The Diversity Paradox: Evidence from College Coeducation*

Andrew T. Hendrickson† Francesca Truffa‡ Ashley Wong§

March 26, 2026

Abstract

How novel ideas are adopted and recognized is crucial to scientific progress, but not all ideas from all groups are equally recognized. This paper studies whether and how increasing gender diversity at universities may lead to greater inclusion and recognition of research traditionally associated with women. Leveraging the transitions to coeducation of 88 all-male universities and novel text-based measures of research content, we show that coeducation led to overall modest shifts toward female-associated research. This aggregate effect masks substantial heterogeneity across fields: in disciplines with higher early female representation, we observe a pronounced increase in female-associated research driven by both existing faculty and new entrants. Male-dominated fields, by contrast, exhibit little change or even declines in female-associated research, primarily due to changes in hiring practices. These findings highlight that while diversity can foster innovation, its effects may only be concentrated in areas already receptive to the new perspectives.

*We thank seminar and conference audiences at Columbia University, University of Michigan, Purdue University, UC Santa Barbara, Barnard-Columbia Gender in the Global Economy Conference, and the 4th Zurich AI+Econ Workshop. We also gratefully acknowledge the financial support of the Women and Science Chair, a Paris Dauphine-PSL University and its Foundation Chair, in partnership with Amundi, Fondation L'Oréal, La Poste, Generali France, Safran and Talan. Ashley Wong received support through the Marie Skłodowska-Curie Postdoctoral Fellowship No. 101148992, funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Horizon Europe. Neither the European Union nor the granting authority can be held responsible for them. We thank Alexandra Antonica, Beatrice Maggio, Sjuul Richters, and Vladislavs Sokolovs for excellent research assistance.

†Tilburg University, Department of Intelligent Systems. Email: A.Hendrickson@tilburguniversity.edu

‡University of Michigan Ross School of Business and CESifo. Email: ftruffa@umich.edu.

§Barnard College of Columbia University, Department of Economics and CESifo. Email: aywong@barnard.edu.

1 Introduction

How novel ideas are adopted and recognized is crucial to scientific progress. Yet not all ideas are equally likely to be recognized by the scientific communities, as visibility and credit often depend on the identity of the researcher (Cheng and Weinberg, 2021). Under the *diversity–innovation paradox*, diverse perspectives tend to generate more novel ideas, but ideas associated with women and minority scholars are often marginalized or overlooked (Hofstra et al., 2020). It remains an open question on what can foster greater acceptance and recognition of ideas from underrepresented groups.

Can creating a more *diverse academic environment* lead scientific communities to be more receptive to diverse perspectives? One important but understudied dimension of diversity comes from changes in the composition of the undergraduate student body, which can alter the social and intellectual environment in which research takes place. Increasing undergraduate gender diversity may influence what ideas are produced and recognized by changing both the research agendas of existing faculty and the broader ecosystem. This impact may depend on whether individuals and institutions are open to engaging with new perspectives: as exposure to diversity increases, some research areas may evolve and incorporate a wider range of ideas, while others remain largely unchanged or even resistant to new influences.

In this paper, we investigate how undergraduate diversity affects research production by studying the historical transition to coeducation in U.S. universities, when elite all-male institutions first enrolled women between 1960 and 1990, leading to a sharp increase in gender diversity on campuses. Specifically, we are interested in understanding whether coeducation shifted the direction of research toward topics historically associated with female researchers. To answer this question, we analyze over four million publications from the Microsoft Academic Graph. We construct two measures of female-associated content at the paper level, using the paper titles and abstracts: (1) a log odds ratio of words appearing more frequently in female- versus male-authored papers within the same field, and (2) a semantic similarity score derived from OpenAI text embeddings of paper titles and abstracts. These measures capture the extent to which new research adopts linguistic and thematic patterns historically associated with women scientists.

Using a generalized difference-in-differences design that exploits the timing of when universities switched to coeducation, we find a modest shift toward female-associated research following coeducation. Instead, coeducation produced polarizing effects across fields. In fields that were already more female-represented before 1960, such as biology, psychology,

and the social sciences, female-associated publications increased, controlling for total publications. Citations to research written by women also increased, suggesting an increase in recognition for female-authored work. By contrast, more male-dominated fields experienced a decline in both total output and recognition to female-associated topics.

This divergence is driven largely by new authors publishing for the first time at the institution, and, to a lesser extent, by incumbents adapting their research agendas and writing more on female-associated topics in female-dominated fields. We find limited evidence of compositional shifts arising from differential attrition or existing researchers switching fields.

These results suggest that increasing gender diversity at the university led to a divergence in the direction of research after coeducation. Coeducation led to new interest and recognition for topics historically associated with women in fields already open to such work, but had limited effects elsewhere. These findings highlight that while diversity can foster innovation, its effects may only be concentrated in areas already receptive to the new entrants.

Our study makes several contributions. First, this paper furthers our understanding of where ideas come from and how scientists choose their topics of research. Prior works that studied this question have focused on incentives and market forces such as competition from new entrants in the field (Borjas and Doran, 2012), perceived barriers to entry due to the presence of a superstar in the field (Azoulay et al., 2019), and the role of research funding (Myers, 2020). We add to this literature by showing that the exposure to diversity in the academic environment is also important for shaping the research agendas of individual scientists.

Second, a growing body of research examines how diversity shapes the direction and content of scientific inquiry. Much of this work focuses on the diversity of researchers themselves. For example, the entry and participation of women and underrepresented groups in science and its implications for the kinds of questions that get asked and the methods that are developed (Schiebinger, 2000; Rosser, 2002; Nielsen et al., 2018; Koning et al., 2019, 2020). This literature highlights how the demographic identity of the researchers can influence which topics are prioritized and which populations are studied. However, there has been comparatively less attention to the effects of *exposure* to diversity, that is, how contact with more diverse peers and environments might shift the research agendas of scientists, regardless of their own identity. A recent exception is Truffa and Wong (2025), who show that exposure to gender diversity in academic settings can shape the production of gender-related research. Building on this line of work, this paper contributes by leveraging new text-based methods to capture broader and subtler shifts in research content, not limited to

gender-related topics.

Third, a related strand of research examines diversity and the recognition of ideas. Several studies document that women and minority scholars are more likely to produce novel or interdisciplinary work (Yang et al., 2022), yet their contributions often receive less recognition in terms of citations or visibility (Huang et al., 2020; Hofstra et al., 2020; Cheng and Weinberg, 2021; Koffi, 2025). These findings underscore that the benefits of diversity may not be evenly distributed across contexts. Our analysis extends this literature by showing that diversity can have unequal effects on the adoption and recognition of women's research, depending on the openness of the field to underrepresented groups.

Finally, we also contribute to the literature on coeducation. A large literature, mostly focused on secondary schools, has examined the impact of coeducation on student outcomes, such as academic engagement, academic achievement, and participation in traditionally female/male subjects, and more recently, faculty research on women (Jackson, 2012; Park et al., 2018; Jackson, 2009; Lavy and Schlosser, 2011; Trickett and Trickett, 1982; Calkins et al., 2020; Truffa and Wong, 2025). We expand on these works by presenting new evidence that coeducation can have indirect effects on the production of research at universities, beyond student outcomes.

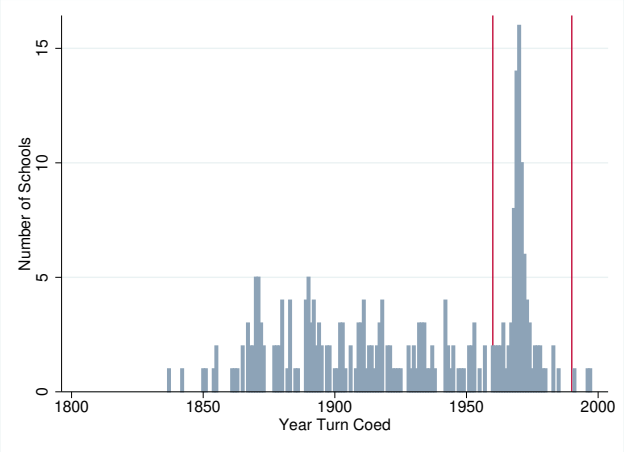
2 Background

2.1 History of University Coeducation in United States

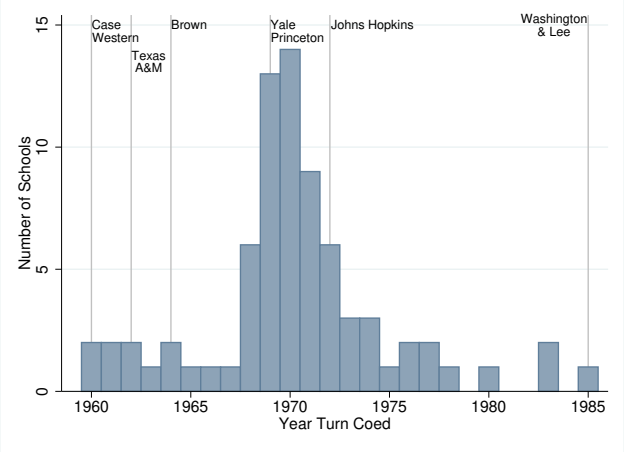
Coeducation in higher education has a long history in the United States, but until the 1960s, women were barred from entering some of the nation's most prestigious colleges and universities. In 1835, Oberlin College became the first coeducational university when it first admitted women (Goldin and Katz, 2011). As shown in Figure 1, the transition of four-year institutions from male-only to coeducation occurred at a steady pace throughout the time period but accelerated during the period between 1960 and 1990, denoted by vertical lines.[1] This latter period forms the time frame of our analysis and is associated with the era with the highest coeducation activities. In total, 88 institutions switched during this period, including prominent research institutions such as the Ivy League universities (Columbia, Dartmouth, Princeton, and Yale) and selective liberal arts colleges (e.g., Amherst, Wesleyan, Williams).

[1]Note that there were concurrent switches from female-only universities to coeducation throughout the entire time frame. However, this paper focuses on the male-only to coeducation switches.

Today, only three universities in the United States remain male-only.



(a) All Years



(b) Restricted to 1960-1990

Figure 1: Distribution of Turn Coed Years

Notes: This figure plots the frequency of coeducation dates for all universities that transitioned from male-only to coeducation for all years from 1800 to 2000 and for the sample period from 1960 to 1990. In (a), the vertical lines denote 1960 and 1990, respectively. In (b), vertical lines denote coeducation dates for a selected number of universities. Data comes from Coeducation College Database (Goldin and Katz, 2011).

It is important to note that while the sample period of interest was considered the celebrated historical period of coeducation, it did not play as important of a role in women's educational attainment as the earlier transitions did (Goldin and Katz, 2011). For instance, by 1924, over 75% of students were already enrolled in coed institutions and the switch to coeducation between 1960 and 1975 only increased the share of undergraduate women

taught in coeducation settings by 4 percentage points (Goldin and Katz, 2011). However, the transitions to coeducation of these universities enabled women to enter some of the most prestigious and research-productive universities.

2.2 Determinants of Coeducation During 1960-1990

The dramatic increase in coeducational universities during this period was driven by a combination of cultural and economic factors. The decades of the '60s and '70s were a period of political and social unrest. Increasingly, students demanded integration both in terms of gender and race, and sought to be educated at coeducational institutions (Miller-Bernal and Poulson, 2004). This shift in demand led to a reduction in enrollment growth and declining student quality at male-only universities compared to coeducational institutions (Goldin and Katz, 2011).

While cultural factors increased the demand for coeducation among students, male-only universities eventually switched to coeducation for financial reasons that were largely unrelated to universities' demand or supply of gender-related research, a feature that we will exploit in our empirical strategy (Miller-Bernal and Poulson, 2004; Goldin and Katz, 2011; Malkiel, 2016). As the historian Nancy Malkiel describes:

> "Coeducation happened because it was in the strategic self-interest of all-male institutions like Princeton and Yale to admit women. By the late 1960s, these schools were beginning to see their applications decline, along with their yields. The high school students they called the 'best boys' no longer wanted to go to all-male institutions [...] It was not the result of a high-minded moral commitment to opening educational opportunities to women, nor was it the result of deep thinking about how to educate women." (Malkiel, 2017, p.31)

This sentiment is also reflected in the way university presidents described the coeducation decision. One such example is a quote by Yale President Kingman Brewster Jr. in 1967 who declared in an address to alumni: "Our concern is not so much what Yale can do for women but what can women do for Yale." (Lozada, 2016, para. 5).

3 Data

We describe in this section the data used in our analysis.

3.1 University Data

To identify colleges that changed from single-sex to coeducation, we use the Coeducation College Database that was compiled and generously provided by Goldin and Katz (2011). The Coeducation College Database contains detailed information on the universe of four-year institutions (around 1,500) that existed from around 1897 to the present (Goldin and Katz, 2011). The data provides the year when each school turned coed, defined by Goldin and Katz (2011) as "one that has classes for men and women together. These classes must include the central ones in a liberal arts college and cannot be limited to a particular school, such as nursing or education." Given that our research purpose is to understand how turning coed changed faculty research interests, we believe this is an appropriate definition in our context.

Data on student enrollment (1968–1998), faculty (1971–1998) and degrees awarded (1965–1998) come from the Integrated Postsecondary Education Data System (IPEDS)[2] and its predecessor Higher Education General Information Survey (HEGIS)[3] collected by the United States Department of Education.

3.2 Publications Data

For data on scientific publications, we use the Microsoft Academic Graph (MAG) database (Sinha et al., 2015; Microsoft Research, 2018) and SciSciNet (Lin et al., 2023). MAG is a large publications database with information on over 207 million papers. The data include information on paper titles, abstracts, fields of study, researchers' names, and affiliation of the researcher at time of publication. SciScNet is a large-scale open data lake that builds upon the MAG dataset, augmenting the publication information on upstream funding sources such as NIH and NSF and downstream public uses, including references of scientific publications in patents, clinical trials, and media and social media mentions. While the publications database does not provide the gender of the researcher, SciSciNet uses the methodology from Van Buskirk et al. (2022) to provide an estimated probability that indicates to what extent a name belongs to an individual gendered female using the full name of the researcher. We assign a researcher as a female if this probability is over 50%.

In total, we have information on 3,931,488 publications written from 1950 to 2000 by researchers affiliated with one of the 1,253 four-year U.S. institutions that were successfully

[2]United States Department of Education. National Center for Education Statistics (1998)
[3]United States Department of Education. National Center for Education Statistics (1986)

matched from the Coeducation College Database to the MAG dataset.

3.3 Measures of Female-Associated Content

To quantify the extent to which research content reflects topics historically associated with women, we construct two complementary measures based on paper titles: a word-based measure using log odds ratios and a semantic measure using text embeddings.

3.3.1 Log Odds Ratio

The first measure identifies words that were disproportionately used by female authors during the baseline pre-period between 1950 and 1959. Within each field, we compute the log odds ratio of word frequencies in paper titles authored exclusively by women relative to those authored exclusively by men. Specifically, for each word w, we compare its empirical frequency in titles written by all-female and all-male teams:

$$\text{Log Odds}(w) = \log\left(\frac{p_w^F/(1-p_w^F)}{p_w^M/(1-p_w^M)}\right).$$

where p_w^g is the empirical share of all words used by teams of gender g in publications written between 1950 and 1950. This statistic measures whether a word appears relatively more frequently in female-authored versus male-authored research titles. Positive values indicate words used more frequently by women, while negative values indicate words used more frequently by men.

In our main analysis, a paper is classified as female-associated if its title contains at least one word among the top 25 percent of the *field*-specific log-odds distribution.[4] Appendix Figure A.1 presents the distribution of log odds at the word-level across fields.[5] Interestingly, in fields such as Materials Science and Mathematics, we observe a bimodal distribution, suggesting limited overlap in the vocabulary used by male- and female-authored papers during the baseline period. In contrast, fields such as Art and Psychology exhibit a more continuous distribution centered around zero, indicating greater linguistic similarity between

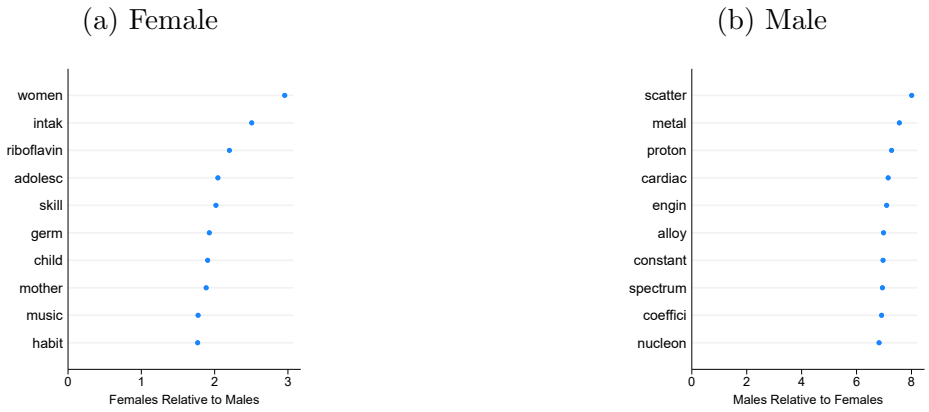
[4]Field refers to the top-level disciplinary classification in the Microsoft Academic Graph (MAG), corresponding to Level 1 "Fields of Study." Examples include Medicine, Economics, Engineering, and Psychology. A subfield refers to the Level 2 Field of Study within a given field. For example, within Medicine, subfields include Public Health, Immunology, and Pharmacology; within Economics, they include Development Economics, Macroeconomics, and Finance.

[5]Appendix Figure A.2 presents the distribution when we include words from abstracts.

male and female authors.

In Figure 2, we explore how language usage differs by gender of the researchers. Panel (a) shows the top 10 female-associated words based on the log-odds ratio across all fields, while Panel (b) shows the top 10 male-associated words. Female researchers are more likely to discuss topics related to women, children, education, and social aspects. In contrast, men are more likely to discuss concepts in physical sciences and engineering disciplines. We present the corresponding analysis by fields in Humanities (Appendix Figure A.3), Social Sciences (Appendix Figure A.4, STEM (Appendix Figure A.9), and Medicine (Appendix Figure A.6).[6]

Figure 2: Log Odds Ratio of Top Female and Male Words, 1950-1959



Notes: Panel (a) shows the top 10 words with the highest female-to-male log odds ratio used the titles of in all publications written between 1950 and 1959. Panel (b) presents the analogous results for male-to-female log odds ratio. Female publications are written by all-female teams and similarly, male publications are written by all-male teams.

3.3.2 Semantic Similarity

The second measure captures semantic similarity to female-authored research using large language models. This approach builds on methods introduced by Garg et al. (2018), who use word embeddings to quantify gender and ethnic stereotypes in written text over time. Specifically, we represent each paper title as a high-dimensional vector using OpenAI text embeddings, which encode semantic content beyond literal word overlap.[7] For each paper,

[6]Appendix Figures A.7, A.8, A.9, and A.10 present the corresponding results when we include words from abstracts.

[7]We also show our results are robust to using both the paper title and abstract.

we compute its average pairwise cosine similarity to every all-female-authored paper and to every all-male-authored paper published in the same field during the 1950s. As in Garg et al. (2018), the *difference* between these two similarity scores reflects how closely the language and themes of a paper resemble those of female versus male authors from the baseline period.

We then classify papers to be of "high semantic similarity towards female papers" if the paper is in the top 25% of the *field-year-specific* distribution of the relative similarity score. The comparison is made within the field and year because both the vocabulary and focus of research are likely to evolve over time from the 1950s baseline. This measure complements the word-based approach by capturing nuanced shifts in framing, language, and research focus that are not limited to specific terms. Appendix Figure A.11 plots the histogram of the relative semantic similarity score by field across all papers. The figure shows there is substantial heterogeneity in similarity to female-authored papers across all fields.

3.4 Female-Associated Research and Recognition

Before turning to our main empirical strategy, we present in this section three key descriptive facts using our measures of female-associated research. First, Figure 3 shows that in the full sample, even though there is significant overlap, female researchers are more likely to produce publications that are semantically more similar to all-female authored papers than all-male authored papers from the baseline period. This suggests that indeed, women are more likely to work on traditionally female-associated research.

Second, 4 shows that while publications with all-female teams tend to receive fewer citations than all-male teams, research that are more *female-associated* receive fewer citations regardless of the gender composition of the team.[8] We measure this association using two measures: (panel a) average log odds, defined as the average log-odds ratio of the specific words appearing in a paper's title based on their gendered frequency and (panel b) semantic similarity score of the title. Across both metrics, a clear citation penalty emerges as research topics align more closely with female-associated themes, though this penalty is notably less severe for mixed-gender teams.

Third, in Figure 5, we show that the words that were historically more female-associated are used more frequently in recent publications. This suggests that female-associated ideas from the 1950s are not obscure or niche topics today, but have instead become increasingly prevalent, even though they are less likely to be cited.

[8]Citations are normalized based on field and year of publication.

