6-16-2022

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

Avery Calkins

Ariel J. Binder

Dana Shaat

Brenden Timpe

Published (2023) American Economic Journal: Applied Economics, 15 (3), pp. 1-34.
DOI: 10.1257/app.20210692

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

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

June 16, 2022

Abstract

We leverage variation in the adoption of coeducation by U.S. women's colleges to study how exposure to a mixed-gender collegiate environment affects women's human capital investments. Our event-study analyses of newly collected historical data find a 3.0-3.5 percentage-point (30-33%) decline in the share of women majoring in STEM. While coeducation caused a large influx of male peers and modest increase in male faculty, we find no evidence that it altered the composition of the female student body or other gender-neutral inputs. Extrapolation of our main estimate suggests that coeducational environments explain 36% of the current 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, Marta Murray-Close, 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 represent the official views of the National Institutes of Health or the U.S. Census Bureau. 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.

While 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), 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). However, the difficulty of isolating plausibly exogenous variation in these non-pecuniary factors has limited efforts to test this hypothesis.

In this paper, we estimate the impact of mixed-gender educational environments on women's field choices by leveraging an important 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 hand-collect information on the timing of schools' transitions to coeducation, merge it to multiple sources of panel data on the near-universe of U.S. baccalaureate institutions, and implement a dynamic difference-in-difference 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. In contrast to the existing literature, which studies marginal changes in the gender composition of classmates or instructors in an otherwise stable campus environment (e.g., Carrell, Page and West, 2010; Huntington-Klein and Rose, 2018), our paper analyzes a more fundamental reform that altered many social and educational aspects of the college experience (Miller-Bernal, 1993).

We first find that coeducation substantially altered the distribution of fields chosen by graduating women.Our event-study estimates show that the share of women majoring in any STEM field declined by 3.0-3.5 percentage points (30-33%) in the ten years following the arrival of coeducational classes. 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 share 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 several different choices of comparison group, including a synthetic control approach.

Coeducation could influence women's field choices through two primary channels. First, a coeducational environment could influence a woman's choice of major, conditional on having enrolled at the college (an *environmental effect*). As might be expected, we find that coeducation had a dramatic effect on the gender composition of peers: The male share of the student body increased steadily, reaching a 21- to 24-percentage-point increase in the latter half of the first decade after the reform. We also find a statistically significant increase in the male share of faculty, although this effect is only one-fifth as large as

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

the male peer effect. In contrast, we find no evidence that the switch to coeducation was accompanied by changes in a range of "gender-neutral" factors that may influence the choice of major, such as class sizes, the menu of courses offered, per-student expenditures, or the level of academic competition.

Second, coeducation could affect STEM-inclined women's propensities to enroll at former women's colleges (a *composition effect*). We investigate these mechanisms using a novel linkage between our hand-collected data and survey data on the career aspirations and academic records of college freshmen. We find no evidence of changes in a rich set of characteristics including *intended* major, intended career, demographic characteristics, or high school GPAs. To maximize our power to detect changes in composition, we use these variables to construct a *predicted* share of women who will major in STEM, based on their freshman characteristics. We find an insignificantly *positive* effect of coeducation on this index, suggesting that if anything, coeducation made the female student body *more* STEM-inclined. Though the data limit the precision of our estimates, the confidence intervals indicate that the composition effect could account for at most 16-30% of the total effect of coeducation on the share of women earning degrees in STEM.

Overall, these results are most consistent with the hypothesis that coeducation reduced the share of women majoring in STEM fields through its effect on the campus environment, and in particular via the increased exposure to male peers. As we discuss, there are several potential mechanisms that could underlie this effect, such as costs associated with mixed-gender competition, shifting gender norms and stereotypes, marriage market signals, and changes to teaching styles and subjective content accompanying the influx of male students. Though we do not have the ability to distinguish among these various channels, our findings suggest that social interventions that target the salience and perception of gender on college campuses may be important in increasing the representation of women in science. In fact, a simple extrapolation of our results suggests that exposure to coeducational environments explains 36% of the contemporary gender gap in STEM major choice.

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 (Goldin, 2015; Astorne-Figari and Speer, 2019; Calkins, 2021), 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 (Bertrand, 2011). 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 the gender composition of peers in educational environments. This literature has relied largely on fluctuations in the sex composition of classmates to study gendered peer effects (Hill, 2015; Huntington-Klein and Rose, 2018; Brenøe and Zölitz, 2020; Shan, 2021; Bostwick and Weinberg, 2022). The evidence generally suggests that more male-dominated peer groups deter women from quantitative fields, although these effects may disappear or even reverse in contexts where gender norms are likely to be less salient than other mechanisms such as social coordination (e.g., Zölitz and Feld, 2018). Our investigation differs in two important ways. First, we exploit a source of variation that induced a more dramatic shift in the male share of peers and the campus environment. In doing so, our estimates capture responses along several policy-relevant margins and shed light on the potential effects of broad-based efforts to change the perception of gender on college campuses. Second, we study exposure to men in a majority-women environment. Our results may be relevant for policies designed to foster access to single-sex educational settings, particularly ongoing initiatives that expose young women to STEM-related content. They may also shed light on the potential effects of the entry of men into women-dominated labor markets, such as nursing and family care.

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] Our paper also builds on work by Billger (2002). In a study of an anonymous liberal arts college that converted from women-only to a coeducational environment, she shows that the reform was accompanied by little or no change in the curriculum or other features of the campus environment. However, relative to aggregate national trends, she finds 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, considering the near-universe of historical women's colleges in the United States, and providing a more detailed exploration of the mechanisms.[4]

[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.

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 315 U.S. colleges served a women-only 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-seven women's colleges remained in operation as of 2022.[5]

Our research design leverages variation in the existence and timing of co-

[5]https://www.womenscolleges.org/

education 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 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 field choices by changing the (perceived or realized) payoffs 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 (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 faculty and student bodies constant, coeducation could influence women's major choices through *gender-neutral peer effects*. For example, coeducation could change the average abilities and attitudes of peers, which could affect the competitiveness of STEM courses and the style or content of instruction.

We interpret our main empirical estimates as the total effect of coeducation on women's field choices, summed across all of these mechanisms. While we cannot quantify every single mechanism, our analyses allow us to rule out some mechanisms in favor of others. Our main goal is to distinguish the composition effect from the environmental effect, which we accomplish via an indirect approach (see Section 5.4). The environmental effect informs our main question of interest: How the field choices of a *fixed group* of women respond to coeducation-induced changes in gendered features of the collegiate environment.

2 Data and Sample Construction

Our analysis relies on a hand-collected list of women's colleges and their dates of transition to coeducation. Our collection effort, which we detail in Appendix B, aimed to identify all baccalaureate-degree-granting women's colleges that transitioned to coeducation beginning in the 1960s. Ours is the only dataset we are aware of that covers this modern period of transition. See

Goldin and Katz (2011) for a similar dataset covering an earlier time period.

We found 154 such institutions, 118 for which we could observe students' field choices at least 4 years before and a decade after coeducation. We then excluded 41 institutions that did not fit into our target population of institutions that offered an arts-and-sciences curriculum and experienced sharp transitions to coeducation.[6] This yielded a main sample of 77 "treated" institutions.

Figure 1 documents the distribution of the years of transition to coeducation at women's colleges in our sample. The modal transition dates are 1969-1971, before the passage of Title IX in 1972. Title IX is unlikely to have meaningfully affected the women's colleges in our sample—the majority of which were private and received little federal funding.[7]

2.1 IPEDS sample on women's fields of degree

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 (National Center for Education Statistics, 2004, 2019). Both sources cover close to the full universe of baccalaureate-degree-granting institutions in the United States, and all of those legally classified as institutes of higher education. For the remainder of the paper we will refer to the HEGIS and IPEDS data as IPEDS for concision.[8]

[6]The excluded schools include 23 that operated coordinate programs or merged with men's colleges, 2 that did not the legal definition of "institutions of higher learning" under Title IV of the Higher Education Act of 1965, 6 that closed shortly after the transition, and 10 that did not appear to offer STEM programs. Our main results are robust to including these schools: the estimated effect of coeducation on women's STEM field choices is similar but slightly attenuated, consistent with the introduction of measurement error into the sample. See Appendix B.2 and Figure A1 for further detail.

[7]Until 1987, Title IX only applied to the particular program receiving federal money at private colleges, rather than *all* programs (Rim, 2020). Aside from student aid grants, colleges in our sample likely received little federal money: direct federal appropriations were unlikely, and the median expenditure on research reported in IPEDS per year was $0.

[8]The data from the HEGIS 1970 issue on degrees awarded has historically not been available in digital form. We digitized this issue for our analysis.

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 the effects of coeducation on the full distribution of degrees awarded across these eleven fields (see Section 4.2).

Figure 2 graphs sex-specific distributions of field choices. For treated schools, we report the distribution of majors for women in the five years prior to coeducation, and for men in the 10 years after coeducation. For comparison, we show the distributions for women and men at always-coeducational schools, weighted to represent the same academic years as those represented among treated schools. Women at treated schools (dark blue) were slightly more predisposed toward high-return fields like STEM and economics than women at coeducational schools (medium blue). However, we see the opposite relationship among men: relative to their counterparts at coeducational schools (gray), the first cohorts of men who entered former women's colleges (light blue) were less likely to choose heavily quantitative majors. Even so, a comparison of men and women at treated schools shows that the newly admitted men were more STEM-inclined than women. This suggests that the gender mix of the classroom in these fields changed substantially in response to coeducation.

Our final analysis tracks changes in $s^G_{f,it}$ over time for the 77 treated schools and 934 potential comparison schools, including 29 "always-women's" colleges that did not adopt coeducation during our sample period.

2.2 HERI data on freshman women's characteristics

To examine the effect of coeducation on the *composition* of newly-enrolled women, we require panel information on the underlying abilities and curricular preferences of women and their peers. We collect these data from the CIRP Freshman Survey, which is produced by the Higher Education Research Institute (HERI) at UCLA (Higher Education Research Institute, 2020). 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 coeducation dates, 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 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 HERI sample to institutions that appear at least once in the five years before and once in the 10 years after coeducation. 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 surrounding the transition to coeducation. Our pool of potential comparison schools consists of 127 that did not transition to coeducation during our sample period and that provided data in the same set of academic years as our treated schools.

Appendix Table A2 compares pre-treatment institutional characteristics, as measured in the IPEDS data, between the main IPEDS sample of switchers and the HERI subsample. The samples are very similar in terms of shares of women majoring in STEM, total enrollment, (very low) number of graduate degrees awarded. In addition, similar shares of each sample are under private control, are affiliated with the Catholic church, and are rated as selective. The table also reports characteristics of treated institutions that were excluded from the main sample. The excluded sample is similar to the main sample, although it contains a lower average share of women majoring in STEM, is less Catholic, and is less selective.

3 Empirical Strategy

Our goal is to estimate the average treatment effect of adopting coeducation on the distribution of fields studied by graduating women at historical women's colleges. We are also interested in the mechanisms underlying any effects on degrees earned. Our difference-in-difference research design compares outcomes before and after coeducation (first difference) to contemporaneous trends at schools that did not alter the gender mix of the campus environment (second difference). Under a parallel trends assumption that we discuss in the next subsection, our estimates can be interpreted as the average treatment effect of coeducation on treated schools (ATT).

Our empirical strategy relies on a version of the difference-in-difference estimator proposed by Callaway and Sant'Anna (2021).[9] Estimation proceeds in two steps.

First, for each school $j \in \mathcal{J}$ that transitioned to coeducation in year t_j^*, we construct event-study estimates of the effect of coeducation on outcome y_{jt} (e.g., the share of women graduating in STEM at school j in year t). From a set of untreated schools, we select a group of comparison schools $\mathcal{C}_j$ that match school j on a vector of baseline covariates. Then, we use observed outcome

[9]We discuss the relationship between our empirical approach and the original Callaway and Sant'Anna (2021) estimator in further detail in Appendix C. A growing body of literature demonstrates the difficulty of interpreting estimates from conventional two-way fixed-effects regressions as ATTs in settings such as ours, where there is variation in treatment timing and where treatment effects are likely to be heterogeneous over time and across treated units (Borusyak, Jaravel and Spiess, 2021; Callaway and Sant'Anna, 2021; Goodman-Bacon, 2021; Sun and Abraham, 2021). Nevertheless, we also report estimates from a two-way fixed-effects regression in Figure A1.

trends at school j and comparison schools $k \in \mathcal{C}_j$, relative to a chosen base period b, to construct the following:

$$\hat{\alpha}_{jt} = \underbrace{(y_{jt} - y_{jb})}_{\substack{\text{Trend at}\\ \text{treated school}}} - \underbrace{\sum_{k \in \mathcal{C}_j} \tilde{\omega}_k \cdot (y_{kt} - y_{kb})}_{\text{Counterfactual trend}}. \tag{2}$$

where $\tilde{\omega}_k$ are non-negative weights that sum to 1. We define the base period b as the five years prior to t_j^*, and y_{kb} as the average outcome observed at school k during this period.[10] We define the weights as the re-scaled number of degrees awarded to women in the school's first year in our sample: $\tilde{\omega}_k = \omega_k / \sum_{l \in \mathcal{C}_j} \omega_l$.

In the second step, we average our school-specific $\hat{\alpha}_{jt}$ estimates across all treated schools by time relative to the year of treatment. That is, if t_j^* is the year that school j adopted coeducation, we construct ATT estimates at different event times $\tau = t - t_j^*$. This yields a flexible estimate of the dynamic response of our outcomes of interest to the adoption of a coeducational campus environment. Specifically, our aggregate event-study parameters are constructed as:

$$\hat{\beta}_\tau = \frac{\sum\limits_{j \in \mathcal{J}} \sum\limits_{t \in \mathcal{T}} \hat{\alpha}_{jt} \omega_j \mathbb{1}\left\{t - t_j^* = \tau\right\}}{\sum\limits_{j \in \mathcal{J}} \sum\limits_{\tau \in \mathcal{T}} \omega_j \mathbb{1}\left\{t - t_j^* = \tau\right\}} \tag{3}$$

In addition, to provide a summary measure and increase statistical power, we construct estimates that pool event-time τ into five-year periods after the transition to coeducation. We refer to these estimates as "short-run" and "long-run" effects:

[10] For example, if school j admitted its first coeducational class in $t_j^* = 1981$, then $y_{kb} = \sum_{t=1976}^{1980} y_{kt}/5$.

$$\hat{\beta}_{SR} = \frac{\sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} \hat{\alpha}_{jt} \omega_j \mathbb{1}\left\{0 \leq t - t_j^* \leq 4\right\}}{\sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} \omega_j \mathbb{1}\left\{0 \leq t - t_j^* \leq 4\right\}} \tag{4}$$

$$\hat{\beta}_{LR} = \frac{\sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} \hat{\alpha}_{jt} \omega_j \mathbb{1}\left\{5 \leq t - t_j^* \leq 9\right\}}{\sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} \omega_j \mathbb{1}\left\{5 \leq t - t_j^* \leq 9\right\}} \tag{5}$$

For all estimated parameters, we conduct inference using a block bootstrap clustered at the school level with 1,000 replications. We report standard 95% pointwise confidence intervals in our event-study figures. In Appendix C, we also report event-study results with more conservative confidence intervals that adjust for multiple hypothesis testing across event-study periods (Callaway and Sant'Anna, 2021).

3.1 Construction of comparison groups

Our school-specific estimates $\hat{\alpha}_{jt}$ are unbiased under a standard parallel trends assumption: that the average outcome trend observed in comparison group $\mathcal{C}_j$ equals the trend we would have observed at school j in the absence of coeducation. In this subsection, we describe the process of choosing comparison groups for which this assumption is most plausible.[11]

Table 1 provides summary statistics that inform our choices of comparison groups. Column 1 presents average institutional characteristics for our main sample of treated schools, calculated in the relevant base period b for each school. In these pre-coeducation years, one-tenth of graduates at treated schools earned degrees in STEM, and the annual growth in STEM was close to zero. The typical treated school had a small enrollment of around $1,200$

[11]The identification assumption underlying our estimate of the ATT parameter β_τ is weaker than this: we can have non-parallel trends at individual schools j, so long as these average out to zero in the population of treated schools J.

and conferred very few graduate degrees. Nearly all treated schools were privately controlled, 64% were affiliated with the Catholic church, and 19% were selective.[12]

Columns 2-5 present statistics for two sets of possible comparison schools. For a given treated school j, our preferred comparison set of **untreated women's colleges** consists of women's colleges that were treated at least 10 years after school j, or never at all. A larger comparison set consists of **all untreated colleges**, i.e., all schools that were coeducational throughout our sample period. All statistics in these columns are re-weighted to represent the same academic years as the treated group.[13] The set of untreated women's colleges (Column 2) is similar in some respects to the sample of treated schools. However, Column 3, which provides formal tests of differences in means, shows that the treated sample has fewer STEM majors, is more likely to be Catholic-affiliated, and is less likely to be private or selective than the comparison set. The differences between the treated sample and the set of all untreated colleges are more stark (Columns 4-5): the untreated schools are much larger on average, award many more graduate degrees, and are less likely to be private or Catholic. In addition, women at these institutions were significantly less likely to major in STEM.

