faced less competition from men in STEM classrooms than would be predicted based on the male share of the student body—implying that our exercise understates the aggregate role of gendered peer effects.

In either case, our findings suggest that interaction with male collegiate peers contributes to meaningful gender disparities in college major choice. While a large-scale return to single-sex education is unlikely, policies that infuse features of single-sex learning environments into coeducational settings may be effective in closing the persistent gender gap in major choice.

References

Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and identity." *The Quarterly Journal of Economics*, 115(3): 715–753.

Altonji, J. G., P. Arcidiacono, and A. Maurel. 2016. "Chapter 7 - The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects." In *Handbook of the Economics of Education*. Vol. 5, , ed. Eric A. Hanushek, Stephen Machin and Ludger Woessmann, 305–396.

Altonji, Joseph G., Erica Blom, and Costas Meghir. 2012. "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers." *Annual Review of Economics*, 4(1): 185–223. _eprint: https://doi.org/10.1146/annurev-economics-080511-110908.

Astorne-Figari, Carmen, and Jamin D. Speer. 2019. "Are Changes of Major Major Changes? The Roles of Grades, Gender, and Preferences in College Major Switching." *Economics of Education Review*, 70: 75–93.

Bertrand, Marianne. 2011. "New Perspectives on Gender." *Handbook of Labor Economics*, 4b: 1543–1590.

Bertrand, Marianne. 2017. "The Glass Ceiling." Becker Friedman Institute for Research in Economics 2018-38.

Bertrand, Marianne, Emir Kamenica, and Jessica Pan. 2015. "Gender identity and relative income within households." *Quarterly Journal of Economics*, 121: 571–614.

Billger, Sherrilyn M. 2002. "Admitting men into a women's college: A natural experiment." *Applied Economics Letters*, 9(7): 479–483.

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

Booth, Alison, and Patrick Nolen. 2012*a*. "Choosing to compete: How different are girls and boys?" *Journal of Economic Behavior & Organization*, 81(2): 542–555.

Booth, Alison L., and Patrick Nolen. 2012*b*. "Gender differences in risk behaviour: does nurture matter?" *The Economic Journal*, 122(558): F56–F78.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2021. "Revisiting event study designs: Robust and efficient estimation." Unpublished working paper.

Bostwick, Valerie K, and Bruce A Weinberg. 2018. "Nevertheless She Persisted? Gender Peer Effects in Doctoral STEM Programs." NBER Working Paper 25028.

Bottia, Martha Cecilia, Elizabeth Stearns, Roslyn Arlin Mickelson, Stephanie Moller, and Lauren Valentino. 2015. "Growing the roots of STEM majors: Female math and science high school faculty and the participation of students in STEM." *Economics of Education Review*, 45: 14–27.

Brown, Charles, and Mary Corcoran. 1997. "Sex-Based Differences in School Content and the Male-Female Wage Gap." *Journal of Labor Economics*, 15(3): 431–465.

Bursztyn, Leonardo, Thomas Fujiwara, and Amanda Pallais. 2017. "Acting Wife: Marriage Market Incentives and Labor Market Investments." *American Economic Review*, 107(11): 3288–3319.

Calkins, Avery. 2020. "Gender, grades, and college major during the dot-com crash."

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with multiple time periods." *Journal of Econometrics*, 225(2): 200–230.

Card, David, and A. Abigail Payne. 2017. "High School Choices and the Gender Gap in STEM." NBER Working Paper 23769.

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

Ceci, Stephen J., Donna K. Ginther, Shulamit Kahn, and Wendy M. Williams. 2014. "Women in Academic Science: A Changing Landscape." *Psychological Science in the Public Interest: A Journal of the American Psychological Society*, 15(3): 75–141.

Ceci, Stephen J., Wendy M. Williams, and Susan M. Barnett. 2009. "Women's underrepresentation in science: sociocultural and biological considerations." *Psychological Bulletin*, 135(2): 218–261.

Chamberlain, Mariam K. 1988. *Women in Academe: Progress and Prospects.* Russell Sage Foundation.

Dasgupta, Nilanjana, and Shaki Asgari. 2004. "Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping." *Journal of Experimental Social Psychology*, 40(5): 642–658.

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review*, 110(9): 2964–2996.

Gneezy, U., M. Niederle, and A. Rustichini. 2003. "Performance in Competitive Environments: Gender Differences." *The Quarterly Journal of Economics*, 118(3): 1049–1074.

Gneezy, Uri, Kenneth L. Leonard, and John A. List. 2009. "Gender Differences in Competition: Evidence From a Matrilineal and a Patriarchal Society." *Econometrica*, 77(5): 1637–1664.

Goldin, Claudia. 2015. "Gender and the Undergraduate Economics Major: Notes on the Undergraduate Economics Major at a Highly Selective Liberal Arts College." Harvard University.

Goldin, Claudia, and Lawrence F. Katz. 2011. "Putting the "Co" in education: Timing, reasons, and consequences of college coeducation from 1835 to the present." *Journal of Human Capital*, 5(4): 377–417.

Gong, Yifan, Todd Stinebricker, and Ralph Stinebrickner. 2019. "Uncertainty about future income: Initial beliefs and resolution during college." *Quantitative Economics*, 10: 604–641.

Goodman-Bacon, Andrew. 2018. "Public Insurance and Mortality: Evidence from Medicaid Implementation." *Journal of Political Economy*, 126(1): 216–262.

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing." *Journal of Econometrics*, 225(2): 254–277.

Harwarth, Irene, Mindi Maline, and Elizabeth DeBra. 1997. *Women's Colleges in the United States: History, Issues, and Challenges.*

Hemelt, Steven W., Kevin M. Stange, Fernando Furquim, Andrew Simon, and John E. Sawyer. 2018. "Why is math cheaper than English? Understanding cost differences in higher education." NBER Working Paper 25314.

Hill, Catherine, Christianne Corbett, Andresse St. Rose, and American Association of University Women. 2010. *Why so few? women in science, technology, engineering, and mathematics.* Washington, D.C:AAUW.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow. 2019. "The Allocation of Talent and U.S. Economic Growth." *Econometrica*, 87(5): 1439–1474.

Huntington-Klein, Nick, and Elaina Rose. 2018. "Gender Peer Effects in a Predominantly Male Environment: Evidence from West Point." *AEA Papers and Proceedings*, 108: 392–395.

Hunt, Jennifer. 2016. "Why do Women Leave Science and Engineering?" *ILR Review*, 69(1): 199–226.

Johnston, Jack, and John DiNardo. 1997. *Econometric Methods.* . 4 ed., The McGraw-Hill Companies, Inc.

Kahn, Shulamit, and Donna Ginther. 2017. "Women and STEM." NBER Working Paper 23525.

Kamas, Linda, and Anne Preston. 2012. "The importance of being confident; gender, career choice, and willingness to compete." *Journal of Economic Behavior & Organization*, 83(1): 82–97.

Kaplan, Susan Romer. 1978. "Women's Education: The Case for the Single-Sex College." In *Higher Education of Women*. 53–67. New York, NY:Praeger Publishers.

Kinzie, Jillian (Jillian L.), 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): 145–165.

Kofoed, Michael S., and Elizabeth McGovney. 2019. "The Effect of Same-Gender or Same-Race Role Models on Occupation Choice Evidence from Randomly Assigned Mentors at West Point." *Journal of Human Resources*, 54(2): 430–467.

Marcus, Michelle, and Pedro H. C. Sant'Anna. 2021. "The Role of Parallel Trends in Event Study Settings: An Application to Environmental Economics." *Journal of the Association of Environmental and Resource Economists*, 8(2): 235–275. Publisher: The University of Chicago Press.

Miller-Bernal, Leslie. 1993. "Single-sex versus coeducational environments: A comparison of women students' experiences at four colleges." *American Journal of Education*, 102(1): 23–54.

Miller-Bernal, Leslie, and Susan L. Poulson. 2004. *Going coed: Women's experiences in formerly men's colleges and universities, 1950-2000.* Vanderbilt University Press.

Niederle, Muriel, and Lise Vesterlund. 2007. "Do Women Shy Away From Competition? Do Men Compete Too Much?" *Quarterly Journal of Economics*, 122(3): 1067–1101.

Niederle, Muriel, and Lise Vesterlund. 2008. "Gender Differences in Competition." *Negotiation Journal*, 24(4): 447–463.

Owen, Stephanie. 2020. "College Field Specialization and Beliefs About Relative Performance: An Experimental Intervention to Understand Gender Gaps in STEM." University of Michigan Working Paper.

Patnaik, Arpita, Matthew Wiswall, and Basit Zafar. 2020. "College majors." Working paper.

Pei, Zhuan, Jorn-Steffen Pischke, and Hannes Schwandt. 2018. "Poorly measured confounders are more useful on the left than on the right." *Journal of Business and Economic Statistics*.

Rim, Nayoung. 2020. "The effect of title IX on gender disparity in graduate education."

Riordan, Cornelius. 1994. "The Value of Attending a Women's College." *The Journal of Higher Education*, 65(4): 486–510.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2020. "IPUMS USA: Version 10.0 [dataset]." IPUMS, Minneapolis, MN.

Schmader, Toni. 2002. "Gender Identification Moderates Stereotype Threat Effects on Women's Math Performance." *Journal of Experimental Social Psychology*, 38(2): 194–201.

Shan, Xiaoyue. 2021. "Does minority status drive women out of male-dominated fields?" Unpublished manuscript.

Shapiro, Casey A., and Linda J. Sax. 2011. "Major selection and persistence for women in STEM." *New Directions for Institutional Research*, 2011(152): 5–18.

Sloane, Carolyn, Erik Hurst, and Dan Black. 2019. "A cross-cohort analysis of human capital specialization and the college gender wage gap." NBER Working Paper 26348.

Smith, Daryl G. 1990. "Women's Colleges and Coed Colleges." *The Journal of Higher Education*, 61(2): 181–197.

Speer, Jamin D. 2021. "Bye Bye Ms. American Sci: Women and the Leaky STEM Pipeline." IZA Institute of Labor Economics IZA DP No. 14676.

Stange, Kevin M. 2012. "An empirical investigation of the option value of college enrollment." *American Economic Journal: Applied Economics*, 4(1): 49–84.

Steele, Claude M. 1997. "A threat in the air: How stereotypes shape intellectual identity and performance." *American Psychologist*, 52(6): 613–629.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225(2): 175–199.

Thomas, Auden D. 2008. "Preserving and strengthening together: Collective strategies of U.S. women's college presidents." *History of Education Quarterly*, 48(4): 565–589.

Tidball, M. Elizabeth. 1980. "Women's Colleges and Women Achievers Revisited." *Signs: Journal of Women in Culture and Society*, 5(3): 504–517.

Tidball, M. Elizabeth. 1989. "Women's colleges: Exceptional conditions, not exceptional talent, produce high acheivers." In *Educating the majority: Women challenge tradition in higher education.* , ed. Carol S. Pearson, Donna L Shavlik and Judith G. Touchton, 157–172. Ace MacMillan.

Turner, Sarah E., and William G. Bowen. 1999. "Choice of Major: The Changing (Unchanging) Gender Gap." *Industrial and Labor Relations Review*, 52(2): 289–313.

Willis, Robert, and Sherwin Rosen. 1979. "Education and Self-Selection." *Journal of Political Economy*, 87(5): S7–36. Publisher: University of Chicago Press.

Wiswall, Matthew, and Basit Zafar. 2015. "Determinants of college major choice: Identification using an information experiment." *The Review of Economic Studies*, 82(2): 791–824.

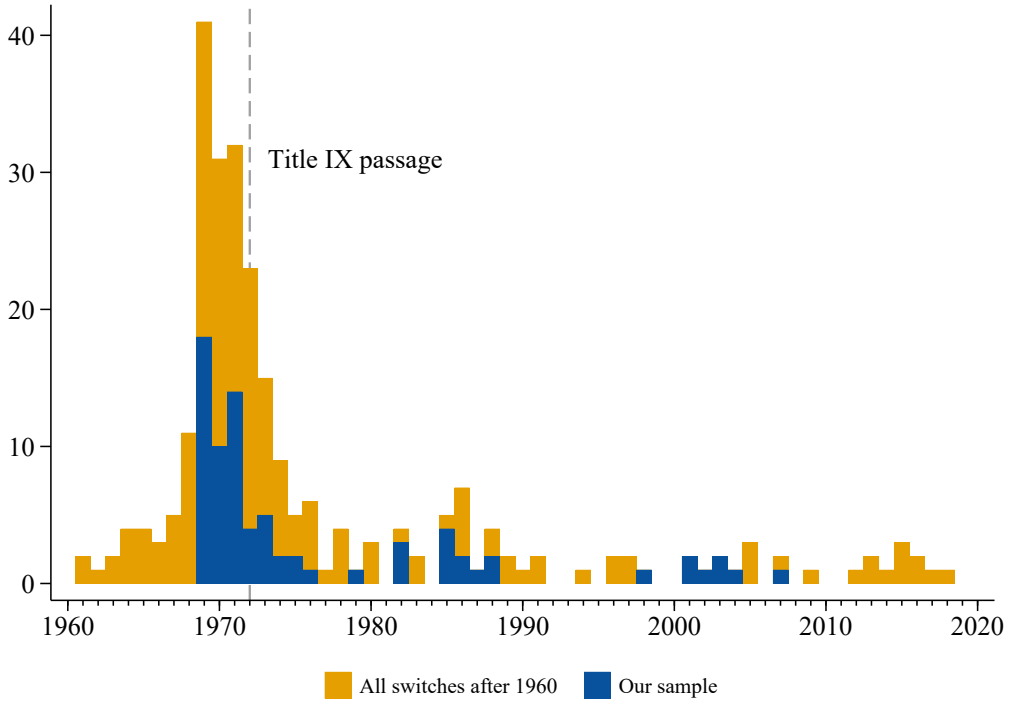
Wiswall, Matthew, and Basit Zafar. 2018. "Preference for the Workplace, Investment in Human Capital, and Gender." *The Quarterly Journal of Economics*, 133(1): 457–507.

Zafar, Basit. 2011. "How do college students form expectations?" *Journal of Labor Economics*, 29(2): 301–348.