Figure 3: Semantic Similarity Score Distribution by Gender

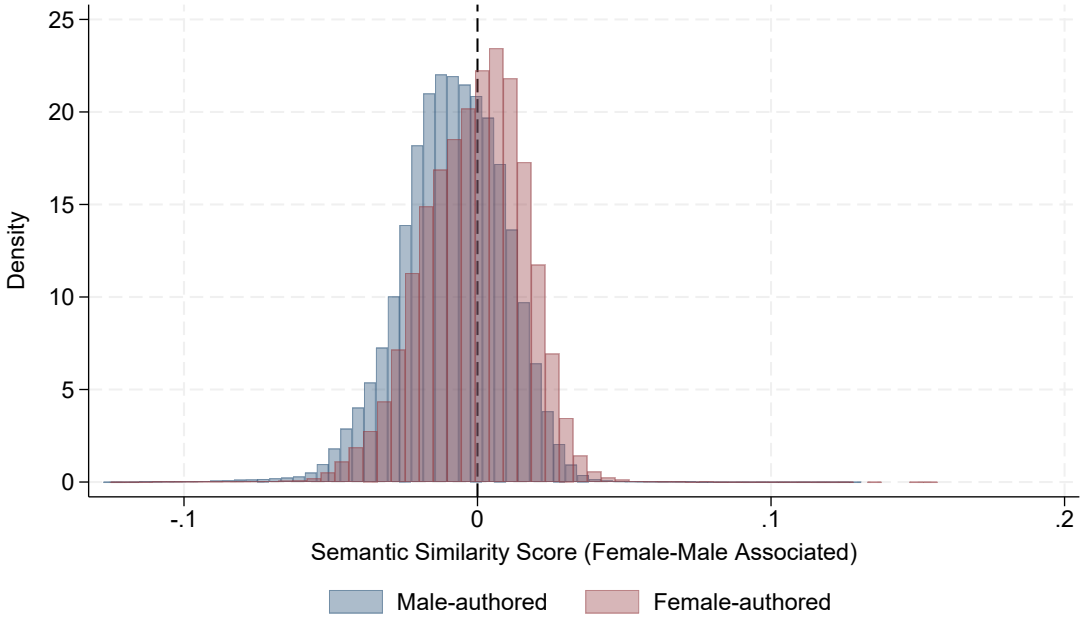
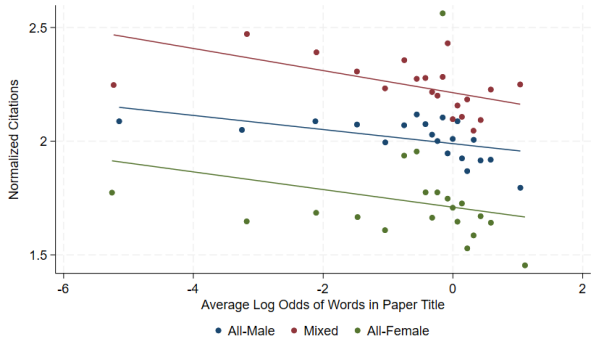


Figure 4: Relationship between Citations and Female-Associated Research by Team Composition

(a) Average Log Odds



(b) Semantic Similarity

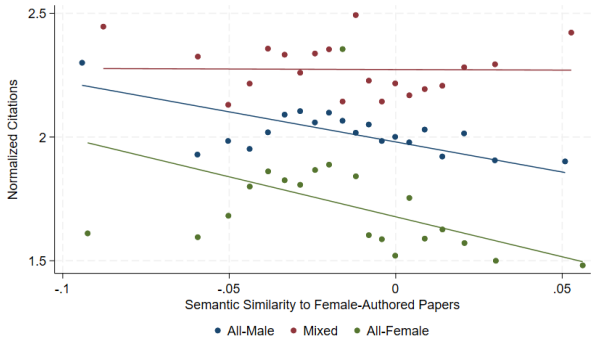
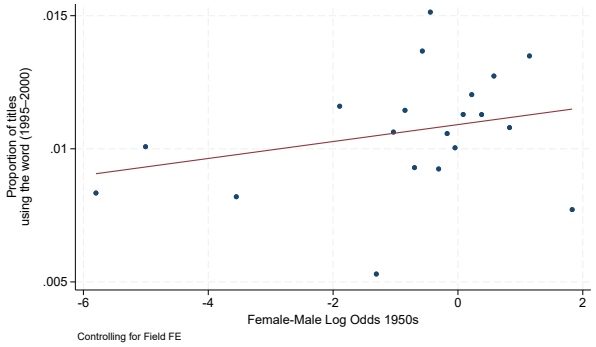
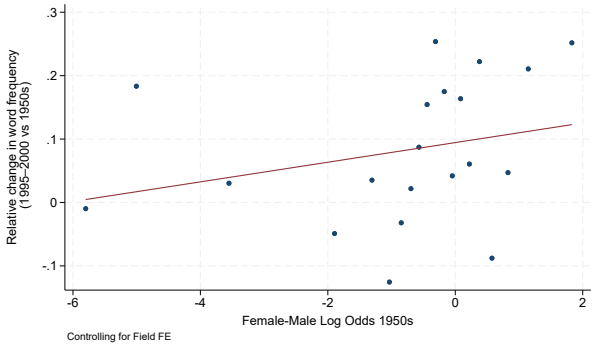


Figure 5: Prevalence of Female-Associated Topics

(a) Word Prevalence in 1995-2000



(b) Growth of Word Usage from 1950s to 1995-2000



4 Empirical Strategy

To identify the causal impact of coeducation on student and research outcomes, we employ a generalized difference-in-differences design with staggered adoption. We estimate the model:

$$y_{ist} = \sum_{\tau \neq -1} \beta_\tau YearRelativeCoed_\tau + \theta_{is} + \delta_t + \gamma_{dt} + \epsilon_{ist} \tag{1}$$

where y_{ist} is the outcome of interest, such as total number of publications at school i in the subfield s in year t.[9] θ_{is} are school-subfield fixed effects, δ_t are year fixed effects, which control for overall trends in the outcome, and γ_{dt} are discipline-by-year fixed effects, which account for changes in research output that are common within disciplines. For each school, we denote $\tau = 0$ as the first full year of coeducation. We include the full set of event time dummies, $YearRelativeCoed_{id\tau}$. The effect at $\tau = -1$ is normalized to zero.

In our specification, we use all available observations between $\tau \geq -20$ and $\tau \leq 20$.[10] In the estimation, we pool all observations prior to $\tau \leq -6$ together as well as for $\tau \geq 11$. We include all event-time dummies between $\tau \geq -5$ and $\tau \leq 10$ in the regression.

In our baseline specification, we use as our control group, the top 100 most productive universities that opened as coeducation. To capture productivity, we used the total number

[9]The subfield is defined as the first sub-classification of the broad field of study (e.g., labor economics is a subfield of economics).

[10]Note that our sample is not balanced because the date ranges for our data are: MAG 1950–2005, enrollment 1968–1998, faculty 1971–1998, and degrees awarded 1965–1998.

of research publications up to 2021. For this group of universities, only the dummy for $\tau = -1$ is turned on and all other event-time dummies are assigned to zero. Including this control group helps us control for any overall changes in research production trends and alleviate concerns related to heterogeneous treatment effects highlighted by the recent difference-in-differences literature (Roth, 2024). Standard errors are clustered at the school level.

Then, exploiting the fact that certain subfields are more likely to be affected by the enrollment of female students, we estimate a modified version of (1):

$$y_{ist} = \sum_{\tau \neq -1} \beta_\tau YearRelativeCoed_\tau \times 1(AboveMedian)_s + \theta_{is} + \delta_t + \gamma_{dt} + \epsilon_{ist} \qquad (2)$$

In this version, we interact the event-time dummies with $1(AboveMedian)_s$, a dummy for whether the subfield is above median in the percentage of female-authored papers from 1950-1959 across all U.S. universities. Throughout the paper, we will refer to this group of subfields as "early female-dominated fields." These early female-dominated fields are likely to be more affected if there is an increase in female students. For example, these may be fields where female students are more likely to enter, which can affect faculty research through classroom interaction. For examples of the fields that we assign to be early female-dominated, in Appendix Table A.1, we present the top 15 subfields with the highest female paper share between 1950 and 1959. The most female-dominated subfields in the decade prior to coeducation are nursing, environmental health, and virology. In contrast, in Table Appendix A.3, we present the subfields with the lowest share. These include many fields in engineering and mathematics.

5 Results

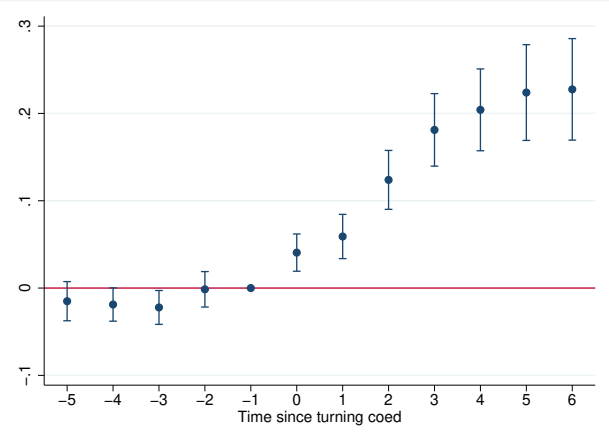
In the first part of our analysis, we show that turning coed had an immediate impact on the gender diversity of the student population at the universities of interest. We then turn to research outcomes and show that turning coed positively affected the research production in early female-dominated fields.

5.1 Effects of Turning Coed on Student Body

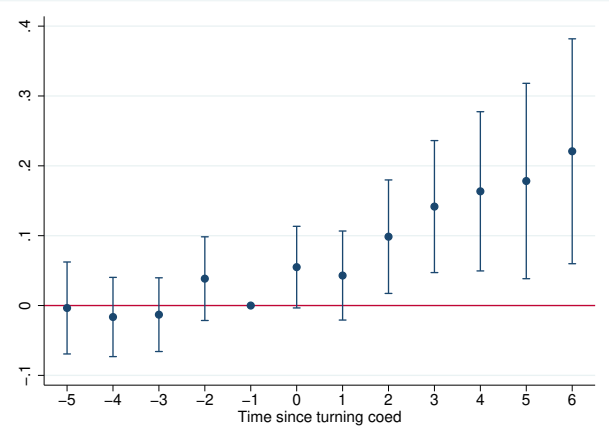
We first consider how transitioning to coeducation altered the composition of students at the universities. Figure 6 plots the event time coefficients and their 95% confidence intervals from estimating equation (1) for the female share of bachelor's degrees awarded and log number of total bachelor's degrees awarded. In Appendix Figure A.12, we present the analogous results using fall enrollment data. We summarize the magnitudes of these findings by presenting the average effects over the years 0 to 2 and 3 to 6 in Table 1. The female share of bachelor's degrees significantly increased by 21 percentage points. Relative to the baseline share of 7%, this indicates that women made up 28% of the graduating classes three to six years after coeducation.

Total degrees increased by 18% during the same period. Given that the baseline class size is 539, this implies an increase of 97 students. Using the earlier estimates on the increase in female enrollment, this suggests that there was an increase of 141 female students and a decrease of seven male students. This means that the schools largely increased the student body size to accommodate the new female students rather than substituting female students for male students. The magnitudes are similar when we consider fall enrollment instead of degrees awarded. We explore heterogeneous effects by field of study in Appendix Section G. We find the largest increase in female degree share in the fields of geography, philosophy, psychology, and sociology.

Figure 6: Effect of Turning Coed on Undergraduate Student Body



(a) Female Bachelor's Degrees Share



(b) Log Bachelor's Degrees Awarded

Notes: Figure 6 plots the event time coefficients and their 95% confidence intervals from estimating (1) for female share of bachelor's degrees awarded and log number of total bachelor's degrees awarded. Regressions are estimated using OLS. We include school fixed effects and year fixed effects. We cluster at the school level.

Table 1: Effect of Turning Coed on Student Body

	Bachelor's Degrees Awarded		Fall Enrollment	
	(1)	(2)	(3)	(4)
	Female Share	Log Total	Female Share	Log Total
Years -5 to -2	-0.014	0.001	0.001	0.014
	(0.009)	(0.025)	(0.010)	(0.033)
Years 0 to 2	0.075	0.065	0.115	0.083
	(0.012)	(0.032)	(0.013)	(0.028)
Years 3 to 6	0.209	0.176	0.251	0.181
	(0.025)	(0.063)	(0.021)	(0.061)
Baseline Mean	0.07	5.99	0.06	7.57
Observations	1789	1789	1184	1184
Estimator	OLS	OLS	OLS	OLS

Notes: Table 1 reports the average effects from estimating equation (1) on student outcomes. Each column reports estimates from a separate regression. Effect at event time $\tau = -1$ is normalized to 0. Baseline mean is the mean of the outcome variable at $\tau = -1$. The coefficient for Years -5 to -2 is the pre-period average of the coefficients for $\tau = -5$ to $\tau = -2$. The coefficient for Years 0 to 2 is the post-period average of the coefficients for $\tau = 0$, $\tau = 1$, and $\tau = 2$. Similarly, the coefficient for Years 3 to 6 is the average of the coefficients for $\tau = 3$ to $\tau = 6$. All regressions include school fixed effects and year fixed effects. All regressions are estimated using OLS. Standard errors in parentheses are clustered at the school level.

5.2 Effect of Turning Coed on Female-Associated Research

Having established that coeducation leads to an increase in gender diversity, we now explore whether this change has an impact on the types of research produced at the universities. Figure 7 presents the regression estimates for female-associated research output using our two measures. We find limited effects on the number of publications using female-associated words (based on the log-odds ratio) and a suggestive increase on the number of papers with high semantic similarity to female papers, though standard errors are large. Because these regressions control for total publications, it suggests that, on average, within subfields, there is a limited shift in research output towards these female-associated research.

We then investigate whether these effects vary based on early female representation of

the field. We hypothesize that fields with higher initial female representation may be more open to ideas associated with women, leading to greater shifts in research content following coeducation. Figure 8 presents the results for total female-associated papers by whether the subfield is above or below the median in the share of female researchers. The patterns reveal that the effects of coeducation were concentrated in fields that already had higher female representation in the 1950s. In these fields, we observe a gradual increase in the number of papers using female-associated words and in those with higher semantic similarity to female-authored work in the years following coeducation. By contrast, fields with below-median early female representation show a slight decline over time.

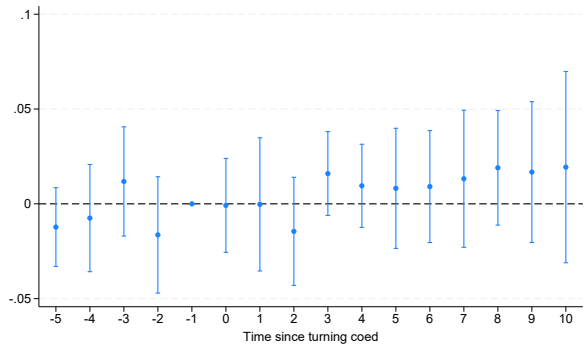
Appendix Figures A.13 and A.14 present the analogous results by gender of the research team for total number of papers using female-associated words and with high semantic similarity towards female papers, respectively. These results show that teams with at least one male researcher are driving much of the increase in the use of female-associated words in early female-represented fields. At the same time, we find an increase in female-associated papers based on semantic similarity for teams of all genders, suggesting a broad shift in language and content towards traditionally female-associated research.

In Figure 9, we show that the diverging pattern between subfields with higher and lower female representation in the 1950s continues to persist even when we rank the subfields *within* fields. This suggests that the findings are not driven by overall changes across different fields of study (e.g. psychology vs. physics), but rather reflect within-field variation in openness to ideas associated with women. In other words, even within the same scientific field, subfields that historically included more women appear more likely to incorporate female-associated research topics following coeducation.

One potential concern is that the shift towards female-associated research may lead to decrease in the overall impact of the research output. We explore whether coeducation impacted the production of "high-impact" female-associated research. We define high-impact papers as those meeting one of the following criteria: (i) being in the top 5% or 10% of the field-specific citation distribution for their publication year (Figure 11), or (ii) being published in a top-tier journal, defined as the top quartile of the Scimago Journal Ranking as of 2025.. Both metrics indicate that in female-dominated subfields, there is a significant increase in high-impact, female-associated research, suggesting that the thematic shift did not come at the expense of scientific impact.

Figure 7: Effect of Turning Coed on Female Associated Research

(a) Total Number of Papers Using Female-Associated Words



(b) Total Number of Papers with High Semantic Similarity Towards Female Papers

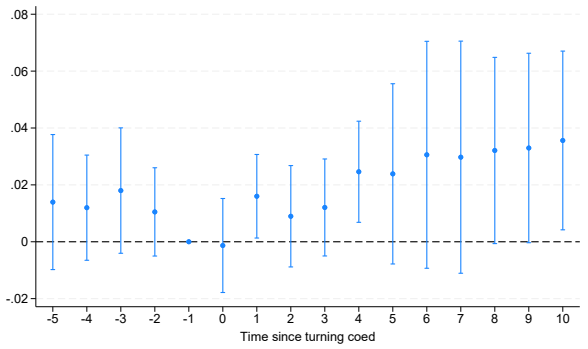
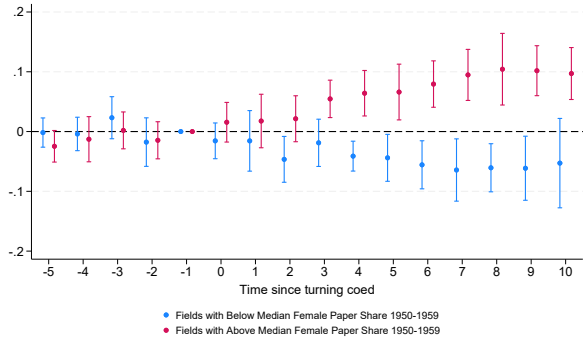


Figure 8: Effect of Turning Coed on Female Associated Research by Early Female Representation

(a) Total Number of Papers Using Female-Associated Words



(b) Total Number of Papers with High Semantic Similarity Towards Female Papers

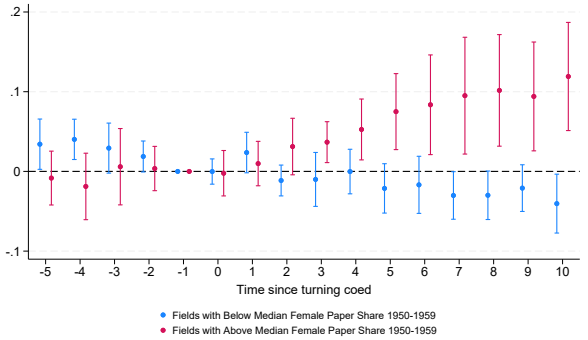
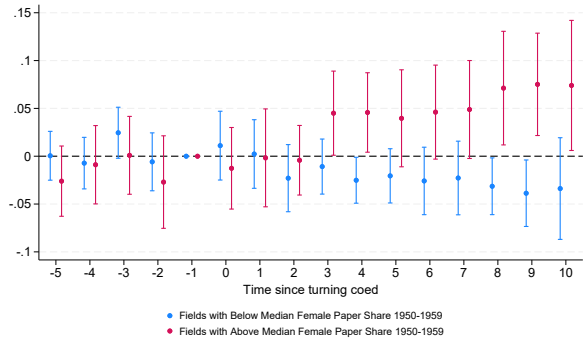


Figure 9: Effect of Turning Coed on Female Associated Research by Early Female Representation within Field

(a) Total Number of Papers Using Female-Associated Words



(b) Total Number of Papers with High Semantic Similarity Towards Female Papers

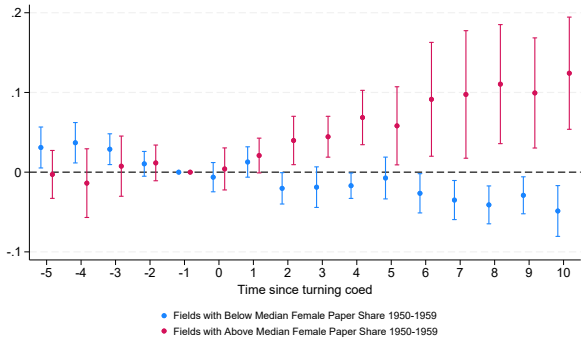
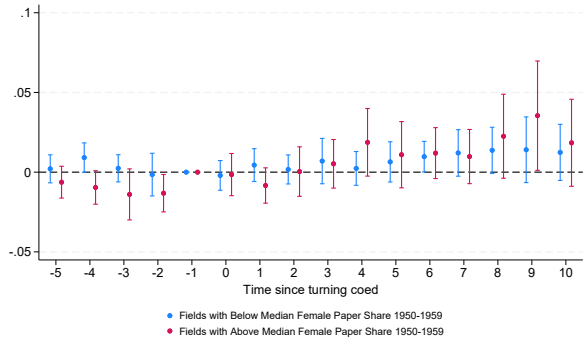
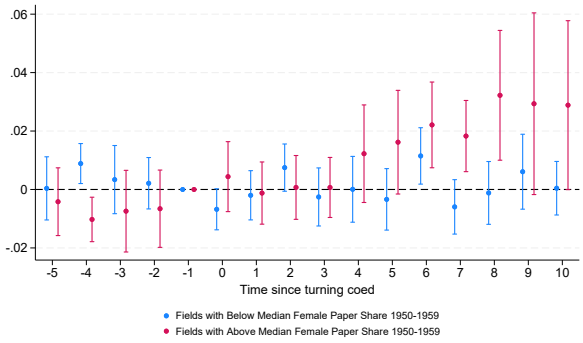


Figure 10: Effect of Turning Coed on Hit Female Associated Research by Early Female Representation

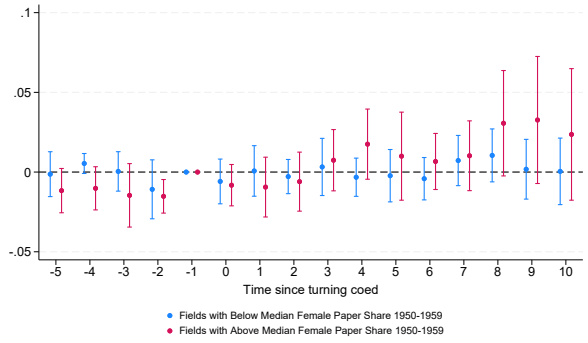
(a) Total Number of Papers Using Female-Associated Words (Top 5% Citations)



(b) Total Number of Papers with High Semantic Similarity Towards Female Papers (Top 5% Citations)



(c) Total Number of Papers Using Female-Associated Words (Top 10% Citations)



(d) Total Number of Papers with High Semantic Similarity Towards Female Papers (Top 10% Citations)

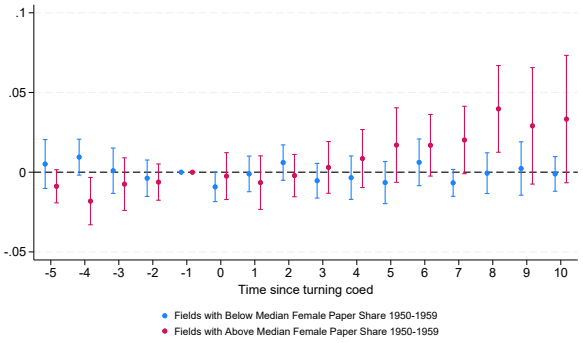
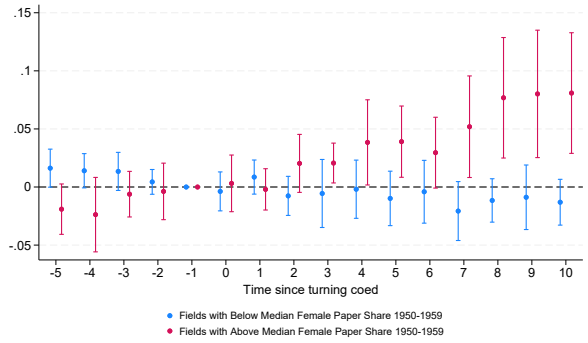
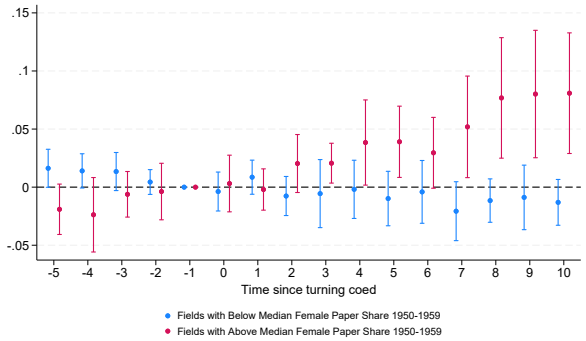


Figure 11: Effect of Turning Coed on Top Journal Publications by Early Female Representation

(a) Total Number of Top Journal Publications Using Female-Associated Words



(b) Total Number of Top Journal Publications with High Semantic Similarity Towards Female Papers

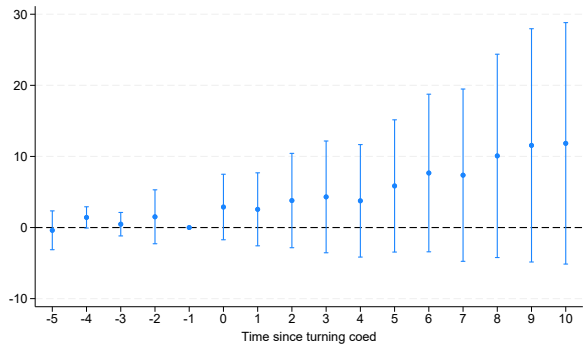


5.3 Effect of Turning Coed on Recognition of Female-Authored Research

We then investigate the impact of coeducation on *recognition* of research authored by female researchers. We use two measures of recognition at the subfield level: (i) the total number of female-authored papers that are cited, and (ii) the total number of female-authored papers that are cited by men. In Figure 12, we find a positive but imprecise increase in both measures, suggesting that female-authored papers received more recognition (also by men) after the schools turned to coeducation. Finally, Figure 13 shows that this increase in recognition is driven by fields with initially more women. Instead, we find no effect in fields with a baseline female paper share below the median.