In columns 6-9 of Table 1, we repeat the exercise using the subset of candidate comparison schools that match our treated institutions on selected covariates. For the untreated women's colleges set, our matching variables are indicators of Catholic affiliation and selectivity—the main dimensions of imbalance shown in Column 3. By construction, the treated and resultant comparison groups are now identical on these characteristics (Columns 6-7). However, the two groups are now also virtually identical on pre-treatment growth in STEM and the share of women choosing STEM. The between-group differences in the

[12] "Selective" refers to a 1972 Barron's rating of 1-3. Some institutions were not rated by Barron's—we code these institutions as non-selective.

[13] We construct these statistics by first matching each treated school j with all comparison-group candidates, e.g., all women's colleges that transitioned at least ten years after j or never at all. We then limit the years of consideration to the relevant base period b, stack comparison groups for all treated schools j, and compute sample means of the resultant stacked sample.

other characteristics remain small in magnitude.

For the comparison group of all untreated colleges, our matching variables are quintiles of school size (i.e., total degrees awarded) and quartiles of base period growth in STEM major choice, in addition to Catholic and selective indicators. Columns 8 and 9 show that this procedure also largely eliminates observable differences between the treated group and the resultant comparison group. The comparison group is still slightly larger and, by construction, contains a higher baseline proportion of male students, but is otherwise very similar to the treated group.

For robustness, we construct two additional matched comparison groups: a **never-treated women's colleges** group that includes only the 29 schools that never adopted coeducation during our sample period, and a group that includes all coeducational or women-only institutions specified as **liberal arts colleges** by the 1987 Carnegie Classification system. Appendix Table A1 provides the details.[14]

It is important to note that pre-treatment balance on observed characteristics is not necessary for the validity of our research design. However, the achievement of balance or near-balance on a variety of characteristics may inspire confidence in the key identification assumption that treated and comparison groups experience similar *trends* in unobserved determinants of major choice. We test this assumption in several ways. First, our event-study approach enables us to estimate β_τ for years prior to coeducation ($\tau < 0$): a series of coefficients that depart significantly from 0 would be suggestive of a violation of parallel trends. We assess the robustness of pre-trend estimates across our 4 comparison groups as well as 6 other specifications presented in Appendix Figure A1. Second, in Appendix D, we estimate a synthetic control specification and verify that the results match our event study specifications. Third, in the next subsection, we test for confounding effects arising from cor-

[14] The vector of matching variables for never-treated women's colleges includes indicators of Catholic affiliation and selectivity. The liberal arts college matching vector additionally includes bins of base-period trends in STEM major choice. Just as in Columns 7 and 9 of Table 1, Columns 7 and 9 of Table A1 show that we achieve balance or near-balance on all relevant baseline variables.

relation between the timing of coeducation and labor market shocks that alter the returns to STEM.

3.2 Tests for confounding labor market shocks

We use the March CPS to construct four measures of labor market conditions at the state-year level (Ruggles et al., 2020), and assign these measures to colleges by state of location and year. Then, we use our event-study framework to test for changes in these measures around the timing of coeducation (Pei, Pischke and Schwandt, 2018). If the resultant estimates depart substantially from 0, it could suggest that our main results are driven not by coeducation, but rather by gender-neutral, external labor-market forces that have previously been shown to impact students' college major choices (Willis and Rosen, 1979; Altonji, Blom and Meghir, 2012; Altonji, Arcidiacono and Maurel, 2016).

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

4 Main Results

Figure 3 presents our main event-study estimates of the effect of coeducation on the share of women graduating in STEM. Consistent with the assumption of parallel trends, we see estimates very close to 0 with little evidence of a trend break in the years prior to the adoption of coeducation. Recall that our data measures baccalaureate degrees *awarded*, so the estimates for the first two years of coeducation primarily represent the field choices of women who were juniors or seniors when the first cohort of male freshmen arrived on campus. Most of these women would have chosen their majors and taken their core classes prior to the reform. In contrast, most of the women who graduated in the *third* year of coeducation (event time $\tau = 2$) would have spent the bulk of their college experience in a mixed-sex environment. Thus, the sharp drop beginning at event time 2 is exactly what we would expect if the arrival of men discouraged women from choosing STEM programs. This pattern is remarkably similar across all 4 of our comparison groups.

Table 2 presents estimates of the "long-run" effect on women's choice of major, constructed using equation 5. Our main specification (Panel A) finds that coeducation induced a 3.0 percentage-point decline in the share of women majoring in STEM. To provide context for this estimate, the second row records the predicted share of women majoring in STEM in the long run (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 (i.e., the second term in equation 2) onto the pre-treatment average share of women students majoring in STEM at treated 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. Our alternative specifications, presented in Panels B-D, show similar but slightly larger estimates. Our estimated effects on STEM are statistically significantly different from 0 in all specifications.

Our main result is robust to a wide variety of alternative methods of constructing the comparison group. Appendix Figure A1 presents several robustness checks. Using the preferred control group specification, Panel A shows that the treatment effect estimates change little when 1) treated schools that were originally excluded are returned to the sample; 2) the main treated sample is restricted to the HERI sub-sample; or 3) early switchers (1969-1971) are excluded from the main treated sample. (In the second case, the pre-trend estimates are slightly different from the main sample, although they display a slight *upward* pre-trend that works against our finding.) Panel B shows robustness of the estimates to: 4) the inclusion of Census division as an additional matching variable; 5) defining liberal arts colleges according to the Carnegie classification rather than our graduate-degree-based measure; 6) using a standard two-way fixed-effects regression estimator; or 7) using a synthetic control estimation approach (see Appendix D for detail).

4.1 Components of STEM

Figure 4 provides estimates of the effect of coeducation on the share of women earning degrees in a more specific set of quantitative fields. Panels A through C of report event study estimates for three main components of STEM: biology, math, and physical sciences. These account for more than 97 percent of all baseline STEM majors in our main treated sample. Panel D considers economics, which is not included in our definition of STEM but is nevertheless highly quantitative. All 4 sets of event studies display similar patterns to the overall STEM event study—a flat pre-trend that breaks downward in the third or fourth year of coeducation before stabilizing in the long run.

Columns 2 through 5 of Table 2 report our estimates of the long-run effect on each field. The preferred 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. The estimates are all statistically significant at the 10% level or lower. Panels B-D report estimates based on our alternative comparison groups, and once again show slightly larger and more statistically significant declines.

4.2 The distribution of field choices

What fields were chosen by women at former 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.[15] We then estimate the long-run effect on each field using equation 5.

Figure 5a presents the estimated long-run effects on each field, ranked in increasing order.[16] 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 statistically significant.

Figure 5b presents semi-elasticity responses that account for the relative sizes of each field by dividing each estimate in Figure 5a by the counterfactual proportion of women majoring in the field. After rescaling by field size in this way, the gender-normative 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. Overall, these results provide strong evidence that the coeducation of women's colleges sparked substantial changes in the degrees earned by women. In the next section, we turn to the mechanisms that may explain these results.

[15] "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. See Appendix E.1 for details on the construction of these groups.

[16] Appendix Figure A5 and Appendix Tables A4 and A5 present the event study and long-run estimates for fields where these estimates are not reported in the main text.

5 Mechanisms

As discussed in Section 1.2, if the switch to coeducation affected women's enrollment decisions, the results reported in section 4 could be attributed to a *composition effect* that altered the preferences or abilities of the female student body. In addition, these responses could be due to *environmental effects*, i.e., women's progress toward degrees in certain fields may have been shaped by the changing campus environment, conditional on having enrolled. We explore these possibilities below. Our analysis suggests that composition effects played a limited role, if any, in driving changes in women's educational choices. Our data and research design are less informative about the precise nature of the environmental effect, which could itself be classified into a number of distinct mechanisms. However, we argue that the evidence is most consistent with environmental mechanisms operating through greater exposure to male peers and faculty as the primary cause of the decline in the share of women majoring in quantitative fields.

5.1 Exposure to male peers and faculty

We begin by exploring the impact of coeducation on the gender mix of the campus environment. While the reform almost mechanically increases the presence of men, the size of this effect is not clear ex ante.

Panel A of Figure 6 reports event-study estimates of the effect of coeducation on the share of freshmen who were male. The male share rose steadily for several years before beginning to level off at one-fifth to one-quarter of entering students. The male share of *graduating* students rose somewhat more slowly: the long-run estimate $\hat{\beta}_{LR}$ of the male share 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 3 of Table A3).

Since these men were disproportionately likely to choose quantitative fields (recall Figure 2), these results suggest that women, particularly in STEM-related classrooms, experienced substantial inflows of male peers. Panel A of Figure A9 shows that the coeducation-driven change in the share of women majoring in a field was negatively correlated with the predicted share of men choosing that field.[17] This pattern holds within our "quantitative" fields of biology, math, physical science, and economics, as well as in general. This negative correlation is merely suggestive, but it is consistent with the theory that women's field-choice responses were partially driven by specific exposure to men in the classroom.

Next, we examine the gender composition of the faculty. In fall 1971, women made up 51% of faculty at women's colleges, compared to only 22% at coeducational institutions in our sample. 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 male share of faculty are presented in Panel B of Figure 6. We observe a small increase, with a statistically significant long-run estimate of 4.5 percentage points (12%, see Table A3). This effect is about one-fifth the size of the observed 21.1-percentage-point increase in the male share of freshmen and 21.6-percentage-point increase in the male share of undergraduates. The increase in male peers is also much larger relative to the base share (of nearly zero percent). Overall, these results suggest a greater potential for gendered peer effects, relative to role model effects, to influence women's field choices.

5.2 Other features of campus environment

We next explore the possibility that the adoption of coeducation was accompanied by changes in gender-neutral features of the college environment. For example, if the admission of men altered the distribution of academic ability among the study body, it may have affected the level of competition (e.g., for grades or other career-advancing opportunities like internships) in STEM-related classrooms, pushing women into other majors (Fischer, 2017).

We study the impact on academic competition by using the HERI Fresh-

[17] The predicted share of men earning degrees in field f is calculated by multiplying the share of men majoring in f (the light blue bar in Figure 2) by the share of degrees granted to men, then dividing by the predicted share of *all* students receiving degrees in field f.

man Survey data to rank *all* entering freshmen in each school and year by their high school grade point averages (GPAs). (GPA is the most consistently observed measure of student ability in the data.) We then estimate the effect of coeducation on *female* students' average rank within their class. As shown in Panel C of Figure 6, we document a statistically significant 4 percentile *increase* in the average woman's ability position. This result suggests that, to the extent that gender-neutral academic competition is important in determining field choices, coeducation would have driven *more* women into competitive majors.[18]

Another way coeducation could have increased gender-neutral pressures is by leading classrooms to become 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 6 suggest 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.

We also consider whether 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 did 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. We construct a proxy for course catalogs by counting the number of individual majors in which a degree was awarded at school j in year t, using

[18] We might still be concerned that these results for the average woman mask a pattern in which the highest-ability men sort into quantitative fields. We therefore repeat the exercise but limit the sample to freshmen who declared an intention to major in STEM. This restriction greatly restricts our sample size and statistical power, but we find no evidence that women saw more competition on academic ability in STEM classes: Our estimates are all statistically indistinguishable from 0 and range from a very small negative effect of -1 percentile to a substantially positive effect of 6 percentiles (Table A6).

the most disaggregated field definitions possible in the IPEDS data.[19] The resulting estimates in Panel E of Figure 6 are small and statistically insignificant, suggesting little change in the menu of degree options available to students. The estimates in Panel F provide a related test of changes in instructional expenditures per student. Once again, we find no evidence of substantial changes in this educational input.

Finally, in Appendix section F, 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.

5.3 Student preferences and characteristics

Another important hypothesis is that coeducation induced a *composition* effect, i.e., that the women enrolling in the newly coeducational environment were *ex ante* less likely to choose STEM fields.

It is important to note that students' field and career choices are far from determined when they enroll: their decisions are shaped substantially by the campus environment and other factors over their collegiate careers (Zafar, 2011; Stange, 2012; Gong, Stinebricker and Stinebrickner, 2019; Owen, 2020; Patnaik, Wiswall and Zafar, 2020). Even among students who formally declare a STEM major, nearly half end up switching to another field or dropping out (Altonji, Arcidiacono and Maurel, 2016). The literature does, however, estab-

[19] Using this proxy, the average number of degrees offered among our switching colleges is 25.

lish moderate relationships between eventual field choices and baseline factors such as preferences for coursework and careers (e.g., Zafar, 2013), expectations about family (e.g., Wiswall and Zafar, 2021), the influence of parents and other family members (Patnaik, Wiswall and Zafar, 2020), and academic ability (e.g., Zafar, 2011; Stinebrickner and Stinebrickner, 2014).

We use the HERI data to estimate the effects of coeducation on several of these factors. We begin with our most direct and powerful predictor: freshman women's field preferences. The first four panels of Figure 7 show event-study estimates of the effect of coeducation on the share of newly-enrolled freshman women *intending* to major in STEM and its largest components: biology, physical sciences, and math.[20] 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 4 and A1.[21] Table 3 summarizes by presenting estimated long-run effects for each major, using each of our three comparison groups. Across all outcomes and comparison groups, our estimates are small and never statistically distinguishable from 0.

The final two panels of Figure 7 present event-study estimates for career and family ambitions. We find no evidence of a decline in the share of freshman women who report an interest in a career in science (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.

Next, we consider changes in other characteristics that may predict students' propensity to choose quantitative majors, such as academic ability and

[20] 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.

[21] 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.

parental background. Panels A and B of Figure 8 test for changes in the HERI data's best measure of academic ability, high school GPA. Following the switch to coeducation, there was no decrease in the share of women earning at least an A, or at least an A-. (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.

Panels C through F of Figure 8 show small and statistically insignificant changes in various other characteristics that we might expect to be correlated with academic choices. In panel C, we see no change in the racial composition of women entering newly coeducational colleges. We also see little evidence of a change in parents' educational background. Entering women were no more likely to be first-generation college attendees (panel D), and no more likely to have parents with college degrees (panels E and F).

5.4 How large could the composition effect be?

As discussed in the previous section, we find no evidence that coeducation systematically altered a variety of characteristics related to students' propensities to choose quantitative majors. An alternative approach would combine a large set of freshman characteristics to construct a *predicted* share who would be expected to major in STEM, and then study the evolution of this measure after the transition to coeducation. We turn to this approach as another test of composition effects and to calculate a reasonable upper bound on the extent to which composition effects could explain our main results.

This approach requires two pieces of information. First, we need to know the effects of coeducation on freshman characteristics. We can estimate these effects using our HERI Freshman Survey data. Second, we need to know the marginal effects of these characteristics on the probability of graduating with a STEM degree. We estimate these effects using a sample of women from the National Longitudinal Study of 1972 (NLS72), which follows students from high school into adulthood between 1972 and 1986 (National Center for Education

Statistics, 1999).[22] The value of these data is that they provide individual measures of pre-college characteristics linked to field of degree earned. In addition, the sample was drawn at a time relatively close to the date of transition to coeducation for the bulk of our sample.

Using our NLS72 sample, we estimate $Y_i = X_i\delta$, where Y_i is an indicator for earning a degree in a STEM field and X_i is a vector of characteristics measured in student i's freshman year or earlier. We then interact the vector of coefficients, $\hat{\delta}$, with the characteristics of students in our HERI Freshman Survey data to predict the share of women in each year and school who will earn a major in STEM. Finally, we use this predicted share as the outcome in our main specification to estimate the long-run effect of coeducation on the share of entering freshman women who would be expected *ex ante* to major in a STEM field.

Panel A of Table 4 presents our estimates of $\hat{\delta}$ from the NLS72. By far the most predictive baseline characteristic of eventually majoring in STEM is intending to major in STEM. This characteristic alone explains about 19% of the variation in attainment of a STEM degree. Our estimates in column 1 suggest that women who intend to major in STEM are nearly 34 percentage points more likely to do so (relative to the overall mean of 8%, or 3% among women who did not plan to major in STEM). In columns 2-4, we add other measures that may play a role in eventual choice of field: career and family aspirations, parents' occupation and education, and high school grades and coursework. The coefficient on STEM intentions changes little, and the R-squared barely budges, rising to only 0.215 in our most detailed specification. These results suggest STEM intentions are a relatively powerful predictor of eventual major, and are also in line with the literature suggesting that students' experiences during college play a large role in choice of major (Arcidiacono, 2004; Zafar, 2011; Stinebrickner and Stinebrickner, 2014; Gong, Stinebricker and Stinebrickner, 2019; Astorne-Figari and Speer, 2019; Patnaik, Wiswall and Zafar, 2020).

