Coeducation and Entrepreneurship

Zhuo Chen *Boston College*

Hanyi Livia Yi* *Boston College*

March 14, 2025

Preliminary and incomplete. Please do not circulate without authors' permission.

We study the effects of gender diversity in undergraduate education on entrepreneurship and subsequent startup characteristics. Leveraging a natural experiment from 1942 to 1997, during which 99 previously all-male universities transitioned to coeducation, we analyze university enrollment and resume data to assess its impact on entrepreneurial outcomes. We find that male students exposed to coeducational environments are 1.2 percentage points more likely to become entrepreneurs. The resulting startups are more likely to operate in industries with higher female representation, such as art, recreation, and retail, and employ more women, consistent with gender-inclusive education fostering diverse entrepreneurial ideas and workforces. Exploring mechanisms, we find evidence consistent with shifting major choices, expanded social networks, and increased risk-sharing within households. Overall, our results suggest that gender diversity in higher education has a lasting impact on entrepreneurial dynamics.

JEL classification: I24, J16, L26
Keywords: Entrepreneurship, Higher Education, Gender Diversity

*We thank Andra Ghent, Edie Hotchkiss, Jonathan M. Karpoff, Zheng Sun, and participants at the University of Washington Fostering Inclusion Workshop and AFA AFFECT Mentoring Workshop for their thoughtful feedback and comments. Emails: Zhuo Chen: zhuo.chen@bc.edu. Hanyi Livia Yi: livia.yi@bc.edu.

1 Introduction

Scholars have long emphasized the importance of interactions among individuals from diverse backgrounds as a catalyst for new idea generation and growth (e.g. Mill, 1965; Glaeser, Kallal, Scheinkman, and Shleifer, 1992; Weitzman, 1998; Jones, 2023; Posch, Schulz, and Henrich, 2023). These interactions may foster the exchange of unique knowledge, skills, approaches, and perspectives, creating environments conducive to entrepreneurship (Mill, 1965; Jacobs, 1969; Glaeser et al., 1992; Posch et al., 2023). Beyond facilitating knowledge exchange, these interactions may expand social networks, potentially influencing family formation patterns and risk tolerance, another crucial determinant of entrepreneurial activity.

However, the presence of diverse environments has evolved significantly over time. A prominent example of this evolution is higher education. Although the vast majority of undergraduate students in the United States today attend coeducational universities, none could do so before 1835, and some of the most prestigious universities remained single sex until the 1960s. The transition to coeducation represented a major shift in higher education, significantly increasing gender diversity among students throughout the 20th century. While this shift has been studied from various socioeconomic angles—such as its effects on education attainment and scientific research directions—little evidence exists on its impact on new business creation (Goldin and Katz, 2011; Calkins, Binder, Shaat, and Timpe, 2023; Truffa and Wong, 2024).

In this paper, we examine the effect of gender diversity in undergraduate education on individuals' entrepreneurial decisions and the characteristics of their startups. Coeducation, by offering a more gender-diverse peer group and fostering interactions at a formative stage, may facilitate the exchange of ideas, influence gender norms, shape the size and composition of one's social networks, and affect household formation and intra-household risk-sharing. These factors may, in turn, impact the formation of new business. We find that after all-male universities adopt coeducation, male students are more likely to become entrepreneurs later in life. Furthermore, their startups are more likely to operate in industries with greater female representations, such as art

and recreation, accommodation, and retail, and employ more women. These results suggest that gender-inclusive educational environments contribute not only to new business formation, but also to the diversification of entrepreneurial ideas and workforce.