Figure 12: Effect of Turning Coed on Recognition of Female-Authored Research

(a) Total Number of Female-Authored Papers Cited at the Subfield Level



(b) Total Number of Female-Authored Papers Cited by Men at the Subfield Level

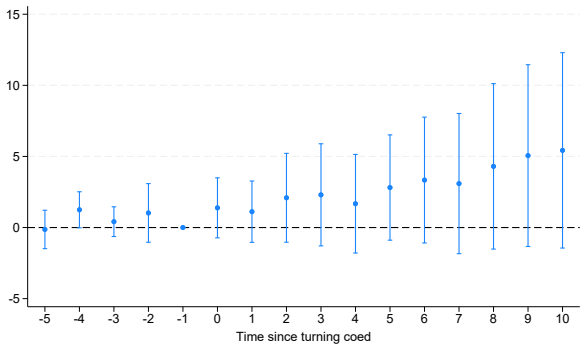
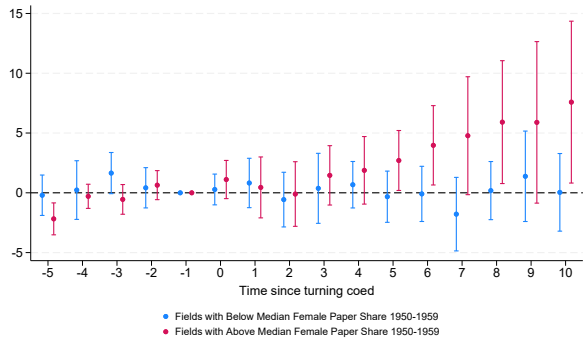
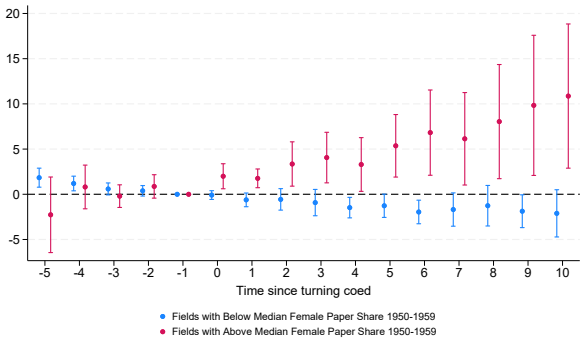


Figure 13: Effect of Turning Coed on Recognition of Female-Authored Research by Fields

(a) Total Number of Female-Authored Papers Cited at the Subfield Level



(b) Total Number of Female-Authored Papers Cited by Men at the Subfield Level



5.4 Additional Outcomes

In this section, we investigate how coeducation affected research output and research impact.

Effect of Turning Coed on Research Output First, we are interested in understanding how coeducation affects research output. Appendix Figure A.15 shows that coeducation had limited effect on the number of total publications. When we categorize fields based on the share of female-authored papers in the 1950s, we find no significant effect in fields where the baseline share of female-authored papers was above the median (Appendix Figure A.16). Fields with a baseline female paper share below the median show a suggestive decline.

Effect of Turning Coed on Research Impact We then explore whether coeducation had any effect on research impact or quality. We use three measures of impact, both at the subfield level: (i) the average normalized number of citations, (ii) the total number of top hit papers defined as the papers at the top 10% of the citation distribution, and (iii) total number of publications published in top-ranking journal, as defined by the top quartile of the Scimago Journal Ranking as of 2025. Appendix Figure A.17 shows no significant effect of coeducation on these two measures. The null effect is true in both sets of subfields, below and above median female-representation in the 1950s, though there is a suggestive increase in total publications in top journals for the female-dominated subfields (Appendix Figure A.18).

5.5 Robustness Checks

We present a series of robustness checks for our results. First, we show our results are robust to using alternative groups of control universities (Appendix Figures A.19 and A.20). These include (i) removing all control schools and estimating a pure event study, (i) universities that were always male-only and never switched, (iii) universities that were always female-only and never switched, and (iv) universities that were always coeducational prior to 1940. Second, we show our results are also robust to using field-by-year fixed effects (Appendix Figure **??**), rather than the baseline discipline-by-year fixed effects.

6 Existing Researchers vs. Compositional Changes

The observed shift toward female-associated research and increased recognition for female-authored research following coeducation could arise due to two potential effects: existing researchers changing the focus of their work, or compositional changes in who is conducting research. In this section, we will show that existing researchers suggestively shifted their research towards female-associated research, while significantly increasing their citations to female-authored research in the early female-represented fields. At the same time, there is a shift in the composition of who is producing research at the university, driven by new entries.

6.1 Existing Researchers

We first explore whether coeducation led to a shift in research interests for incumbent researchers. To do so, we construct an author-level dataset in which we assign researchers to

their modal fields of study based on their publication history. We then restrict the sample to those who have published at least one paper prior to coeducation. Note that for this analysis, we do not include control universities as we cannot define incumbents for researchers at universities that never switch. We then estimate equation (1) at the author level with individual fixed effects to control for any systematic differences across researchers in propensity to produce female-associated research.

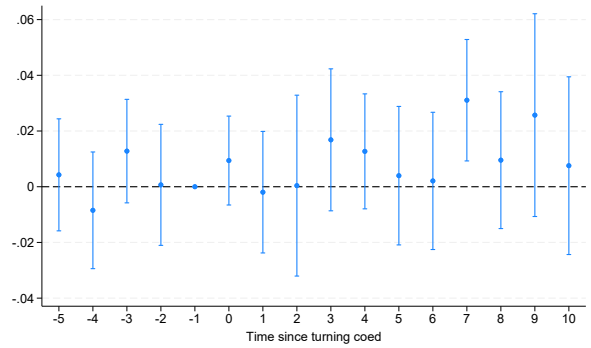
Figure 14 presents suggestive evidence that incumbent researchers shifted their research output towards female-associated research. Interestingly, unlike the subfield level results, we do not find substantial heterogeneity based on the early representation of the subfield (Appendix Figure A.22). Appendix Figures A.23 and A.24 report results separately for male and female incumbents, using both the log-odds and semantic similarity measures of female-associated research. We find the positive shift in research towards female associated research comes from both male and female researchers with somewhat larger effects for female researchers.

Figure 15 examines whether these shifts extend to the recognition of female-authored research by existing researchers, as measured by citations to female-authored papers. The results indicate a gradual increase in citations to female-authored work after coeducation, particularly in fields with higher early female representation. In Appendix Figure A.25, we further disaggregate the results by the gender of incumbent researchers and the early female representation of their fields. Among male incumbents, we observe the divergent pattern across subfields. Male researchers increase citations to female-authored work following coeducation, but only in early female-represented subfields. Instead, we observe a decline in citations in fields with below-median early female representation. In contrast, there is a limited effect for female incumbents in both types of subfields. These results suggest that exposure to greater gender diversity may have encouraged male researchers in more inclusive fields to engage more with female-authored work, whereas those in traditionally male-dominated fields are less responsive or even resistant to such shifts.

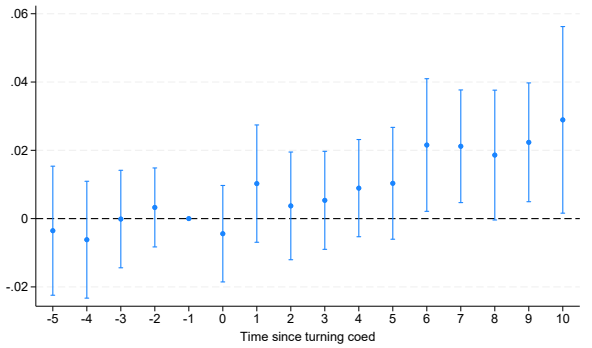
Together, these results suggest coeducation increased adoption and recognition for women's research for incumbent faculty, though the effects for female-associated research are imprecisely estimated.

Figure 14: Effect of Turning Coed on Female Associated Research by Early Female Representation for Incumbent Researchers

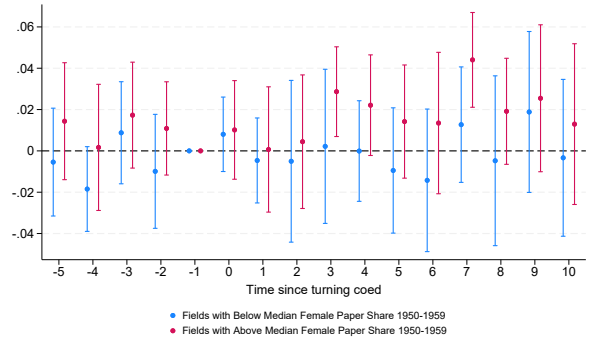
(a) Total Number of Papers Using Female-Associated Words



(b) Total Number of Papers with High Semantic Similarity Towards Female Papers



(c) Total Number of Papers Using Female-Associated Words



(d) Total Number of Papers with High Semantic Similarity Towards Female Papers

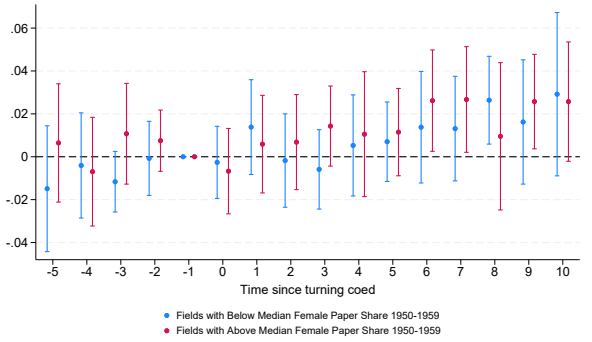
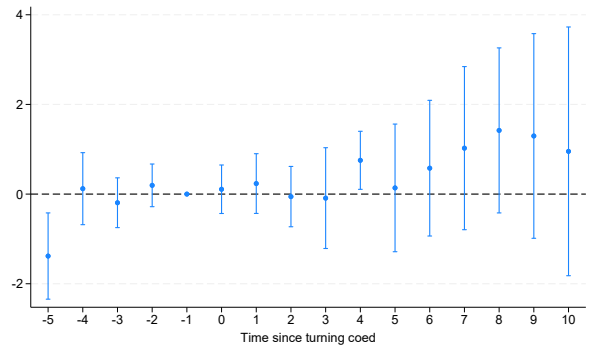
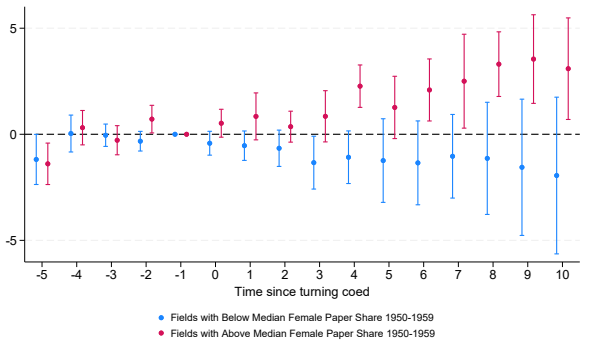


Figure 15: Effect of Turning Coed on Recognition of Female-Authored Research for Incumbent Researchers

(a) Total Number of Female-Authored Paper Cited



(b) Total Number of Female-Authored Paper Cited by Early Female Representation



6.2 Compositional Effects

We next examine how much of the observed effects can be attributed to compositional changes within universities. In addition to shifts in research interests of existing researchers, increases in female-associated research may also arise from changes in who conducts research and where. First, coeducation may have altered the distribution of researchers across fields, that is, individuals who are more interested in female-associated research may be more likely to move into more female-represented areas after coeducation. Such re-sorting would mechanically generate an increase in female-associated output in above-median fields and a decline in below-median ones. Second, there may be differential attrition, where researchers more interested in female-associated topics are more likely to remain in fields with higher early female representation. Third, the effects could reflect changes in hiring, if universities began recruiting scholars whose research interests were more aligned with female-associated topics following coeducation.

6.2.1 Researcher Mobility Across Fields

We first examine whether the observed increase in female-associated research stems from mobility across fields after coeducation. In particular, we focus on researchers who were already active at the university, but began publishing in a new subfield after coeducation, whom we refer to as "switchers." Switchers are defined as individuals who publish in an above- or below-median field (based on early female representation) for the first time, but who had published previously in a below- or above-median field at the same university.

Appendix Figure A.26 presents event-study estimates for this group. Panel (a) shows the number of new authors entering above- or below-median fields, while Panels (b) and (c) display the number of female-associated papers produced by switchers using the log-odds and embedding-based measures, respectively. Across all measures, we find no systematic increase in the number of switchers or in their production of female-associated research following coeducation. The divergence observed in overall research output therefore does not appear to be driven by researchers switching into more female-represented fields.

These findings suggest that the effects of coeducation on research direction are unlikely to result from mobility across fields within universities. However, as our data capture publications rather than researchers' field affiliations or intentions, these results should be interpreted with some caution.

6.2.2 Differential Attrition

We next examine whether the observed changes in research direction reflect differential attrition among researchers. To assess this, we focus on departures, defined as researchers who publish for the last time at a given university in a particular year.

Appendix Figure A.27 presents event-study estimates for departing researchers. Panel (a) shows the total number of researchers leaving the university each year, while Panels (b) and (c) report the number of departing researchers who had previously written female-associated papers, based on the log-odds and embedding-based measures, respectively. Across all specifications, we find no evidence of a differential effect in departures following coeducation. The number and composition of exiting researchers remain stable before and after coeducation, suggesting that the overall divergence in female-associated research output is not driven by attrition.

6.2.3 New Entrants and Hiring

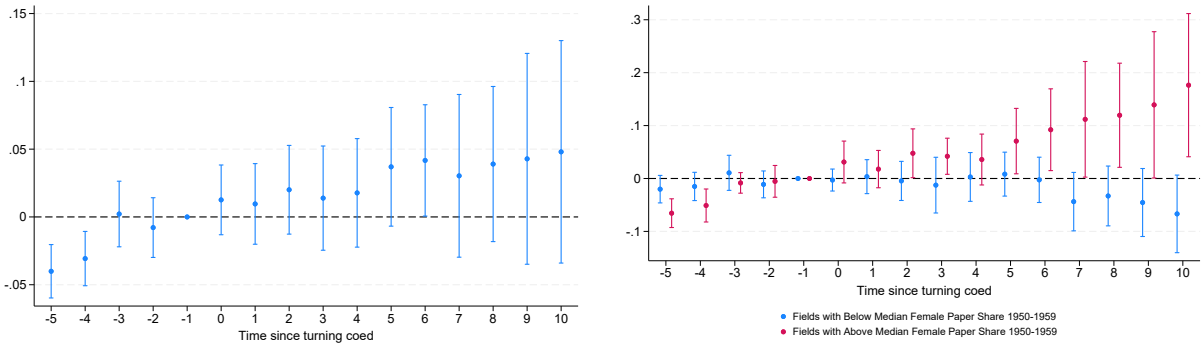
We next investigate whether the observed changes in research direction are driven by new entrants to the university following coeducation.

First, we show that after coeducation, there is a suggestive increase in female researchers. Figure 16 presents event-study estimates for the total number of female authors, controlling for total number of authors at the subfield level. We find a suggestive increase in the number of female authors, especially in early-female-represented fields. Instead, fields with a baseline female paper share below the median show a small but negative effect. Since female researchers are more likely to investigate female-associated topics, this pattern can explain the increase in female-associated research in these subfields.

We then examine whether new entrants, those publishing for the first time at a university, contribute to shifts in research direction. Figure 17 displays the corresponding event-study results. Panel (a) shows number of new authors, while Panels (b) and (c) report female-associated papers using the log-odds and embedding-based measures, respectively. The results reveal that, while there is no significant effect on the number of new authors, first-time authors in early-female fields are more likely to produce female-associated research, whereas new authors in more male-dominated fields are less likely to do so. This divergence mirrors the patterns observed at overall subfield level.

Together, these results highlight that coeducation led to hiring practices that altered the research direction of the subfields, differentially by openness of the field.

Figure 16: Effect of Turning Coed on Gender Composition of Scientists

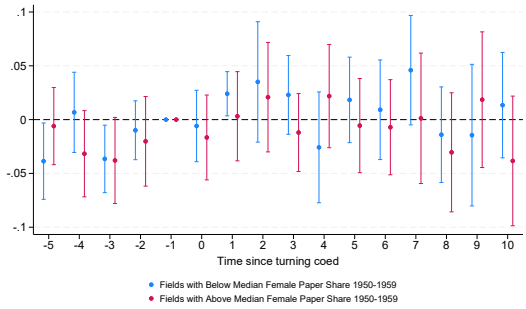


(a) Total Female Authors at the Subfield Level (b) Total Female Authors at the Subfield Level

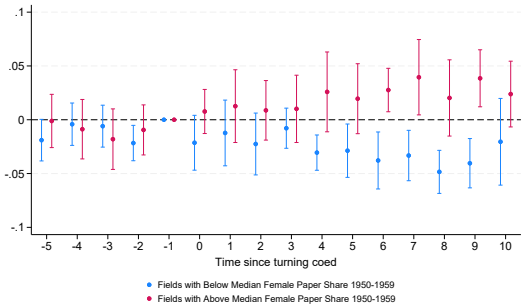
Notes: Figure 16 plots the event time coefficients and their 95% confidence intervals from estimating (1) for total number of female authors. Regressions are estimated using OLS. We include school fixed effects and year fixed effects. We cluster at the school level.

Figure 17: Effect of Turning Coed on Female-Associated Research by New Authors

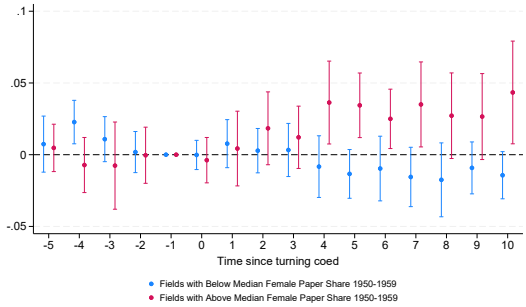
(a) Number of New Authors



(b) Female-Associated Papers by New Authors (Top 25% Log Odds)



(c) Female-Associated Papers by New Authors (Embeddings)



7 Mechanisms

In this section, we discuss potential explanations for why we observe positive effects for early female-dominated fields and suggestive negative effects for the more male-dominated fields. We first rule out alternative explanations related to differential trends in early female-dominated fields and changes in graduate student population. Then we consider two potential explanations for our effects: exposure to gender diversity and potential backlash due to the change in student composition.

7.1 Alternative Explanations

Differential Trends in Early Female-Dominated Fields One potential explanation is the early female-dominated fields happen to be fields that are already growing more at these universities prior to coeducation, and as a result, turning coed is coincidental to the

overall trends. We can rule out this explanation, however, due to the lack of pretrends in our difference-in-differences results.

Changes in Graduate Student Population Alternatively, we may think that at the same time as turning coed, schools begin to increase graduate enrollment or enroll female graduate students, especially in early female-dominated fields. As discussed in Truffa and Wong (2025), for some of these universities, graduate programs were coeducational for many years before undergraduate admissions.[11] To capture whether there is a differential effect on graduate student composition by field, we utilize the Degrees Awarded files from HEGIS/IPEDS. This data is only available at the major level, which can be matched to the field, not subfield, in the publications database. As a result, for these analyses, we use the female paper share in the 1950s at the broad field level to assign degrees to above or below median fields. Figure A.28 shows there are no significant differences in the effect of turning coed on either the size of the graduate student body (including both masters and doctorate degrees) or share of female students. This suggests that the differential effect we observe for research production is unlikely to be explained by differences across fields in terms of graduate student composition.

Broader cultural changes: Robustness to Using Region-by-Year FE We may also be concerned that the timing of coeducation is associated with broader cultural changes that could contemporaneously lead to shift in research agendas. Appendix Figure A.29 shows that our results are robust to the inclusion of region-by-year fixed effects which could control for broader changes in cultural changes at the regional level.

Differential Field Growth and Hot Topics A potential alternative explanation is that historically female-represented fields became trendier over time, experiencing faster growth and generating more "hot" topics that also contain female-associated language. However, we find little support for this mechanism. In Appendix Figure A.30, we present the relationship between the growth of publications in each subfield and early female representation. The correlation between subfield growth and the share of women in the 1950s is very small ($\rho = 0.07$), suggesting that differential field growth is unlikely to drive our results.