In the lower panels of Table 4, we evaluate the effect of coeducation on

[22]See Appendix G for more details on the NLS72 and construction of our sample.

the predicted share of entering freshman women who will major in STEM, which is calculated using intent to major in STEM and the characteristics specified as covariates in panel A. Panel B reports estimates using our preferred comparison group. Using only intended major, we find a slight *increase* in the predicted STEM share of one-half of one percentage point. Recall that our linked HERI-IPEDS subsample suggested a 3.4 percentage-point *decrease* in STEM degree attainment; this implies that composition changes explain $0.5/-3.4$ or -16% of the overall effect on coeducation.[23] A more conservative approach to quantifying the composition effect would be to consider the lower bound on the 95% confidence interval for the effect on the predicted share in STEM. In this case, the composition effect can explain a decrease of at most 1.1 percentage points, or 32% of the overall decrease in STEM.

In panel C of Table 4, we conduct the same exercise using estimates from our comparison group comprised of all colleges that did not adopt coeducation during our sample period. In this case, we bound the composition effect at 16%-29% of the overall effect. We report the same exercise for our other two comparison groups in Appendix Table A7 and find similar results.

In sum, our main point estimates of the composition effect are close to zero or are even *positive*, which could occur if the arrival of men on campus caused a shift toward more STEM-inclined women who were undeterred from mixed-gender competition. A more conservative interpretation of our results, however, is that we can rule out moderately negative composition effects. Overall, the evidence suggests that the decline in women majoring in STEM at former women's colleges can be attributed primarily to changes in the campus environment – the most important of which was the increase in male peers.

[23] Note that this calculation assumes no correlation between the effects on women's STEM intentions and STEM degree receipt. Using a bootstrap on our linked HERI-IPEDS subsample, we estimate a correlation of -0.02 between the two statistics. This correlation near 0 is exactly what we would expect in the absence of composition effects.

6 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 mixed-gender 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 produced relatively little evidence on how major choices actually respond to changes in non-pecuniary features of the environment.

Drawing on a newly assembled historical dataset, we report event-study estimates 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 3.0-3.5 percentage points (30-33%) relative to control colleges. We also estimate negative effects of coeducation on women's likelihoods of earning degrees in economics and business.

While our data and research design cannot pinpoint the precise mechanisms behind this shift, we are able to rule out some important channels. We find little evidence that coeducation altered the STEM preferences or preparedness of matriculating women. Our bounding exercises suggest that even under the most conservative assumptions, composition changes can explain less than one-third of the overall observed effect. Examining "gender-neutral" features of the campus environment, we find no evidence that coeducation 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. However, we show that coeducation dramatically altered the gender mix of students and – to a lesser extent – faculty. We therefore conclude that the bulk of the evidence suggests that women's exit from STEM fields was driven by greater exposure to male peers and a small decrease in opportunities for women to interact with same-gender faculty role models.

What do our estimates imply about the overall gender gap in STEM observed in recent years? 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. We can construct an estimate of the contribution of coeducation to this gap with a simple extrapolation of our main results. Note that in 2016, 57 percent of total degrees were awarded to women, so the average woman's potential peer group was 43 percent male. As discussed in Section 5.1, 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. This 43/21.6 scaling factor assumes that our main estimate of the effect of coeducation on women's STEM majoring is driven exclusively by greater exposure to male peers, but is consistent with our finding that coeducation caused negligible changes to student composition or other environmental features. [24]

Of course, this counterfactual exercise must be qualified in several respects. Our main samples of treated and comparison 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 coeducational environments. 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 coeducational environments.

[24] A related interpretation would assume that male peer share is a sufficient statistic for exposure to coeducational environments, and thus that other environmental factors adjust in proportion to the change in the male share of peers. For example, we found that coeducation caused a 4.5 percentage-point increase in the male share of faculty along with the 21.6 percentage-point increase in male peers.

In either case, our findings suggest that the way gender is understood and performed[25] within coeducational college environments contributes meaningfully to gender disparities in field choice. Our results have implications for ongoing initiatives that infuse features of single-sex learning environments into coeducational settings, such as coding academies for girls, mentoring for freshman women by upperclasswomen, and the creation of academic clubs and conferences for women in STEM. For instance, to increase the proportion of women in computer engineering, Harvey Mudd College has provided first-year students opportunities to attend the Grace Hopper Celebration of Women in Computing conference (Corbett and Hill, 2015). While a large-scale return to single-sex education is unlikely, policies such as these may be effective in closing the persistent gender gap in college 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.

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

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.

[25] See Butler (1990).

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. 2022. "Nevertheless She Persisted? Gender Peer Effects in Doctoral STEM Programs." *Journal of Labor Economics*, 40(2): 397–436.

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.

Brenøe, Anne Ardila, and Ulf Zölitz. 2020. "Exposure to more female peers widens the gender gap in STEM participation." *Journal of Labor Economics*, 38(4): 1009–1054.

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.

Butler, Judith. 1990. *Gender Trouble: Feminism and the Subversion of Identity.* New York, NY:Routledge.

Calkins, Avery. 2021. "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.

Corbett, Christianne, and Catherine Hill. 2015. *Solving the Equation: The Variables for Women's Success in Engineering and Computing.* Washington, DC:AAUW.

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.

Fischer, Stefanie. 2017. "The downside of good peers: How classroom composition differentially affects men's and women's STEM persistence." *Labour Economics*, 46: 211–226.

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. 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.

Higher Education Research Institute. 2020. "CIRP Freshman Survey." https://heri.ucla.edu/cirp-freshman-survey/.

Hill, Andrew J. 2015. "The Girl Next Door: The Effect of Opposite Gender Friends on High School Achievement." *American Economic Journal: Applied Economics*, 7(3): 147–77.

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.

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.

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.

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.

National Center for Education Statistics. 1999. "National Longitudinal Study of the Class of 1972." Inter-University Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/ICPSR08085.v1.

National Center for Education Statistics. 2004. *Higher Education General Information Survey 1965-1980.* Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

National Center for Education Statistics. 2019. "Integrated Postsecondary Education Data System 1980-2016." https://nces.ed.gov/ipeds/use-the-data.

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.

Stinebrickner, Ralph, and Todd R. Stinebrickner. 2014. "A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout." *The Review of Economic Studies*, 81(1 (286)): 426–472.

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.

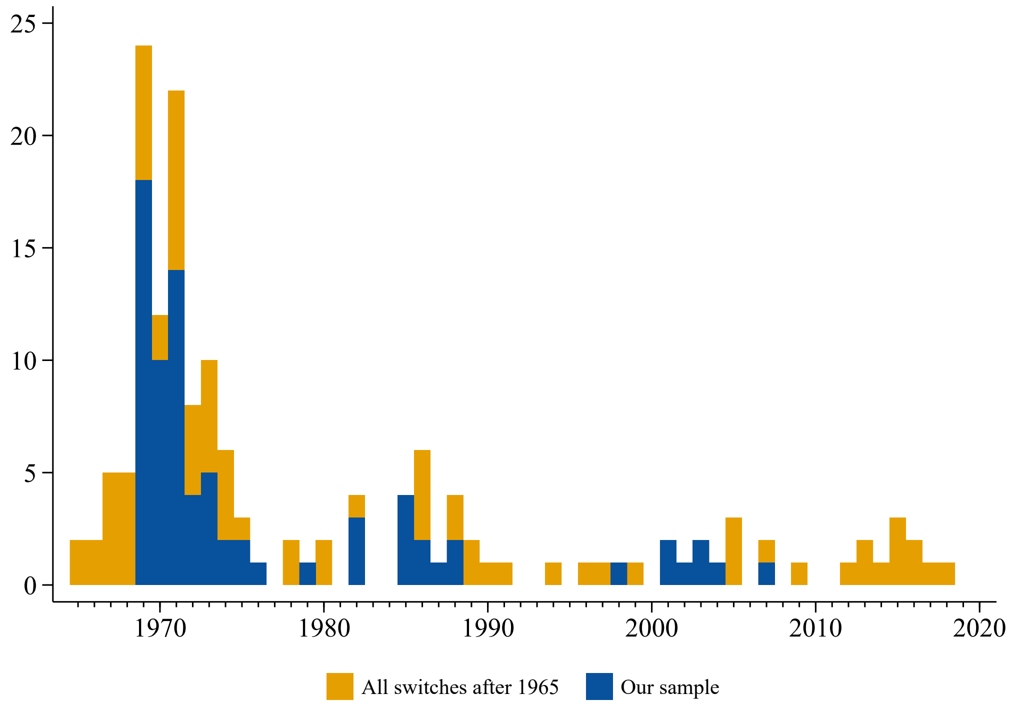
Wiswall, Matthew, and Basit Zafar. 2021. "Human Capital Investments and Expectations about Career and Family." *Journal of Political Economy*, 129(5): 1361–1424.

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.

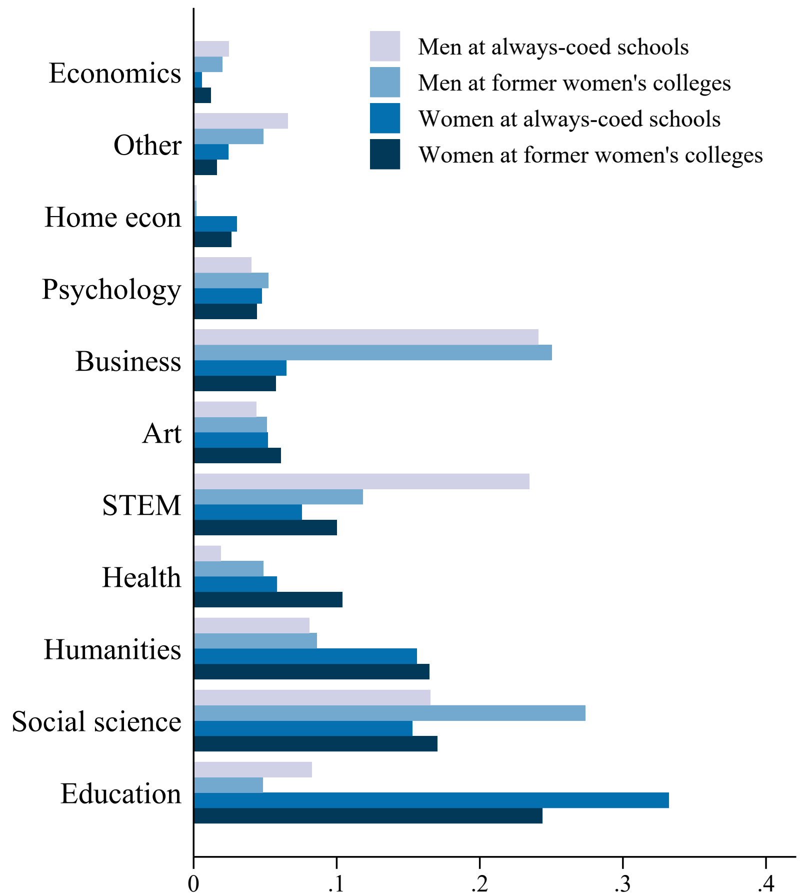
Zölitz, Ulf, and Jan Feld. 2018. "The effect of peer gender on major choice." University of Zurich Department of Economics Working Paper 270.

Figure 1: Distribution of the year of switch to coeducation



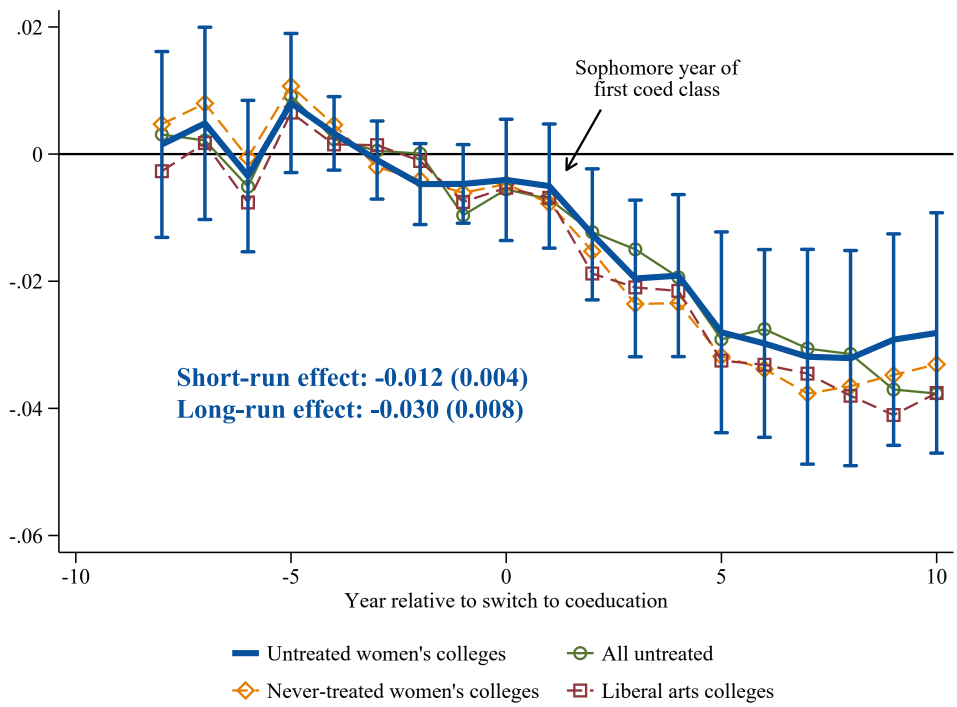
Source: Hand-collected data on the years that former women-only institutions switched to coeducation. Our analysis sample omits schools that closed shortly after switching to coeducation, operated coordinate programs with nearby men's colleges, did not offer STEM fields, or that we do not observe for 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



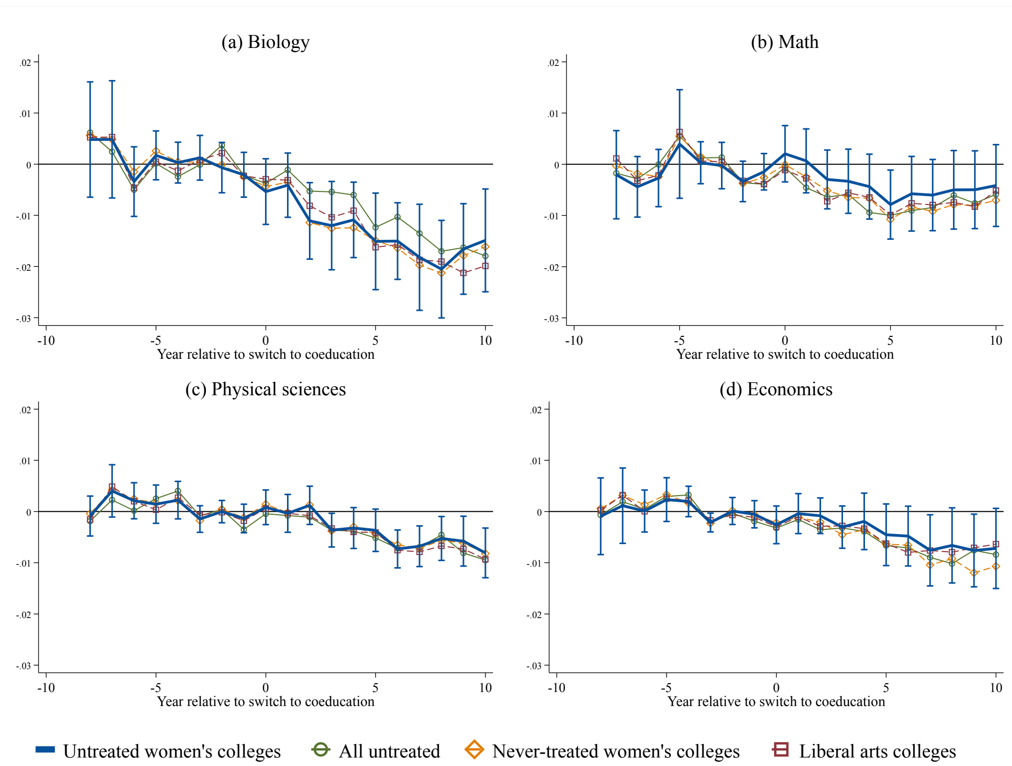
Notes: Data drawn from the IPEDS surveys, spanning 1966-2016, linked to hand-collected data on the dates of transition to coeducation by institution. Each bar shows the fraction of graduates of a given sex and college type earning degrees in the corresponding field. The distribution of majors among women at former women's colleges is calculated in the five years preceding coeducation. 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 distributions of majors among women and men at "always-coed" schools is re-weighted to match the academic years represented by the figures for their same-sex counterparts at former women's colleges.

Figure 3: Effect of coeducation on the share of women graduating 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 an estimate of β_k from equation 3, 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.