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

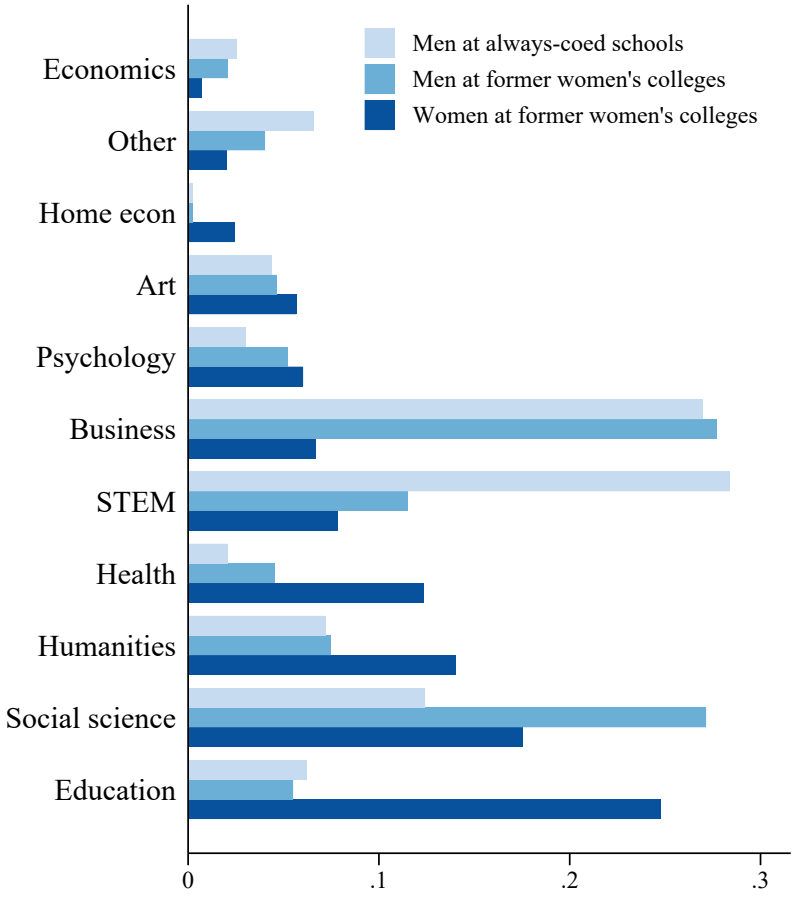
Zölitz, Ulf, and Jan Feld. 2018. "The effect of peer gender on major choice." University of Zurich Department of Economics Working Paper 270.

Figure 1: Distribution of the year of switch to coeducation



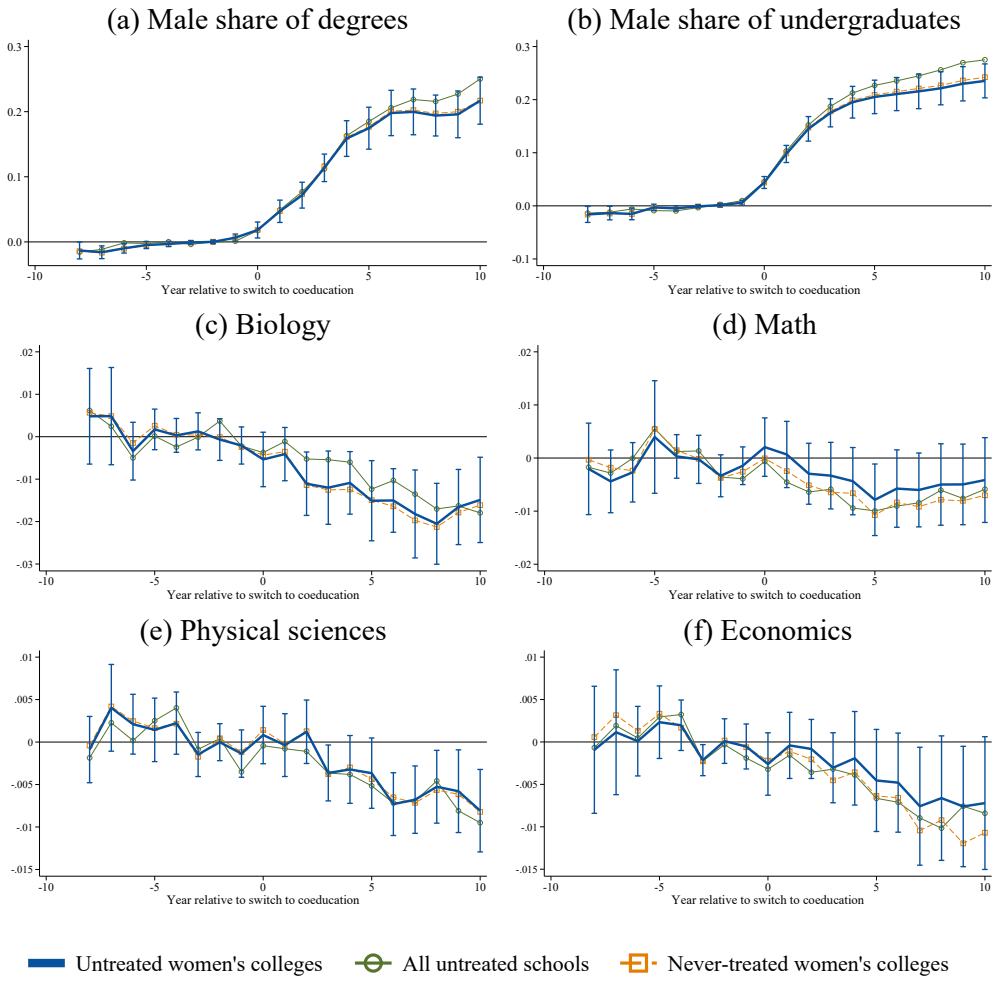
Source: Hand-collected data on the years that former women-only institutions switched to coeducation. Our analysis sample omits schools that closed shortly after switching to coeducation, operated coordinate programs with nearby men's colleges, did not offer STEM fields, or that we do not observe for at least 4 years prior to and 10 years after the switch to coeducation. See Section 2 for further background on how this list was compiled, and Appendix B for a comprehensive list of formerly women's colleges and sample inclusion criteria.

Figure 2: Field of major among students at coeducational and former women's colleges



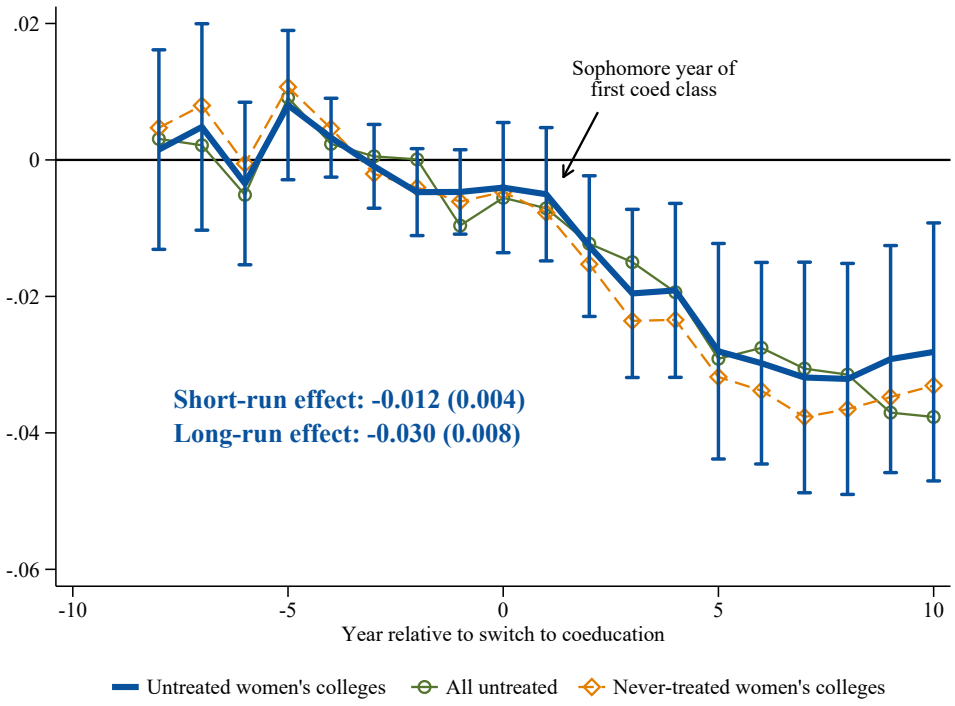
Notes: Data drawn from the IPEDS surveys, spanning 1966-2016, linked to hand-collected data on the dates of transition to coeducation by institution. Each bar shows the fraction of graduates of a given sex and college type earning degrees in the corresponding field. The distribution of majors among men at former women's colleges is calculated among men graduating in the first 10 years after the transition to coeducation. The distribution of majors among men at "always-coed" schools is calculated among men at schools that were coeducational throughout our sample period, weighted so that the distribution of years represented matches the distribution of years represented among men at former women's colleges. The distribution of majors among women is calculated among women graduating in the last 5 years before the transition to coeducation.

Figure 3: Effect of coeducation on gender composition and women's choice of quantitative majors



Notes: Data drawn from IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. See Section 2 for further detail. Panels display estimate of β_k from equation 2. Dependent variable is the share of degrees earned in STEM among all degrees earned by women in the academic year. Error bars show 95% confidence intervals constructed using a block bootstrap clustered at the institution level.

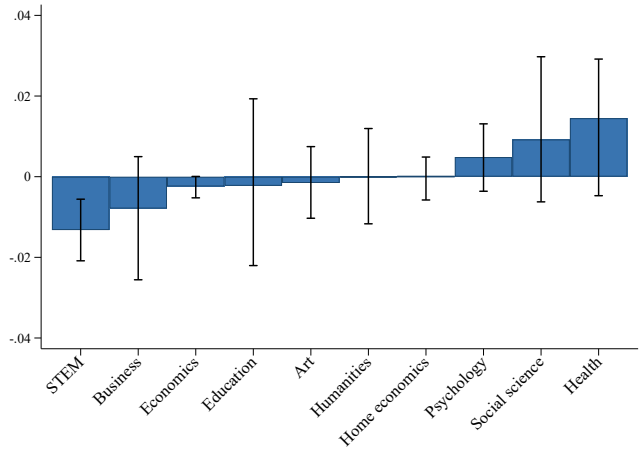
Figure 4: Effect of coeducation on the share of women majoring in STEM



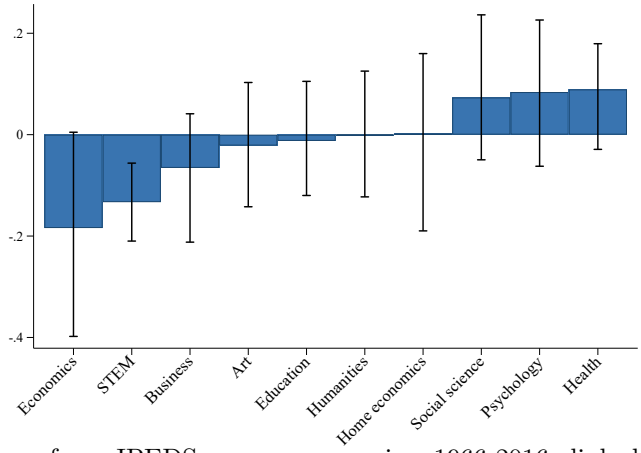
Notes: Figure displays estimated effect of coeducation on the share of female graduates earning a degree in a STEM field. Each point shows and estimate of β_k from equation 2, using the comparison group specified in the legend. Error bars show 95% confidence interval constructed using block bootstrap clustered at the institution level. Data drawn from IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. See Section 2 for further detail.

Figure 5: The long-run effect of coeducation on the full distribution of women's major choices