Differential R&D Expenditures At the same time coeducation occurred, universities may promote female-associated research and provide subfields with higher baseline female representation with additional funding and resources. However, this is unlikely, given that the primary reason for coeducation was financial rather than gender equality concerns. In

[11] For example, although the first female undergraduate arrived at Yale College in 1969, Yale's School of Art opened to female graduates in 1869 (https://guides.library.yale.edu/c.php?g=871411&p=6256097).

Appendix Figure A.31, we also find no evidence that R&D expenditures disproportionately shifted towards female-dominated departments within universities, based on data from the National Science Foundation.

Spillovers based on Same City as Treated Schools Lastly, in Appendix Figure A.32, we provide evidence that spillovers across universities are unlikely to bias our results. We find limited evidence of spillover effects of turning coed on universities outside of the sample in the same city, suggesting that our results are unlikely to be driven by spillovers across universities in the sample.

7.2 Exposure to Gender Diversity in the Academic Environment

Role of Undergraduate Gender Diversity

A natural explanation for the stronger effects in early-female fields is that coeducation increased exposure to gender diversity more meaningfully in these settings. Appendix Figure A.33 shows early-female-represented fields experience a larger rise in female undergraduate representation after coeducation. This suggests that researchers after coeducation in these fields are more likely to interact in a more gender-diverse environment.

Student interactions

There are several reasons why interaction with students may lead to a shift in research interests. First, as a result of coeducation and the change in the undergraduate gender composition, student demand for new topics may change, especially in these subfields. For example, in Truffa and Wong (2025), one observable shift was the increased discussion of gender-related topics in the university curricula. With the increase in female students, departments may have responded by introducing courses that reflect female students' interests, which are more likely to be on topics that are traditionally female-associated. Through classroom interactions, faculty may be exposed to these new ideas and become interested in these topics. It can also generate greater demand for faculty researchers who can teach these classes and produce research in these areas. To test this hypothesis, we collected and parsed course catalogues with full course descriptions for 22 universities that transitioned to coeducation during our sample period (Truffa and Wong, 2024).[12] We then apply our measure of female-associated language to course titles and descriptions, identifying whether a course contains words that fall within the top quartile of the female-associated log-odds distribution defined

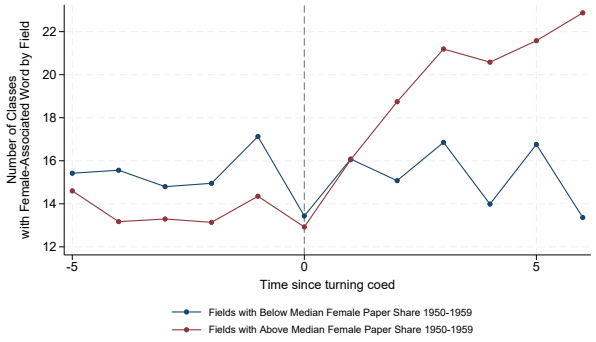
[12]These universities are selected because they made available scanned images of their full course catalogues online.

in the main analysis. Figure 18 shows an increase in the use of female-associated words in course titles and descriptions following the transition to coeducation, consistent with shifts in curricular content toward topics historically associated with female scholars.

Moreover, an additional mechanism through which coeducation may have influenced research is student involvement in the research process. Expanding student enrollment may increase the availability of student research assistants, who play a key role in faculty research productivity (Borjas et al., 2017). The influx of students into particular fields may have created greater complementarities in research production, where faculty could take on more ambitious research projects due to the expanded pool of student labor. This suggests that there may be expanded research capacity driven by the growing student body in female-associated topics in these early-female subfields. In line with this hypothesis, Figure 19 shows that the effects in our analysis are driven by universities that experienced the largest increase in arrival of female undergraduates.

Figure 18: Effect of Turning Coed on Classes Using Female-Associated Words

(a) Trends in Number of Classes Using Female-Associated Words



(b) Total Number of Classes Using Female-Associated Words

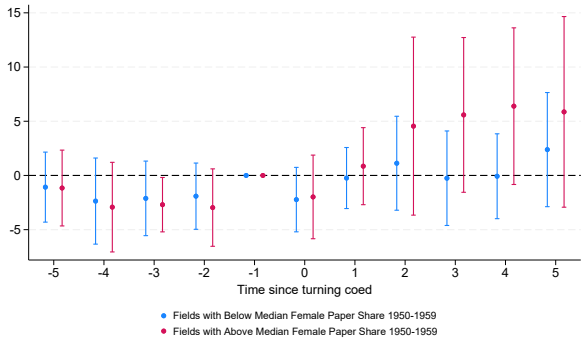
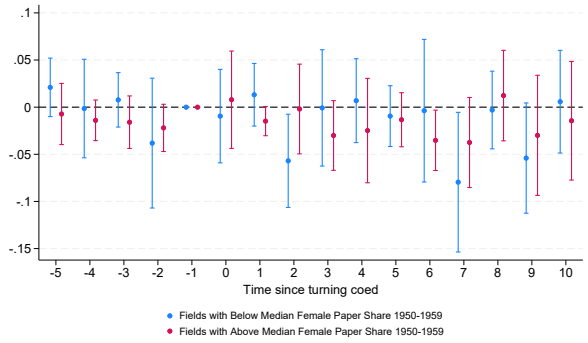
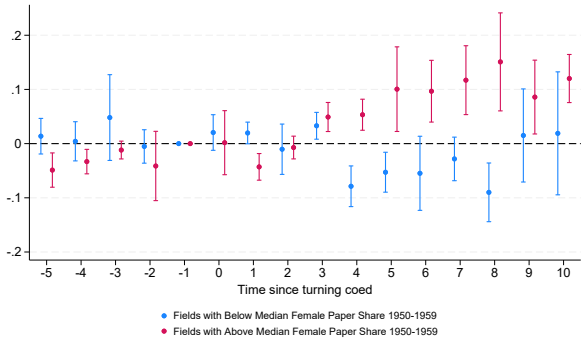


Figure 19: Heterogeneous Effect of Turning Coed on Female Associated Research by Change in Female Share of Bachelor's Degrees Awarded from $\tau = 1$ to $\tau = 3$

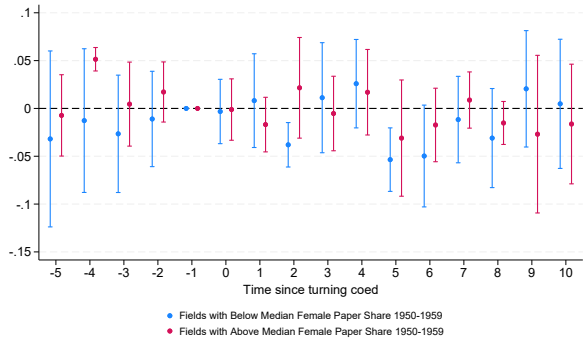
(a) Total Number of Papers Using Female-Associated Words (Universities with Below Median Change)



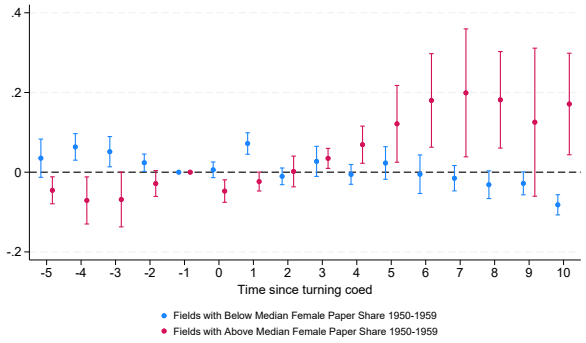
(b) Total Number of Papers Using Female-Associated Words (Universities with Above Median Change)



(c) Total Number of Papers with High Semantic Similarity Towards Female Papers (Universities with Below Median Change)



(d) Total Number of Papers with High Semantic Similarity Towards Female Papers (Universities with Above Median Change)

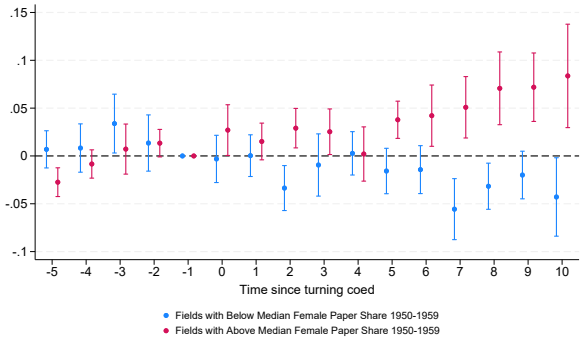


Peer interactions

Coeducation also altered the composition of academic peers. Early-female fields saw a larger influx of female and first-time authors who were themselves more likely to produce female-associated research. This change may have fostered more discussions among researchers, potentially leading to new research interests and collaborations in these topics. Figure 20 shows a clear divergence in the number of papers with at least one female coauthor following the transition to coeducation. Specifically, fields with an above-median female paper share in the 1950s experienced a sustained and significant increase in female coauthorship, beginning roughly two years after the policy change. In contrast, fields that were historically male-dominated show no such increase, and even experience a slight downward

trend in female coauthorship relative to the pre-coeducation period.

Figure 20: Effect of Turning Coed on Number of Papers with a Female Coauthor by Early Female Representation



7.3 Potential Backlash

One potential explanation for the suggestive negative effect in early non-female subfields is that coeducation may have led to a potential backlash response. Social psychology research has documented efforts to increase diversity can provoke defensive reactions among members of dominant groups (See Iyer (2022) for a review). Under the *resource threat* hypothesis, dominant groups may react defensively when increased diversity is perceived as a challenge to their control over valued positions, status, or resources (Blumer, 1958; Iyer, 2022).

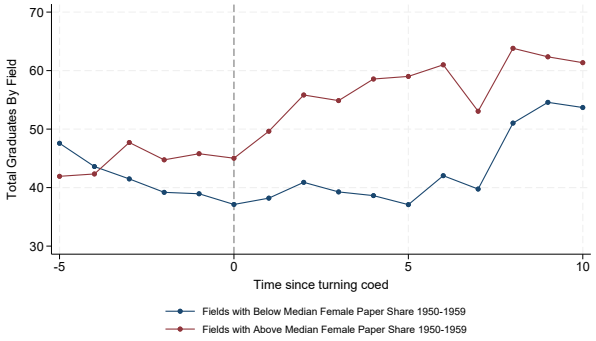
Consistent with this interpretation, we find that the effects of coeducation also changed the size and composition of academic fields. Figure 21 shows after coeducation, fields with above-median female paper share differentially grew more.

These results suggest that as a result of the admissions policy change, early female-dominated fields became larger departments, in terms of student enrollment, and have a greater increase in gender diversity. In response, under the resource threat hypothesis, more male-dominated subfields may have hired new faculty whose research interests are more aligned with traditional field topics or less oriented toward female-associated topics.

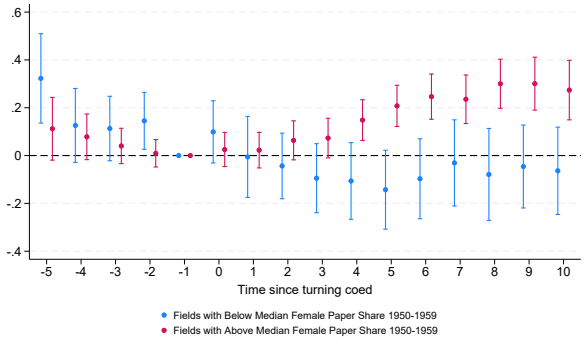
In ongoing work, because the HEGIS/IPEDS data series do not provide statistics on faculty at the department-level and we do not observe faculty if they are not publishing, we are digitizing historical faculty directories to explore these channels.

Figure 21: Effects of Turning Coed on Undergraduate Size by Field

(a) Total Bachelor's Degree by Field



(b) Log Bachelor's Degree by Field



8 Conclusion

This study provides new empirical evidence on whether and how diversity fosters the inclusion and recognition of women's research. While prior work has focused primarily on the role of female scientists in shaping research agendas, we highlight the broader impact of students on the production of research historically associated with female researchers and the recognition of female-authored work. Using the historical transition of U.S. universities from male-only to coeducation as a natural experiment, we document how the influx of women into higher education contributed to shifts in research composition, productivity, and recognition across fields.

Our findings indicate that the introduction of coeducation led to a significant increase in female enrollment. Although we find limited evidence of an overall shift toward female-associated research, coeducation produced divergent effects across fields. In disciplines that were already more female-represented in the 1950s, the share of female-associated publications increased, even after controlling for total research output. Citations to work authored by women also increased, suggesting greater recognition of female work. By contrast, in more male-dominated fields, both total research output and citations to female-associated topics declined.

This divergence appears to be driven primarily by new authors publishing for the first time at the institution, and to a lesser extent, by incumbent researchers reshaping their research agendas to include more female-associated topics in female-dominated fields. We find little evidence that these patterns reflect compositional changes due to differential attrition or

researchers switching fields.

A key mechanism underlying these effects was the change in gender diversity, especially in early female-represented fields. The influx of female students likely contributed to the increasing interest in female-associated research topics. Additionally, complementarities in research production, such as increased availability of student research assistants, may have further supported the observed rise in research output. Finally, interactions between incumbent and new researchers may have led to new discussions and interest in female-associated research.

In ongoing work, we will explore whether the decline in female-associated research in the more male-dominated subfields is due to backlash.

Overall, our study highlights the broader role of diversity on the direction of research after coeducation. Coeducation not only expanded access to higher education for women but also reshaped academic research. The findings suggest that changes in student demographics can have diverging effects on research agendas, faculty hiring, and recognition dynamics, based on the openness of the field to female research.

References

Azoulay, P., C. Fons-Rosen, and J. S. Graff Zivin (2019, August). Does science advance one funeral at a time? *American Economic Review 109*(8), 2889–2920.

Blumer, H. (1958). Race prejudice as a sense of group position. *Pacific sociological review 1*(1), 3–7.

Borjas, G. J. and K. B. Doran (2012). The Collapse of the Soviet Union and the Productivity of American Mathematicians. *Quarterly Journal of Economics 127*(3), 1143–1203.

Borjas, G. J., K. B. Doran, and Y. Shen (2017). Ethnic Complementarities after the Opening of China: How Chinese Graduate Students Affected the Productivity of Their Advisors. *The Journal of Human Resources 53*(1), 1–31.

Calkins, A., A. J. Binder, D. Shaat, and B. Timpe (2020). When sarah meets lawrence: The effect of coeducation on women's major choices.

Cheng, W. and B. A. Weinberg (2021). Marginalized and Overlooked: Minoritized Groups and the Adoption of New Scientific Ideas. *NBER Working Paper No. 29179*.

Garg, N., L. Schiebinger, D. Jurafsky, and J. Zou (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences 115*(16), E3635–E3644.

Goldin, C. and L. F. Katz (2011). Coeducation college database.

Goldin, C. and L. 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.

Hofstra, B., V. V. Kulkarni, S. Munoz-Najar Galvez, B. He, D. Jurafsky, and D. A. McFarland (2020). The diversity–innovation paradox in science. *Proceedings of the National Academy of Sciences 117*(17), 9284–9291.

Huang, J., A. J. Gates, R. Sinatra, and A.-L. Barabási (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the National Academy of Sciences 117*(9), 4609–4616.

Iyer, A. (2022). Understanding advantaged groups' opposition to diversity, equity, and inclusion (dei) policies: The role of perceived threat. *Social and Personality Psychology Compass 16*(5), e12666.

Jackson, K. (2012). Single-sex schools, student achievement, and course selection: Evidence from rule-based student assignments in Trinidad and Tobago. *Journal of Public Economics 96*(1).

Jackson, M. O. (2009). Student demographics, teacher sorting, and teacher quality: evidence from the end of school desegregation. *Journal of Labor Economics 27*(2), 213–256.

Koffi, M. (2025, July). Innovative ideas and gender (in)equality. *American Economic Review 115*(7), 2207–36.

Koning, R., S. Samila, and J.-P. Ferguson (2019). Female inventors and inventions. *Working Paper*.

Koning, R., S. Samila, and J.-P. Ferguson (2020, May). Inventor gender and the direction of invention. *AEA Papers and Proceedings 110*, 250–54.

Lavy, V. and A. Schlosser (2011). Mechanisms and Impacts of Gender Peer Effects at School. *American Economic Journal: Applied Economics 3*(2), 1–33.

Lin, Z., Y. Yin, L. Liu, and D. Wang (2023, June). Sciscinet: A large-scale open data lake for the science of science research. *Scientific Data 10*(315).

Lozada, C. (2016, October). How harvard, princeton and yale discovered women. *The Washington Post*.

Malkiel, N. (2016). Keep the Damned Women Out: The Struggle for Coeducation.

Malkiel, N. W. (2017, March). 'keep the damned women out: The struggle for coeducation in the ivy league, the seven sisters, oxford, and cambridge. *Proceedings of the American Philosophical Society 161*(1).

Microsoft Research (2018). Microsoft academic graph. `https://www.microsoft.com/en-us/research/project/microsoft-academic-graph/`. Accessed: 2018-04.

Miller-Bernal, L. and S. L. Poulson (2004). *Going Coed: Women's Experiences in Formerly Men's Colleges and Universities, 1950-2000*. Vanderbilt University Press.

Myers, K. (2020, October). The elasticity of science. *American Economic Journal: Applied Economics 12*(4), 103–34.

Nielsen, M. W., J. P. Andersen, L. Schiebinger, and J. W. Schneider (2018). One and a half million medical papers reveal a link between author gender and attention to gender and sex analysis. *Nature Human Behaviour 1*, 791–796.

Park, H., J. R. Behrman, and J. Choi (2018). Do single-sex schools enhance students' STEM (science, technology, engineering, and mathematics) outcomes? *Economics of Education Review 62*, 35–47.

Rosser, S. V. (2002). An overview of women's health in the US since the mid-1960s. *History and Technology 18*, 355–369.

Roth, J. (2024). Interpreting event-studies from recent difference-in-differences methods. Working paper.

Schiebinger, L. (2000). Has Feminism Changed Science? *Signs 25*(4), 1171–1175.

Sinha, A., Z. Shen, Y. Song, H. Ma, D. Eide, B. june (Paul) Hsu, and K. Wang (2015). An Overview of Microsoft Academic Service (MAS) and Applications. *Proceedings of the 24th International Conference on World Wide Web (WWW 2015 Companion)*, 243–246.

Trickett, E. and P. Trickett (1982). The independent school experience: aspects of the normative environment of single-sex and coed secondary schools. *Journal of Educational Psychology 73*(3), 374–381.

Truffa, F. and A. Wong (2024). Course catalog data. Data and Code for: "Undergraduate Gender Diversity and the Direction of Scientific Research." American Economic Association. Nashville, TN; distributed by Inter-university Consortium for Political and Social Research, Ann Arbor, MI.

Truffa, F. and A. Wong (2025). Undergraduate Gender Diversity and the Direction of Scientific Research. *American Economic Review*.

United States Department of Education. National Center for Education Statistics (1965–1986). Higher education general information survey (hegis). Inter-university Consortium for Political and Social Research.

United States Department of Education. National Center for Education Statistics (1984–1998). Integrated postsecondary education data system (ipeds). Inter-university Consortium for Political and Social Research.

Van Buskirk, I., A. Clauset, and D. B. Larremore (2022). An open-source cultural consensus approach to name-based gender classification. *arXiv preprint*.

Yang, Y., T. Y. Tian, T. K. Woodruff, B. F. Jones, and B. Uzzi (2022). Gender-diverse teams produce more novel and higher-impact scientific ideas. *Proceedings of the National Academy of Sciences 119*(36), e2200841119.