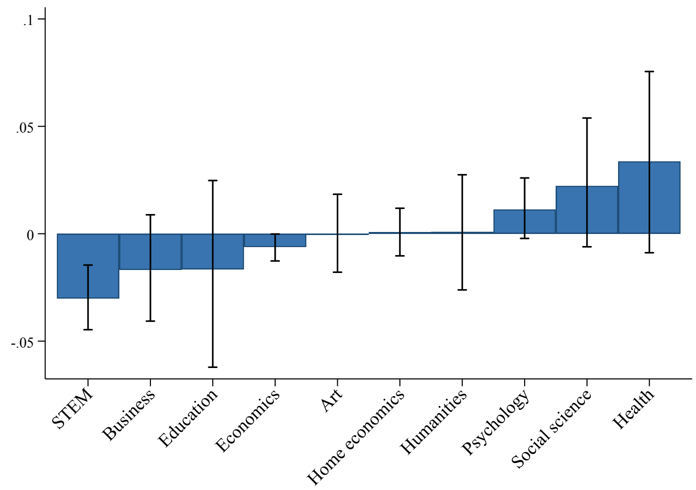
Figure 4: Effect of coeducation on women's choices of quantitative majors



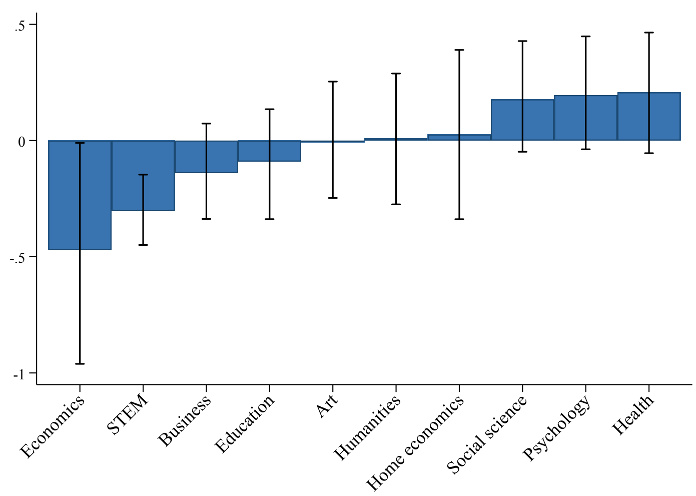
Notes: Data drawn from IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. Panels display estimate of β_l from equation 3. Dependent variable is the share of degrees earned in STEM among all degrees earned by women in the academic year. Error bars show 95% confidence intervals constructed using a block bootstrap clustered at the institution level.

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

(a) Percentage-point changes by major

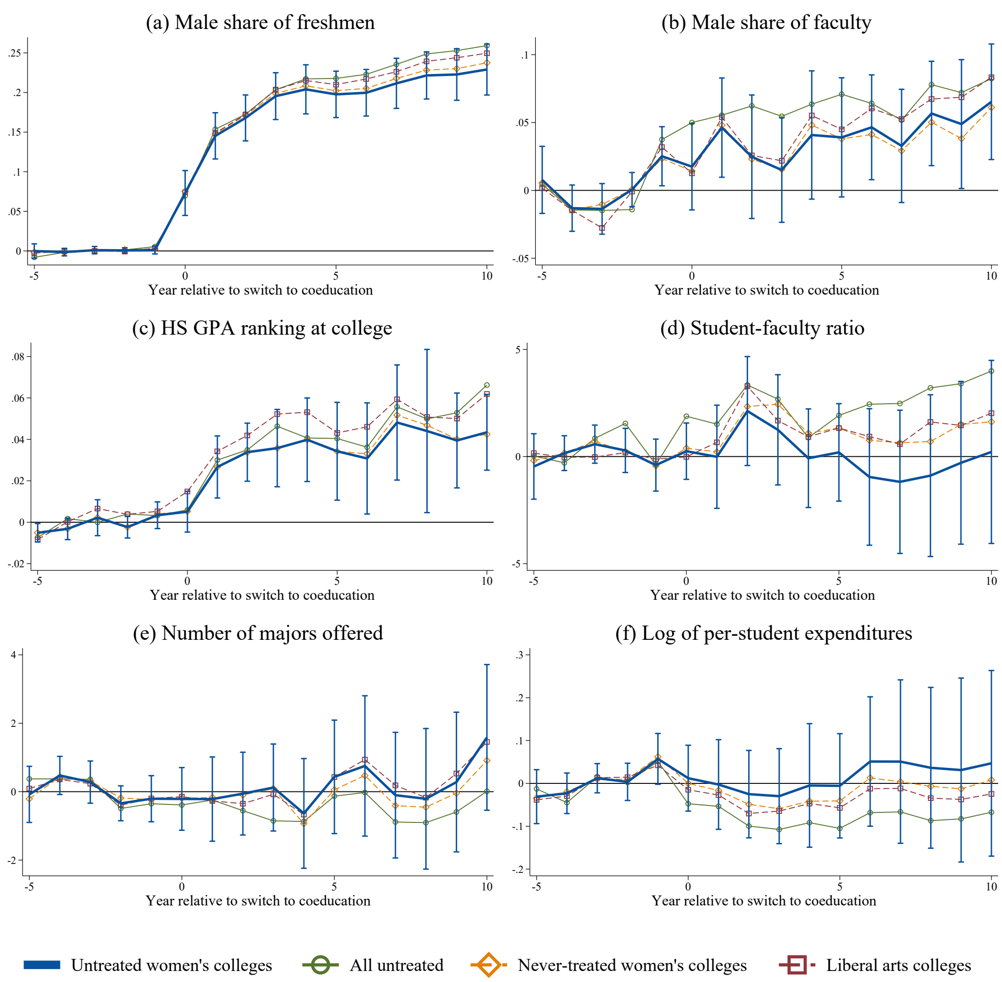


(b) Percent changes by major



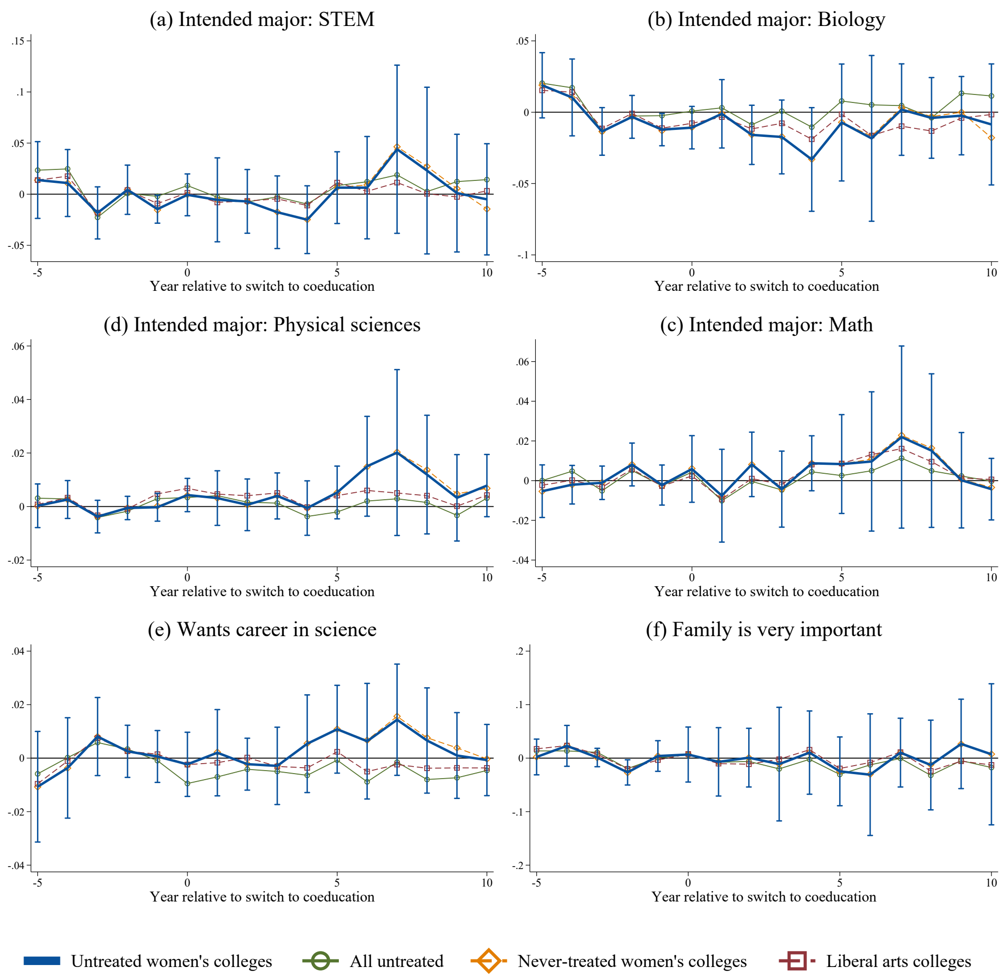
Notes: Data drawn from IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. The majors included in each field are described in Appendix E.1.2. Each panel displays the long-run effect of coeducation on the share of women majoring in each field using equation 5. In panel B, the estimates are scaled by the counterfactual mean. 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, and match treated school on selectivity and Catholic affiliation. Long-run effects are calculated as impacts in the second five years after the transition to coeducation. 95% confidence intervals are constructed using a block bootstrap with 1,000 replications. For scale, the "other" major category is omitted; estimates for that group are -0.001 [-0.034,0.034] and 0.037 [-0.88,0.87], respectively.

Figure 6: Effect of coeducation on features of the campus environment



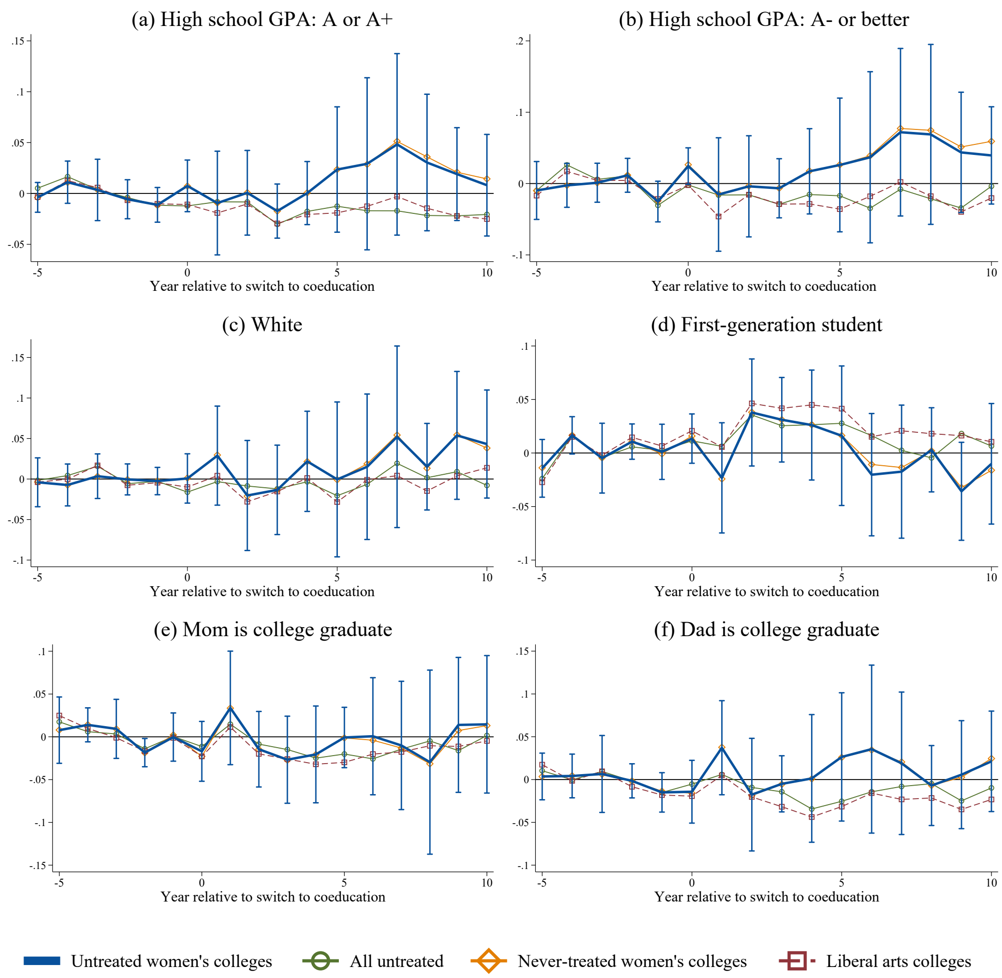
Notes: Each point shows and estimate of β_τ from equation 3, using the comparison group specified in the legend. Error bars show 95% confidence interval constructed using block bootstrap clustered at the institution level. Data in panel C drawn from HERI, spanning 1966-2006, linked to hand-collected dates of transitions to coeducation by institution. Data in all other panels drawn from HEGIS and IPEDS files 1966-2016. See Section 2 for further detail.

Figure 7: Effect of coeducation on preferences of entering freshman women



Notes: Each point shows and estimate of β_k from equation 3, 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 8: Effect of coeducation on the composition of entering freshman women



Notes: Each point shows and estimate of β_τ from equation 3, using the comparison group specified in the legend. Students are coded as first-generation if neither parent attended college. 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.

Table 1: Summary statistics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Treated	Candidate comparison groups				Matched comparison groups			
		Untreated wom. clg.	Difference (1)-(2)	All untreated	Difference (1)-(4)	Untreated wom. clg.	Difference (1)-(6)	All untreated	Difference (1)-(8)
STEM share of women's degrees	0.10	0.12	-0.02**	0.08	0.02***	0.10	-0.00	0.09	0.01
	(0.05)	(0.05)	[0.04]	(0.05)	[0.00]	(0.05)	[0.98]	(0.06)	[0.14]
Annual growth rate, STEM	-0.005	-0.003	-0.002	-0.003	-0.003**	-0.004	-0.001	-0.006	0.000
	(0.011)	(0.010)	[0.15]	(0.015)	[0.03]	(0.010)	[0.34]	(0.013)	[1.00]
Total enrollment	1226	1162	64	6020	-4794***	1149	77	1537	-304**
	(917)	(732)	[0.64]	(7654)	[0.00]	(675)	[0.56]	(1283)	[0.05]
Female share of all degrees	0.97	1.00	-0.02***	0.47	0.50***	0.99	-0.02**	0.59	0.38***
	(0.07)	(0.02)	[0.01]	(0.13)	[0.00]	(0.03)	[0.04]	(0.24)	[0.00]
Graduate degrees awarded	27	15	12*	233	-206***	15	12*	38	-11
	(51)	(43)	[0.07]	(513)	[0.00]	(43)	[0.06]	(88)	[0.34]
Private college	0.92	0.99	-0.07**	0.56	0.35***	0.99	-0.07**	0.92	0.00
	(0.27)	(0.11)	[0.02]	(0.50)	[0.00]	(0.10)	[0.02]	(0.28)	[1.00]
Ever Catholic-affiliated	0.64	0.43	0.21***	0.07	0.57***	0.64	0.00	0.63	0.00
	(0.48)	(0.49)	[0.01]	(0.25)	[0.00]	(0.48)	[1.00]	(0.48)	[1.00]
Selective admission	0.19	0.45	-0.26***	0.15	0.04	0.19	0.00	0.18	0.00
	(0.39)	(0.50)	[0.00]	(0.35)	[0.41]	(0.39)	[1.00]	(0.38)	[1.00]
Observations	77	57	134	934	1,011	54	131	496	571

Notes: Column 1 shows means (standard deviations) for our main treated sample of women's colleges, averaged over the five years prior to each treated school j's transition to coeducation. Comparison group means in column 2 are constructed by matching each treated school j with all women's colleges that never transitioned, or transitioned at least ten years after j. We then stack these comparison groups across all treated schools j and compute sample means for the resultant "grand" group. Column 3 restricts to comparison schools that match school j on selectivity and Catholic affiliation. Columns 4-5 show estimated differences in means [p-values] between treated and comparison samples. STEM share of women's degrees is weighted by total degrees awarded to women, female share of degrees is weighted by total degrees awarded, and all other means are unweighted. Columns 6-9 show analogous statistics for the full comparison group of colleges. Column 7 matches on enrollment and annual growth rate of STEM, in addition to selectivity and Catholic affiliation. Data are drawn from IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. See Section 2.1 for further information on sample construction and characteristics.

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	-0.030***	-0.017***	-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	5,505	5,505	5,505	5,505	5,505
Panel B: All-college comparison group					
Long-run effect	-0.031***	-0.014***	-0.0064***	-0.0082***	-0.0081**
	(0.0067)	(0.0045)	(0.0022)	(0.0031)	(0.0033)
Counterfactual mean	0.103	0.057	0.017	0.022	0.014
Observations	27,618	27,618	27,618	27,618	27,618
Panel C: Never-treated comparison group					
Long-run effect	-0.035***	-0.018***	-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
Panel D: Liberal arts college comparison group					
Long-run effect	-0.035***	-0.018***	-0.0067***	-0.0082***	-0.0074***
	(0.0059)	(0.0039)	(0.0018)	(0.0028)	(0.0028)
Counterfactual mean	0.105	0.062	0.017	0.021	0.014
Observations	20,954	20,954	20,954	20,954	20,954

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 5. 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	0.016	-0.0064	0.011	0.011
	(0.026)	(0.015)	(0.015)	(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	0.011	0.0056	0.0052	0.0002
	(0.013)	(0.013)	(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	0.018	-0.0052	0.012	0.012
	(0.027)	(0.016)	(0.016)	(0.0097)
Counterfactual mean	0.092	0.081	0.005	0.006
Observations	1,400	1,400	1,400	1,400
Panel D: Liberal arts college comparison group				
Long-run effect	0.0052	-0.0087	0.0099	0.0040
	(0.013)	(0.013)	(0.0086)	(0.0042)
Counterfactual mean	0.088	0.076	0.003	0.008
Observations	9,083	9,083	9,083	9,083

Notes: Table displays estimates long-run effect effect of women's colleges' switch to coeducation on the degree intentions of entering freshman women. Effects are derived from the doubly-robust estimator of Callaway and Sant'Anna (2021) using equation 12 and aggregating cohort-level effects using equation 5. 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.