(a) Percentage-point effects by major

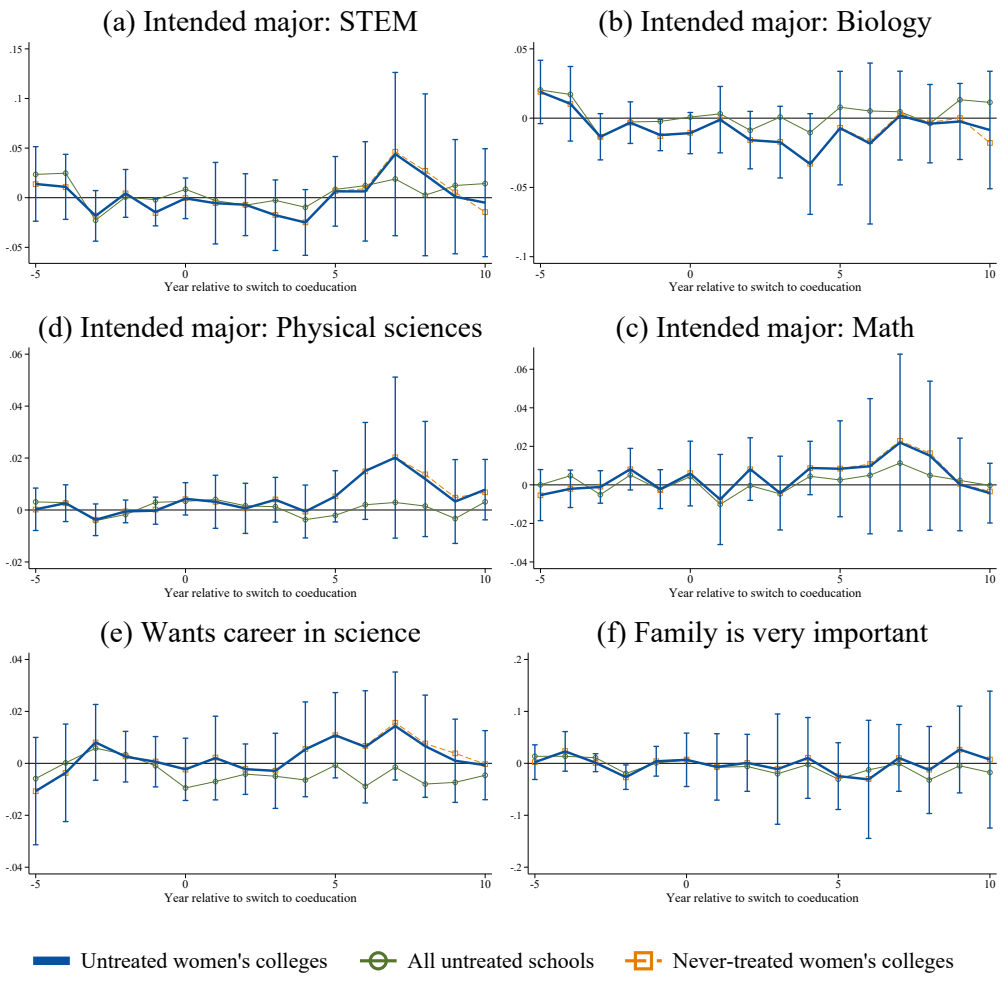


(b) Semi-elasticity by major



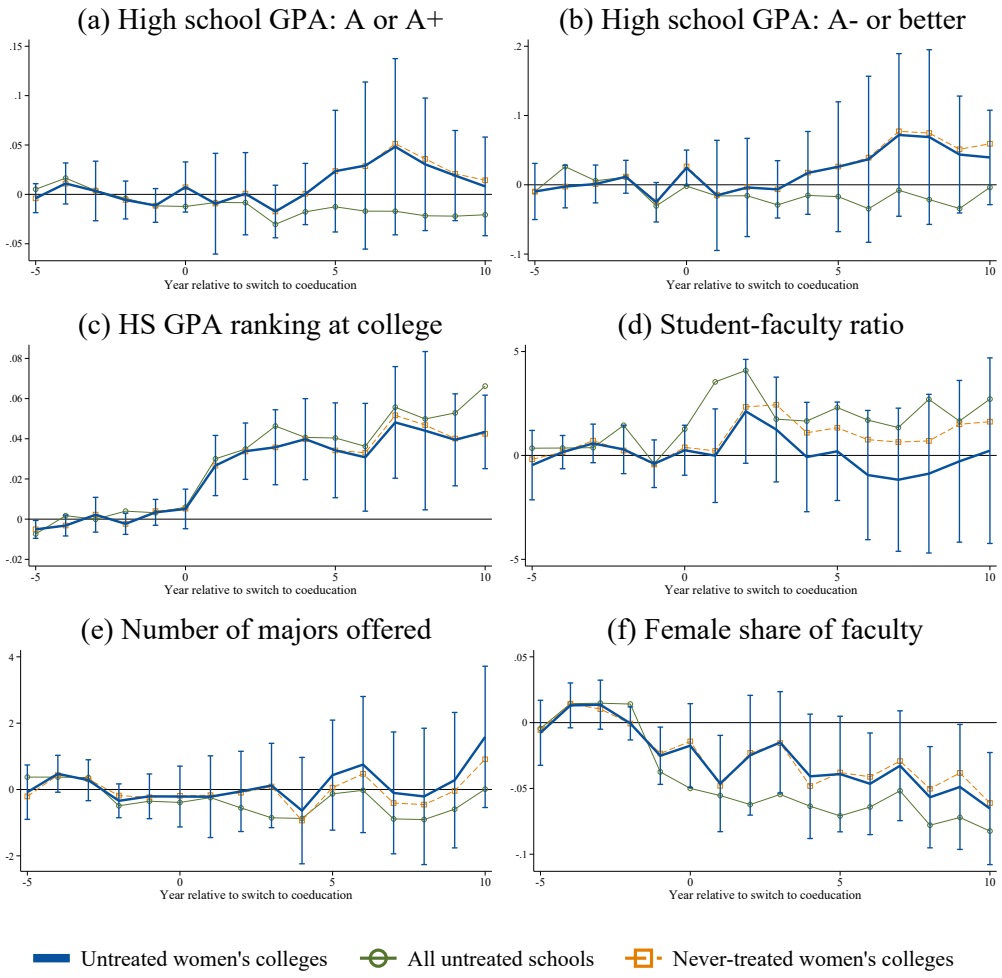
Notes: Data drawn from IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. The majors included in each field are described in Appendix E.1.2. Each panel displays the long-run effect of a 10-percentage-point increase in the male share of graduating students (scaled by the counterfactual mean in panel B), estimated using equation 2 and aggregating cohort-level effects using equation 4. For each treated school, pool of comparison institutions is all women's colleges that switch to coeducation 10 or more years later, or never switch during our sample period. Estimates are conditional on selectivity and Catholic affiliation. Long-run effects are calculated as impacts in the second five years after the transition to coeducation. Confidence intervals correspond to the 2.5th and 97.5th percentiles of estimates from a block bootstrap routine with 1,000 replications. For scale, the "other" major category is omitted; estimates for that group are -0.001 [-0.016,0.017] and -0.029 [-0.41,0.43], respectively.

Figure 6: Effect of coeducation on intended major of entering freshman women



Notes: Each point shows and estimate of β_k from equation 2, using the comparison group specified in the legend. Error bars show 95% confidence interval constructed using block bootstrap clustered at the institution level. Data drawn from HERI, spanning 1966-2006, linked to hand-collected dates of transitions to coeducation by institution. See Section 2 for further detail.

Figure 7: Effect of coeducation on other characteristics of the college environment



Notes: Each point shows and estimate of β_k from equation 2, using the comparison group specified in the legend. Error bars show 95% confidence interval constructed using block bootstrap clustered at the institution level. Data in panels a-c drawn from HERI Freshman Survey, spanning 1966-2006, linked to hand-collected dates of transitions to coeducation by institution. Data in panels d-f drawn from IPEDS data on degree completions from 1966-2016, enrollment from 1969-2016, and faculty characteristics from 1971-2015. See Section 2 for further detail.

Table 1: Summary statistics

	Switchers	Women's colleges	All non-switchers	HERI linked sample
Share of women's degrees in STEM	0.09	0.11	0.09	0.09
	(0.04)	(0.04)	(0.08)	(0.04)
Total enrollment	1600.67	1228.14	5937.85	1255.73
	(1243.80)	(559.81)	(7380.44)	(560.13)
Female share of degrees awarded	0.97	0.99	0.50	0.97
	(0.06)	(0.02)	(0.13)	(0.05)
Graduate degrees awarded	58.09	20.53	246.66	41.14
	(84.96)	(35.84)	(467.53)	(58.85)
Private college	0.80	1.00	0.56	0.93
	(0.41)	(0.00)	(0.50)	(0.26)
Ever Catholic-affiliated	0.50	0.31	0.06	0.49
	(0.50)	(0.47)	(0.24)	(0.51)
Selective admission	0.25	0.45	0.14	0.34
	(0.44)	(0.51)	(0.35)	(0.48)
Observations	77	29	927	30

Notes: Columns 1 and 4 show average characteristics among women's colleges in the year prior to their transition to coeducation. Columns 2 and 3 are constructed from the pool of potential comparison schools, reweighted to represent average characteristics in the years prior to the transitioning schools' conversion to coeducation. Data in columns 1-3 are drawn from IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. Data in column 4 is additionally linked to data from the HERI Freshman Survey and limited to transitioning schools that we observe in the HERI data at least once in the 5 years before and once in the 10 years after coeducation. See Section 2 for further detail. Catholic affiliation is coded as 1 if the school was ever affiliated with the Catholic Church. Schools are coded as having selective admission if they received a Barron's rating of 1, 2, or 3 in 1972. The majors included in the STEM field are described in Section 2 and Appendix B. Sample statistics are weighted by the number of degrees awarded to women.

Table 2: Long-run effects of coeducation on the shares of women choosing STEM and economics majors

	(1) STEM	(2) Biology	(3) Phys science	(4) Math	(5) Economics
Panel A: Not-yet-treated comparison group					
Long-run effect	-0.0302***	-0.0171***	-0.0058***	-0.0059*	-0.0062*
	(0.0077)	(0.0040)	(0.0017)	(0.0033)	(0.0032)
Counterfactual mean	0.099	0.060	0.016	0.019	0.013
Observations	5,455	5,455	5,455	5,455	5,455
Panel B: All-college comparison group					
Long-run effect	-0.0312***	-0.0139***	-0.0064***	-0.0082***	-0.0081**
	(0.0067)	(0.0045)	(0.0022)	(0.0031)	(0.0033)
Counterfactual mean	0.102	0.057	0.017	0.021	0.015
Observations	23,967	23,967	23,967	23,967	23,967
Panel C: Never-treated comparison group					
Long-run effect	-0.0349***	-0.0180***	-0.0060***	-0.0088***	-0.0089**
	(0.0092)	(0.0046)	(0.0019)	(0.0041)	(0.0037)
Counterfactual mean	0.103	0.060	0.016	0.022	0.016
Observations	5,164	5,164	5,164	5,164	5,164

Notes: Table displays the estimated effect of the switch to coeducation on graduating female students' choice of major, estimated using $\hat{\beta}_{LR}$ from equation 4. Each panel uses the specified pool of institutions to construct a comparison group and estimate a counterfactual trend in major choices, conditional on college selectivity and historical affiliation with the Catholic Church. In panel B, we additionally condition on school size, as measured by number of degrees granted, and the trend in STEM choice among all students. Data drawn from IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. Standard errors are estimated using a block bootstrap with 1,000 replications that accounts for intracluster correlation at the institution level. Counterfactual mean is the share of women that would have chosen each major at treated schools if choices at those schools had followed trends at the comparison group of institutions.

Table 3: Long-run effects of coeducation on freshman women's *intended* choices of major

	(1) STEM	(2) Biology	(3) Math	(4) Phys science
Panel A: Not-yet-treated comparison group				
Long-run effect	0.0157	-0.0064	0.0111	0.0111
	(0.0260)	(0.0151)	(0.0149)	(0.0092)
Counterfactual mean	0.090	0.080	0.005	0.006
Observations	1,426	1,426	1,426	1,426
Panel B: All-college comparison group				
Long-run effect	0.0109	0.0056	0.0052	0.0002
	(0.0132)	(0.0130)	(0.0068)	(0.0024)
Counterfactual mean	0.082	0.065	0.005	0.012
Observations	4,680	4,680	4,680	4,680
Panel C: Never-treated comparison group				
Long-run effect	0.0182	-0.0052	0.0118	0.0116
	(0.0265)	(0.0157)	(0.0155)	(0.0097)
Counterfactual mean	0.092	0.081	0.005	0.006
Observations	1,400	1,400	1,400	1,400

Notes: Table displays estimates long-run effect effect of women's colleges' switch to coeducation on the degree intentions of entering freshman female students. Effects are derived from the doubly-robust estimator of Callaway and Sant'Anna (2021) using equation 10 and aggregating cohort-level effects using equation 4. Each panel uses the specified pool of institutions to construct a comparison group and estimate a counterfactual trend in major choices, conditional on college selectivity and historical affiliation with the Catholic Church. For each treated school, comparison group is also limited to schools that participated in the HERI Freshman Survey in the same years from 5 years before to 10 years after the transition to coeducation. In panel B, we additionally condition on pre-reform school size, as measured by number of entering freshman students in our sample. Counterfactual mean is the estimated share of women intending to choose each major at treated schools if those schools had followed trends at the comparison group of institutions.

For online publication

A A formal model of the effect of coeducation on women's STEM majoring

We use a very simple Roy model of college and major choice to illustrate the possible effects of transition to co-education on subsequent women's outcomes. We assume there are 3 collegiate institutions in the market: h, j and k. There are two time periods: 0 and 1, which are separated by a substantial number of years. At $t = 0$, institutions h and j are women-only while k is co-educational. Between $t = 0$ and $t = 1$, institution j transitions to co-education. All institutions in each time period offer two majors: STEM (S) and non-STEM (NS). We assume away capacity constraints. (In Section G we show that evidence consistent with this assumption.)

Each time period consists of two stages. In the first stage, women make enrollment decisions η under uncertainty about the values of attending each college. In the second stage, women who have chosen to enroll in a college choose a major μ in which to graduate, with full information about major-specific payoffs. We assume that every woman enrolls in college, and that every woman who starts college completes a degree at her starting institution.

Consider a hypothetical high school senior w making decisions in period t. A given enrollment choice η_{wt} returns the expected payoff $V_{wt}(\eta_{wt})$. She chooses the enrollment choice η^*_{wt} that maximizes this function:

$$V_{wt}(\eta^*_{wt}) = \max\left\{V_{wt}(h), V_{wt}(j), V_{wt}(k)\right\}. \tag{5}$$

After making her enrollment choice, woman w realizes her major-specific payoffs and chooses her major μ_{wt}. We represent her payoff from choosing major μ at institution η as $v_{wt}(\mu_{wt}; \eta)$. Her major choice μ^*_{wt} thus satisfies:

$$v_{wt}(\mu^*_{wt}; \eta) = \max\left\{v_{wt}(S; \eta), v_{wt}(NS; \eta)\right\}, \ \eta \in \{h, j, k\}. \tag{6}$$

Woman w's expected payoff from enrolling at institution η is simply equal to the expected payoff from choosing her most-preferred major at η:

$$V_{wt}(\eta) = E\left[v_{wt}(\mu^*_{wt}; \eta)\right] \tag{7}$$

Assume there are many women w in the market with varying preferences for colleges and majors. Consider the students who chose to enroll at women's institution j in period t. Denote each enrolled woman as belonging to the set A_{jt}. The share of this student body graduating from h with a STEM degree is given by $s_{STEM,jt}$:

$$s_{STEM,jt} = \frac{\sum_{w \in A_{jt}} \mathbf{1}\left\{S = \text{argmax}\left\{v_{wt}(S; j), v_{wt}(NS; j)\right\}\right\}}{\sum_{w} \mathbf{1}\left\{j = \text{argmax}\left\{V_{wt}(h), V_{wt}(j), V_{wt}(k)\right\}\right\}} \tag{8}$$

Suppose that, aside from institution j transitioning to co-education, nothing else changes between periods 0 and 1. Then, the object

$$\Delta = s_{STEM,j1} - s_{STEM,j0}$$

describes the treatment effect of co-education on the production of women STEM majors at institution j.

Two channels determine Δ. First, suppose that the set of women enrolling at institution j, A_j, does not change between time periods 0 and 1. Then, Δ simply depends on how the transition to co-education alters the payoffs to majoring in STEM ($v_w(S; j)$), relative to majoring in non-STEM ($v_w(N; j)$), for this population of women. We call this the "environmental effect." See Section 1.2 for a discussion of the various channels determining this effect.

Second, the transition to co-education might induce a change in the enrolled set of students A_j. To see why this might be the case, plug (7) into (5) and re-express the optimal enrollment decision:

$$\eta_{wt} = \text{argmax}\left\{E\left[v_{wt}(\mu^*_{wt}; h)\right], E\left[v_{wt}(\mu^*_{wt}; j)\right], E\left[v_{wt}(\mu^*_{wt}; k)\right]\right\} \tag{9}$$

That is, women forecast their (major-specific) payoffs from attending each

institution, and use those expectations to guide their enrollment decisions. When institution h transitions to co-education, the women that strongly desire a single-sex environment may experience a reduction in $E\left[v_{wt}(\mu^*_{wt}; j)\right]$ and may substitute from j to women's college h. Additionally, the women that strongly desire a co-educational environment may experience an improvement in $E\left[v_{wt}(\mu^*_{wt}; h)\right]$, and may substitute from co-educational college k to j. If the women who most desire a single-gender environment also have the highest expected payoffs from majoring in STEM (say, because they are the most prepared for STEM coursework), then j's transition to co-education causes its subsequent population of women to become more negatively selected on expected STEM payoffs: plausibly leading to a reduction in STEM majoring. We call this channel the "composition effect."