Online Appendix for "How Female Representation Changes Science: Evidence from College Coeducation" Drew Hendrickson, Francesca Truffa, and Ashley Wong

A Additional Figures and Tables

Figure A.1: Distribution of Log Odds Ratio by Field (Words from Titles)

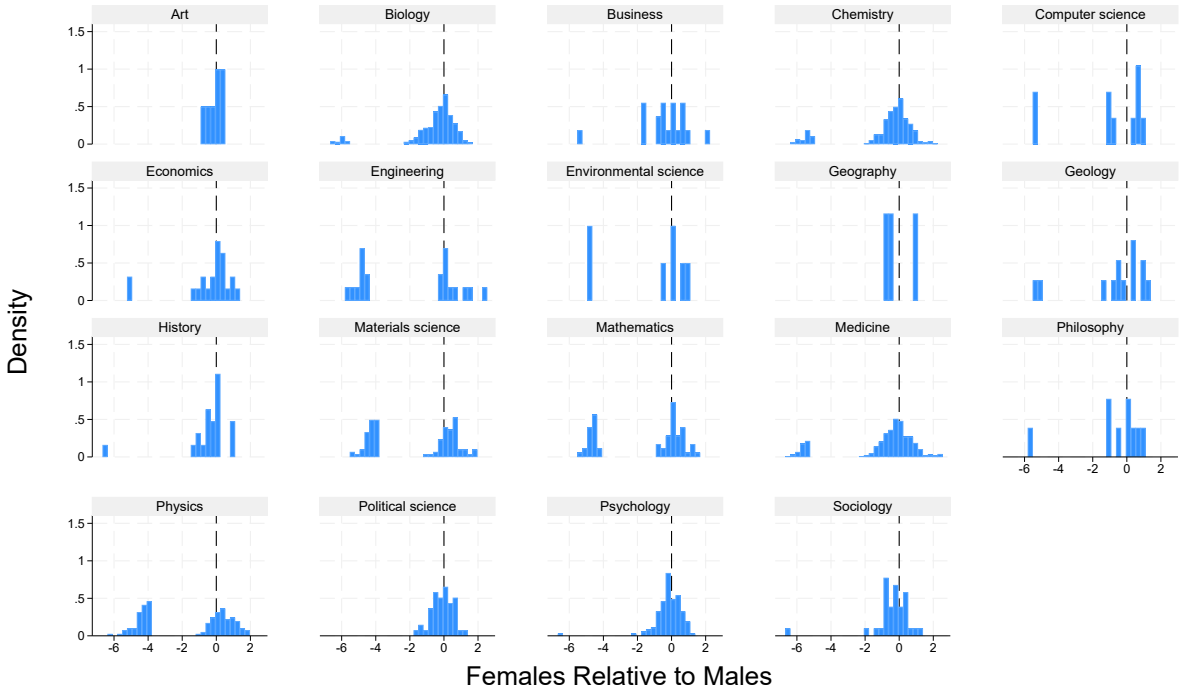


Figure A.2: Distribution of Log Odds Ratio by Field (Words from Titles and Abstracts)

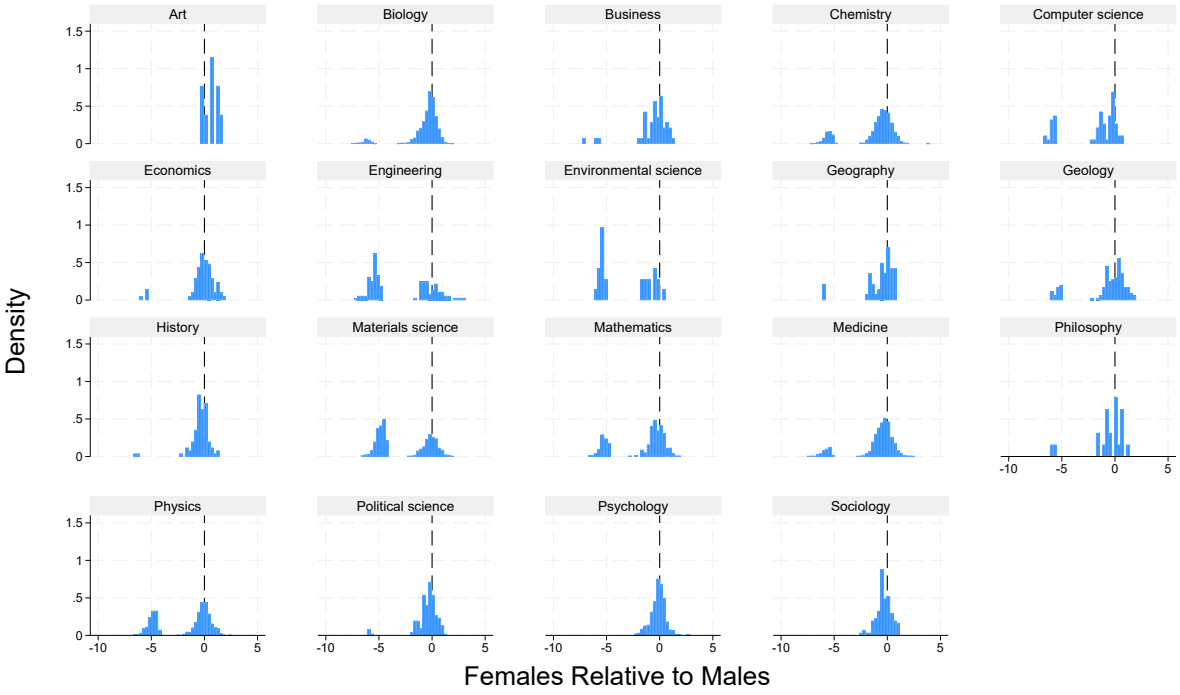
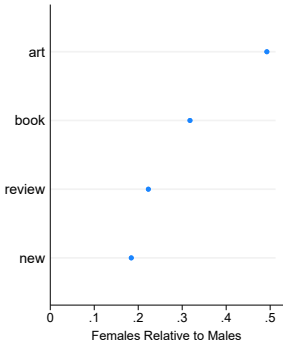
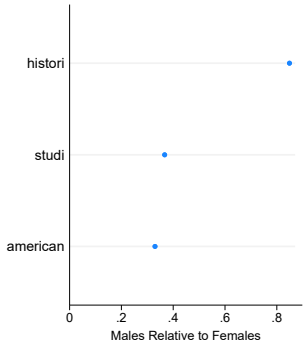


Figure A.3: Log Odds Ratio of Top Female and Male Words by Fields in Humanities, 1950-1959 (Words from Titles)

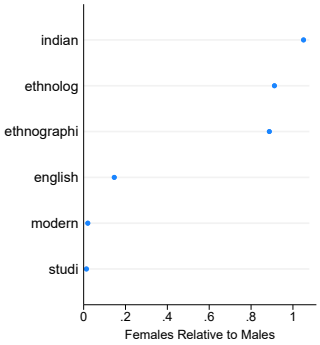
(a) Art - Female



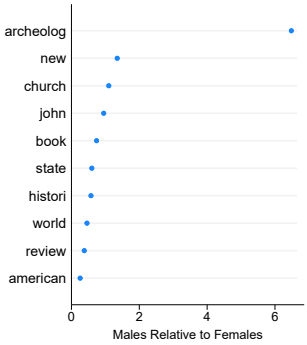
(b) Art - Male



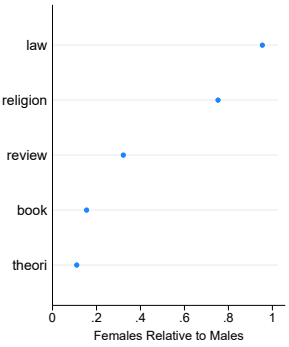
(c) History - Female



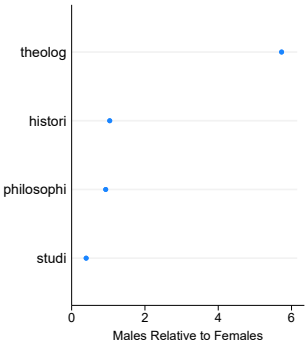
(d) History - Male



(e) Philosophy - Female



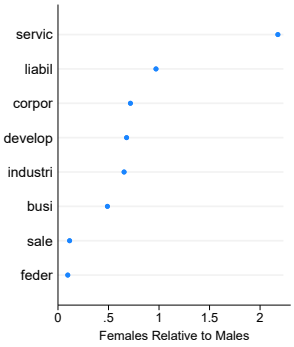
(f) Philosophy - Male



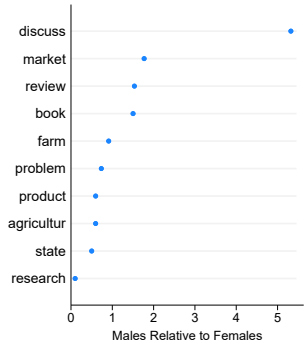
Notes: Panel (a) shows the top 10 words with the highest female-to-male log odds ratio used the titles of in all publications written between 1950 and 1959. Panel (b) presents the analogous results for male-to-female log odds ratio.

Figure A.4: Log Odds Ratio of Top Female and Male Words by Fields in Social Sciences, 1950-1959 (Words from Titles)

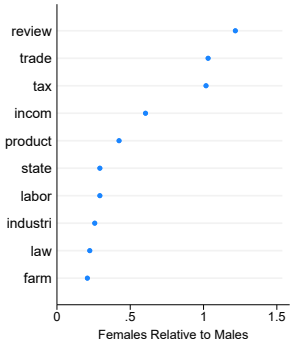
(a) Business - Female



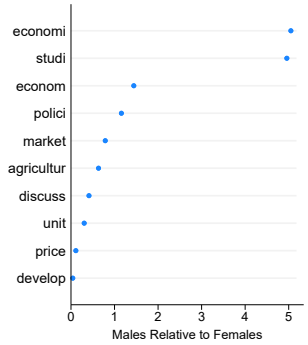
(b) Business - Male



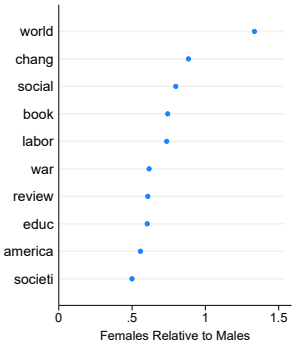
(c) Economics - Female



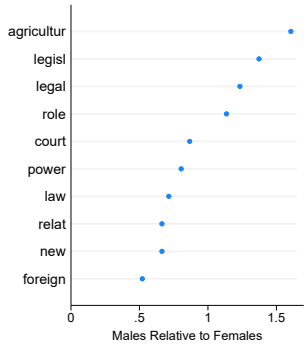
(d) Economics - Male



(e) Political Science - Female



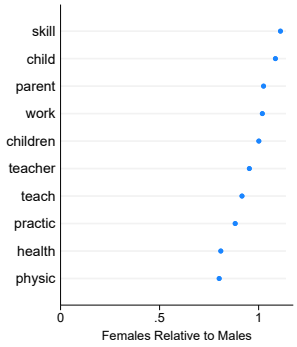
(f) Political Science - Male



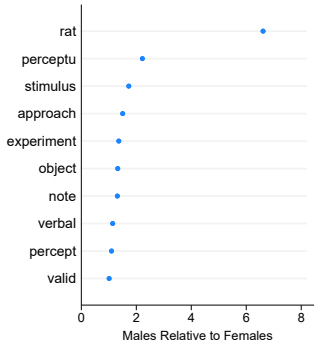
Notes: Panel (a) shows the top 10 words with the highest female-to-male log odds ratio used the titles of in all publications in the specified field written between 1950 and 1959. Panel (b) presents the analogous results for male-to-female log odds ratio.

Figure A.4: Log Odds Ratio of Top Female and Male Words by Fields in Social Sciences, 1950-1959 (Words from Titles) – Continued

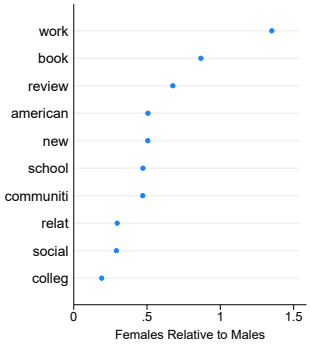
(g) Psychology - Female



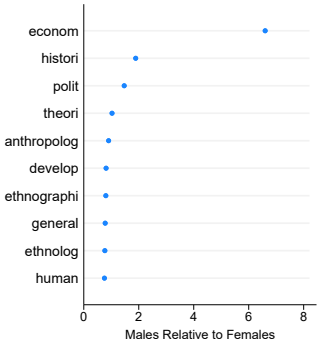
(h) Psychology - Male



(i) Sociology - Female



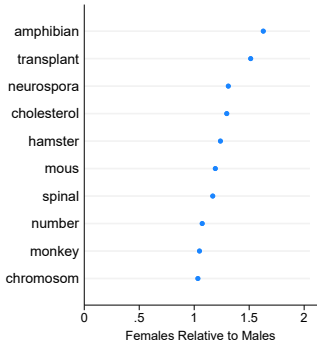
(j) Sociology - Male



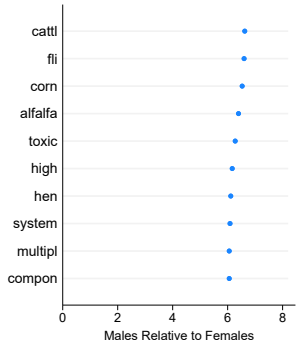
Notes: Panel (a) shows the top 10 words with the highest female-to-male log odds ratio used the titles of in all publications in the specified field written between 1950 and 1959. Panel (b) presents the analogous results for male-to-female log odds ratio.

Figure A.5: Log Odds Ratio of Top Female and Male Words by Fields in STEM, 1950-1959 (Words from Titles)

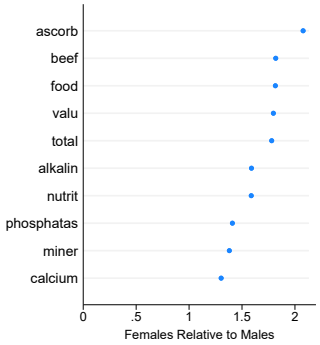
(a) Biology - Female



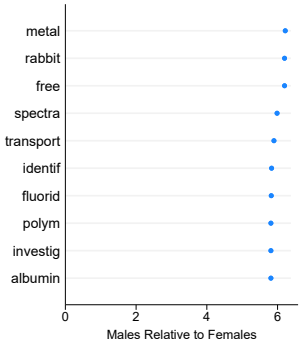
(b) Biology - Male



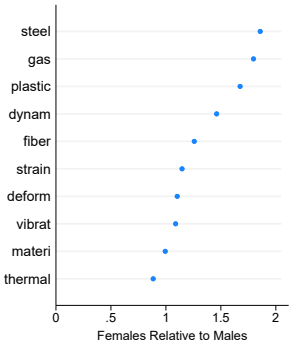
(c) Chemistry - Female



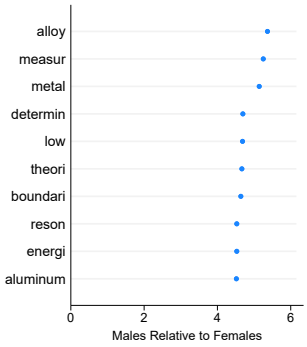
(d) Chemistry - Male



(e) Materials Science - Female



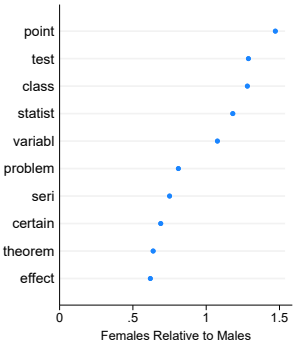
(f) Materials science - Male



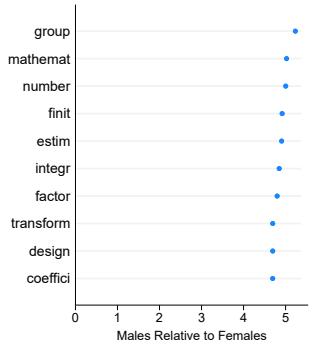
Notes: Panel (a) shows the top 10 words with the highest female-to-male log odds ratio used the titles of in all publications in the specified field written between 1950 and 1959. Panel (b) presents the analogous results for male-to-female log odds ratio.

Figure A.5: Log Odds Ratio of Top Female and Male Words by Fields in STEM, 1950-1959 (Words from Titles) — Continued

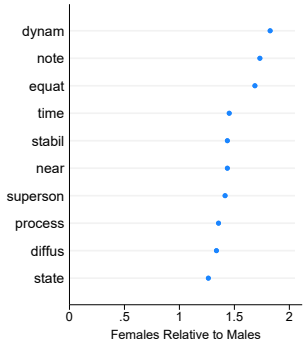
(g) Mathematics - Female



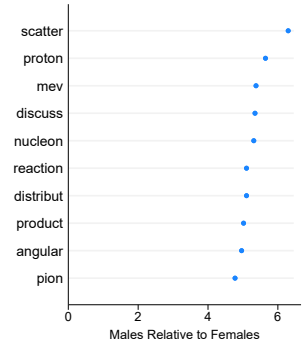
(h) Mathematics - Male



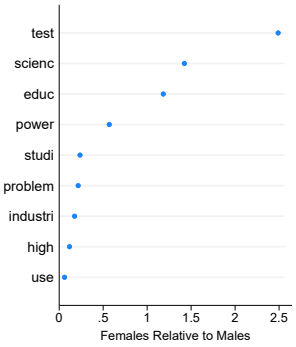
(i) Physics - Female



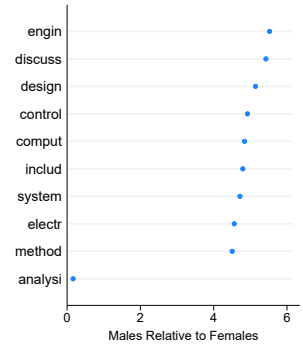
(j) Physics - Male



(k) Engineering - Female



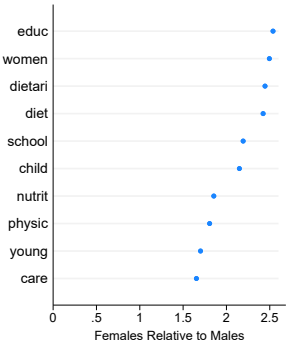
(l) Engineering - Male



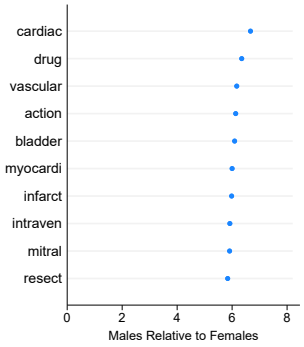
Notes: Panel (a) shows the top 10 words with the highest female-to-male log odds ratio used the titles of in all publications in the specified field written between 1950 and 1959. Panel (b) presents the analogous results for male-to-female log odds ratio.

Figure A.6: Log Odds Ratio of Top Female and Male Words in Medicine, 1950-1959 (Words from Titles)

(a) Medicine - Female



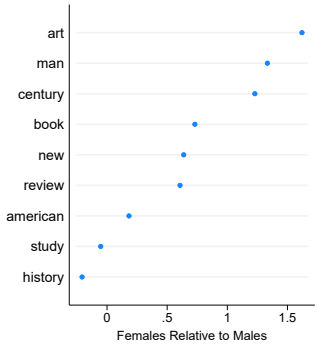
(b) Medicine - Male



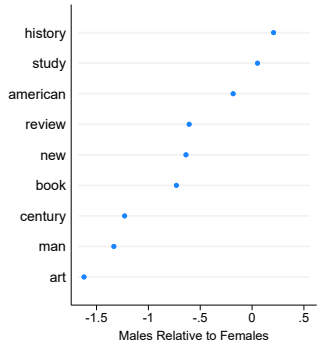
Notes: Panel (a) shows the top 10 words with the highest female-to-male log odds ratio used the titles of in all publications in the specified field written between 1950 and 1959. Panel (b) presents the analogous results for male-to-female log odds ratio.

Figure A.7: Log Odds Ratio of Top Female and Male Words by Fields in Humanities, 1950-1959 (Words from Titles and Abstracts)

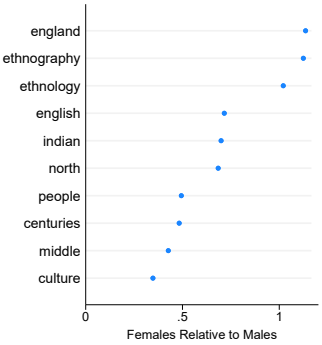
(a) Art - Female



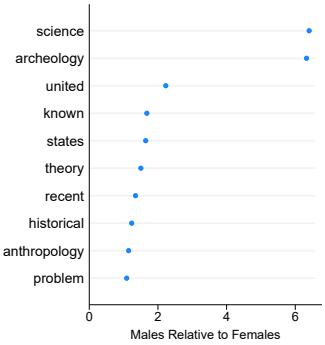
(b) Art - Male



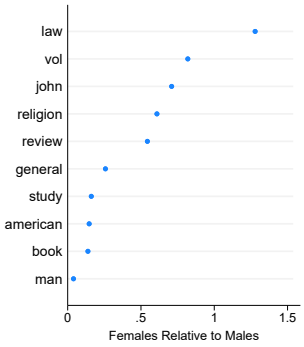
(c) History - Female



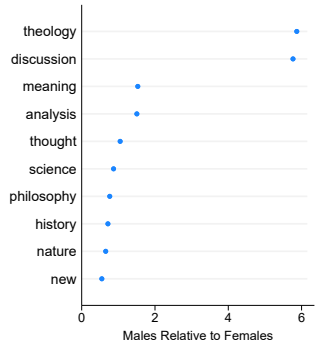
(d) History - Male



(e) Philosophy - Female



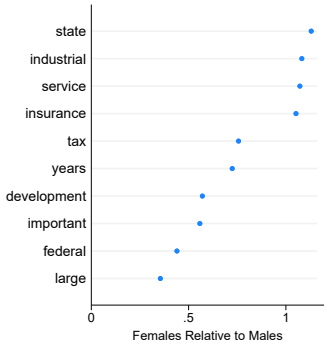
(f) Philosophy - Male



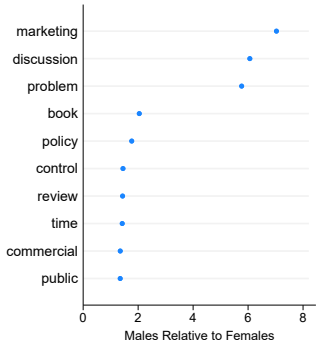
Notes: Panel (a) shows the top 10 words with the highest female-to-male log odds ratio used the titles of in all publications written between 1950 and 1959. Panel (b) presents the analogous results for male-to-female log odds ratio.

Figure A.8: Log Odds Ratio of Top Female and Male Words by Fields in Social Sciences, 1950-1959 (Words from Titles and Abstracts)

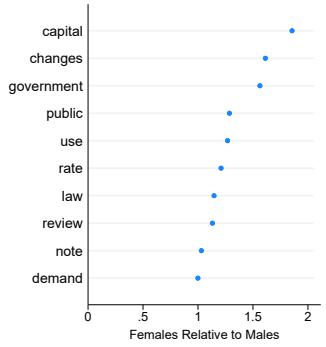
(a) Business - Female



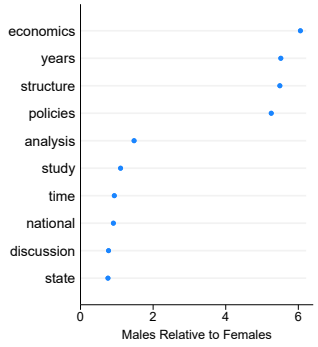
(b) Business - Male



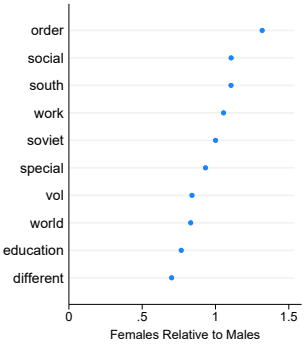
(c) Economics - Female



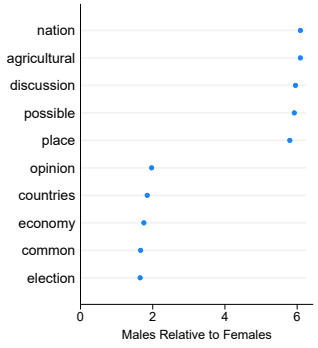
(d) Economics - Male



(e) Political Science - Female



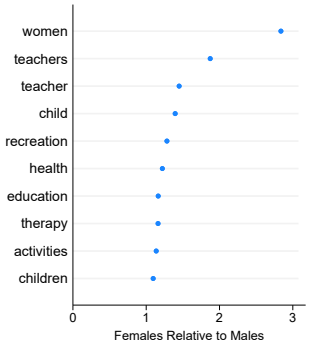
(f) Political Science - Male



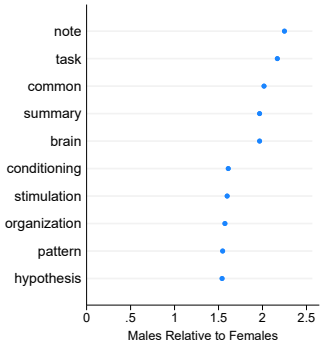
Notes: Panel (a) shows the top 10 words with the highest female-to-male log odds ratio used the titles of in all publications in the specified field written between 1950 and 1959. Panel (b) presents the analogous results for male-to-female log odds ratio.