Table 4: Bounding the composition effect of coeducation on STEM degree receipt

	(1)	(2)	(3)	(4)
Panel A: Effect of freshman characteristics on women's likelihood of earning STEM degree				
Effect of intent to major in STEM	0.336***	0.333***	0.332***	0.317***
	(0.040)	(0.040)	(0.040)	(0.041)
Covariates:				
Career, family aspirations		X	X	X
Parental education, occupation			X	X
High school grades, coursework				X
R-squared	0.191	0.199	0.205	0.215
Observations	1,235	1,235	1,235	1,235
Panel B: Effect of coeducation on predicted share of female freshmen who will major in STEM, preferred comparison group				
Estimated composition effect	0.005	0.009	0.009	0.012
	(0.008)	(0.010)	(0.010)	(0.011)
Composition effect / Total effect of coeducation on STEM major choice	-16%	-28%	-27%	-37%
Composition effect upper bound	32%	32%	31%	29%
Panel C: Effect of coeducation on predicted share of female freshmen who will major in STEM, all-college comparison group				
Estimated composition effect	0.004	0.001	0.002	0.001
	(0.005)	(0.006)	(0.006)	(0.006)
Composition effect / Total effect of coeducation on STEM major choice	-11%	-4%	-5%	-4%
Composition effect upper bound	16%	29%	28%	29%

Notes: Panel A reports regression estimates of the effect of intention to major in STEM as of freshman year on share of students earning STEM degree, derived from sample of women in National Longitudinal Study of 1972. Panels B and C report implied long-run effect on the predicted share of freshman women at newly coeducational colleges who will major in STEM, estimated using equation 5 and sample of women from the HERI Freshman Survey. Predicted share in STEM is constructed by interacted coefficients from the regressions in panel A with characteristics of entering freshman women in the HERI data. Share of total effect explained by composition is constructed by dividing predicted STEM effect by estimated effect of coeducation on the share of women earning STEM degree from our linked IPEDS-HERI data (-0.034, see Figure A1a). Upper bound on composition effect is constructed by dividing lower bound of 95% confidence interval of predicted STEM effect by -0.034. See Appendix Table A7 for estimates drawing on alternative comparison groups.

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 F 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:

$$V_{wt}(\eta^*_{wt}) = \max\left\{V_{wt}(h), V_{wt}(j), V_{wt}(k)\right\}. \qquad (6)$$

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:

$$v_{wt}(\mu^*_{wt}; \eta) = \max\left\{v_{wt}(S; \eta), v_{wt}(NS; \eta)\right\}, \ \eta \in \{h, j, k\}. \qquad (7)$$

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:

$$V_{wt}(\eta) = E\left[v_{wt}(\mu^*_{wt}; \eta)\right] \tag{8}$$

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 $s_{STEM,jt}$:

$$s_{STEM,jt} = \frac{\sum_{w \in A_{jt}} \mathbf{1}\left\{S = \text{argmax}\left\{v_{wt}(S; j), v_{wt}(NS; j)\right\}\right\}}{\sum_w \mathbf{1}\left\{j = \text{argmax}\left\{V_{wt}(h), V_{wt}(j), V_{wt}(k)\right\}\right\}} \tag{9}$$

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 (8) into (6) and re-express the optimal enrollment decision:

$$\eta_{wt} = \text{argmax}\left\{E\left[v_{wt}(\mu^*_{wt}; h)\right], E\left[v_{wt}(\mu^*_{wt}; j)\right], E\left[v_{wt}(\mu^*_{wt}; k)\right]\right\} \tag{10}$$

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 4, 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 5 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

Our research design requires a comprehensive timeline of the process by which historical women's colleges converted to coeducation in the latter half of the 20th century and first two decades of the 21st century. Since to the best of our knowledge there did not exist a comprehensive list of this nature, we collected the information by hand.

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 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 HEGIS/IPEDS data and investigated these institutions by hand using a variety of resources, including Howe, Howard and Strauss (1982) and institutions' own websites. Over 90% of our transition dates were found on *.edu* websites. The three 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. This procedure identified 211 institutions that were women-only in the first year they were observed, 154 of which eventually transitioned to coeducation.

We thank Claudia Goldin and Lawrence Katz for sharing a similar, independently collected dataset that covers a partially overlapping time period (the late 1800s to roughly 1990; see Goldin and Katz (2011)). The transition dates for most former women's colleges are consistent across the datasets; where they disagree, the discrepancies are usually only 1-2 years or can be attributed to differing definitions of coeducation.

B.2 Constructing our sample

We are interested in studying the effect of a rapid influx of male students into a historically female-only college campus. Our original sample consists of 154 “switching” institutions and 3,663 potential comparison schools. Many of these institutions are outside the population of interest for this paper (e.g., junior colleges, art schools) or did not offer a setting that provides a “clean” transition from single-sex to coeducation (e.g., coordinate schools that had long allowed the female-only study body to take classes at a nearby male-only college). After making a number of restrictions to narrow our sample to the population of interest, we are left with a treatment group of 77 schools and 934 comparison schools. The sample restrictions, and their impact on the eventual analysis sample, are detailed below. The restrictions' impact on our sample size applies if these restrictions are implemented in order; some schools may satisfy multiple criteria for exclusion.

1. First, we restrict the sample to institutions that were female-only or coeducational in the 1965-66 school year, the first year in which we observe degree completions. We also drop schools that had converted from male-only to coeducation in the period shortly before our sample begins. This eliminates 136 potential comparison schools. (Resulting sample includes $N_c = 3,527$ potential comparison schools and $N_t = 154$ treated schools.)

2. To ensure we observe a reasonably lengthy pre-period for our event-study estimates, we eliminate treated schools that we see for fewer than 4 years prior to the switch to coeducation or that are completely missing from the data durign this pre-period. This means the earliest transition to coeducation in our analysis sample is the 1969-1970 school year. These restrictions eliminate 24 treated schools from the sample, as well as 7 potential comparison schools that had transitioned between 1954 and the start of our sample. ($N_c = 3,520$, $N_t = 130$)

3. To allow us to observe at least a decade of post-transition outcomes, we remove women's colleges that adopted coeducation after 2007 from our treatment group. We retain these institutions as potential *comparison* schools. This restriction removes 12 colleges from our treatment group

but adds them to the pool of comparison schools. ($N_c = 3,532$, $N_t = 118$)

4. To ensure our sample is limited to schools that switched from female-only to mixed-gender environments, we eliminate schools that were ever classified as coordinate institutions or merged with a men's college. We made this restriction because we suspect that classes on campus were coeducational long before mergers occurred, as is common with coordinate institutions. This restriction eliminated 23 treated schools and 27 untreated schools from the sample. ($N_c = 3,505$, $N_t = 95$)

5. We drop schools that entered the data after 1987. The IPEDS data dramatically expanded the sample at this time to include schools that had not been classified as "institutions of higher education" under Title IV of the Higher Education Act of 1965 and the response rate of those new institutions was much lower than the response rate of institutions included in HEGIS. Most of these schools were community colleges or other similar institutions. See the ICPSR documentation of the 1986-1987 academic year finance data for further details. This restriction eliminates 2 treated institutions and 1,703 untreated institutions from the sample. ($N_c = 1,802$, $N_t = 93$)

6. We eliminate for-profit institutions, once again to focus on traditional liberal arts programs. This restriction eliminates 55 untreated schools. ($N_c = 1,747$, $N_t = 93$)

7. We drop schools that closed fewer than 10 years after the switch to coeducation, as well as untreated schools that were in the data for fewer than 15 years. This restriction eliminates 6 treated and 322 untreated schools. ($N_c = 1,425$, $N_t = 87$)

8. Since our focus is on the choice of major for women at traditional liberal arts colleges, and in particular on the share majoring in quantitative

fields, we eliminate schools that did not grant any degrees in STEM fields in the first year we observe them in the data. This restriction eliminates 10 treated schools and 491 untreated schools. ($N_c = 934$, $N_t = 77$)

Returning coordinate schools and mergers (item 4), post-1987 entrants (item 5), for-profit institutions (item 6), institutions that closed shortly after the transition to coeducation (item 7), and institutions that may not have offered STEM degrees (item 8) does not substantially change our main results. We reported estimated effects on the share of women majoring in STEM using this larger sample in Appendix Figure A1. The estimated effects on the share of females majoring in STEM is very similar, although slightly smaller. This is accords with our expectations based on our reasoning for excluding these groups. The exclusion of coordinate colleges and mergers (item 4) is 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, muting the effects of the transition to coeducation. Adding schools that did not appear to have STEM programs (item 8) would also be expected to attenuate our estimates, since changes in the STEM share of female degrees would be 0 or positive by construction.

C Implementation of the 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 evolve dynamically, 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 (j) and time (t) fixed effects to estimate the impact of a policy for each time period relative to the date t_j^* of implementation of the reform:

$$y_{jt} = \alpha_j + \theta_t + \sum_{k=m}^{M} \beta_k \mathbb{1}\left\{t - t_j^* = k\right\} + \epsilon_{jt} \tag{11}$$

Recent studies have revealed that the TWFE specification may provide 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 un-modeled heterogeneous effects across units.

We instead adopt a slightly modified version of the estimator proposed by Callaway and Sant'Anna (2021).[26] 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), then aggregating those cohort-specific effects into an overall estimate of the average treatment effect on the treated. By estimating the effects one cohort at a time, the procedure facilitates transparency in the choice of comparison group used for each treatment group (e.g., the researcher can ensure the comparison group is not polluted by a recently treated unit that may still be adjusting to the reform) and allows potentially heterogeneous effects to be aggregated using the choice of weights best suited for estimating target parameter of interest (i.e., it avoids weighting by the inverse of the variance of exposure to the treatment, as is the default in regression-based methods).

[26]We are indebted to Brantly Callaway and Pedro Sant'Anna for their generous and illuminating correspondence about the finer details of their estimator.

The doubly-robust estimator developed by Callaway and Sant'Anna (2021) relies on two strategies to construct appropriate counterfactual trends for units that adopted the treatment of interest. First, to limit the comparison group to schools that "look like" the treated group, each school j gets a cohort-g-specific propensity score $\hat{p}_g(X_j)$. In addition, the counterfactual trend in outcome y_{jt} between time t and some base period b is estimated by regressing changes $\Delta y_{jt,b} = y_{jt} - y_{jb}$ on the same vector of covariates X_j in a sample made up solely of the comparison group, and then using these regression estimates to predict changes $\Delta\hat{y}_{jt,b}(X_j)$ for the treated cohort g.

Formally, for a sample made up of schools $j \in \{1, 2, ..., J\}$, the estimator is constructed with the following sample analog of Callaway and Sant'Anna (2021) equation 2.4 (weights are omitted here for parsimony, but it is straightforward to add them):

$$\hat{\alpha}_{g,t} = \frac{1}{J}\sum_{j\in J}\left(\frac{G_{jg}}{\frac{1}{J}\sum_j G_{jg}} - \frac{\frac{\hat{p}_g(X_j)C_{jg}}{1-\hat{p}_g(X_j)}}{\frac{1}{J}\sum_j \frac{\hat{p}_g(X_j)C_{jg}}{1-\hat{p}_g(X_j)}}\right)(\Delta y_{jt,b} - \Delta\hat{y}_{jt,b}(X_j)) \qquad (12)$$

where G_{jg} is a binary indicator for school j belonging to treatment cohort g and C_{jg} is an indicator for belonging to the pool of candidate comparison schools for group g.

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_{jb} as the average of the outcome variable in the five years immediately preceding the switch to coeducation, $y_{jb} = \frac{1}{5}\sum_{s=g-5}^{g-1} y_{js}$. 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 – defines $\Delta y_{t,b} = y_t - y_b$ for $t \geq t_j^*$, but as single-year differences $\Delta y_{t,b} = y_t - y_{t-1}$ for $t < t_j^*$. We instead adopt the former definition for all periods, i.e., $\Delta y_{t,b} = y_t - y_b \; \forall \; t$. While the two approaches are very similar conceptually, we believe our approach to reporting event-study results will be more familiar and intuitive for our readers, most of whom are accustomed to interpreting event-study coefficients as changes in an outcome relative to an omitted period, rather than as the first derivative of those changes.

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 (Albert and Anderson, 1984; Firth, 1993). It also allows us to focus on what we believe are the most important conditioning variables in our setting, which happen to be discrete.

These three changes result in a simplified version of equation 12. To see this, note that because our vector X_j is made up entirely of discrete variables, the propensity score for any group g will either 0 or a constant $\bar{p}_g$, where $\bar{p}_g$ is the share of treated schools among all institutions where X_j is identical to the treated school in question. In addition, because we define group g as a single school, the formula simplifies to

$$\hat{\alpha}_{g,t} = (\Delta y_{jt,b} - \Delta \hat{y}_{jt,b}(X_j)) - \sum_{j \in J} \omega_j(X_j) \left(\Delta y_{jt,b} - \Delta \hat{y}_{jt,b}(X_j) \right) \tag{13}$$

where $\omega_j(X_j)$ sums to 1 and represents school j's share of the sample for which $C_{jg} = 1$ *and* $p_g(X_j) > 0$, i.e., its share of the comparison group for treatment group g. Since $\Delta \hat{y}_{jt,b}(X_j)$ is calculated using only candidate comparison schools with strictly positive propensity scores, two of the final three terms cancel out and equation 13 simplifies to equation 2.

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_j 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_j: 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.

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 A6 and A7 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.

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 A8 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.[27] This estimate is slightly larger in magnitude than the one

[27] We conduct inference by randomly reassigning treatment status and estimating the effect

we obtain in our main event study model.

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.[28] Revisions of the CIP occurred in 1985, 1990, 2000, and 2010.[29] 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,[30] though the 1990, 2000, 2010 revisions did move, split, and combine some two-digit codes.[31]

---

of the transition to coeducation on the placebo institutions, using 250 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.

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

[29] 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.

[30] 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.

[31] 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.

For this paper, all 6-digit codes were crosswalked to the 4-digit 1990 CIP. 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.[32] Four-digit majors with fewer than 950 school-by-year observations were combined with majors that cover similar material[33] 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.

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

[33]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.

Art Visual and performing arts, architecture and related services

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 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 5 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.[34]

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 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.

[34]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.

G The National Longitudinal Study of the Class of 1972

We conduct additional analyses using the National Longitudinal Study of the Class of 1972 (NLS72) to determine which baseline (i.e., pre-freshman) characteristics are most important for predicting that a student will complete a STEM degree (National Center for Education Statistics, 1999). We chose the NLS72 for this analysis because it provides information on both characteristics before the beginning of college and STEM degree completion for a cohort who attended college during the part of our sample period when most of our switching colleges transitioned. NLS72 is a nationally representative longitudinal survey that followed high school seniors in the graduating class of 1972 for twelve years following completion of high school. The baseline survey, conducted in spring 1972 (right before students graduated high school), collected a substantial amount of information on students' backgrounds and plans for the future. The five follow-up surveys (conducted between 1973 and 1984) focused on what students had been doing since the previous survey (including degree completion and college major). We focus on female students who responded to both the baseline survey and the fourth follow-up (conducted between October 1979 and May 1980), who indicated on the baseline survey that they planned to attend a four-year college starting in Fall 1972 and who had completed a degree by October 1979.

We used the NLS72 data to estimate the correlation between STEM degree completion and baseline intention to major in a STEM field, defining STEM to include biology, computer science, engineering, mathematics, and physical science (similarly to our analyses of IPEDS and HERI data). We estimate the following equation using OLS:

$$STEMBA_i = \alpha STEMIntent_i + X_i\beta + \varepsilon_i \tag{14}$$

where $STEMBA_i$ is a 0/1 indicator that the student's completed bachelor's degree was in a STEM field and $STEMIntent_i$ is an indicator variable that the student planned to major in STEM at baseline, and X_i is a vector of control variables. In our first specification, X_i includes only a constant. In our second specification, we add a set of controls for students' occupation plans (professional, homemaker, or other) and whether student considered marriage and family to be "very important." In our third specification, we add a set of controls for students' background, including an indicator that students are white, an indicator for being a first-generation college student (i.e., neither parent attended college), an indicator that the student's father completed college, an indicator that the student's mother completed college, and information on parents' occupations (indicators for a professional occupation fathers and indicators for professional occupation or homemaker for mothers). In our fourth specification, we add an indicator that students had a GPA of A- or better in high school and the number of years of high school math and science that the student completed.[35] In our subsequent analysis, we use the coefficient estimates from these regressions to predict the share of entering freshman women in our HERI Freshman Survey data who will major in STEM, and then evaluate the effect of coeducation on this predicted share.[36] We interpret this analysis as a test of composition effects.