In Section 3, we estimate the overall treatment effect Δ. Because the assumption that nothing else about the collegiate environment changes between periods 0 and 1 is likely false, we apply difference-in-difference methodologies to estimate Δ. That is, we compare the evolution of women's major choices at colleges that transitioned to coeducation to the evolution of major choices at comparable colleges that did not transition. Section 4 attempts to decompose Δ into composition versus environmental effects.

B Data Collection and Sample Construction

B.1 Data collection on years of the switch to coeducation

We construct our full dataset via the following procedure. First, we identify all institutions that awarded over 90% of their degrees to women in the first year they they are observed in the HEGIS/IPEDS data. Because of concerns about response rates from institutions that entered the data after the transition to IPEDS, we restrict our sample to HEGIS-participating institutions

that began offering four-year degrees before 1987.[20] For each of these institutions, we gathered information on whether it was once a women's college. If it was, we recorded the date of transition to coeducation based on the institution's website, the "Everywoman's Guide to Colleges and Universities" (Howe, Howard and Strauss, 1982), or other online sources. Where possible, we relied on institutions' own websites for this information. Over 90% of our transition dates were found on *.edu* websites. This procedure identified 219 institutions that were women-only in the first year they were observed, 155 of which eventually transitioned to coeducation.

We define the first year of coeducation as the first year that men were admitted to traditional four-year undergraduate programs with coeducational courses. Schools where men were admitted to these programs only as commuter students are counted as coeducational, but schools where men could only participate in evening or adult education classes or graduate programs are not. We exclude all coordinate institutions, that is, institutions such as Columbia and Barnard where a men's college and a women's college share a campus and allow cross-registration in classes. We also exclude cases where a women's college merged with a men's college.

We sourced the years that single-sex institutions switched to coeducation in three different ways. The first source of information was a comprehensive check of the top 120 liberal arts colleges and the top 80 universities in the 2018 *U.S. News and World Report* for the gender of the student body in 1966 and a date of switch to coeducation. The second source of information was a list of current and former women's colleges from the Women's College Coalition, including a date of switch to coeducation. Finally, we generated a list of institutions that awarded more than 90% of their degrees to women in the first year they appeared in the data and used a research assistant to track down which of those institutions are current or former women's colleges. The RA also found the date of the switch to coeducation for former women's colleges. The three

[20]The ICPSR documentation for the IPEDS financial data for fall 1987 (study 2220) discusses concerns about response rates and imputation for institutions which enter the data after the transition to IPEDS.

lists were then compared. Institutions that appeared on multiple lists with matching switch dates were considered confirmed. Institutions with conflicts between the switch dates or that appeared on only one list were independently verified.

Our classification of the "gender" of an institution is based on the gender of the student body in 1966. Institutions that did not appear on any of the lists noted above were assigned a gender based on the HEGIS or IPEDS "sex code" variable in the first year they appeared in the data.

In total, we found 136 former women's colleges that switched to coeducation ("switchers"). Of these, 118 were never coordinate, entered the data before 1987, did not merge with a men's college, and were never for-profit.

B.2 Problems in the data

Sixty-four of our "switchers" had at least one man graduate before the official date of the full switch to coeducation. In the vast majority of cases, this seems to have been either occasional one-off male students or the introduction of a small coeducational adult education program. In other cases, schools either opened a men-only college on campus or had male commuter students before completely switching to coeducation. We denote problematic cases with the following flags:

1. A small number of men ($\leq$ 10 per year) graduate from the institution before the switch to coeducation, or $>$ 10 men graduate from the institution before the switch and we can verify the existence of a coeducational adult education program. In these cases, we believe that there are unlikely to be errors with our coding of the dates of transition to coeducation and that these male students probably did not interact with traditional undergraduates. ($n = 46$)

2. A large number of men ($>$ 10 per year) graduate from the institution before the switch to coeducation, and we cannot verify the existence of a coeducational adult education program that did not interact with traditional undergraduates. In these cases, we became concerned that

there was an error in the date of transition to coeducation that we had found, which might mean that the institution was not truly single-sex for traditional undergraduates during the time we had coded it as such. ($n = 10$)

3. Men were allowed as commuter students long before the official date of coeducation. In these cases we were usually able to identify the true date that men were allowed to register in traditional undergraduate coursework and continued to include them in the sample. ($n = 8$)

4. Rather than becoming fully coeducational, the school opened a men's college (or men-only program) on campus – basically becoming coordinate rather than coeducational. We reclassified these institutions as coordinate. ($n = 2$)

We exclude group 2 from the main specification. We also exclude institutions that awarded zero bachelor's degrees in STEM fields to women in 1966 and institutions that closed shortly after a switch to coeducation. We believe these restrictions help us comply with our identifying assumption, that there is no residual determinant of women's STEM majoring that occurred at the same time as the transition to coeducation, especially in the case of removing institutions that did not regularly offer STEM programs before transitioning to coeducation, as all but two of the institutions dropped for this reason began offering STEM programs only when men began attending the institution.

B.3 Defining our sample

Our sample includes institutions that:

1. Were woman-only or coeducational in 1966 ($N = 3,707$; $N_{switch} = 155$)[21]

2. Entered the data in 1987 or earlier. This cut was made to deal concerns noted by ICPSR about the accuracy of imputation for nonrespondents

[21]Note that the sample sizes here are meant to note how each subsequent restriction decreases the size of the sample, and different restrictions can overlap.

after the switch to the IPEDS data in 1987. The IPEDS data dramatically expanded the sample to include schools that had not been classified as "institutions of higher learning" under Title IX and the response rate of those new institutions was much lower than the response rate of institutions included in HEGIS. See the ICPSR documentation of the 1986-1987 academic year finance data for further details. ($N = 1,850$; $N_{switch} = 150$)

3. If switched to coeducation, did not close 9 years after the switch to coeducation (but not omitting schools that switched to coeducation in 2007 or later); if untreated, was in data for at least 15 years. This cut was made for balance. ($N = 1,508$; $N_{switch} = 126$)

4. Had at least one woman complete a STEM (life or physical science, math, engineering, computer science) degree in 1966. This cut was made because of concerns that the establishment of STEM degree programs may have been a move that went hand-in-hand with the transition to coeducation. ($N = 971$; $N_{switch} = 102$)

5. Were never classified as a coordinate institution (a women's college sharing a campus with a men's college) and were not part of a merger with a men's college after 1966. We made this cut because we suspect that classes on campus were coeducational long before mergers occurred, as is common with coordinate institutions. ($N = 971$; $N_{switch} = 94$)

6. Were never a for-profit institution. This cut was made because our data provides information only on degrees awarded, and for-profit institutions tend to have very low graduation rates. However, many for-profit institutions were already removed by the requirement that institutions enter the data prior to 1987. ($N = 963$; $N_{switch} = 91$)

7. Had fewer than 10 male students per year graduate before coeducation started or had coeducational adult education program that we could

verify dates of existence for – that is, was not part of group 2 in the previous section. We made this cut because were concerned about the accuracy of the date of transition we had found. ($N = 957$; $N_{switch} = 87$)

Returning post-1987 entrants (item 2), institutions that closed shortly after the transition to coeducation (item 3), and institutions that may not have offered STEM degrees (item 4) do not substantially change our main results. The results of returning early closures and schools that did not offer STEM, compared to our main analysis, are reported in Appendix Figure A1. Returning these institutions to the analysis leaves us with 124 switchers, 44 always women's colleges, and 3,478 always coeducational institutions. We believe that all of the restrictions we have made are key to either maintaining our identifying assumption or to avoid spurious effects created by poor imputation procedures in the post-1987 entrants. We believe that the exclusion of coordinate colleges and mergers are particularly important, as it is not clear that there was truly a transition from women-only courses to coeducational courses at either time. Especially at institutions where there was a merger between a women's college and either a men's college or an institution which was already coeducational, we think it is likely that a number of other changes came about at the same time, and coordinate colleges likely had coeducational courses before the transition to coeducation.

We also make a number of cuts that only apply to our comparison groups. To promote sample balance, we omit institutions that were in the data for less than 15 years from these comparison groups. We also omit military academies, institutions that were men's colleges in 1966, coordinate (or partner) colleges, institutions where no women were awarded STEM degrees in the first year we observe them,[22] and for-profit institutions.

[22]Many men's colleges transitioned to coeducation program-by-program, often establishing a nursing or teaching program before allowing women into academic programs. STEM programs were often the last to open to women. We therefore believe the inclusion of these institutions might contaminate our control group.

C Implementation of the doubly-robust Callaway and Sant'Anna (2021) estimator

Our research design exploits variation in the timing of women's colleges' switch to coeducation, as well as variation in the decision to switch at any time, to study the effect of the gender mix of the collegiate environment on women's choice of major. Because we expect the effect of this reform to vary over time, we present event-study estimates that show the evolution of changes in choice of major at switching colleges relative to the comparison group.

The conventional event study model is based on a two-way fixed effects (TWFE) design, which implements a pooled panel regression with controls for unit and time fixed effects. But recent studies have revealed that the TWFE specification often provides misleading estimates of treatment effects when there is variation in treatment timing across units, as there is in our setting (Callaway and Sant'Anna, 2021; Sun and Abraham, 2021; de Chaisemartin and D'Haultfœuille, 2020; Borusyak, Jaravel and Spiess, 2021; Goodman-Bacon, 2021). These issues are particularly pronounced in the presence of heterogeneous effects across units.

We instead adopt a slightly modified version of the estimator proposed by Callaway and Sant'Anna (2021).[23] The estimator avoids the shortcomings of TWFE models by estimating event-study-style treatment effect parameters separately for each treatment "cohort" g (e.g., schools that switched in 1969-1970 school year are part of cohort $g = 1969$) and acadamic year t. These cohort-by-year estimates then serve as building blocks that can be used to construct parameters of interest.

Any difference-in-difference research design must adopt a strategy to impute counterfactual trends in the outcome variable for units that adopted the treatment of interest. The doubly-robust estimator developed by Callaway and Sant'Anna (2021) relies on two strategies to construct these counterfactual trends. First, to limit the comparison group to schools that "look like" the

[23]We think Brantly Callaway and Pedro Sant'Anna for their generous and illuminating correspondence about the finer details of their estimator.

treated group, it weights observations by the inverse of the cohort-g-specific propensity score $p_g(X)$. In addition, it imputes a counterfactual trend in outcome variable Y_t between time t and some base period b by regressing changes $\Delta Y_{t,b} = Y_t - Y_b$ on the same vector of covariates X in a sample made up solely of the comparison group, and then predicting changes for the treated cohort g.

Formally, the estimator is constructed as follows:

$$\hat{\alpha}_{g,t} = \frac{1}{n_{g,t}} \sum \left(\frac{G_g}{\frac{1}{n_g}\sum G_g} - \frac{\frac{\hat{p}_g(X)C_g}{1-\hat{p}_g(X)}}{\frac{1}{n_g}\sum \frac{\hat{p}_g(X)C_g}{1-\hat{p}_g(X)}} \right) \left(\Delta Y_{t,b} - \Delta \hat{Y}_{t,b}(X) \right) \tag{10}$$

where G_g is a binary indicator for belonging to cohort g, C_g is an indicator for belonging to the pool of comparison schools for group g, and $\Delta Y_{t,b} = Y_t - Y_b$ is the change in outcome Y between year t and some pre-reform base period b. In our preferred specification, our comparison group consists of all women's colleges that switch to coeducation at least 10 years after cohort g – or never switch at all. Our vector X includes indicators for affiliation with the Catholic church and having a selectivity ranking of 1-3 in the 1972 Barron's ratings. Our estimates rely on the assumption that, conditional on these observed characteristics, trends in women's choice of major at our comparison group of schools accurately reflects the counterfactual trends that would have occurred at our treated schools in the absence of a switch to coeducation. As robustness checks, we also estimate $\alpha_{g,t}$ using only never-treated women's colleges as the comparison group, and then by using *all* four-year colleges that did not switch the gender mix of their student body during our sample period. Because the latter exercise adds a large number of schools to the comparison group – many of which are very different from our treated group of historical women's colleges – we add two additional characteristics to our vector of covariates X: A measure of school size (proxied by discrete categories of the number of degrees awarded in pre-reform years) and pre-reform linear trends in the share of degrees among all students that are in STEM fields. All estimates are weighted by the number of female students in the school in its first year in our

dataset.

Our implementation differs in a few minor ways from the procedure outlined by Callaway and Sant'Anna (2021). The first is the choice of base period b. The estimator used by Callaway and Sant'Anna (2021) sets the base period as the year before researchers believe treatment effects would be expected to arise in their setting. In cases without anticipation effects, this would generally mean $b = g - 1$. We instead define Y_b as the average of the outcome variable in the five years immediately preceding the switch to coeducation, $Y_b = \frac{1}{5}\sum_{s=g-5}^{g-1} Y_s$. This choice requires slightly stronger assumptions about parallel trends (Marcus and Sant'Anna, 2021), but should improve efficiency and reduce the impact of noise on our estimates (Borusyak, Jaravel and Spiess, 2021). Second, the conventional estimator – as implemented in the "did" package for R – takes advantage of the parallel-trends assumption to define $\Delta Y_{t,b} = Y_t - Y_b$ for post-treatment years, but as single-year differences $Y_t - Y_{t-1}$ in pre-reform years. We instead adopt the former definition for all periods. While the two approaches are nearly identical conceptually, we believe our approach to reporting event-study results provides a more familiar and intuitive way for our readers to evaluate the possibility of violations of the parallel trends assumption. Finally, rather than estimating a propensity score with a logit or probit model for each cohort g, we use discrete variables (or discretized versions of continuous variables) to find exact matches, school by school, for each institution in the treatment group before aggregating our estimates. This is equivalent to defining a propensity score using fully saturated OLS. This approach avoids the pitfalls of estimating logit and probit models in situations with few treated units. It also allows us to focus on what we believe are the most important conditioning variables in our setting, which happen to be discrete.

Callaway and Sant'Anna (2021) also propose an inference procedure that accounts for multiple hypothesis testing across time periods in an event-study figure. While the results in our main appendix report pointwise 95% confidence intervals, Figures A5 and A6 show that our estimates are noisier but, for the most part, still statistically distinguishable from 0 when using this procedure.

Multiple-testing corrections have minimal impact on our estimates of "long-run" effects.

C.1 Testing for differences in intended, completed major

One key question in our analysis is whether the decreases in the share of females earning STEM degrees can be explained by changes in the composition of entering freshman. In particular, if women who are interested in STEM also chose to enroll elsewhere when colleges adopted coeducation, then our estimated effects could be explained by sorting across colleges rather than changes in gendered peer or role-model effects. In this case, we might see a decrease in STEM degree attainment even if the reform had no effect on the degree choice of any individual student.