Figure A.8: Log Odds Ratio of Top Female and Male Words by Fields in Social Sciences, 1950-1959 (Words from Titles and Abstracts) – Continued

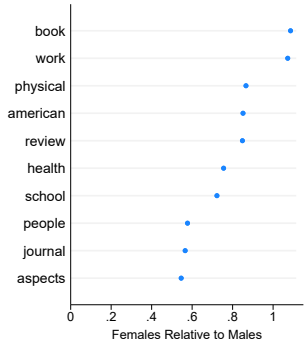
(g) Psychology - Female



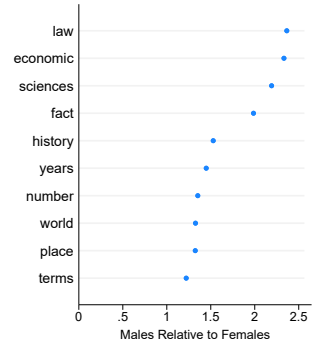
(h) Psychology - Male



(i) Sociology - Female



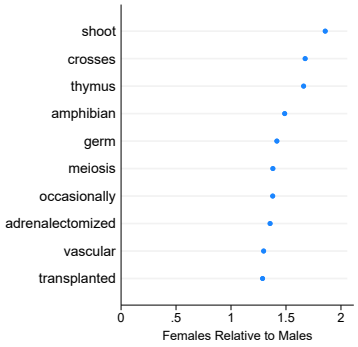
(j) Sociology - Male



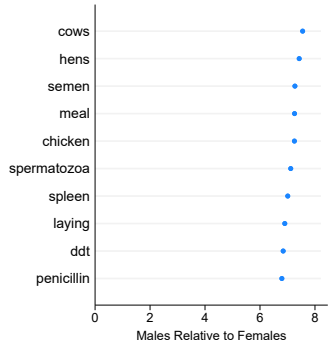
Notes: Panel (a) shows the top 10 words with the highest female-to-male log odds ratio used the titles of in all publications in the specified field written between 1950 and 1959. Panel (b) presents the analogous results for male-to-female log odds ratio.

Figure A.9: Log Odds Ratio of Top Female and Male Words by Fields in STEM, 1950-1959 (Words from Titles and Abstracts)

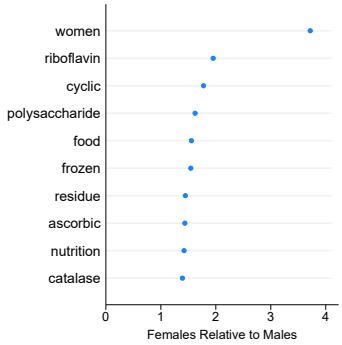
(a) Biology - Female



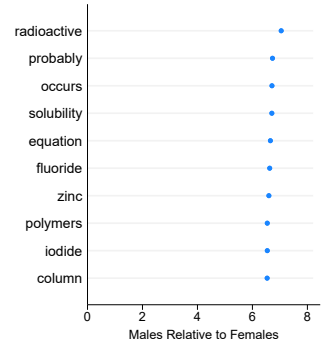
(b) Biology - Male



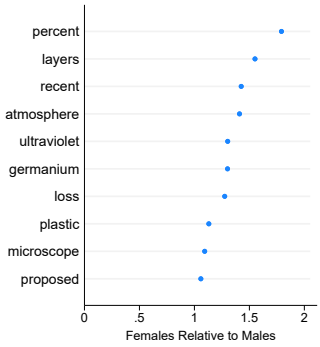
(c) Chemistry - Female



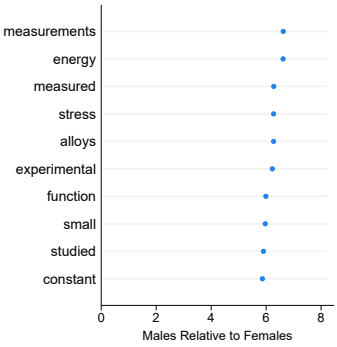
(d) Chemistry - Male



(e) Materials Science - Female



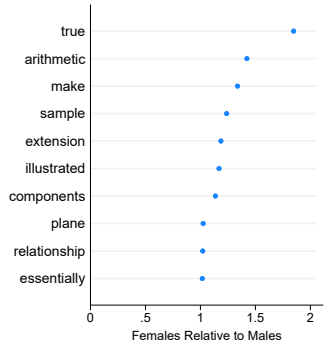
(f) Materials science - Male



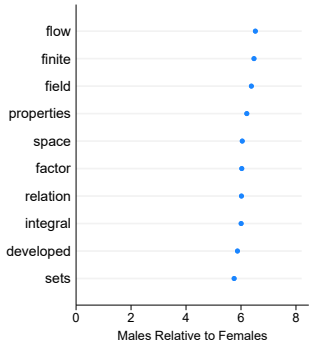
Notes: Panel (a) shows the top 10 words with the highest female-to-male log odds ratio used the titles of in all publications in the specified field written between 1950 and 1959. Panel (b) presents the analogous results for male-to-female log odds ratio.

Figure A.9: Log Odds Ratio of Top Female and Male Words by Fields in STEM, 1950-1959 (Words from Titles and Abstracts) — Continued

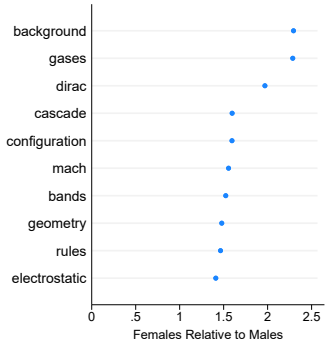
(g) Mathematics - Female



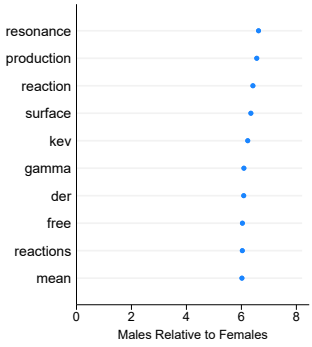
(h) Mathematics - Male



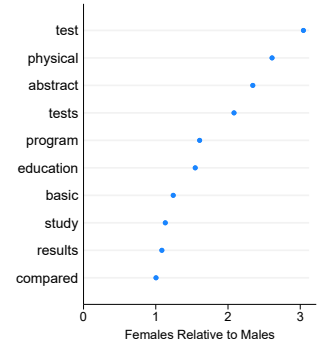
(i) Physics - Female



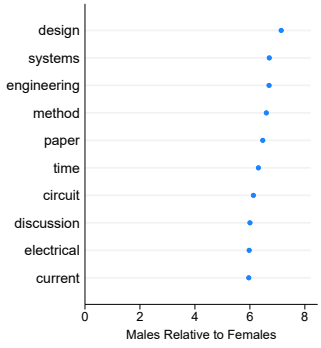
(j) Physics - Male



(k) Engineering - Female



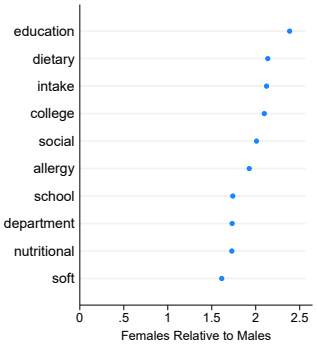
(l) Engineering - Male



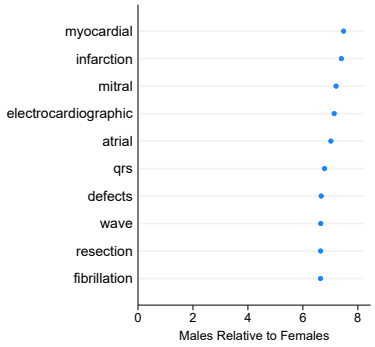
Notes: Panel (a) shows the top 10 words with the highest female-to-male log odds ratio used the titles of in all publications in the specified field written between 1950 and 1959. Panel (b) presents the analogous results for male-to-female log odds ratio.

Figure A.10: Log Odds Ratio of Top Female and Male Words in Medicine, 1950-1959 (Words from Titles and Abstracts)

(a) Medicine - Female

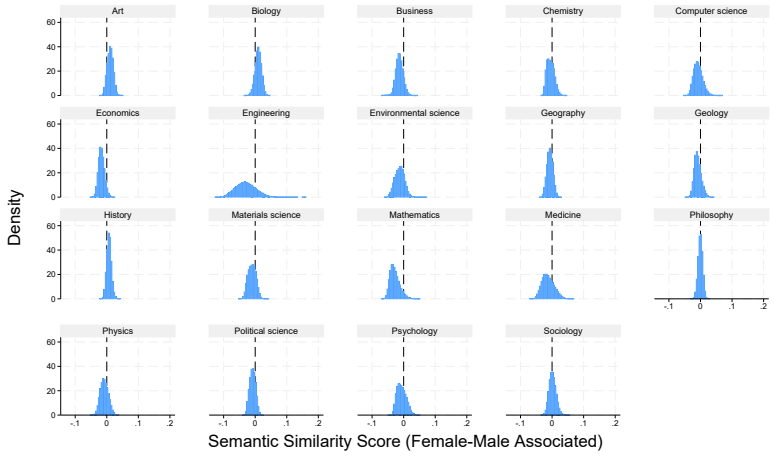


(b) Medicine - Male



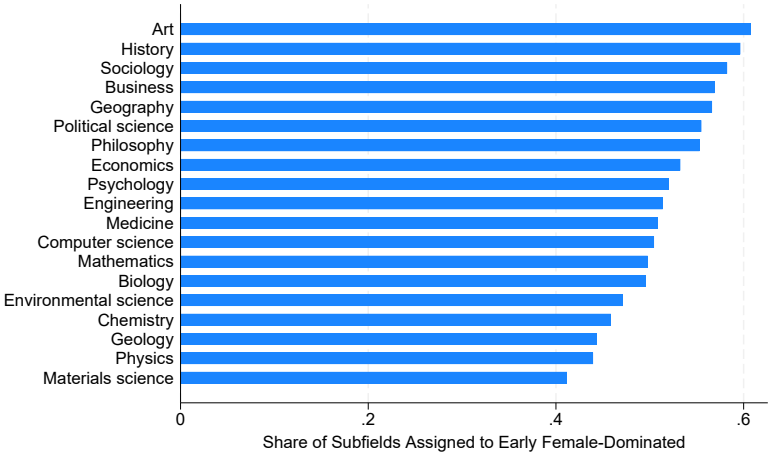
Notes: Panel (a) shows the top 10 words with the highest female-to-male log odds ratio used the titles of in all publications in the specified field written between 1950 and 1959. Panel (b) presents the analogous results for male-to-female log odds ratio.

Figure A.11: Distribution of Relative Semantic Similarity Score by Field



Notes: This figures plots the histogram of the relative semantic simliarity score by field. The score is defined as the difference between a paper's cosine similarity to all-female versus all-male papers from the 1950s, computed within each field. A positive value suggests a greater bias towards all-female papers relative to all-male papers. The distribution includes all papers in the MAG dataset from 1960–2000.

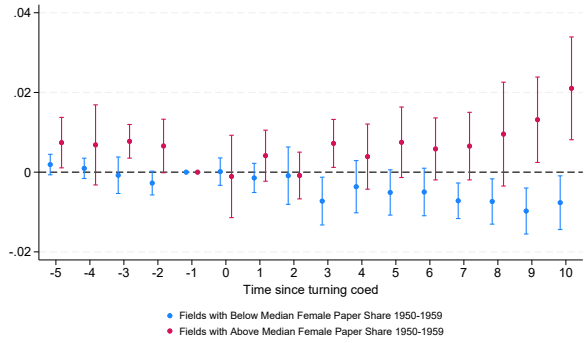
Figure A.12: Early-Female Dominated Subfields by Field



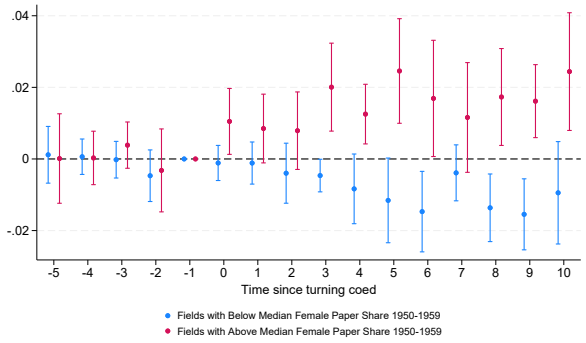
Notes: This figures plots the share of subfields within each field that is classified as being above the median in share of publications with all female authors from 1950-1959.

Figure A.13: Effect of Turning Coed on Total Number of Papers Using Female-Associated Words by Early Female Representation and Gender of the Researcher

(a) Total Number of Female-Associated Papers by All-Female Teams



(b) Total Number of Female-Associated Papers by Mixed-Gender Teams



(c) Total Number of Female-Associated Papers by All-Male Teams

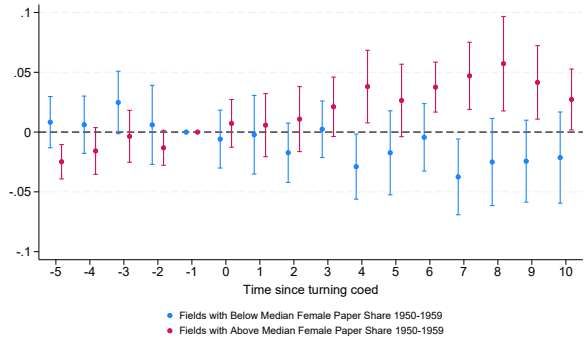
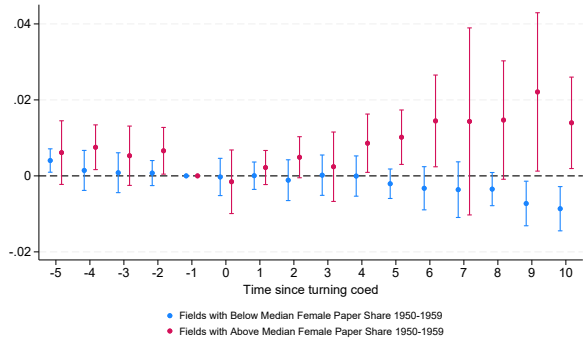
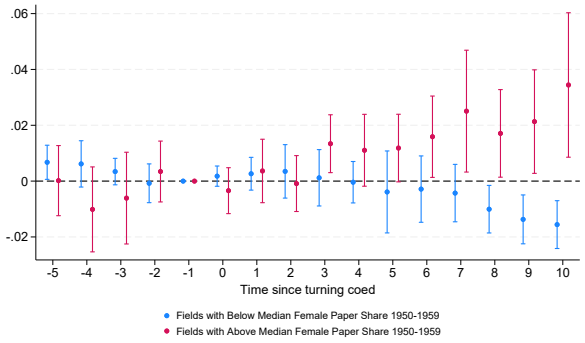


Figure A.14: Effect of Turning Coed on Total Number of Papers with High Semantic Similarity Towards Female Papers by Early Female Representation and Gender of the Researcher

(a) Total Number of Female-Associated Papers by All-Female Teams



(b) Total Number of Female-Associated Papers by Mixed-Gender Teams



(c) Total Number of Female-Associated Papers by All-Male Teams

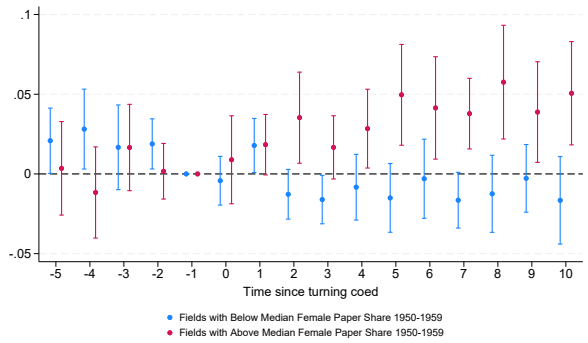


Figure A.15: Total Publications at the Subfield Level

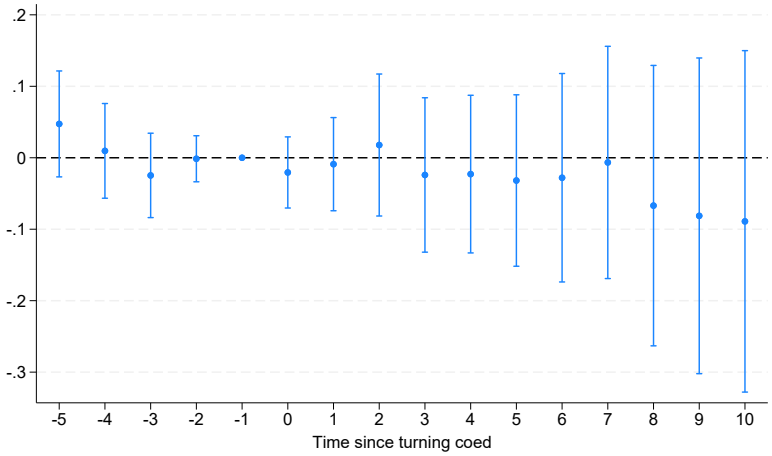


Figure A.16: Total Publications at the Subfield Level

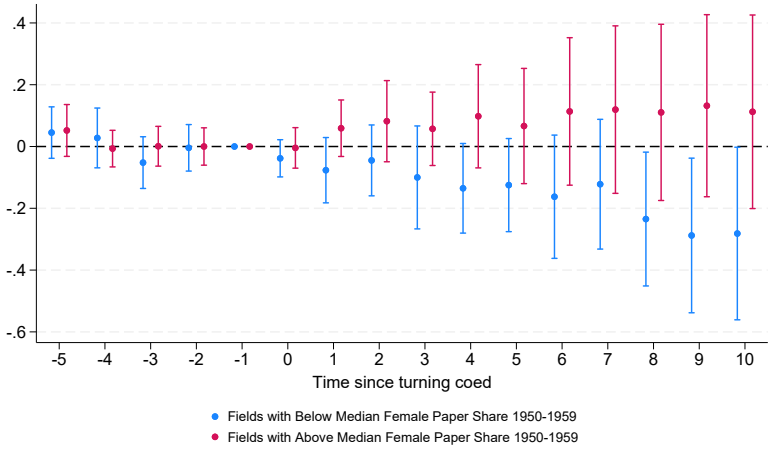
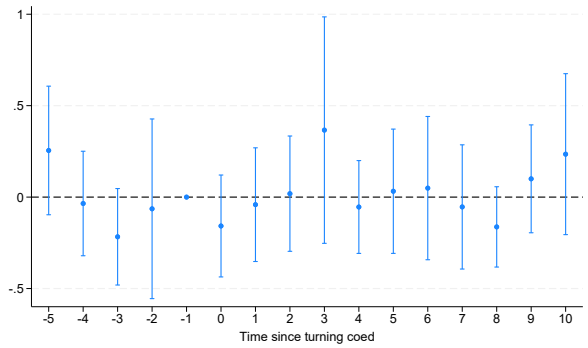
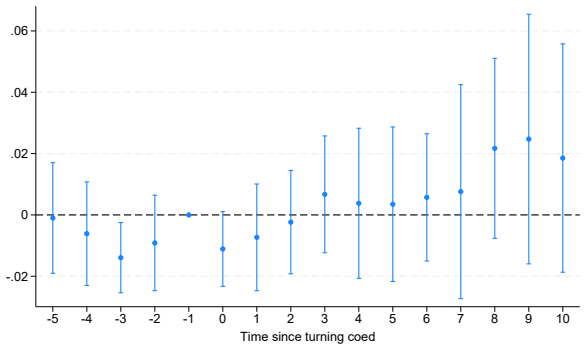


Figure A.17: Effect of Turning Coed on Research Quality

(a) Average Normalized Citations at the Subfield Level



(b) Total Hit Papers (Top 10%) at the Subfield Level



(c) Total Number of Top Journal Publications

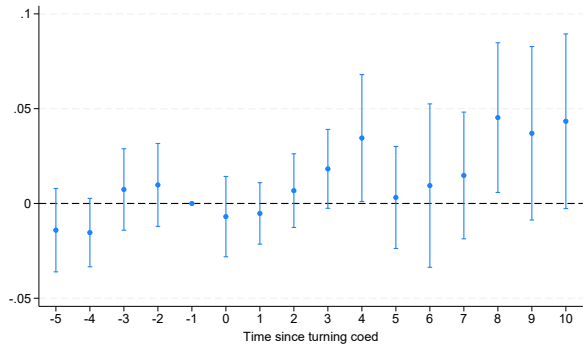
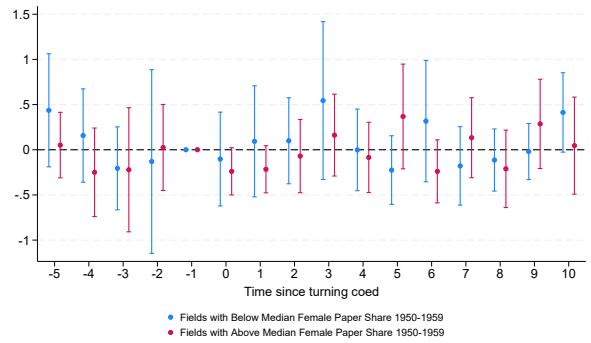
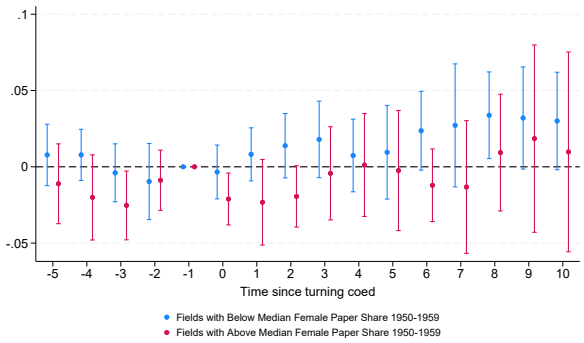