We use the R^2 values from these regressions to determine the relative importance of each baseline characteristic to completing a STEM degree. See Panel A of Table 4 for the results. Plans to major in STEM right before high school graduation have a R^2 value of 0.191, indicating that baseline preferences account for approximately 19% of variation in STEM degree completion in the NLS72 sample. Adding each successive set of control variable shifts the R^2 by no more than 0.012, with the largest shift coming from the addition of high school grades and coursework.[37] We take these results as confirmation

[35]We recoded the control variables in NLS72 to match information available in TFS as closely as possible.

[36]In our HERI data, school-year observations sometimes have missing values for some of the characteristics measured in the NLS72 analysis. In those cases, we use all available characteristics to predict the share of women who will major in STEM.

[37]Further analyses suggest that coursework is more important than grades, consistent with Card and Payne (2017). However, information on high school coursework is only available in TFS after 1984.

that preferences are the main characteristic of interest in determining whether shifts in the composition of female students could be responsible for the effects of coeducation on future STEM majoring.

H Alternative approach to quantifying the environmental effect

Our main analysis in section 5.4 relies on estimates from our linked HERI data and the NLS72 to quantify the potential role played by composition effects in our main findings that coeducation reduced the share of women majoring in STEM. In this section, we present results from an alternative approach that relies on the assumption that women who matriculated *before* coeducation was adopted may comprise a sample that is relatively free from selection into the college. In particular, the sophomore class during the first year of coeducation would have experienced a relatively pure "environmental" effect – since, as underclasswomen, they would likely share classrooms and social spaces with the new men – but had already chosen their college. In contrast, freshman cohorts may have been more prone to composition effects, while juniors and seniors would have been both less likely to interact with the entering men and more likely to be locked into an academic program by the time those men arrived.

Our IPEDS data does not provide information on the time to degree, limiting our ability to measure cohorts precisely. Our best proxy is to focus on women who graduated in year $t_j^* + 2$, i.e., the third year of coeducation at school j. Assuming four years to completion of the degree for most women, this cohort should be made up primarily of women who were sophomores when coeducation was implemented.

Table A8 reports estimates of $\beta_{\tau=2}$ and β_{LR}, constructed from equation 3 with the share of women majoring in STEM as the outcome. Column 1 presents estimates of the effect on the sophomore class. These estimates are generally less precise than our main estimates of the long-run effect, but they can be statistically distinguished from 0 in every comparison group.

What does this imply for the question of whether the main effects are driven by composition or features of the campus environment? Note that the smaller magnitudes in column 1 of Table A8 are consistent with our conclusions that the decreasing share of women in STEM is most likely driven by increasing interactions with men, because the male share of men at former women's colleges also increased gradually. The sophomore class was thus exposed to a lower "dose" of men. If we re-scale our estimates by the male share of graduates, we find that that share of women in STEM falls by about 1.8 percentage points for every 10-percentage-point increase in men (column 1 of panel A). This figure is very similar to the estimate of 1.6 percentage points (column 2, panel A). These very similar magnitudes are far from conclusive, but they provide yet more evidence that is consistent with the hypothesis that composition effects were negligible – and, if anything, wrong-signed – in this setting.

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.

Albert, A, and JA Anderson. 1984. "On the existence of maximum likelihood estimates in logistic regression models." *Biometrika*, 71(1): 1–10.

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.

Firth, David. 1993. "Bias reduction of maximum likelihood estimates." *Biometrika*, 80(1): 27–38.

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.

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

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.

National Center for Education Statistics. 1999. “National Longitudinal Study of the Class of 1972.” Inter-University Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/ICPSR08085.v1.

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.

I Tables and Figures

Table A1: Summary statistics, treated and alternative comparison schools

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Treated	Candidate comparison groups				Matched comparison groups			
		Never-treated wom. clg.	Difference (1)-(2)	Liberal arts colleges	Difference (1)-(4)	Never-treated wom. clg.	Difference (1)-(6)	Liberal arts colleges	Difference (1)-(8)
STEM share of women's degrees	0.10	0.11	-0.01	0.11	-0.00	0.09	0.01	0.11	-0.00
	(0.05)	(0.05)	[0.27]	(0.06)	[0.79]	(0.05)	[0.31]	(0.06)	[0.77]
Annual growth rate, STEM	-0.005	-0.003	-0.003*	-0.003	-0.003**	-0.004	-0.001	-0.005	-0.000
	(0.011)	(0.009)	[0.07]	(0.014)	[0.04]	(0.009)	[0.53]	(0.013)	[1.00]
Total enrollment	1226	1250	-24	1500	-274**	1254	-28	1422	-196
	(917)	(653)	[0.88]	(1177)	[0.03]	(585)	[0.85]	(773)	[0.20]
Female share of all degrees	0.97	0.99	-0.02**	0.54	0.44***	0.99	-0.02*	0.60	0.37***
	(0.07)	(0.02)	[0.01]	(0.17)	[0.00]	(0.03)	[0.07]	(0.23)	[0.00]
Graduate degrees awarded	27	19	9	10	17***	14	13*	14	13
	(51)	(42)	[0.30]	(40)	[0.00]	(38)	[0.08]	(56)	[0.11]
Private college	0.92	1.00	-0.08**	0.89	0.03	1.00	-0.08**	0.97	-0.05
	(0.27)	(0.00)	[0.01]	(0.32)	[0.38]	(0.00)	[0.01]	(0.17)	[0.10]
Ever Catholic-affiliated	0.64	0.30	0.33***	0.06	0.58***	0.64	0.00	0.64	0.00
	(0.48)	(0.46)	[0.00]	(0.23)	[0.00]	(0.48)	[1.00]	(0.48)	[1.00]
Selective admission	0.19	0.45	-0.27**	0.20	-0.02	0.19	0.00	0.19	0.00
	(0.39)	(0.50)	[0.01]	(0.40)	[0.74]	(0.39)	[1.00]	(0.39)	[1.00]
Observations	77	29	106	362	439	29	106	350	427

Notes: Column 1 shows means (standard deviations) for our main treated sample of women's colleges, averaged over the five years prior to each treated school j's transition to coeducation. Comparison group means in column 2 are constructed by matching each treated school j with all women's colleges that never transitioned. We then stack these comparison groups across all treated schools j and compute sample means for the resultant "grand" group. Column 3 reports result of test of difference in means between treated (column 1) and never-treated comparison group (column 2). Column 4 reports summary statistics for comparison group of untreated liberal arts college using Carnegie Classification from 1987. Column 5 presents results of test of difference in means between treated group and untreated liberal arts colleges. Columns 6-9 report analogous summary statistics using matched samples as described in Section 3.1. STEM share of women's degrees is weighted by total degrees awarded to females, female share of degrees is weighted by total degrees awarded, and all other means are unweighted. Data are drawn from IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. See Section 2 for further information on sample construction and characteristics.

Table A2: Summary statistics for schools that switched to coeducation

	(1) All switchers	(2) Excluded switchers	(3) Main sample	(4) HERI subsample
STEM share of women's degrees	0.096	0.070***	0.103*	0.106
	(0.053)	(0.050)	(0.051)	(0.055)
Annual growth rate, STEM	-0.004	-0.002***	-0.005**	-0.003***
	(0.010)	(0.008)	(0.011)	(0.007)
Total enrollment	1190	1136	1226	1072
	(1069)	(1265)	(917)	(484)
Female share of all degrees	0.96	0.93	0.97	0.98
	(0.08)	(0.10)	(0.07)	(0.04)
Graduate degrees awarded	49	89	27	26
	(251)	(419)	(51)	(46)
Private college	0.95	1.00	0.92	0.97
	(0.23)	(0.00)	(0.27)	(0.18)
Ever Catholic-affiliated	0.56	0.40	0.64	0.63
	(0.50)	(0.49)	(0.48)	(0.48)
Selective admission	0.14	0.05	0.19	0.23
	(0.34)	(0.21)	(0.39)	(0.42)
Institutions	118	41	77	30

Notes: Table shows sample means (standard deviations) calculated in the five years prior to the switch to coeducation. Column 1 includes all women's colleges that adopted coeducation after 1965-66 and that we observe during the five years before and decade after the transition. Column 2 includes schools we drop because they did not fit into our target population of institutions that offered an arts-and-sciences curriculum and experienced sharp transitions to coeducation. Column 3 shows our main sample of colleges that switched between 1969 and 2007. Column 4 show summary statistics for the subset of column 3 that can be linked to schools in the HERI Freshman Survey. In columns 2-4, the designation of 1, 2, or 3 stars indicates that a test of differences between each subset of switchers and the sample of all switchers (column 1) results in a p-value below 0.10, 0.05, or 0.01, respectively. Data drawn from IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. Trends in STEM and total degrees are the estimated linear trend in the five years prior to the switch to coeducation. 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. See Section 2 for further detail on sample construction.

Table A3: Long-run effect of coeducation on presence of men at former women's colleges

	(1) Male share of freshmen	(2) Male share of students	(3) Male share of degrees	(4) Male share of faculty
Panel A: Not-yet-treated comparison group				
Long-run effect	0.211***	0.216***	0.192***	0.045***
	(0.007)	(0.016)	(0.016)	(0.005)
Counterfactual mean	0.030	0.034	0.036	0.385
Observations	5,158	5,164	5,505	2,428
Panel B: All-college comparison group				
Long-run effect	0.236***	0.251***	0.210***	0.059***
	(0.001)	(0.015)	(0.021)	(0.009)
Counterfactual mean	0.005	0.000	0.009	0.387
Observations	25,880	26,015	27,621	12,318
Panel C: Never-treated comparison group				
Long-run effect	0.217***	0.232***	0.195***	0.039***
	(0.004)	(0.016)	(0.017)	(0.007)
Counterfactual mean	0.025	0.028	0.034	0.390
Observations	4,839	4,844	5,164	2,186
Panel D: Liberal arts college comparison group				
Long-run effect	0.227***	0.240***	0.212***	0.059***
	(0.006)	(0.012)	(0.016)	(0.004)
Counterfactual mean	0.013	0.010	0.013	0.380
Observations	19,636	19,695	20,954	12,387

Notes: Table displays the estimated effect of the switch to coeducation on male share of freshmen (column 1), male share of undergraduate students (column 2), male share of degrees earned (column 3), and male share of faculty (column 4), estimated using equation 5. Freshman and undergraduate enrollment is available only beginning in 1968-69 school year. Faculty data available in selected years beginning in 1971. 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 pre-reform 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.

Table A4: Long-run effect of coeducation on the share of women choosing other majors

	(1) Art	(2) Business	(3) Education	(4) Health
Panel A: Not-yet-treated comparison group				
Long-run effect	-0.001	-0.017	-0.017	0.034
	(0.009)	(0.013)	(0.022)	(0.022)
Counterfactual mean	0.073	0.121	0.184	0.163
Observations	5,505	5,505	5,505	5,505
Panel B: All-college comparison group				
Long-run effect	0.006	-0.027**	0.008	0.053**
	(0.008)	(0.012)	(0.019)	(0.021)
Counterfactual mean	0.074	0.130	0.161	0.129
Observations	27,618	27,618	27,618	27,618
Panel C: Never-treated comparison group				
Long-run effect	0.003	-0.013	-0.011	0.016
	(0.012)	(0.012)	(0.025)	(0.026)
Counterfactual mean	0.069	0.116	0.178	0.180
Observations	5,164	5,164	5,164	5,164
Panel D: Liberal arts college comparison group				
Long-run effect	-0.005	-0.029***	-0.003	0.065***
	(0.008)	(0.011)	(0.016)	(0.019)
Counterfactual mean	0.079	0.136	0.171	0.124
Observations	20,954	20,954	20,954	20,954

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 5. 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 A5: 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	5,505	5,505	5,505	5,505	5,505
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	0.107	0.042	0.066	0.145
Observations	27,618	27,618	27,618	27,618	27,621
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
Panel D: Liberal arts college comparison group					
Long-run effect	0.003	-0.006	0.004	0.004	0.009
	(0.005)	(0.010)	(0.012)	(0.008)	(0.011)
Counterfactual mean	0.028	0.102	0.036	0.066	0.138
Observations	20,954	20,954	20,954	20,954	20,954

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 5. 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 A6: Long-run effect of coeducation on women's high school GPA and ranking within class

	(1) GPA rank in class	(2) GPA rank in STEM
Panel A: Not-yet-treated comparison group		
Long-run effect	0.0388***	0.0552
	(0.0108)	(0.0785)
Counterfactual mean	0.391	0.336
Observations	1,426	1,363
Panel B: All-college comparison group		
Long-run effect	0.0463***	-0.0033
	(0.0116)	(0.0113)
Counterfactual mean	0.366	0.350
Observations	4,680	4,250
Panel C: Never-treated comparison group		
Long-run effect	0.0405***	0.0567
	(0.0114)	(0.0786)
Counterfactual mean	0.390	0.333
Observations	1,400	1,339
Panel D: Liberal arts college comparison group		
Long-run effect	0.0495***	-0.0152
	(0.0105)	(0.0123)
Counterfactual mean	0.371	0.362
Observations	9,084	8,639

Notes: Table displays the estimated effect of the switch to coeducation on female students' average high-school GPA ranking among their college freshman classmates (column 1), and female students' average high-school GPA ranking among college freshman classmates who intended to major in STEM (column 2), estimated using $\hat{\beta}_{LR}$ from equation 5. 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. Data drawn from the HERI Freshman Survey, spanning 1966-2006, 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.

Table A7: Bounding the composition effect of coeducation on STEM degree receipt, alternative comparison groups

	(1)	(2)	(3)	(4)
Panel A: Effect of freshman characteristics on women's likelihood of earning STEM degree				
Effect of intent to major in STEM	0.336***	0.333***	0.332***	0.317***
	(0.040)	(0.040)	(0.040)	(0.041)
Covariates:				
Career, family aspirations		X	X	X
Parental education, occupation			X	X
High school grades, coursework				X
R-squared	0.191	0.199	0.205	0.215
Observations	1,235	1,235	1,235	1,235
Panel B: Effect of coeducation on predicted share of female freshmen who will major in STEM, never-treated comparison group				
Estimated composition effect	0.006	0.010	0.010	0.014
	(0.009)	(0.011)	(0.011)	(0.012)
Composition effect / Total effect of coeducation on STEM major choice	-18%	-31%	-30%	-40%
Composition effect upper bound	37%	34%	33%	27%
Panel C: Effect of coeducation on predicted share of female freshmen who will major in STEM, liberal arts college comparison group				
Estimated composition effect	0.002	0.000	0.001	0.001
	(0.004)	(0.004)	(0.004)	(0.004)
Composition effect / Total effect of coeducation on STEM major choice	-5%	-1%	-2%	-3%
Composition effect upper bound	20%	24%	23%	21%

Notes: Panel A reports regression estimates of the effect of intention to major in STEM as of freshman year on share of students earning STEM degree, derived from sample of women in National Longitudinal Study of 1972. Panels B and C report implied long-run effect on the predicted share of freshman women at newly coeducational colleges who will major in STEM, calculated using equation 5 and sample of women from the HERI Freshman Survey. Predicted share in STEM is constructed by interacted coefficients from the regressions in panel A with characteristics of entering freshman women in the HERI data. Share of total effect explained by composition is constructed by dividing predicted STEM effect by estimated effect of coeducation on the share of women earning STEM degree from our linked IPEDS-HERI data (-0.034, see Figure A1a). Upper bound on composition effect is constructed by dividing lower bound of 95% confidence interval of predicted STEM effect by -0.034. See Appendix Table A7 for estimates drawing on alternative comparison groups.