In fact, using data from HERI on freshmen's degree intentions by school, year, major, and sex, we find no evidence of changes in the share of entering freshman women who intended to major in STEM fields – an estimated long-run effect of 0.016 (s.e. 0.026). Our novel linkage between switching schools in HERI and IPEDS degree-completion data allows us to estimate a STEM degree-completions effect using the same subset of schools: an effect of -0.034 (s.e. 0.008), as shown in Figure A1. To formally test the difference between these two estimates, $\hat{\delta} = \hat{\beta}_{LR}^{HERI} - \hat{\beta}_{LR}^{IPEDS}$, we would require an estimate of their covariance. However, we cannot estimate the covariance because we lack a linkage between *non-switching* schools in the two datasets.

We instead employ a parametric bootstrap for statistical inference (Johnston and DiNardo, 1997; Goodman-Bacon, 2018). This routine exploits the fact that $\hat{\beta}_{LR}^{HERI}$ and $\hat{\beta}_{LR}^{IPEDS}$ can be viewed as draws from a bivariate normal distribution with an unknown (off-diagonal) covariance structure. While we do not know their covariance, we can use this theoretical insight to calculate confidence intervals under a range of plausible values of the correlation between them. We undertake this exercise and calculate 90% and 95% confidence intervals using the procedure described below. The results are reported

in Figure A9.

Furthermore, we can use the institutional details of our current setting to narrow the range of plausible values of the correlation coefficient. If composition effects (as measured by preferences for STEM) are entirely absent, then we would expect to see no correlation between bootstrapped estimates of $\hat{\beta}_{LR}^{HERI}$ and $\hat{\beta}_{LR}^{IPEDS}$, since idiosyncratic changes in the share of women intending to major in STEM are necessarily unrelated to effects on STEM degree completions under this assumption. On the other hand, if composition effects explain a meaningful share of our estimated effects on degree completions, then a bootstrap draw that delivers a large estimated effect on degree intentions should also tend to deliver a large estimated effect on degree completions. In other words, the correlation between $\hat{\beta}_{LR}^{HERI}$ and $\hat{\beta}_{LR}^{IPEDS}$ should be positive. We therefore hypothesize that the relevant range of correlation coefficients should be between 0 and 1.

To implement the parametric bootstrap procedure, we first hypothesize a potential value of the correlation coefficient ρ. We then draw 10,000 times from a bivariate normal distribution, where the means and standard deviations of the marginal distributions are equal to the means and standard errors of our estimated effects on degree intentions and degree completions, and the correlation between the two is ρ. For each of these 10,000 realizations, we calculate the simple difference $\hat{\delta}_b = \hat{\beta}_{LR,b}^{HERI} - \hat{\beta}_{LR,b}^{IPEDS}$. Following Goodman-Bacon (2018), we then use a modified percentile method to construct the confidence intervals. Specifically, the lower bound of the 95% confidence interval is the 5th percentile of $\hat{\delta}_b$ among realizations where $\hat{\delta}_b < \hat{\delta}$. The upper bound of the 95% confidence interval is the 95th percntile of $\hat{\delta}_b$ among realizations where $\hat{\delta}_b \geq \hat{\delta}$. The 90% confidence interval is constructed analogously. This procedure is then constructed at intervals of 0.05 for possible values of $\rho \in [-0.95, 0.95]$.

The results of this exercise are plotted in Figure A9. Under the assumption that there are no composition effects – and so effects on STEM intentions and degree copmletions are uncorrelated – the 90% confidence interval for the difference between these two effects excludes 0. We do not have enough precision to rule out equal effects at the 5% level. However, assuming some positive

correlation between the two effects, the bounds tighten further, such that our 95% confidence interval can reject 0 for any correlation larger than about 0.2. In sum, we conclude that this exercise suggests that there is little reason to believe that composition effects contribute substantially to the decrease in women majoring in STEM at newly coeducational colleges.

D Robustness check: the synthetic control method

As a robustness check on our main result, we use the synthetic control method to estimate the effects of transitioning to coeducation on women's STEM major choices. The synthetic control method offers a data-driven procedure to construct a control group that matches our treatment group based on pre-treatment characteristics. Thus, it may provide a valid comparison group even if our identification assumption fails in the standard difference-in-differences methodology used above.

One complication of our setting is that we have multiple "treated" schools rather than the single treated unit that is standard in synthetic control settings (e.g. Abadie, Diamond and Hainmueller, 2010). We adjust the standard procedure in two ways to incorporate this complication. First, we group schools that switched to coeducation in the same year, so that the "treated" groups are effectively school-cohort combinations. Second, we construct a synthetic control group separately for each cohort of treated schools and then average the effects by year relative to the switch (Cavallo et al., 2013; Acemoglu et al., 2016).

Our baseline specification constructs a synthetic control group for each treated school-cohort observation by matching on the entire set of pre-*transition* outcome variables (Ferman, Pinto and Possebom, 2020). Appendix Figure A7 reports the results of this estimation procedure. For consistency with our event-study results, time 0 corresponds to the effect on the female STEM share in the junior year the first coeducational cohort. Note that because years -2 and -1 are not used in the matching procedure, the fact that they remain near 0 provides an informal cross-validation test and some reassurance of the validity

of our design. In fact, we see little evidence of a departure from 0 effect until the graduating year of the first coeducational senior class. The synthetic control event study traces a similar path as did our standard event study (Figure A2): it shows a 2 percentage point decrease in the share of women majoring in STEM by five years after the transition to coeducation and a 3 percentage point decrease by nine years after the transition to coeducation. We calculate a "difference-in-differences" estimate by averaging the post-treatment coefficients and subtracting them from the average pre-treatment coefficients. The estimate of -0.025 is an outlier in the distribution of placebo effects, with a p-value of 0.01.[24] This estimate is slightly larger in magnitude than the one we obtain in our main event study model.

By matching on the entire set of pre-transition outcome variables, this baseline approach reduces concerns about specification searching (Ferman, Pinto and Possebom, 2020). In addition, in cases with sufficiently long pre-periods, it reduces the bias of the synthetic control estimator (Abadie, Diamond and Hainmueller, 2010; Kaul et al., 2018). However, its disadvantage is that it runs the risk of overfitting, particularly for our early-switching schools with few observed pre-treatment periods. To assuage these concerns, Appendix Figure A8 reports estimates from three other specifications (Ferman, Pinto and Possebom, 2020). The first specification replicates the baseline. The second specification matches on only the second half of pre-treatment outcomes as matching variables, rather than the full pre-period. The third specification, labeled "college characteristics," matches on a five-year average of the pre-treatment period plus some features of the school: A five-year average of the ratio of total PhD and professional degrees to bachelor's degrees, and the share of *all* students at a school that majored in humanities, social science, physical science, and business in the year before the reform. Finally, the fourth specification, labeled "School major distribution," replaces the bachelor's-degree shares from event-year -1 with those same shares for the last half of the pre-

[24] We conduct inference by randomly reassigning treatment status and estimating the effect of the transition to coeducation on the placebo institutions, using 500 replications (Abadie, Diamond and Hainmueller, 2015). If our estimated effect is either below the 2.5th percentile or above the 97.5th percentile of placebo effects, the effect is statistically significant.

period. While these robustness checks are generally noisier than our baseline estimates, the results are relatively consistent across specifications.

E Other data processing notes

E.1 Major codes

E.1.1 Coding scheme and crosswalks

This paper uses consistent 4-digit, 2-digit, and grouped 2-digit versions of major codes. The consistent coding scheme is based on the 1990 version of the Classification of Instructional Programs (CIP) from the National Center from Education Statistics (NCES).

Codes to describe college majors have been revised several times over our sample period. There were two sets of major codes in the HEGIS data, with a revision in 1970, and coding switched to the CIP in the early 1980s.[25] Revisions of the CIP occurred in 1985, 1990, 2000, and 2010.[26] Crosswalks between the 1970s HEGIS codes and the CIP, and between different versions of the CIP, are available from NCES, but they are not complete.

Similar to occupation codes, the CIP has 2-, 4-, and 6-digit versions of codes, while the HEGIS codes have only 2- and 4-digit versions. Revisions of the CIP only rarely move major categories across 2-digit codes,[27] though the 1990, 2000, 2010 revisions did move, split, and combine some two-digit codes.[28]

For this paper, all 6-digit codes were crosswalked to the 4-digit 1990 CIP.

[25]The first version of the CIP was constructed in 1980, but HEGIS seems not to have adopted it until 1983.

[26]There seems to have been late adoption of the new coding schemes in the IPEDS data – the switches seem to have occurred in 1987, 1992, 2002, and 2012, and may not have occurred uniformly across schools. Revisions of the CIP vary in how many changes were made, with the 1985 revision being much smaller than subsequent revisions.

[27]Exceptions include clinical versions of the life sciences, materials science, and educational psychology, all of which could be considered to be part of multiple two-digit codes.

[28]For instance, the 1990 revision of the CIP combined category 17, Allied Health, with category 18, Health Sciences, into category 51, Health Professions and Related Sciences. Most of the 4-digit categories were preserved but re-numbered in the revision.

Where crosswalks provided by the NCES were incomplete, they were supplemented by lists and descriptions of CIP codes created by the NCES. When majors were not included in the NCES crosswalks, they were matched to the major of the most similar title and description in the 1990 CIP. If two 4-digit codes were combined in any version of major codings after 1970, they were combined in the consistent coding scheme. The same is true for the 2-digit codes. Six-digit majors that were created or deleted at any point were assigned to the same 4-digit code in the "other" category, and 4-digit codes that were ever created or deleted were assigned the the 4-digit code for "other" within the same 2-digit code.[29] Four-digit majors with fewer than 950 school-by-year observations were combined with majors that cover similar material[30] or with the "other" category within their two-digit code. Smaller 2-digit codes, such as Law, Library Science, and Military Science, were treated as a single 4-digit code.

For the main result, majors were combined into groups of 2-digit codes, with the most important of those groups being STEM. STEM in this case includes the 2-digit codes for Life Sciences, Physical Sciences, Engineering, Computer Science, and Mathematics. Alternative specifications also included Health Professions.

E.1.2 Categories of majors

The following list is the two-digit categories of majors in each group of 2-digit codes. Groups are in bold and the two-digit categories are listed afterward. Where the two-digit sets of codes are not informative, four-digit codes are included in parentheses. Some groups contain only one two-digit code. The "other" group includes majors that generally cannot be found at small liberal arts colleges or that are generally very small.

Art Visual and performing arts, architecture and related services

[29]For instance, African Languages were not included in the 1990 CIP and were therefore assigned to the 4-digit code for Other Foreign Languages.

[30]For instance, Architectural Engineering and Civil Engineering, Business Administration and Enterprise Management, and the health categories such as medicine, dentistry, and others which require a professional degree.

Business Business, marketing

Education All education fields (including math education)

Economics Economics (4-digit code)

Health Health professions and clinical services

Home Economics Home economics/family and consumer sciences

Humanities Area and group studies (e.g. gender studies, Hispanic Studies), English, foreign languages and linguistics, philosophy and religious studies

Psychology Psychology

Other Social Sciences Social sciences except economics (general social science, anthropology, criminology, demography, geography, history, international relations, political science, social science, urban studies), communications

STEM Life sciences, physical sciences, mathematics and statistics, computer and information science, engineering, engineering technology, science technology

Other Agriculture, forestry, law, trades/vocational, military science, library science, multi- and inter-disciplinary, theology and religious vocations, protective services, public administration and social services

E.2 School Codes

NCES uses two different coding schemes for individual schools at different points in the data. HEGIS identifies schools using FICE codes, which is a six-digit identification code assigned to schools doing business with the Office of Education in the 1960s. IPEDS uses the UnitID, which is also a six-digit code. Our data uses the FICE as a consistent identifier throughout the survey, with some modifications as detailed below.

Not every institution has a FICE code. Institutions that do not have a FICE code are those that entered the IPEDS data after the Institutional Characteristics file stopped listing FICE codes (which was during the 1990s). We drop those institutions from our sample, as according to the ICPSR files for IPEDS financial characteristics between 1988 and 1990, institutions that entered the sample after the beginning of the IPEDS have a much lower response rate than institutions in the HEGIS sample. However, the data set itself has the UnitID entered in place of the FICE code for those institutions.

Some institutions have multiple FICE codes. In most of these cases, a public institution originally reported all branches under one observation, and then switched to reporting each branch separately. The vast majority of cases where all degrees awarded are reported under the main campus occur in 1966, with a few additional cases between 1967 and 1969. We do not link such cases together. In other cases, an institution switched FICE codes in the middle of the sample. We are generally not sure why this occurs. We do link these cases together so that we have a single FICE code for all years the institution was in the data. Finally, there are a few institutions (notably Cornell and Columbia) with several different administrative units that separately report degrees awarded to IPEDS and HEGIS. We treat these institutions as a single observation and collapse them to a single FICE code.

Some FICE codes apply to multiple institutions. In these cases, all institutions are part of the same system, and the majority of these cases occur among institutions who enter the data in 1987 and later, especially among for-profit institutions with multiple campuses nationwide (e.g. the University of Phoenix). There are some cases where a public college with several branches (e.g. the University of Pittsburgh) reported degrees separately from each branch but reported the same FICE from each school. Where we could, we assigned these institutions to separate codes for each branch, but the rest of them are collapsed to the FICE level. We have also dropped schools that are ever classified as for-profit schools from our sample, which removes many of these cases from our analysis.

F Substitution of STEM majors to other women's colleges

Another indicator of a "composition effect" of the transition to coeducation would be a substitution to the remaining single-sex colleges of women interested in quantitative fields. To test for this possibility, we draw on our IPEDS/HEGIS dataset and estimate the following equation:

$$s^{W}_{STEM,i,t} = \theta_i + \rho CW_{it} \times Exposure_{it} + \delta_{r(i)t} + \psi_{c(i)t} + \varepsilon_{it} \tag{11}$$

This equation contains fixed controls for the institution θ_i, region $\delta_{r(i)t}$, and control of the institution $\psi_{c(i)t}$. The independent variable of interest is $Exposure_{it}$, which equals

$$\frac{\text{Number of comparable coeducational colleges}}{\text{Number of comparable coeducational and women-only colleges}}$$

where "comparable colleges" are colleges within the same Census region and 1972 Barron's selectivity band as college i. The indicator CW_{it} is 1 if the institution is currently women-only and 0 otherwise. The parameter ρ captures the extent to which women's STEM majoring increases at current women's colleges as comparable women's colleges increasingly transition to coeducation. We consider two samples: women's colleges only, and women's colleges plus always-coeducational colleges.[31]

Our results are reported in Table A4. When considering only colleges that were women-only in a given year, we find a coefficient of -0.00064. When including institutions that were always coeducational, which better controls for regional- and selectivity-driven trends in women's STEM majoring, we find a coefficient of -0.0064. That is, a ten percentage point increase in the share of

[31] When excluding always-coed schools, ρ is identified off of differential rates of transition *across* region-by-selectivity cells. Including always-coed schools allows ρ to be identified from co-variation in transition events and women's STEM majoring at women's colleges *within* a region-by-selectivity cell, relative to trends in women's STEM majoring at always-coed colleges.

comparable colleges which were coeducational led to a 0.06 percentage point decrease in women's STEM majoring. This estimate is neither statistically nor economically significant, especially considering that each women's college in our sample had on average 8.5 peer institutions within a region-by-selectivity band in the first year they appeared in the data. This evidence further indicates that the composition effect, if it exists, is quite small.