Figure A.18: Effect of Turning Coed on Research Quality by Fields

(a) Average Normalized Citations at the Subfield Level



(b) Total Hit Papers (Top 10%) at the Subfield Level



(c) Total Number of Top Journal Publications

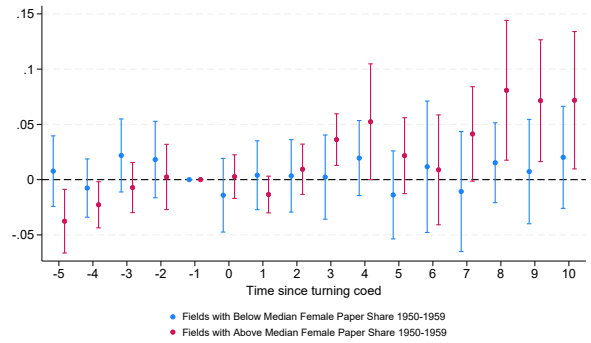
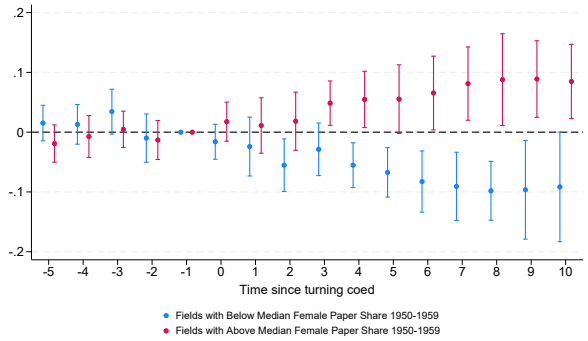
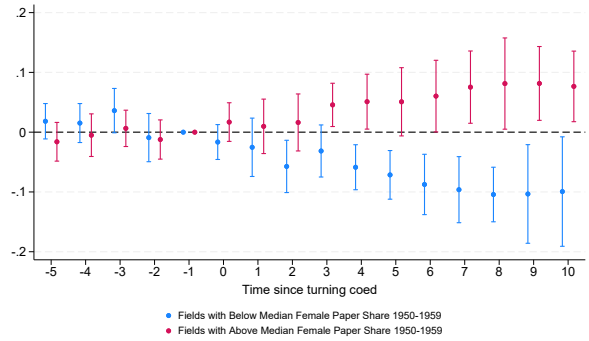


Figure A.19: Total Number of Papers Using Female-Associated Words - Alternative Control Universities

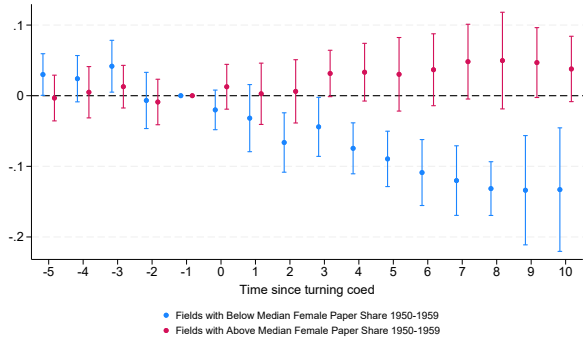
(a) No Control Schools



(b) Always Male



(c) Always Female



(d) Always Coed

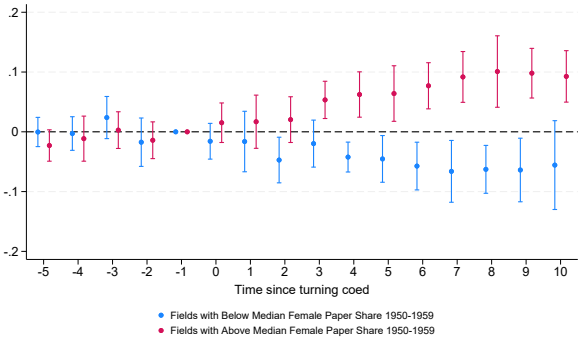
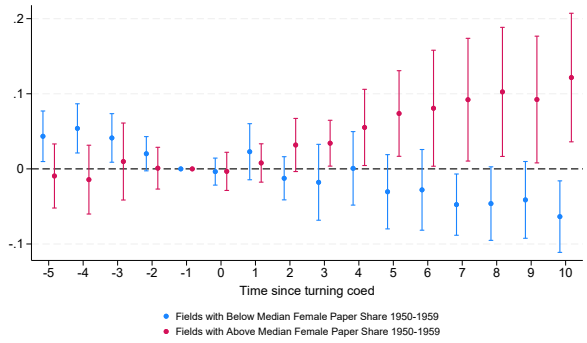
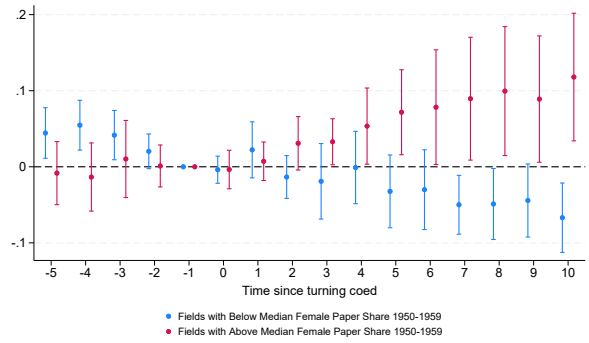


Figure A.20: Total Number of Papers with High Semantic Similarity Towards Female Papers - Alternative Control Universities

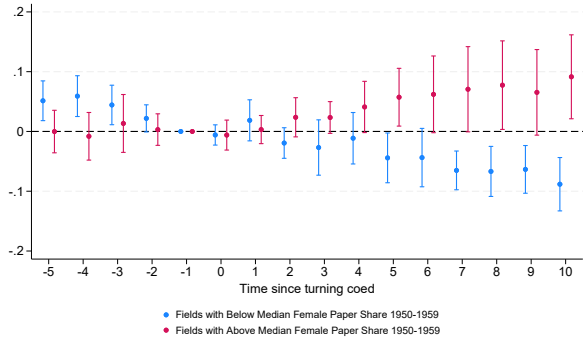
(a) No Control Schools



(b) Always Male



(c) Always Female



(d) Always Coed

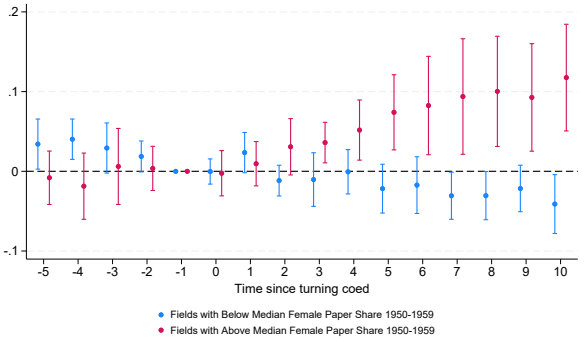
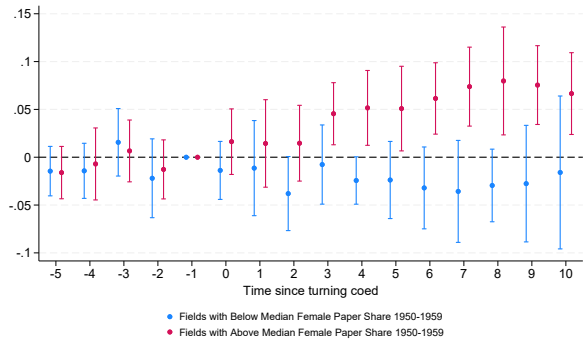


Figure A.21: Robustness to Using Field-by-Year FE

(a) Total Number of Papers Using Female-Associated Words



(b) Total Number of Papers with High Semantic Similarity Towards Female Papers

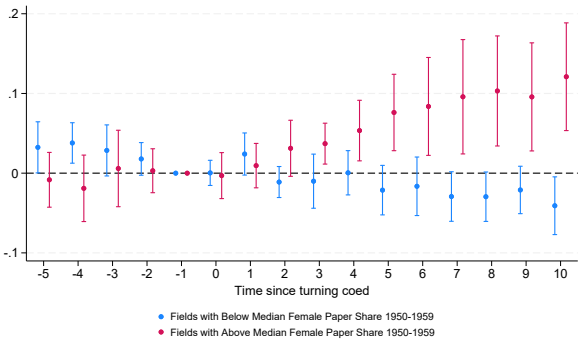
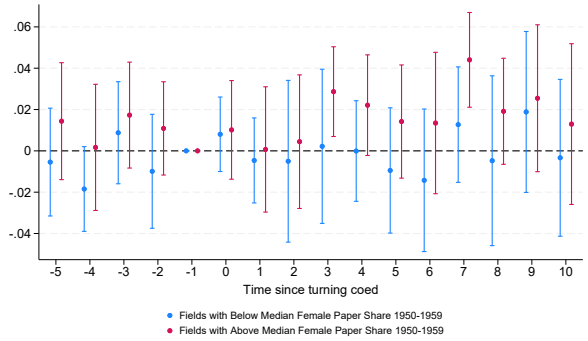
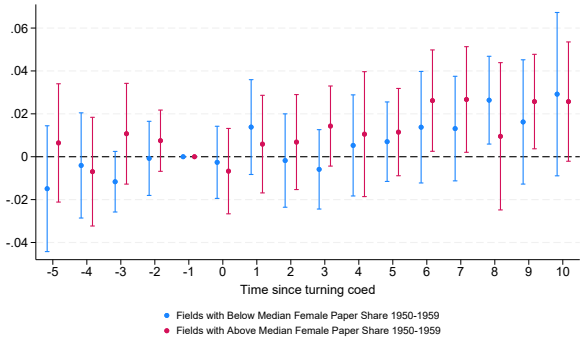


Figure A.22: Effect of Turning Coed on Female Associated Research by Early Female Representation for Incumbent Researchers

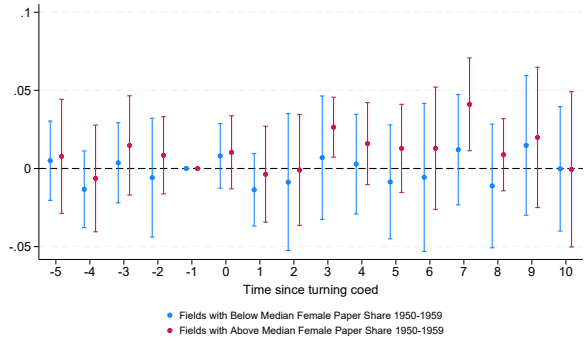
(a) Total Number of Papers Using Female-Associated Words



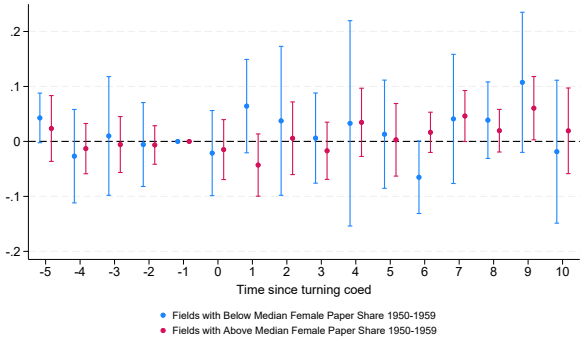
(b) Total Number of Papers with High Semantic Similarity Towards Female Papers



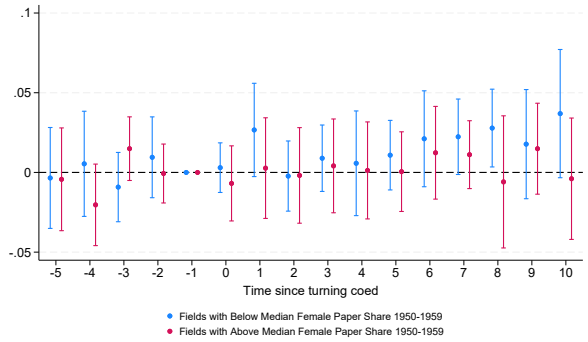
(c) Total Number of Papers Using Female-Associated Words - Male



(d) Total Number of Papers Using Female-Associated Words - Female



(e) Total Number of Papers with High Semantic Similarity Towards Female Papers - Male



(f) Total Number of Papers with High Semantic Similarity Towards Female Papers - Female

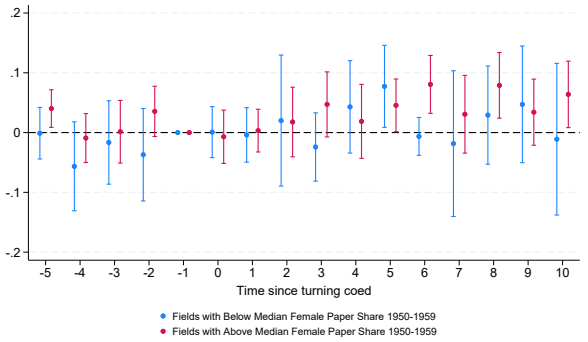
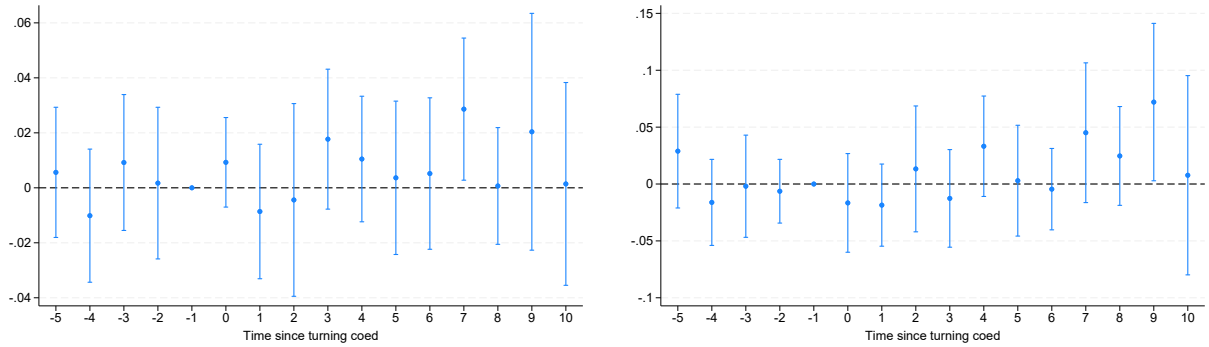


Figure A.23: Effect of Turning Coed on Female Associated Research (Log Odds) by Early Female Representation for Incumbent Researchers by Gender

(a) Male

(b) Female



(c) Male

(d) Female

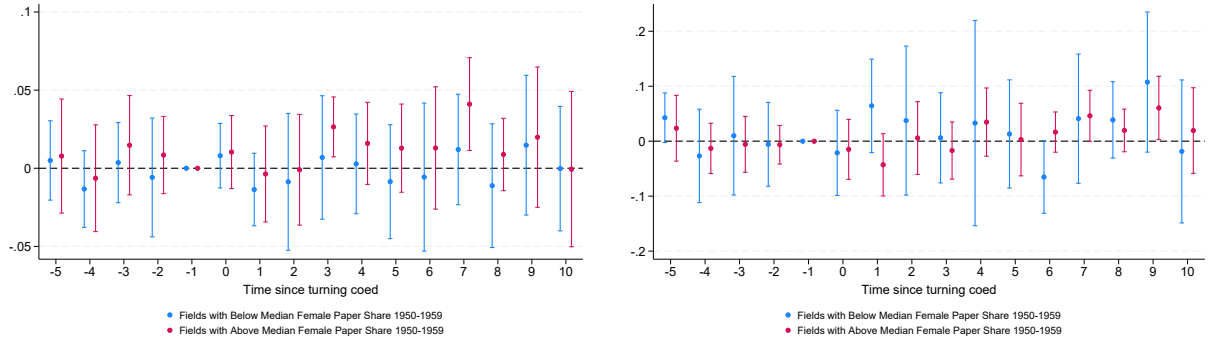
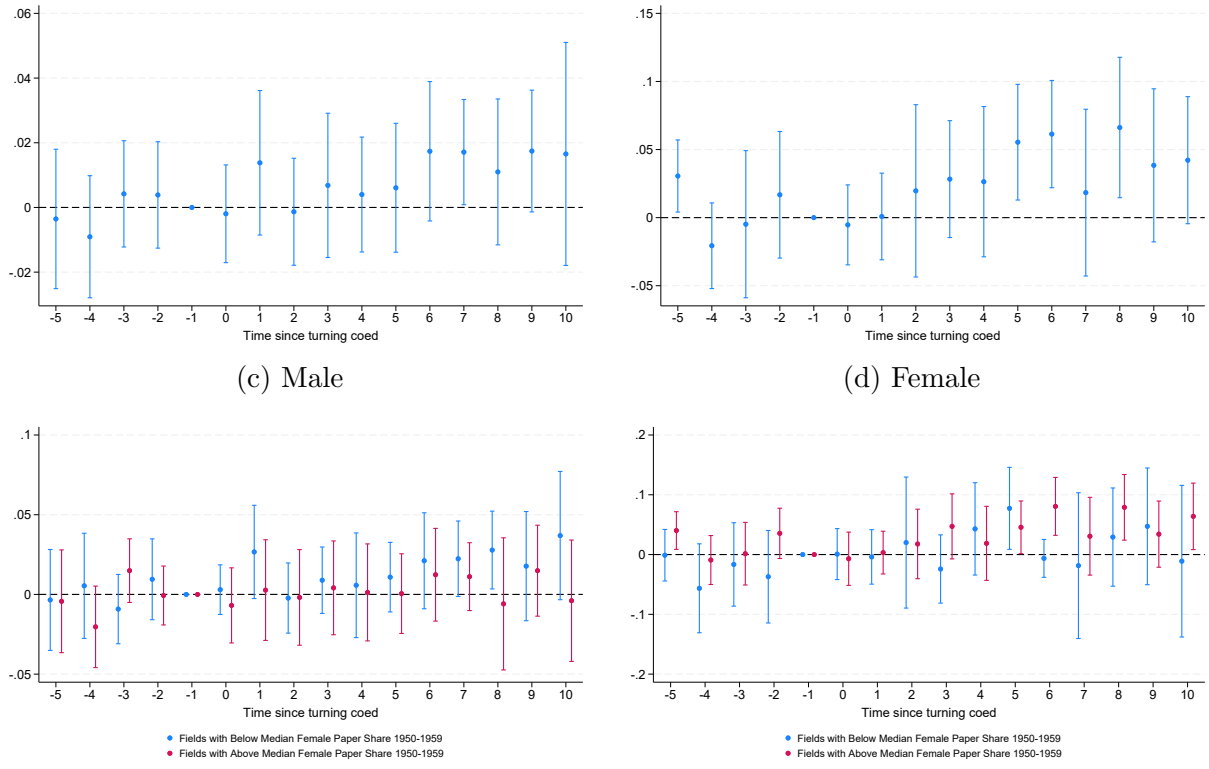


Figure A.24: Effect of Turning Coed on Female Associated Research (Semantic Similarity) by Early Female Representation for Incumbent Researchers by Gender

(a) Male

(b) Female



(c) Male

(d) Female

Figure A.25: Effect of Turning Coed on Recognition of Female-Authored Research by Early Female Representation for Incumbent Researchers by Gender

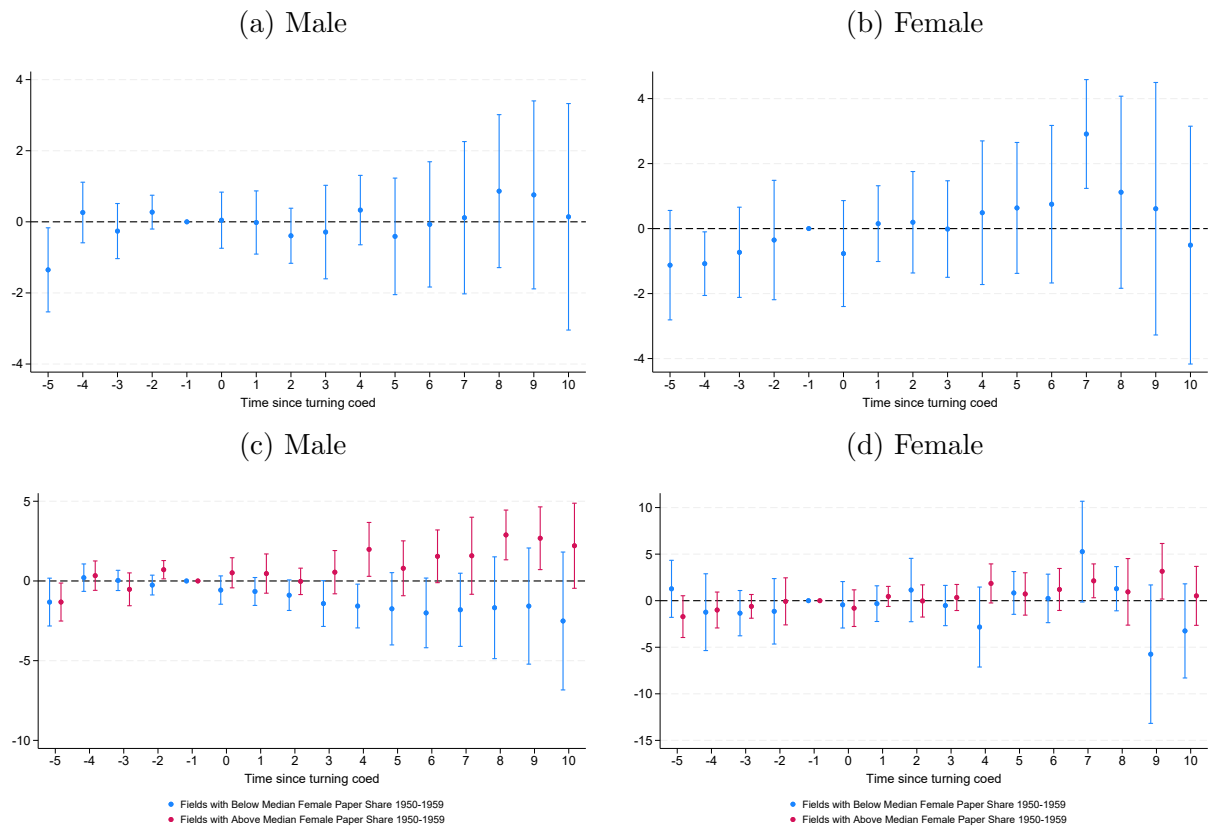


Figure A.28: Effects of Turning Coed on Graduate Degrees

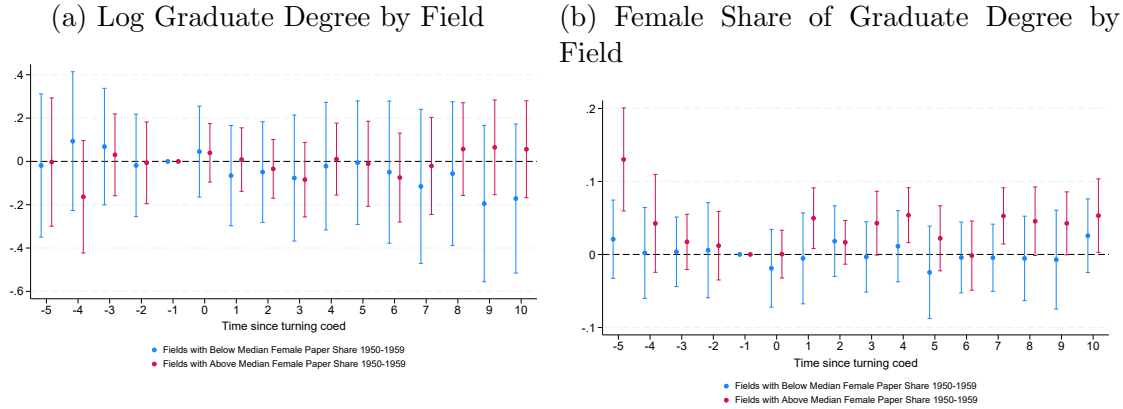
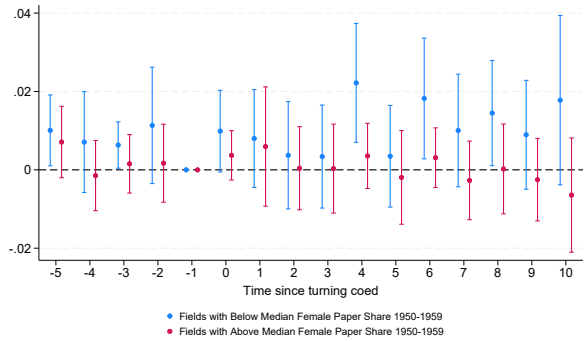
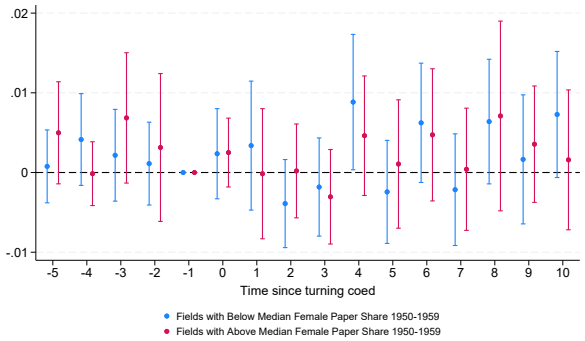