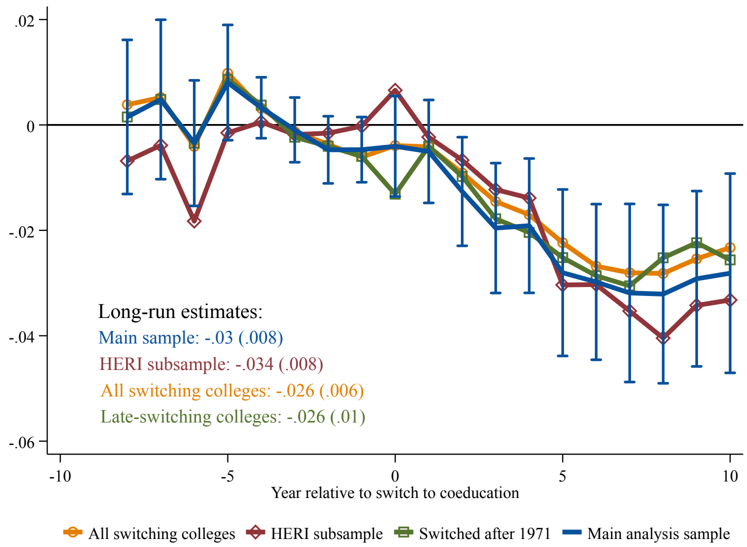
Table A8: Effect of coeducation on the share of women in the sophomore class majoring in STEM

	(1) Sophomore class	(2) Long-run effect
Panel A: Not-yet-treated comparison group		
Effect on STEM	-0.0126***	-0.0302***
	(0.0053)	(0.0077)
Effect / Δ male grads	-0.176	-0.157
Panel B: All-college comparison group		
Effect on STEM	-0.0123***	-0.0312***
	(0.0053)	(0.0067)
Effect / Δ male grads	-0.159	-0.148
Panel C: Never-treated comparison group		
Effect on STEM	-0.0153***	-0.0349***
	(0.0059)	(0.0092)
Effect / Δ male grads	-0.208	-0.179
Panel D: Liberal arts college comparison group		
Effect on STEM	-0.0162***	-0.0349***
	(0.0058)	(0.0059)
Effect / Δ male grads	-0.206	-0.164

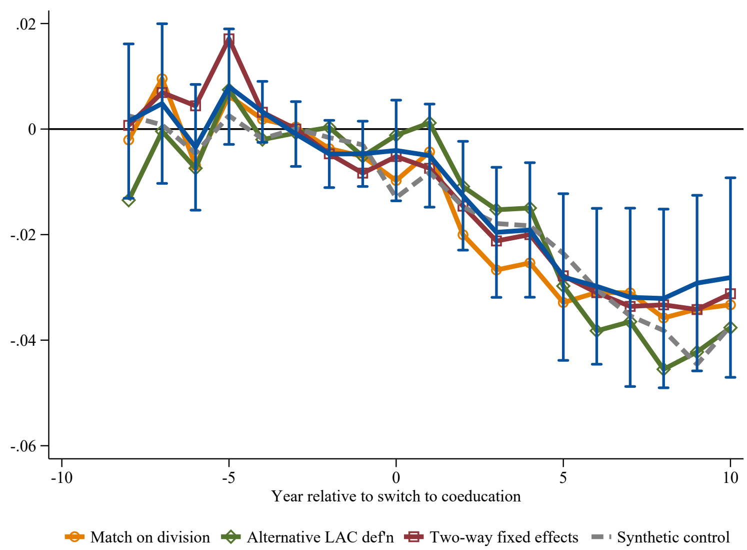
Notes: Table displays the estimated effect of the switch to coeducation on the share of women majoring in a STEM field. Estimates in column 1 correspond to $\hat{\beta}_{\tau=2}$ from equation 3, and estimates in column 2 correspond to equation 5. Each panel uses the specified pool of institutions to 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. Third row of each panel rescales the estimated effect by the effect on male share of graduates in event-year $\tau = 2$ (column 1) or in event-years 5 through 9 (column 2).

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

(a) Robustness to selection of treatment group

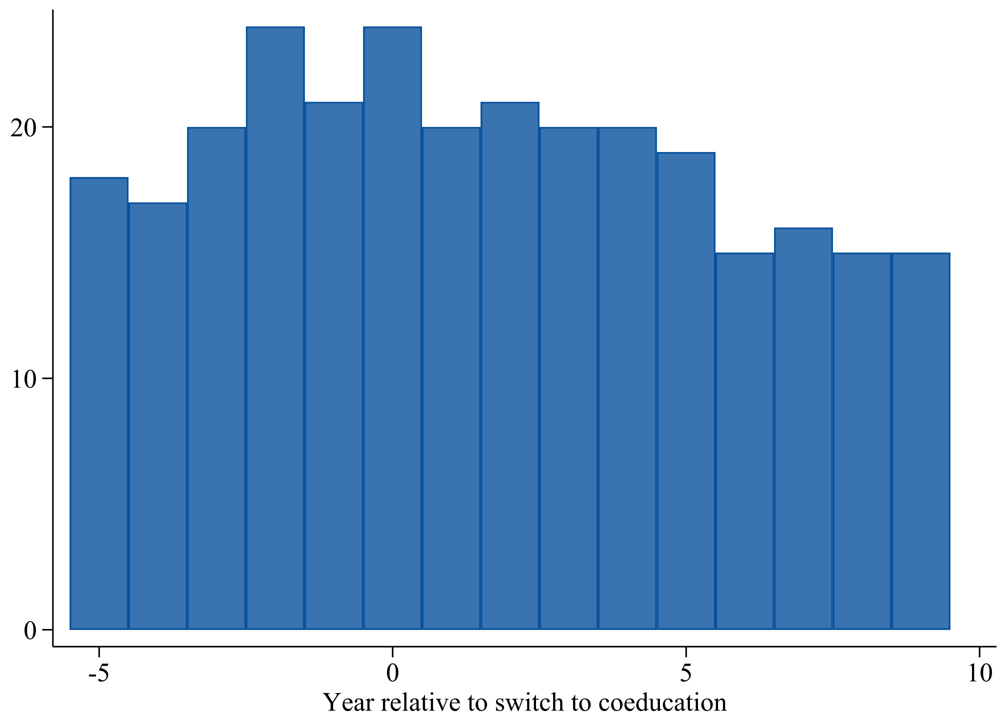


(b) Robustness to construction of comparison group



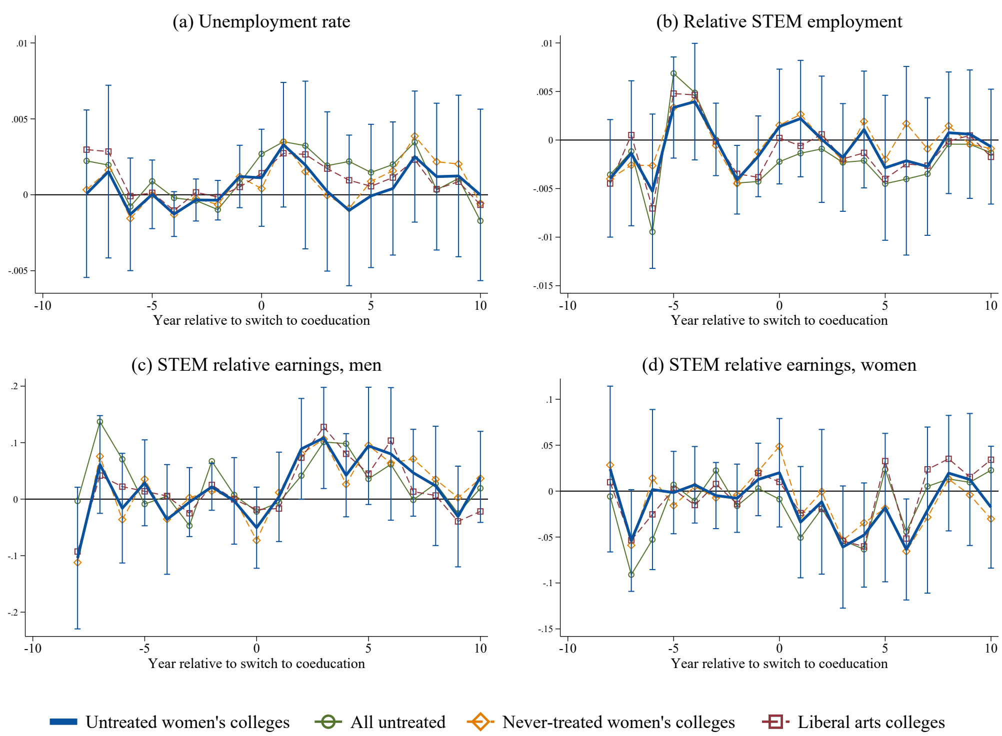
Notes: Figures show event-study estimates from equation 3. In each figure, blue line replicates estimates from Figure 3, which is based on a sample of 77 treated schools. Other lines show estimated effect on share of women in STEM using alternative treatment (Figure A1a) or comparison (Figure A1b) groups.

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



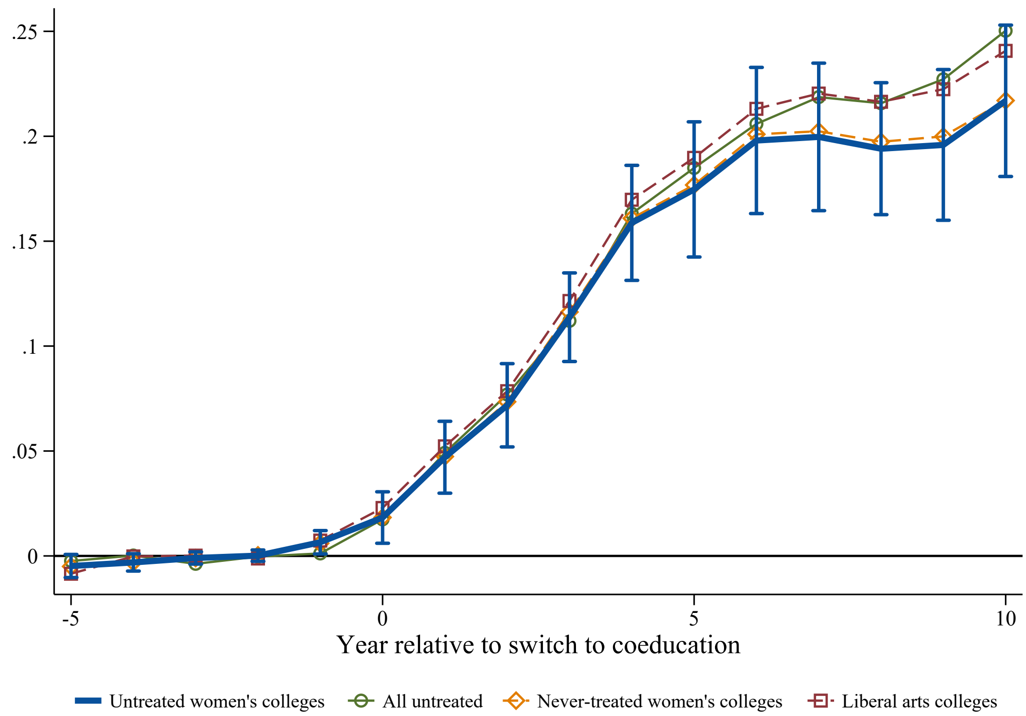
Notes: Data drawn from 1966-2006 versions of The Freshman Survey administered by HERI, linked to hand-collected dates of transitions to coeducation by institution's state. Each bar shows the number of treated schools that appear in the survey in each year relative to the switch to coeducation. Sample 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



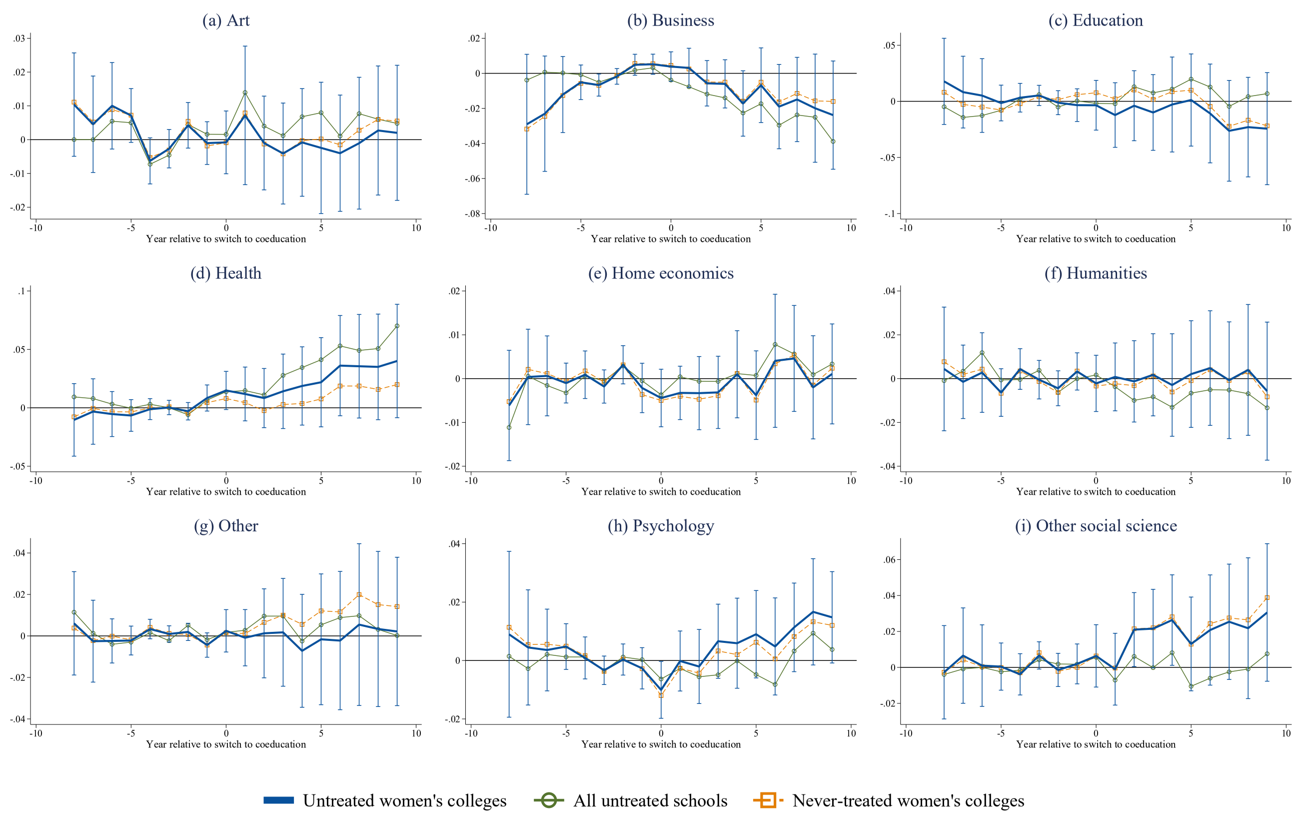
Notes: Data drawn from 1966-2016 CPS data accessed via IPUMS (Ruggles et al., 2020), linked to hand-collected dates of transitions to coeducation by institution's state. See Section 2 for further detail. Unemployment rate is measured among individuals age 18-64. Relative STEM employment is constructed 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 3. Standard errors are constructed from a block bootstrap clustered at the institution level.

Figure A4: Effect of coeducation on male share of graduates



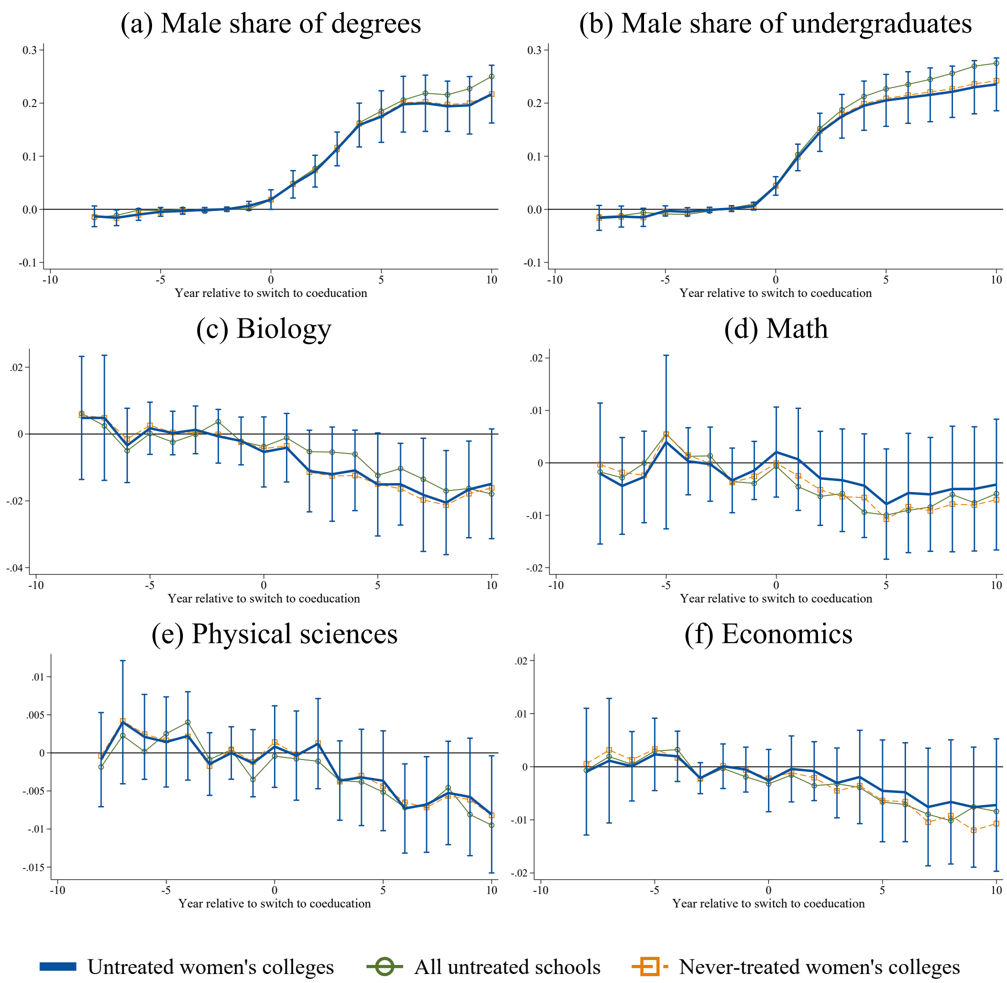
Notes: Figure shows event-study estimates from equation 3, using comparison group specified in legend. Data on degrees granted comes from HEGIS and IPEDS surveys, 1966-2016, linked to hand-collected information on dates of transition to coeducation. 95% confidence intervals are constructed using block bootstrap clustered at the institution level.