G The role of capacity constraints

One possible explanation for our finding of a reduction in the share of women majoring in STEM at newly coeducational colleges is that such colleges hit capacity constraints to a larger extent in STEM than in non-STEM fields. For example, if STEM is both more costly for colleges to provide and more popular among men, students might be more crowded out of STEM majors than non-STEM majors after the transition to coeducation. This would suggest a negative relationship between field-specific cost of instruction and growth in degrees earned.

To test this hypothesis, we examine growth in degrees earned (by both men and women) in a wide range of fields and compare these figures with the marginal cost of instruction, as reported by Hemelt et al. (2018). In the first step, we use equation 4 to estimate the long-run effect on the number of degrees awarded in each field. We then link these results to estimates of marginal costs. Where the fields in our data did not overlap exactly with those reported in Hemelt et al. (2018), we aggregate the marginal cost estimates by calculating the simple average.[32]

Results are shown in Figure A11. The four main fields of interest in our paper are shown as orange triangles, while all others are in blue. Among the three STEM fields of math, biology, and physical sciences, we see a slight *positive* relationship, suggesting that the costliest-to-teach fields were also the

[32]Hemelt et al. (2018) report a range of cost estimates by field. We rely on marginal costs estimated with program fixed effects, and using only schools without graduate programs to maximize comparability with the treated schools in our sample. For their estimates, see Table 5, column 5 of their NBER working paper.

fields where growth was largest. Indeed, despite the relatively low cost of adding students to math or economics courses, we find significant negative effects on the share of women majoring in these fields, suggesting that these effects cannot be explained by women being physically crowded out of the classroom.

Appendix references

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association*, 105(490): 493–505.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2015. "Comparative politics and the synthetic control method." *American Journal of Political Science*, 59(2): 495–510.

Acemoglu, Daron, Simon Johnson, Amir Kermani, James Kwak, and Todd Mitton. 2016. "The value of connections in turbulent times: Evidence from the United States." *Journal of Financial Economics*, 121: 368–391.

Cavallo, Eduardo, Sebastian Galiani, Ilan Noy, and Juan Pantano. 2013. "Catastrophic natural disasters and economic growth." *Review of Economics and Statistics*, 95(5): 1549–1561.

Ferman, Bruno, Cristine Pinto, and Vitor Possebom. 2020. "Cherry picking with synthetic control." *Journal of Policy Analysis and Management*, 39(2): 510–532.

Howe, Florence, Suzanne Howard, and Mary Jo Boehm Strauss. 1982. *Everywoman's Guide to Colleges and Universities.* The Feminist Press.

Kaul, Ashok, Stefan Klößner, Gregor Pfeifer, and Manuel Schieler. 2018. "Synthetic control methods: Never use all pre-intervention outcomes as economic predictors." MPRA Working paper 83790.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2020. "IPUMS USA: Version 10.0 [dataset]." IPUMS, Minneapolis, MN.

H Tables and Figures

Table A1: Determinants of the transition to coeducation

	(1) Switched	(2) Switched	(3) Switched before 1972	(4) Switched before 1972
Region = Midwest	-0.0246	0.0565	-0.572	-0.484
	(0.130)	(0.127)	(0.163)	(0.177)
Region = South	0.156	0.195	-0.377	-0.283
	(0.0937)	(0.0909)	(0.128)	(0.131)
Region = West	-0.143	-0.172	-0.376	-0.471
	(0.136)	(0.135)	(0.177)	(0.167)
Ever Catholic	0.191	0.108	-0.174	-0.333
	(0.101)	(0.117)	(0.128)	(0.111)
Log enrollment	0.00456	0.00734	-0.0702	-0.0723
	(0.0901)	(0.0859)	(0.104)	(0.0973)
Selective		-0.219		-0.317
		(0.116)		(0.155)
Observations	112	110	85	83
R-squared	0.122	0.153	0.169	0.218

Notes: Data drawn from HEGIS/IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. Linear probability model estimates based on transitioning to coeducation (for all initial women's colleges) or transitioning to coeducation before 1972 (for all switchers) as the dependent variable. Robust standard errors in parentheses. Region refers to U.S. Census regions. Omitted region category is Northeast. "Selective" refers to a Barron's rating of 1, 2, or 3 in 1972.

Table A2: Long-run effect of coeducation on male share of degrees and the share of women choosing other majors

	(1) Male share	(2) Art	(3) Business	(4) Education	(5) Health
Panel A: Not-yet-treated comparison group					
Long-run effect	0.192***	-0.001	-0.017	-0.017	0.034
	(0.016)	(0.009)	(0.013)	(0.022)	(0.022)
Counterfactual mean	0.036	0.072	0.121	0.184	0.163
Observations	5,455	5,455	5,455	5,455	5,455
Panel B: All-college comparison group					
Long-run effect	0.210***	0.006	-0.027**	0.008	0.053**
	(0.021)	(0.008)	(0.012)	(0.019)	(0.021)
Counterfactual mean	0.013	0.070	0.129	0.164	0.141
Observations	23,969	23,967	23,967	23,967	23,967
Panel C: Never-treated comparison group					
Long-run effect	0.195***	0.003	-0.013	-0.011	0.016
	(0.017)	(0.012)	(0.012)	(0.025)	(0.026)
Counterfactual mean	0.034	0.069	0.116	0.178	0.180
Observations	5,164	5,164	5,164	5,164	5,164

Notes: Table displays the estimated effect of the switch to coeducation on male share of degrees earned (column 1) and graduating female students' choice of major (columns 2-5), estimated using $\hat{\beta}_{LR}$ from equation 4. Each panel uses the specified pool of institutions to construct a comparison group and estimate a counterfactual trend in major choices, conditional on college selectivity and historical affiliation with the Catholic Church. In panel B, we additionally condition on school size, as measured by number of degrees granted, and the trend in STEM choice among all students. Data drawn from HEGIS/IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. Standard errors are estimated using a block bootstrap with 1,000 replications that accounts for intracluster correlation at the institution level. Counterfactual mean is the share of women that would have chosen each major at treated schools if choices at those schools had followed trends at the comparison group of institutions.

Table A3: Long-run effect of coeducation on the share of women choosing other majors

	(1) Home ec	(2) Humanities	(3) Other	(4) Psychology	(5) Soc sci
Panel A: Not-yet-treated comparison group					
Long-run effect	0.001	0.001	0.001	0.011	0.022
	(0.006)	(0.014)	(0.017)	(0.007)	(0.015)
Counterfactual mean	0.030	0.095	0.039	0.058	0.126
Observations	5,455	5,455	5,455	5,455	5,455
Panel B: All-college comparison group					
Long-run effect	0.004	-0.007	0.005	0.001	-0.002
	(0.006)	(0.010)	(0.012)	(0.007)	(0.012)
Counterfactual mean	0.029	0.102	0.034	0.068	0.145
Observations	23,967	23,967	23,967	23,967	23,967
Panel C: Never-treated comparison group					
Long-run effect	0.001	-0.001	0.015	0.008	0.026**
	(0.006)	(0.015)	(0.012)	(0.007)	(0.013)
Counterfactual mean	0.030	0.098	0.026	0.061	0.122
Observations	5,164	5,164	5,164	5,164	5,164

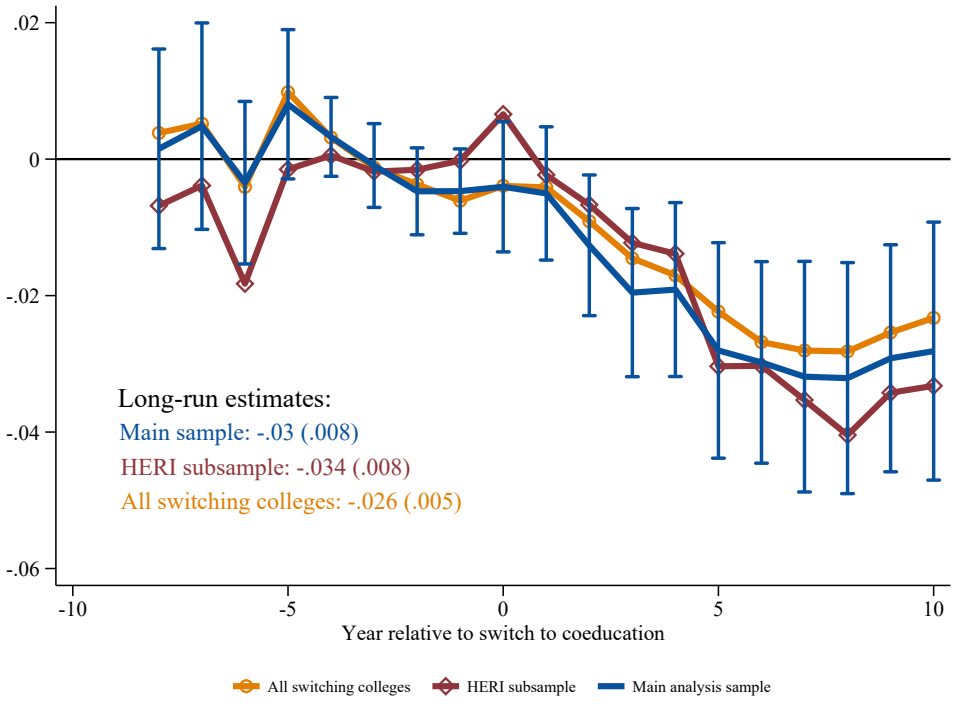
Notes: Table displays the estimated effect of the switch to coeducation on graduating female students' choice of major, estimated using $\hat{\beta}_{LR}$ from equation 4. Each panel uses the specified pool of institutions to construct a comparison group and estimate a counterfactual trend in major choices, conditional on college selectivity and historical affiliation with the Catholic Church. In panel B, we additionally condition on school size, as measured by number of degrees granted, and the trend in STEM choice among all students. Data drawn from HEGIS/IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. Standard errors are estimated using a block bootstrap with 1,000 replications that accounts for intracluster correlation at the institution level. Counterfactual mean is the share of women that would have chosen each major at treated schools if choices at those schools had followed trends at the comparison group of institutions.

Table A4: Spillover effects analysis: the effect of coeducation on STEM majoring at comparable colleges remaining women-only

	Current women's colleges	Current women's + always coed
$CW_{it} \times Exposure_{it}$	-0.00064	-0.00639
	(0.0187)	(0.0218)
Observations	2,313	47,297
R-squared	0.738	0.721
Colleges	112	1,027

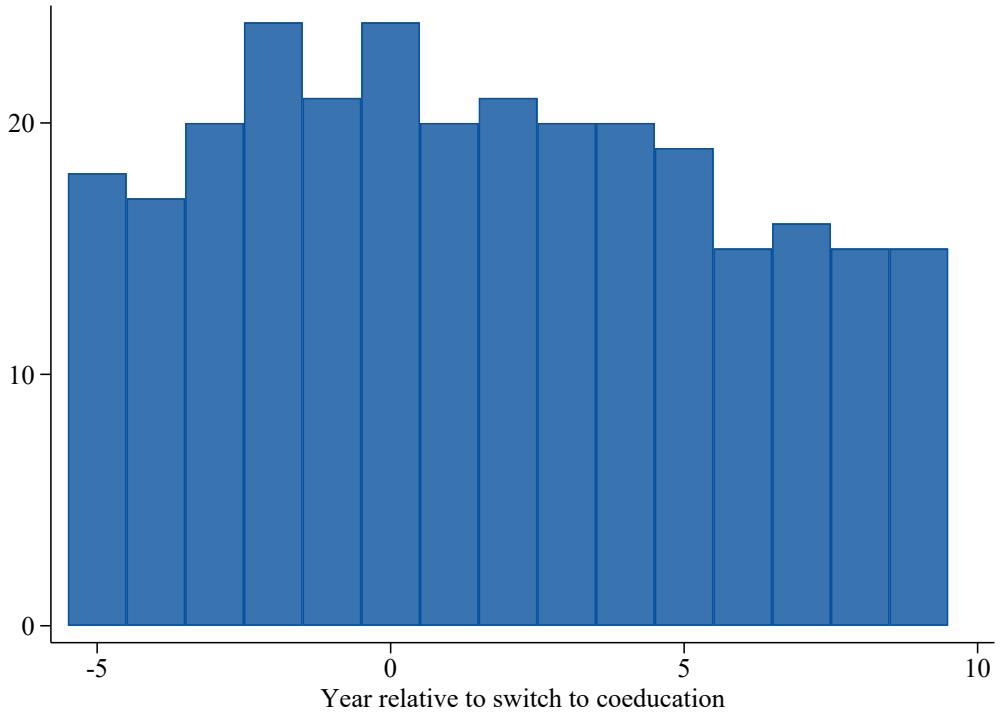
Notes: Data drawn from HEGIS/IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. See Section 2 for further detail. The majors included in the STEM concentration are described in Section 2 and Appendix B. The table presents estimates of the effect of exposure to coeducational colleges within the same region and selectivity band on the share of women graduating from women's colleges with STEM degrees. See Section 4 for further detail. Standard errors are clustered at the institution level. A small number of institutions were dropped due to missing Barron's selectivity ratings.