We build a large dataset, combining information on universities' enrollment and graduation, transition dates to coeducation, and individuals' resumes. Our sample includes 99 previously all-male universities that shifted to coeducation between 1942 and 1997. We match the universities turning coeducation with the educational background and career history of student cohorts enrolled around the transition years. We use a stacked difference-in-differences (DiD) research design to leverage variations in the timing of universities' transition to coeducation. This design is akin to a collection of event studies.[1]

The key identifying assumption underlying our empirical approach is that individuals' likelihood of becoming entrepreneurs in treated and control universities would have moved in parallel in the absence of coeducation adoption. This assumption is likely to hold if universities' timing of adoption is independent of factors that might otherwise affect students' choice to become entrepreneurs. The historical context and motivations behind coeducation transitions largely support this. Many universities pursued coeducation to boost applications and cost efficiency (Malkiel, 2016; Truffa and Wong, 2024). Anecdotal evidence also suggests that the liberalization wave prompted by the Second Vatican Council (1962-1965) advocated for women's right to education and encouraged Catholic universities to transition afterward, with others following to stay competitive (Poulson, 1995; Goldin and Katz, 2011; Calkins et al., 2023). Furthermore, the specific timing of adoption was also influenced by factors such as alumni resistance and proximity to schools of the other gender (Malkiel, 2016). These factors were less likely to be correlated with students' future entrepreneurship rates. When examining factors predicting the timing of coeducation adoption, only a university's establishment year plays a significant role, likely because older universities have more influential alumni and thus face stronger opposition. We control for establishment years using

[1] As documented in previous literature, the two-way fixed-effects model in the staggered DiD setting has several drawbacks, such as implicit weighting and biased estimates (Goodman-Bacon, 2021). Following the literature, we use stacked DiD methods to avoid these issues (Cengiz, Dube, Lindner, and Zipperer, 2019; Deshpande and Li, 2019; Butters, Sacks, and Seo, 2022; Wing, Freedman, and Hollingsworth, 2024).

university fixed effects. Moreover, our results remain similar when controlling for location-by-cohort fixed effects, alleviating concerns that local economic conditions influence the timing of adoption.

We first examine the impact of coeducation on students' likelihood of becoming entrepreneurs. When all-male universities transition to coeducation, male students are 1.2 percentage points more likely to become entrepreneurs—a 9% increase relative to the sample's unconditional entrepreneurship probability of 13%. The main specification includes university, entry year, and event fixed effects to control for institution-specific factors (e.g., location, selectivity) and national trends that may affect entrepreneurship rates, and to restrict comparisons within each coeducation event. This specification ensures that the identifying variation stems from comparisons between treated and control universities within the same event, after purging the influence of time-invariant university characteristics and macroeconomic shocks. The results are similar under stricter specifications with event-by-university and event-by-entry year fixed effects, which further account for potential heterogeneity across events (Butters et al., 2022).

Next, we study whether coeducation affects the characteristics of the subsequent startups. Using ChatGPT to classify startups' industries based on their names, we find that post-coeducation startups are more likely to operate in sectors with historically greater female presence, such as accommodation and hospitality, arts and recreation, and retail and wholesale. This could be consistent with coeducation fostering a broader exchange of information, ideas, and perspectives across genders. Furthermore, these startups employ, on average, 6.9% more female workers– an economically large effect of 17.8% increase relative to the mean. This effect persists after controlling for industry, which implies that the rise in female share is not solely due to changes in industry composition. These findings suggest that coeducation affects not only startup creation but also the industry and gender composition of the founded firms, which could potentially reflect a long-term shift in gender norms.

Exploring the mechanisms behind the rise in entrepreneurship rates, we find evidence consistent with coeducation influencing college major choices, professional networks, and household risk-

sharing. After coeducation, male students are more likely to major in business and science and less likely to major in humanities, consistent with a shift toward (away from) stereotypically male-dominated (female-dominated) fields. This evidence mirrors the STEM crowding-out effects documented when female schools turn coeducation (Calkins et al., 2023). Such changes in majors may contribute to higher entrepreneurship rates, as startup founders disproportionately come from business and science backgrounds.[2]

Professional social networks, especially alumni networks, are crucial for entrepreneurial financing by facilitating information flow (Garfinkel, Mayer, Strebulaev, and Yimfor, 2021). Since coeducation often coincides with an increase in student enrollment, we examine its impact on the size of students' future professional networks, measured by LinkedIn connections. We find that after coeducation, the likelihood of having more than 500 LinkedIn connections increases by 1.2 percentage points, a 5.4% rise relative to the unconditional mean. While the expansion of professional networks is not unique to coeducation, it may still serve as a pathway through which coeducation fosters higher entrepreneurship rates, highlighting the importance of alumni connections in startup formation.