Figure A5: Effect of coeducation on share of women choosing other degrees



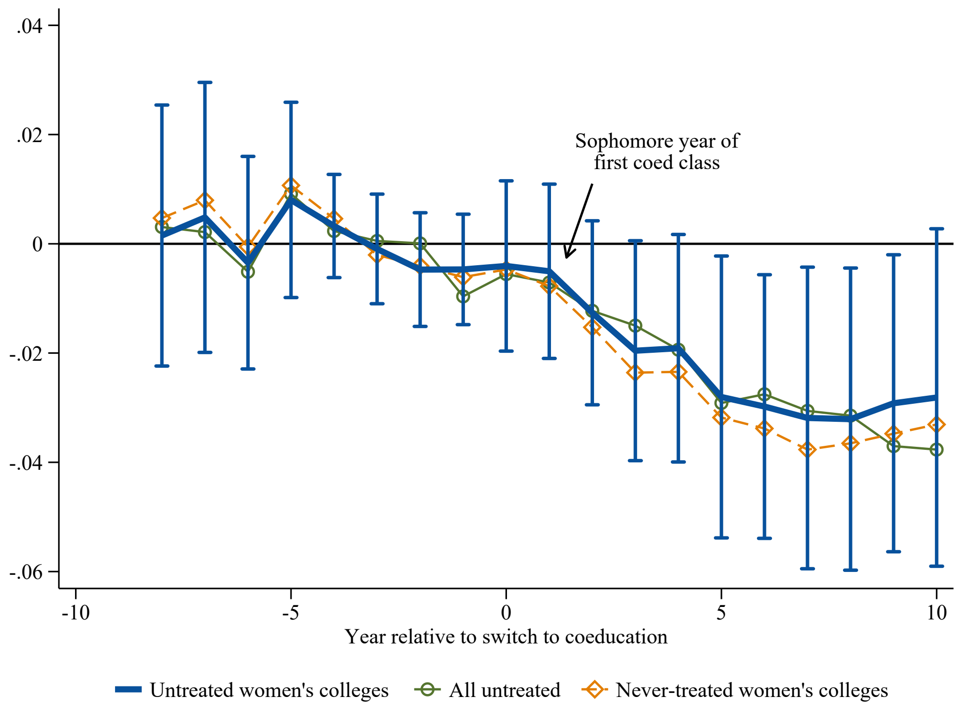
Notes: Figure shows event-study estimates from equation 12, aggregated using equation 3, using comparison group specified in legend. Data on degrees granted comes from HEGIS and IPEDS surveys, 1966-2016, linked to hand-collected information on dates of transition to coeducation. 95% confidence intervals are constructed using block bootstrap clustered at the institution level.

Figure A6: Effect of coeducation on gender composition and women's choice of quantitative majors, with uniform confidence intervals



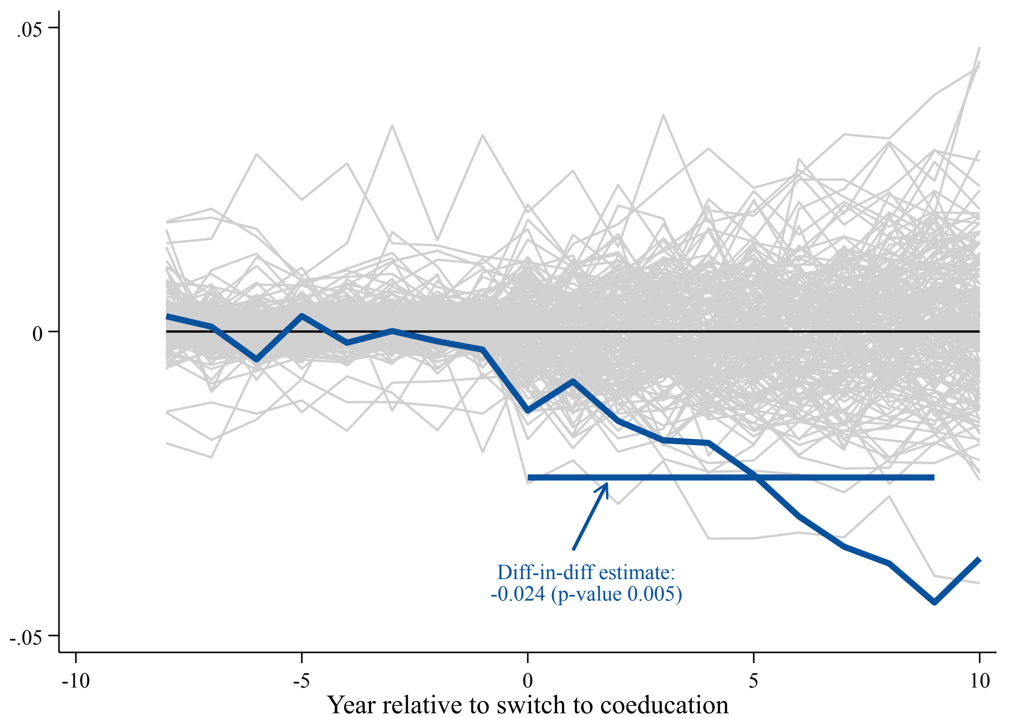
Notes: Figure repeats point estimates from Figure 4 with confidence intervals that account for multiple testing across event-time periods.Data drawn from HEGIS/IPEDS surveys, spanning 1966-2016, linked to hand-collected dates of transitions to coeducation by institution. See Section 2 for further detail. Panels display estimate of β_k from equation 3. Dependent variable is the share of degrees earned in STEM among all degrees earned by women in the academic year.

Figure A7: Effect of coeducation on the share of women majoring in STEM, 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 3. 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.

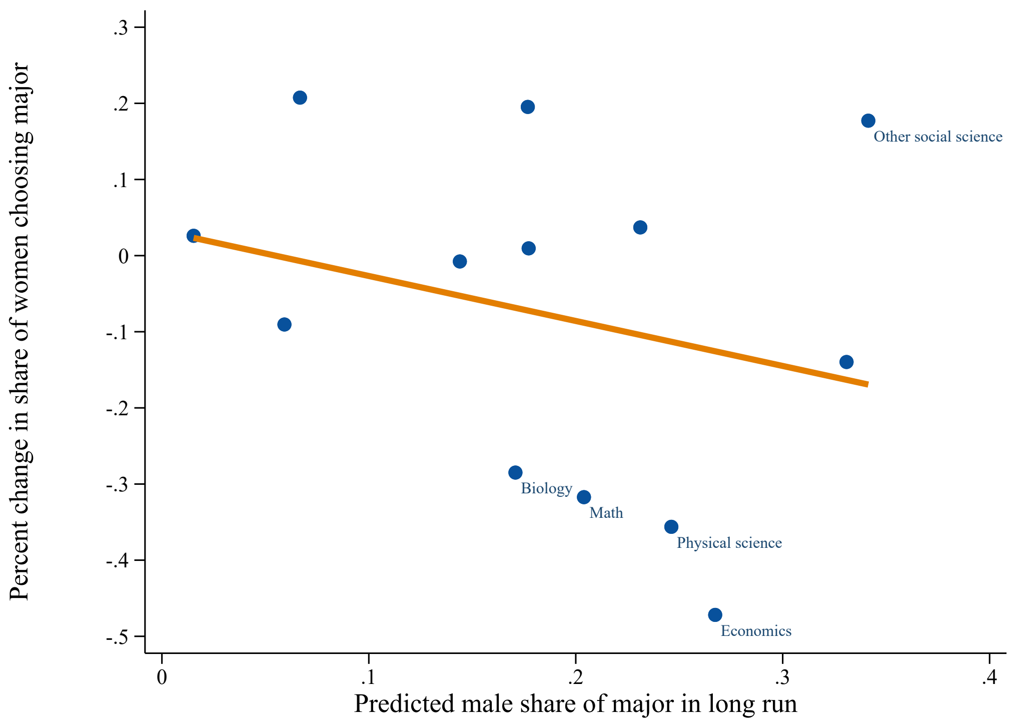
Figure A8: The effect of coeducation on the STEM share of degrees awarded to women: synthetic control specification



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 250 replications.

Figure A9: Effect of male inflows on distribution of women's choices of major

(a) Effect on women's choice of major vs. male inflow to major



(b) Elasticity of women's choice of major as share of total male inflow

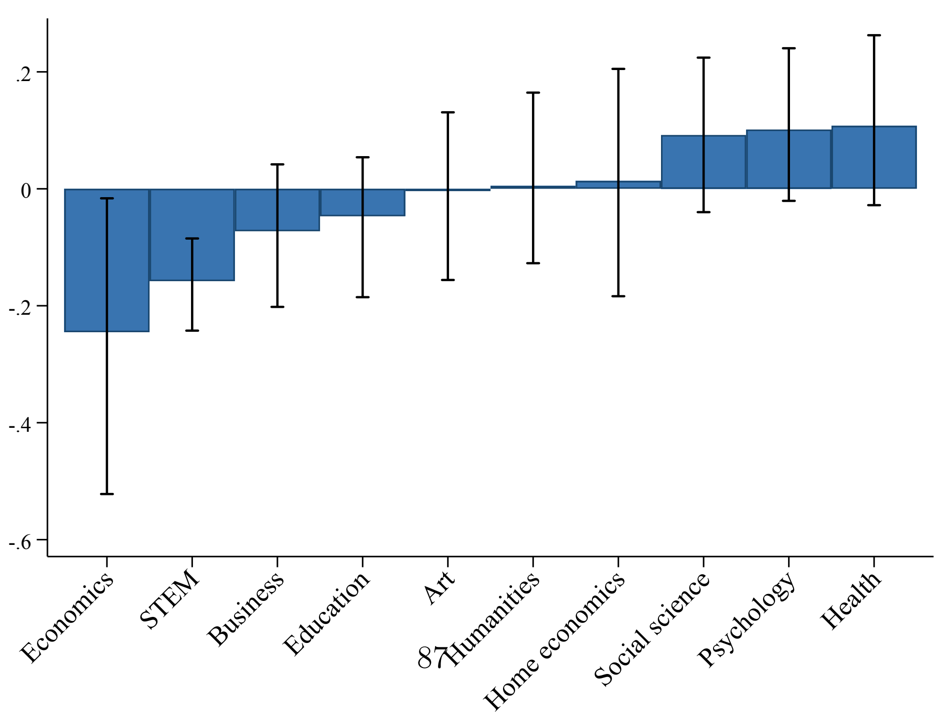
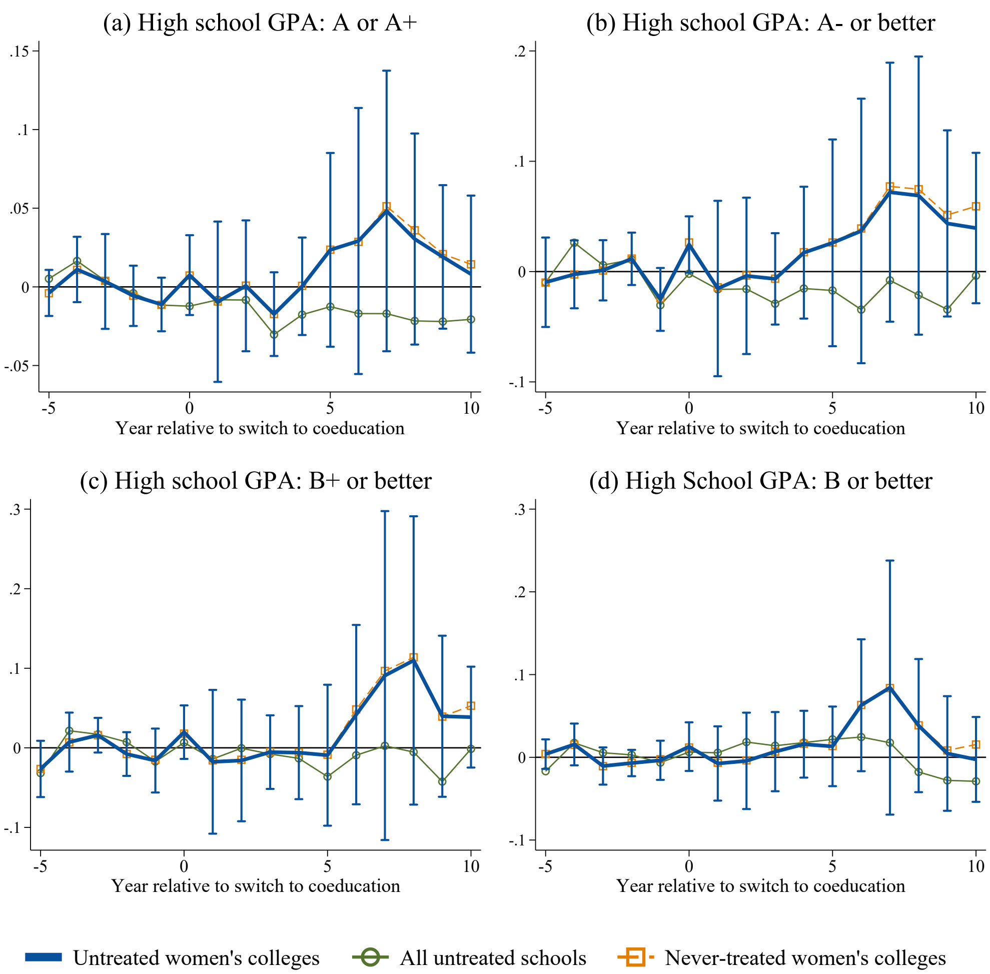


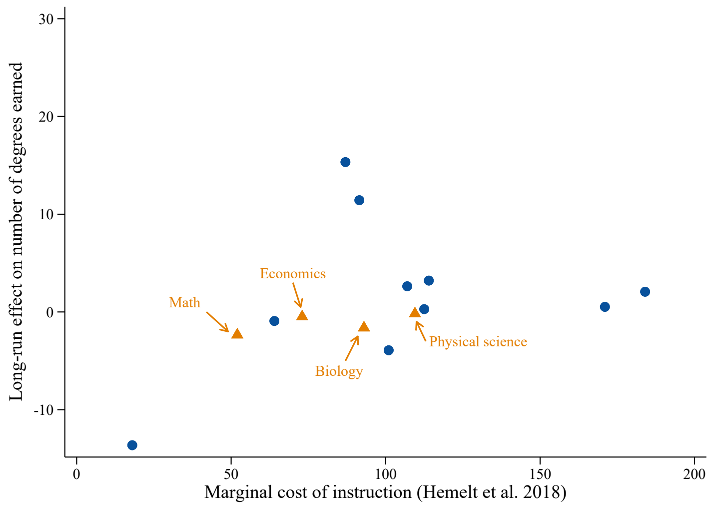
Figure A10: The effects of coeducation on the distribution of female matriculants' high school GPA



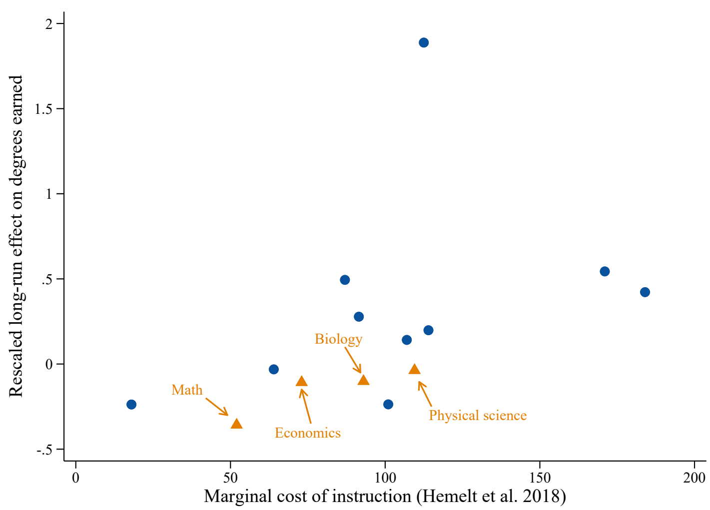
Notes: Each point shows and estimate of β_k from equation 3, using the comparison group specified in the legend. Error bars show 95% confidence interval constructed using block bootstrap clustered at the institution level. Data drawn from HERI, spanning 1966-2006, linked to hand-collected dates of transitions 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 5 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).