Figure A1: Robustness of estimated effect of coeducation on share of women majoring in STEM to sample criteria



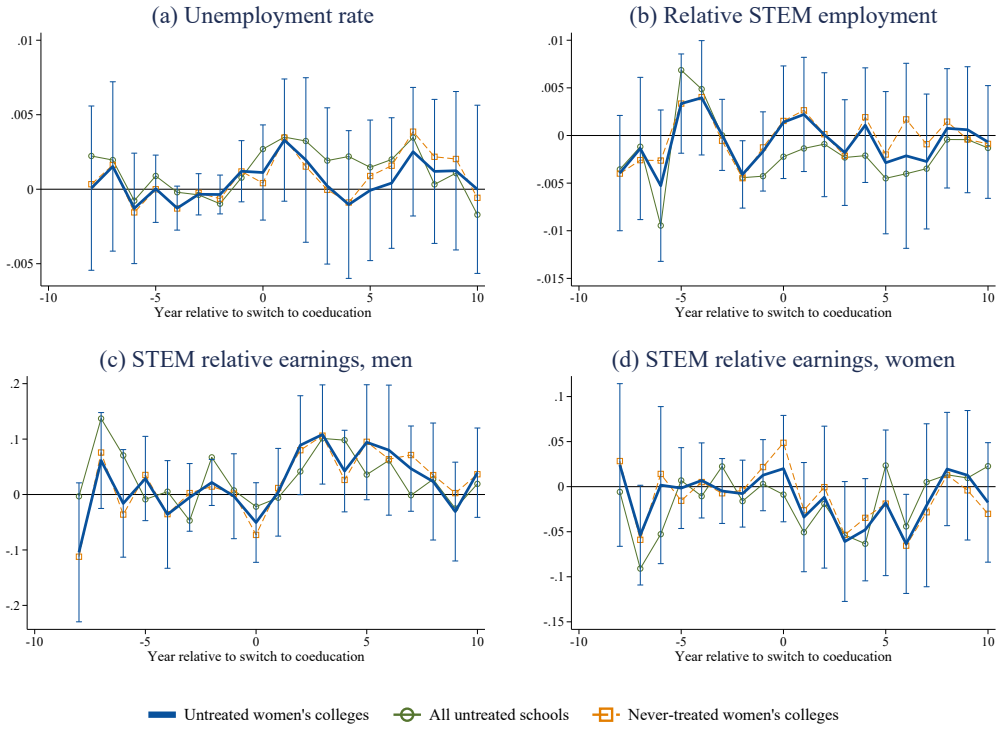
Notes: Figure shows event-study estimates from equation 2. Blue line replicates estimates from Figure 4, which is based on a sample of 77 treated schools. Orange line with circle markers shows estimates from the larger sample that includes another 42 treated schools that are excluded from the main sample because they did not have a "clean" transition to coeducation, as detailed in Appendix B.3. Maroon line with diamond markers shows estimates drawn from subsample of IPEDS/HEGIS schools that can be linked to data from The Freshman Survey.

Figure A2: Representation of former women's colleges in The Freshman Survey



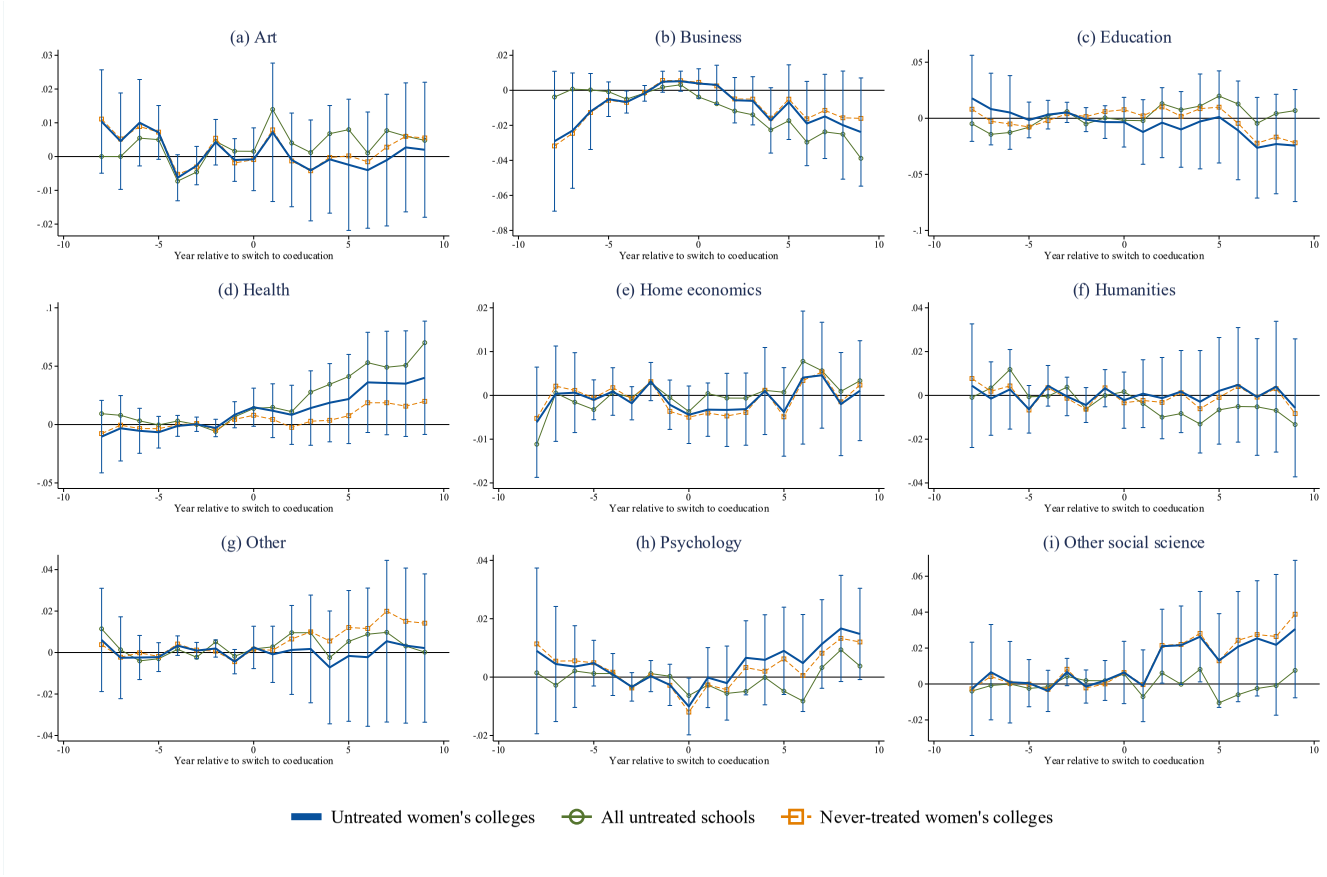
Notes: Data drawn from 1966-2006 versions of The Freshman Survey administered by HERI, linked to hand-collected dates of transitions to coeducation by institution's state. Each bar shows the number of treated schools that appear in the survey in each year relative to the switch to coeducation. Sample of treated schools is limited to 30 institutions that were surveyed at least once in the five years prior and once in the 10 years after the reform.

Figure A3: Tests for coinciding labor-market shocks



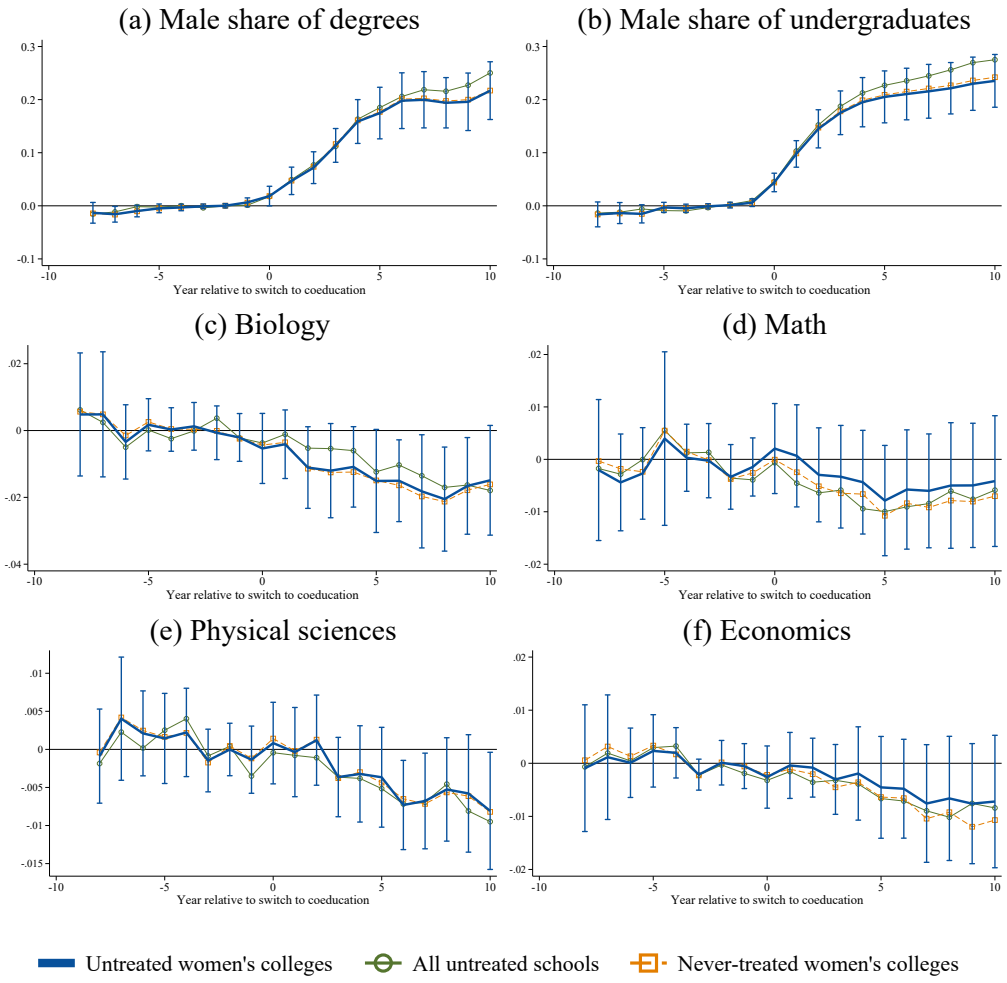
Notes: Data drawn from 1966-2016 CPS data accessed via IPUMS (Ruggles et al., 2020), linked to hand-collected dates of transitions to coeducation by institution's state. See Section 2 for further detail. Unemployment rate is measured among individuals age 18-64. Relative STEM employment is constructed as the ratio of college-educated workers in STEM occupations to workers in non-STEM occupations. Relative income among men is constructed as the ratio of average annual income among college-educated men currently working in a STEM occupation to average annual income among college-educated men currently working in a non-STEM occupation. Relative income for women in STEM is constructed in the same manner, except that we include individuals with 0 earnings in the previous year. Panels display estimates of β_k from equation 2. Standard errors are constructed from a block bootstrap clustered at the institution level.

Figure A4: Effect of coeducation on share of women choosing other degrees



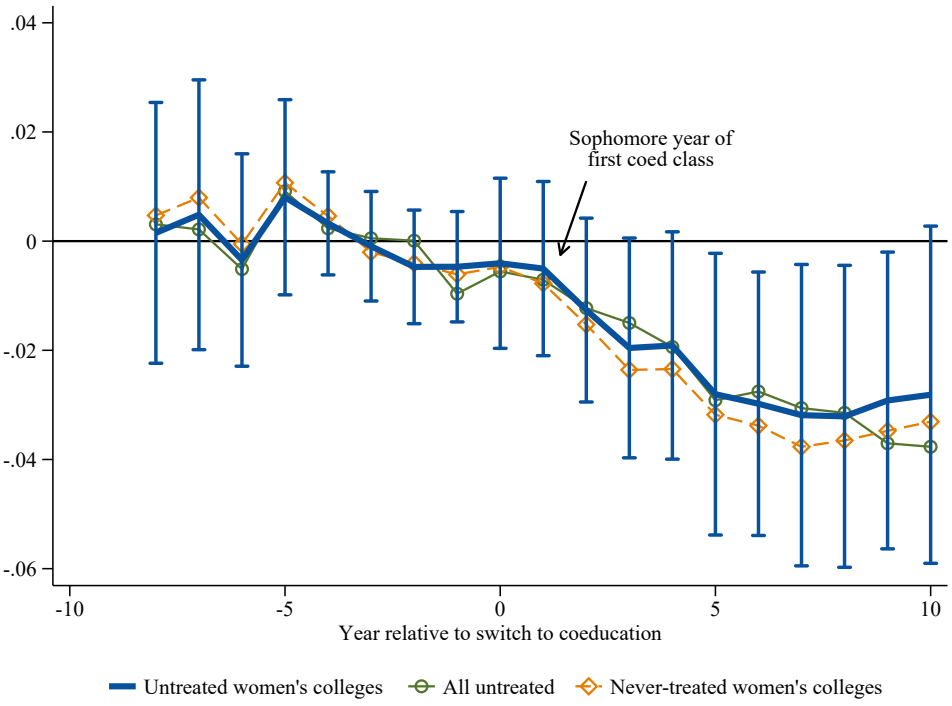
Notes: Figure shows event-study estimates from equation 10, aggregated using equation 2, using comparison group specified in legend. Data on degrees granted comes from HEGIS and IPEDS surveys, 1966-2016, linked to hand-collected information on dates of transition to coeducation. 95% confidence intervals are constructed using block bootstrap clustered at the institution level.

Figure A5: Effect of coeducation on gender composition and women's choice of quantitative majors, with uniform confidence intervals



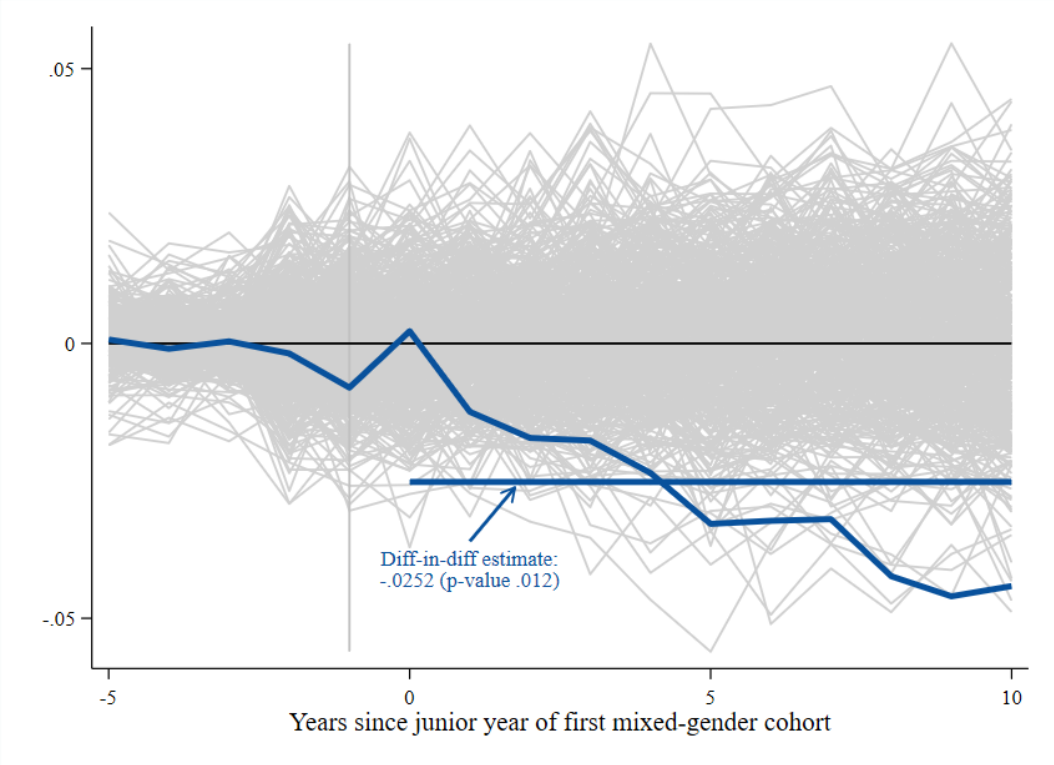
Notes: Figure repeats point estimates from Figure 3 with confidence intervals that account for multiple testing across event-time periods.Data drawn from HEGIS/IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. See Section 2 for further detail. Panels display estimate of β_k from equation 2. Dependent variable is the share of degrees earned in STEM among all degrees earned by women in the academic year.