Figure A.26: Effect of Turning Coed on Female-Associated Research for Switchers

(a) Number of New Authors to the Field but Already at University



(b) Female-Associated Papers by Switchers (Top 25% Log Odds)



(c) Female-Associated Papers by Switchers (Embeddings)

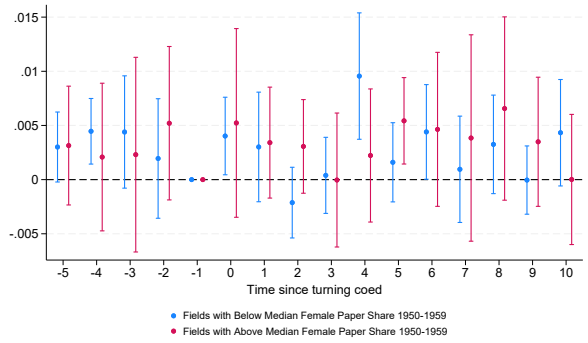
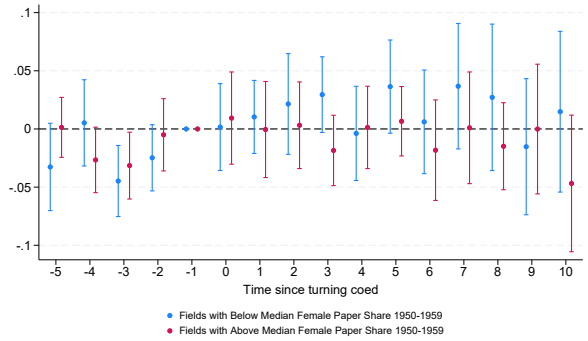
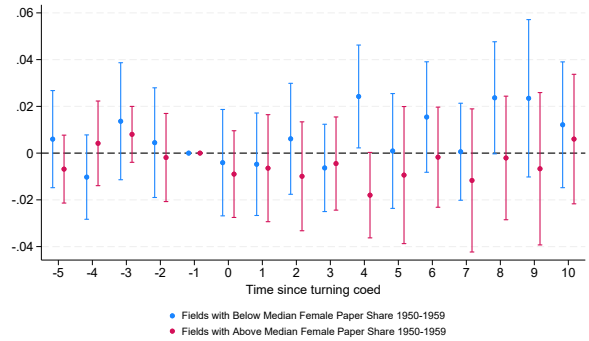


Figure A.27: Effect of Turning Coed on Female-Associated Research for Departing Researchers

(a) Number of Departing Researchers



(b) Number of Departing Researchers Who Had Written a Female-Associated Paper (Top 25% Log Odds)



(c) Number of Departing Researchers Who Had Written a Female-Associated Paper (Embeddings)

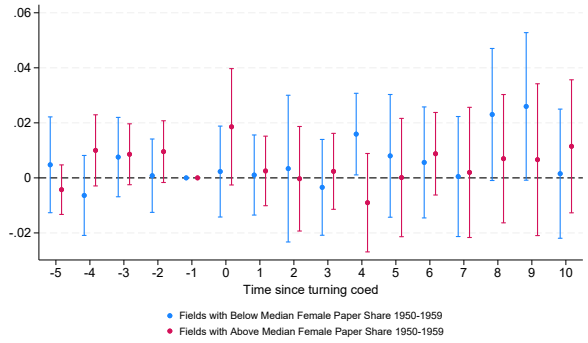
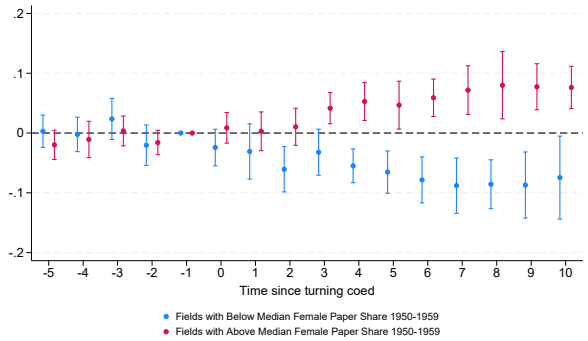


Figure A.29: Robustness to Using Region-by-Year FE

(a) Total Number of Papers Using Female-Associated Words



(b) Total Number of Papers with High Semantic Similarity Towards Female Papers

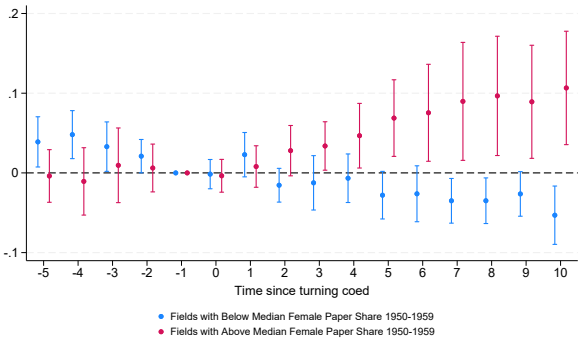


Figure A.30: Relationship between Growth of Subfields and Early Female Representation

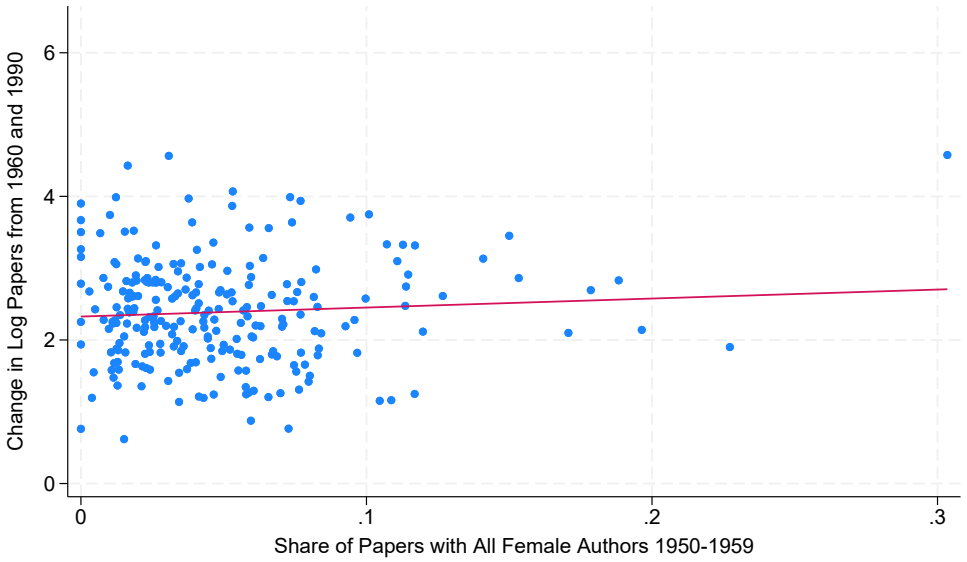


Figure A.31: Effects of Turning Coed on Research Expenditures

(a) Trends in Log R&D Expenditures by Field (b) Effects on Log R&D Expenditures by Field

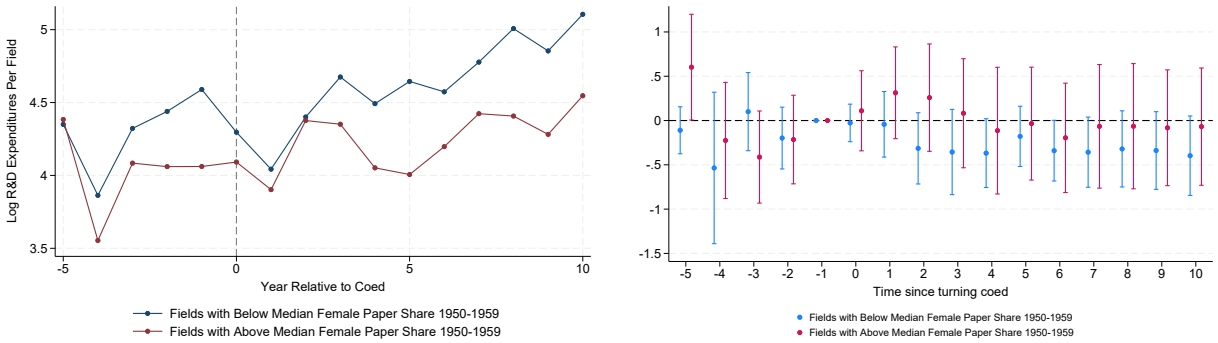
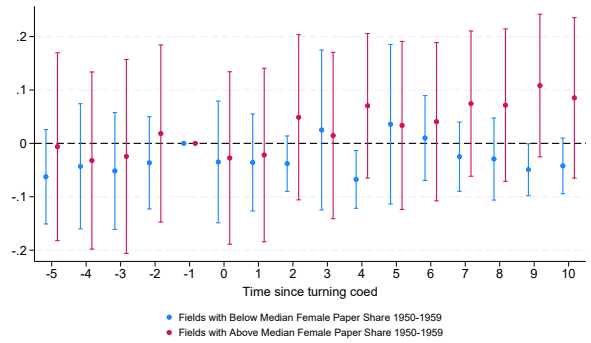


Figure A.32: Spillover Effects

(a) Total Number of Papers Using Female-Associated Words



(b) Total Number of Papers with High Semantic Similarity Towards Female Papers

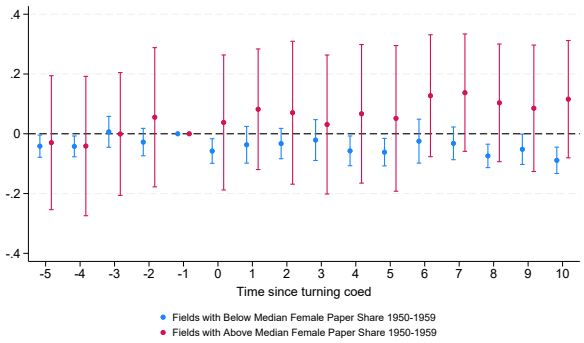
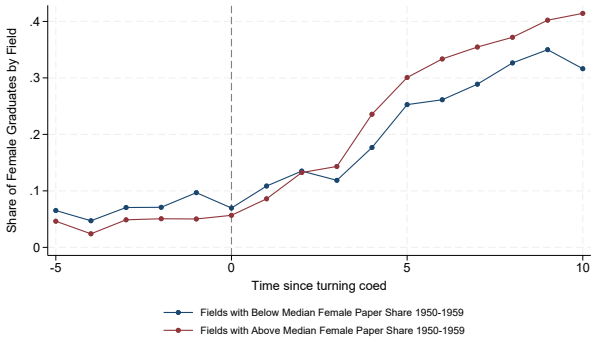


Figure A.33: Effects of Turning Coed on Undergraduate Gender Composition by Field

(a) Female Bachelor's Degree Share by Field



(b) Female Bachelor's Degree Share by Field

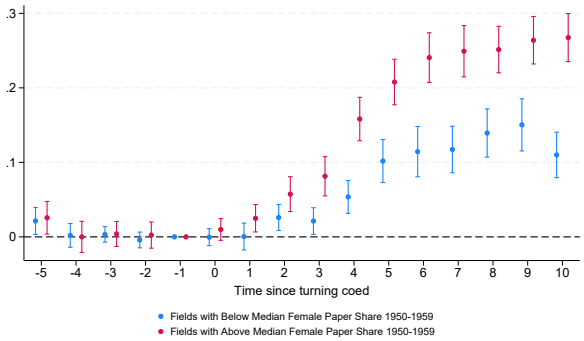


Table A.1: Top 15 Subfields with Highest Share of Publications with 1+ Female Authors, 1950-1959

Field	% Women
nursing	35.2
environmental health	31.4
virology	29.0
molecular biology	27.8
immunology	26.2
endocrinology	25.6
microbiology	25.4
andrology	24.6
visual arts	23.9
medical education	23.7
food science	23.7
optometry	23.2
developmental psychology	22.7
biochemistry	22.5
psychiatry	22.4

Table A.2: Top 15 Subfields with Highest Share of Publications with All Female Authors, 1950-1959

Field	% Women
nursing	30.3
visual arts	22.7
optometry	19.6
medical education	18.8
environmental health	17.9
socioeconomics	17.1
gender studies	15.3
gerontology	15.0
developmental psychology	14.1
pedagogy	12.7
mathematics education	12.0
family medicine	11.7
food science	11.7
cell biology	11.5
criminology	11.4

Table A.3: Bottom 15 Subfields with Lowest Share of Publications with 1+ Female Authors, 1950-1959

Field	% Women
microeconomics	0.0
electronic engineering	0.0
control engineering	0.7
electrical engineering	0.7
aerospace engineering	1.1
discrete mathematics	1.2
pure mathematics	1.4
control theory	1.6
seismology	1.6
polymer science	1.7
telecommunications	1.8
mathematical economics	1.8
financial economics	1.9
geotechnical engineering	1.9
computational physics	2.0

Table A.4: Bottom 15 Subfields with Lowest Share of Publications with All Female Authors, 1950-1959

Field	% Women
microeconomics	0.0
control engineering	0.0
telecommunications	0.0
pattern recognition	0.0
seismology	0.0
polymer science	0.0
electronic engineering	0.0
mathematical optimization	0.0
nuclear engineering	0.0
discrete mathematics	0.3
geotechnical engineering	0.4
electrical engineering	0.4
computational physics	0.5
control theory	0.7
particle physics	0.8

Table A.5: Summary Statistics, $\tau = -1$

	All	Low Female Share	High Female Share	Difference p-value in par.
Total Researchers	2.24	2.34	2.23	0.11
	(7.17)	(6.29)	(8.16)	(0.63)
Female Researchers	0.24	0.13	0.35	-0.22**
	(1.29)	(0.76)	(1.69)	(0.00)
Share of Female Researchers	0.09	0.04	0.14	-0.09**
	(0.18)	(0.12)	(0.23)	(0.00)
Total Papers	0.87	0.95	0.83	0.11
	(2.34)	(2.32)	(2.41)	(0.11)
Share of Papers with All-Female Teams	0.06	0.03	0.10	-0.07**
	(0.20)	(0.14)	(0.25)	(0.00)
Share of Papers with Mixed Teams	0.09	0.06	0.12	-0.06**
	(0.23)	(0.19)	(0.27)	(0.00)
Share of Papers with All-Male Teams	0.85	0.92	0.79	0.13**
	(0.29)	(0.23)	(0.34)	(0.00)
Average Normalized Citations	2.11	2.16	2.04	0.12
	(5.64)	(4.31)	(6.84)	(0.71)
Total Papers in Top 10% of Citations in Field	0.19	0.19	0.19	-0.00
	(0.66)	(0.64)	(0.69)	(1.00)
Total Female-Authored Paper Cited	4.23	2.61	6.15	-3.53**
	(27.55)	(13.70)	(37.38)	(0.00)
... By All-Male Teams	0.29	0.20	0.40	-0.20**
	(1.70)	(1.35)	(2.04)	(0.00)
Total Papers with Female-Associated Word in Title	0.31	0.26	0.37	-0.11**
	(1.23)	(0.81)	(1.57)	(0.00)
... By All-Female Teams	0.02	0.01	0.03	-0.02**
	(0.18)	(0.10)	(0.23)	(0.00)
... By Mixed Teams	0.04	0.02	0.07	-0.04**
	(0.38)	(0.19)	(0.52)	(0.00)
... By All-Male Teams	0.19	0.16	0.23	-0.07**
	(0.75)	(0.52)	(0.94)	(0.00)
Total Papers with High Semantic Similarity to Female Papers	0.21	0.15	0.28	-0.13**
	(0.91)	(0.68)	(1.12)	(0.00)
... By All-Female Teams	0.02	0.01	0.03	-0.02**
	(0.15)	(0.08)	(0.20)	(0.00)
... By Mixed Teams	0.03	0.01	0.05	-0.04**
	(0.31)	(0.11)	(0.43)	(0.00)
... By All-Male Teams	0.13	0.10	0.16	-0.06**
	(0.57)	(0.46)	(0.68)	(0.00)
Observations	4463	2238	2113	4351

Notes: This table presents means and standard deviations (parentheses) at the school-subfield level in the year prior to coeducation across all turn-coed universities by those in above or below the median in share of publications with all female authors from 1950 to 1959.

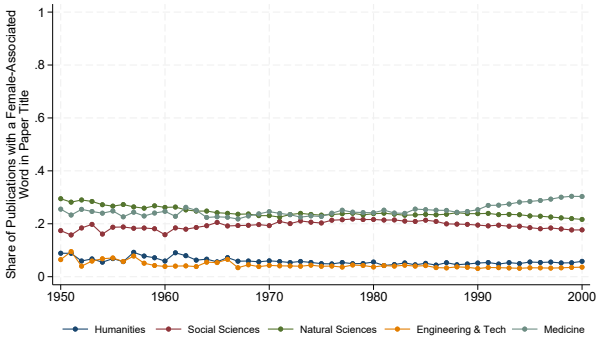
Table A.6: Characteristics of Subfields by Share of Publications with All Female Authors, 1950-1959

	All	Low Female Share	High Female Share	Difference p-value in par.
Total Papers	0.32	0.32	0.33	-0.01**
	(0.97)	(0.88)	(1.05)	(0.00)
Share of Papers with All-Male Teams	0.89	0.94	0.85	0.08**
	(0.29)	(0.23)	(0.33)	(0.00)
Share of Papers with Mixed Teams	0.06	0.04	0.07	-0.03**
	(0.22)	(0.19)	(0.23)	(0.00)
Share of Papers with All-Female Teams	0.05	0.02	0.08	-0.05**
	(0.21)	(0.13)	(0.25)	(0.00)
Total Researchers	0.64	0.64	0.66	-0.03**
	(2.47)	(2.21)	(2.69)	(0.01)
Share of Female Researchers	0.07	0.04	0.10	-0.06**
	(0.21)	(0.14)	(0.24)	(0.00)
Share of Papers with Female-Associated Word in Title	0.31	0.27	0.35	-0.08**
	(0.43)	(0.41)	(0.45)	(0.00)
Share of Papers Citing an All-Female Paper	0.03	0.02	0.04	-0.03**
	(0.10)	(0.07)	(0.11)	(0.00)
Average Normalized Citations	2.08	2.37	1.91	0.46**
	(19.85)	(21.76)	(18.74)	(0.01)
Share of Papers in Top 10% of Citations in Field	0.19	0.19	0.18	0.01**
	(0.36)	(0.36)	(0.35)	(0.00)
Observations	302640	118020	175650	293670

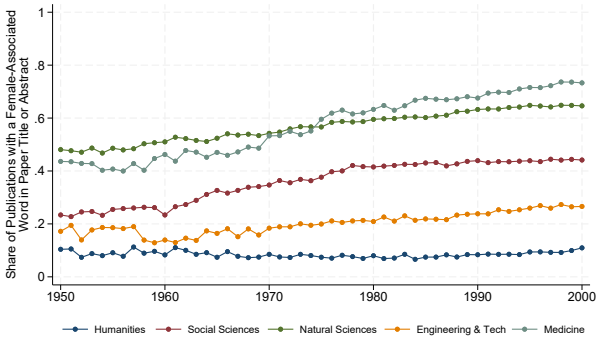
Notes: This table presents means and standard deviations (parentheses) at the school-subfield level from 1950-1959 across all universities by those in above or below the median in share of publications with all female authors from 1950 to 1959.

Figure A.34: Trends in Gendered Word Use by Discipline

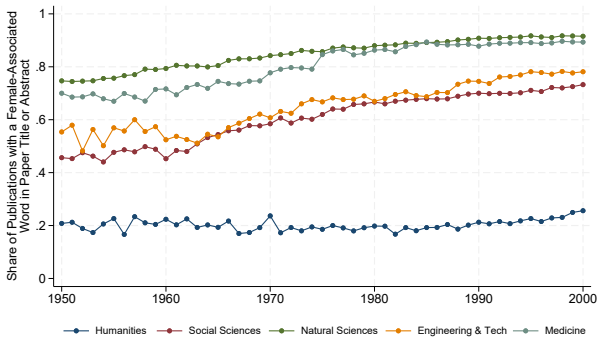
(a) Words from Titles – Applied to Titles



(b) Words from Titles – Applied to Titles and Abstracts



(c) Words from Titles and Abstracts



Notes: Figure shows the share of papers by discipline with at least one female-associated word in either the paper title or abstract. Female-associated word is a word in the top 25% of the log-odds distribution in that field.