Figure A6: Effect of coeducation on the share of women majoring in STEM, with uniform confidence intervals



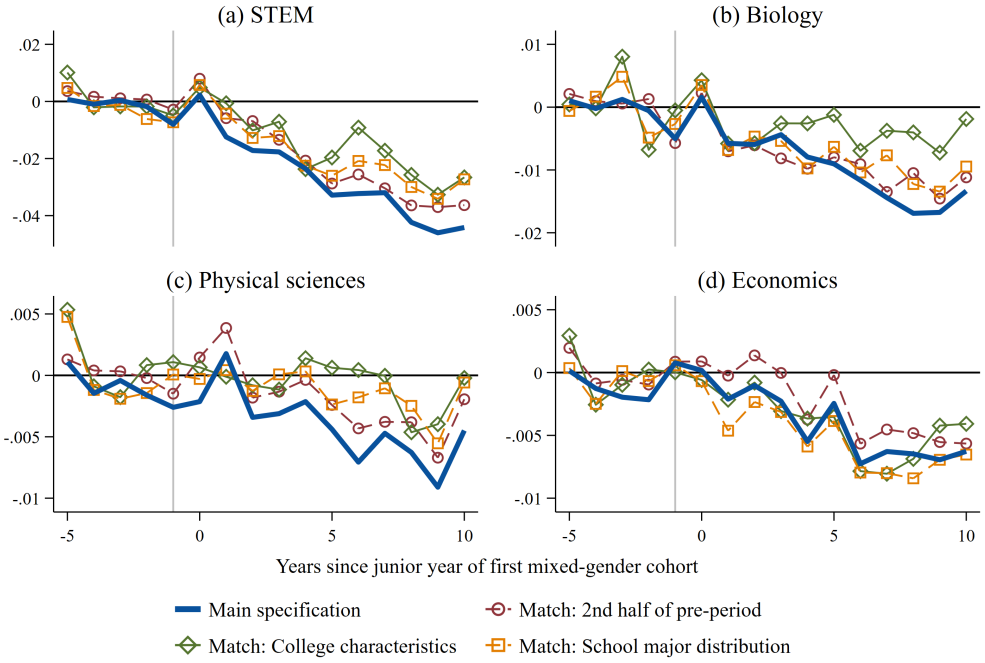
Notes: Figure repeats point estimates from Figure 4 with confidence intervals that account for multiple testing across event-time periods.Data drawn from HEGIS/IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. See Section 2 for further detail. Panels display estimate of β_k from equation 2. Dependent variable is the share of degrees earned in STEM among all degrees earned by women in the academic year. STEM fields include math, biology, physical sciences, engineering, engineering technology, and computer science.

Figure A7: The effect of coeducation on the STEM share of degrees awarded to women: synthetic control specification



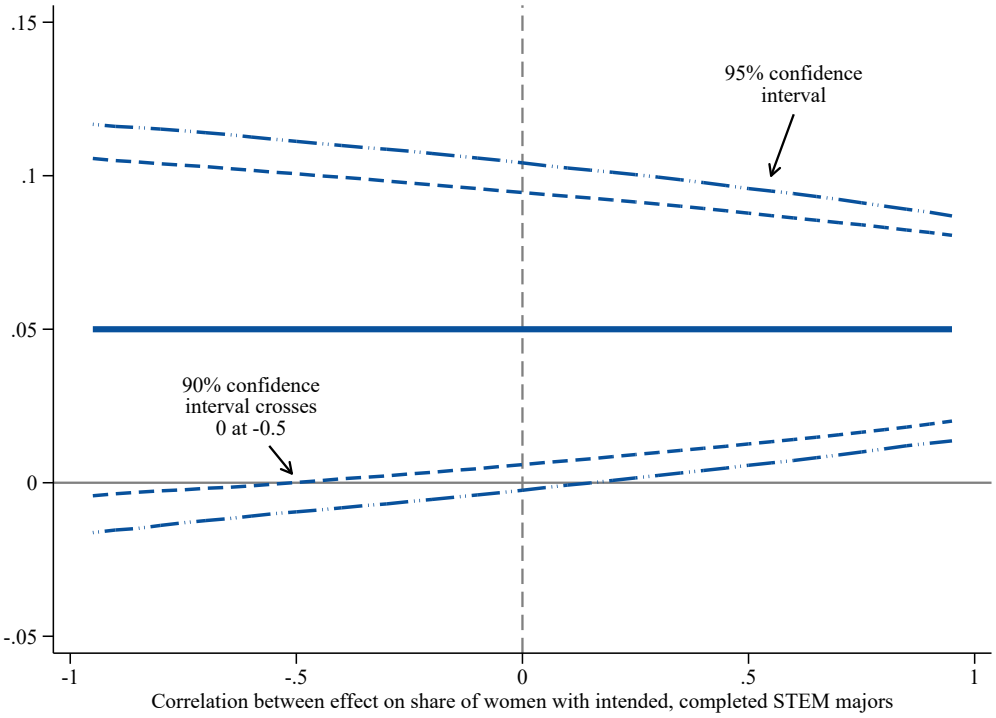
Notes: Data drawn from HEGIS/IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. See Section 2 for further detail. The majors included in the STEM concentration are described in Section 2 and Appendix B. See Appendix D for description of the synthetic controls procedure. Dark line reports the main estimate, while grey lines report the results of a randomization inference procedure with 500 replications.

Figure A8: Synthetic control estimates using alternate specifications and majors



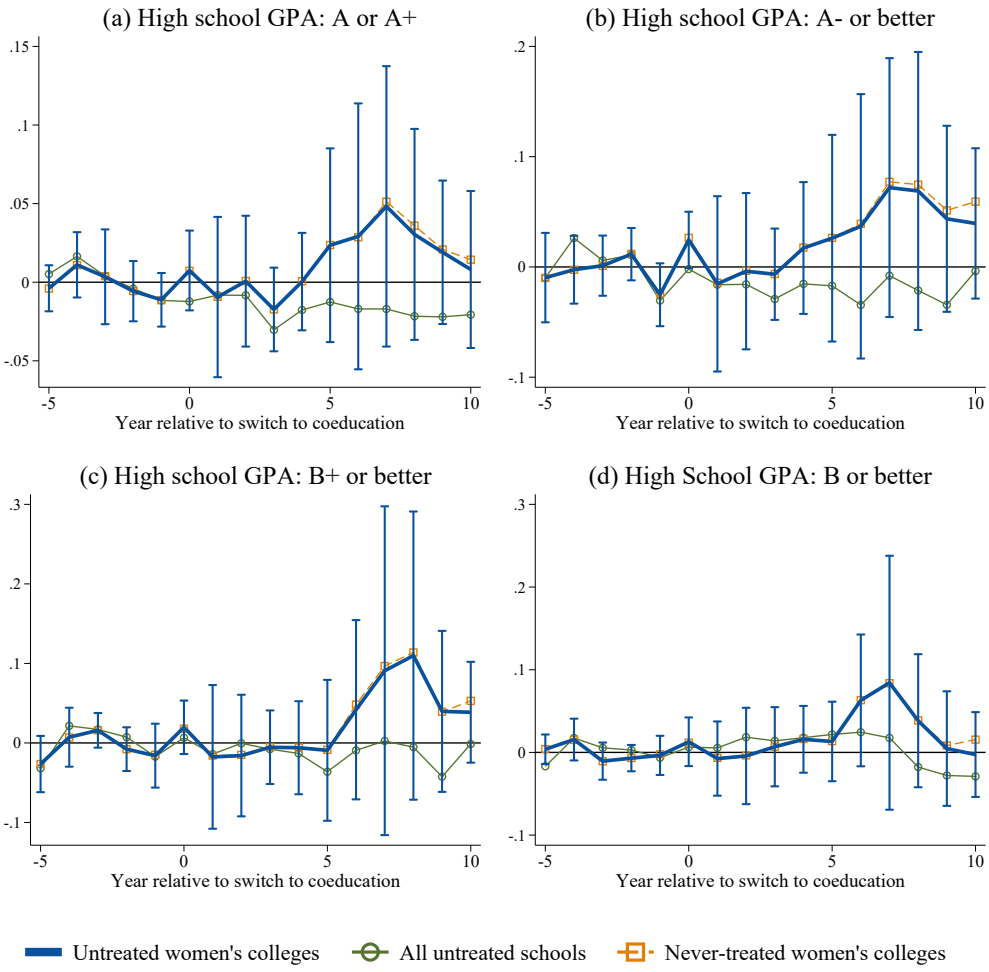
Notes: Data drawn from HEGIS/IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. See Section 2 for further detail. See Appendix B.2 for information on the specific majors included within each concentration. Appendix D provides further detail on the synthetic control estimation procedure. Main specification matches on outcome variable for all pre-treatment years, i.e. all years before the freshman year of the first coeducational cohort. Specification in red (circle markers) matches on second half of pre-period. Green line (diamond markers) matches on average of outcome over 5 years period to the transition to coeducation, plus five-year average of ratio of professors and PhD students to undergraduates, share of all students majoring in humanities, social science, physical science, and business. Orange line (square markers) matches on 5-year average of pre-treatment outcomes, ratios of professors and PhD students to undergraduates, and share of the following majors in the last half of the pre-period: Humanities, social sciences, business, physical science.

Figure A9: Estimated difference between share of women intending to major in STEM and completing degrees in STEM



Notes: Figure plots estimated difference between effect on share of women intending to major in STEM and share of women earning degrees in STEM, $\hat{\delta} = \hat{\beta}_{LR}^{HERI} - \hat{\beta}_{LR}^{IPEDS}$, where $\hat{\beta}_{LR}^{HERI}$ is the effect on the share of women intending to major in STEM from our HERI data and $\hat{\beta}_{LR}^{IPEDS}$ is the effect on the share of women earning degrees in STEM from our IPEDS data, calculated using equation 4. Confidence intervals are obtained using a parametric bootstrap with 10,000 draws of $\hat{\beta}_{LR}^{HERI}$ and $\hat{\beta}_{LR}^{IPEDS}$ from a bivariate normal distribution with means and standard errors drawn from our estimates for the HERI (0.016, s.e. 0.026) and IPEDS (-0.034, s.e. 0.008) results (Johnston and DiNardo, 1997). The correlation between the draws of HERI and IPEDS estimates is given by the value on the horizontal axis. See Appendix C.1 for a detailed description of the parametric bootstrap routine.

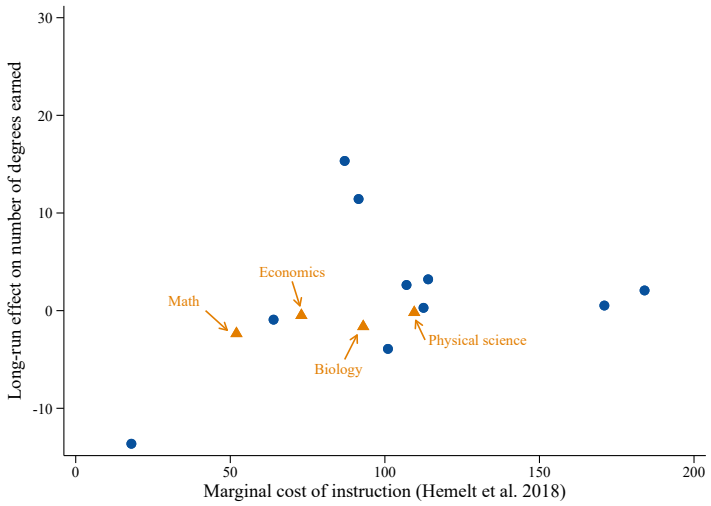
Figure A10: The effects of coeducation on the distribution of female matriculants' high school GPA



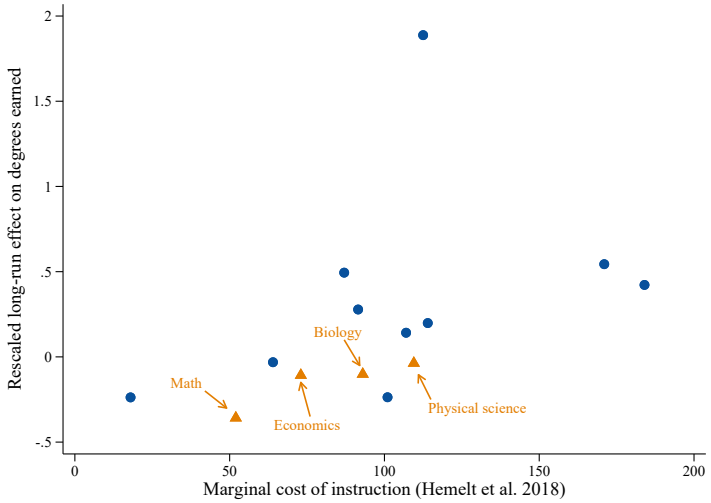
Notes: Each point shows and estimate of β_k from equation 2, using the comparison group specified in the legend. Error bars show 95% confidence interval constructed using block bootstrap clustered at the institution level. Data drawn from HERI, spanning 1966-2006, linked to hand-collected dates of transitions to coeducation by institution. See Section 2 for further detail.

Figure A11: Relationship between growth in total degrees earned and marginal cost of instruction by field

(a) Long-run effect vs marginal cost of instruction



(b) Rescaled long-run effect vs marginal cost of instruction



Notes: Long-run effects on degrees earned are estimated using equation 4 and data drawn from HEGIS/IPEDS surveys. In Figure A11b, effects are rescaled by the counterfactual mean of total degrees earned in the field. Marginal cost of instruction come comes from Hemelt et al. (2018) estimates among colleges with no graduation programs (see Table 5, column 5).