Finally, coeducation may also influence entrepreneurship by affecting household formation and risk-sharing. If coeducation increases the likelihood of marrying a college-educated spouse, it could increase spousal income, improve downside protection, and provide greater risk-sharing capacity, thus encouraging entrepreneurial risk-taking (Becker, 1974; Benham, 1974; Wong, 1986; Özcan, 2011; Hombert, Schoar, Sraer, and Thesmar, 2020; Barrios, Hochberg, and Yi, 2022; Denes, Lagaras, and Tsoutsoura, 2025). If this hypothesis holds, we would expect larger effects in areas where the employment and income gaps between college-educated and non-college-educated married women are wider. Consistent with this, we find that coeducation leads to a 1.9% (1.5%) increase in entrepreneurship in regions with above-median income (employment) gaps, compared to only 0.3% (0.8%) in regions with below-median gaps.

[2]For instance, Wadhwa, Freeman, and Rissing (2008) show that out of 652 surveyed U.S.-born tech company founders with companies established from 1995 through 2005, 33% of them hold a terminal degree in Business, 47% in STEM fields, and 3% in Arts, Humanities, Social Sciences.

Our main results are also robust to alternative definitions of founders, sample windows, control groups, and fixed effects (Cengiz et al., 2019; Deshpande and Li, 2019). A potential concern is that coeducational universities may attract, rather than influence, entrepreneurial students. To alleviate this concern, we conduct a subsample analysis focusing on public university students, who are less likely to select institutions based on their coeducational status, and find consistent results. Additionally, to address right-censoring in resume data where earlier cohorts have longer career histories, we standardize founder measures to a fixed post-college career length and find similar results.

Our paper contributes to several strands of literature. First, in the coeducation literature, previous papers study its impact on women's educational attainment, faculties' direction of research, and student traits, perceptions, and major choices (Goldin and Katz, 2011; Booth, Cardona-Sosa, and Nolen, 2014; Lee, 2021; Lee and Nakazawa, 2022; Calkins et al., 2023; Park, 2023; Truffa and Wong, 2024). For instance, Goldin and Katz (2011) establishes that coeducation opens more possibilities for women going to college and is associated with increased women's educational attainment. Truffa and Wong (2024) find that coeducation leads to an increase in gender-related research publications. In the setting of randomly assigned high schools in South Korea, researchers find that coeducation in high school affects gender stereotypes, levels of education, earnings, and coworker relationships (Lee, 2021; Lee and Nakazawa, 2022; Park, 2023). In the U.K., Booth et al. (2014) find that random exposure to single-sex college courses increases risk-taking among women, suggesting that observed gender differences might reflect social learning rather than inherent traits. Importantly, Calkins et al. (2023) explore college major choices after female universities adopt coeducation and find that female students are crowded out of STEM fields and shift toward more stereotypically female-oriented disciplines. Our findings on changing major choices could reflect a similar crowding-out effect for male students when male universities adopt coeducation.

Our paper contributes to this literature by documenting that coeducation influences not only the students of the newly admitted gender but also the outcomes for students of the existing gender when the opposite gender joins campus. Moreover, we focus on a previously understudied aspect of

students' career choice—namely, their likelihood of becoming entrepreneurs and the characteristics of the startups they create. This aspect is important not only because entrepreneurship has long been considered a driver of economic development (Schumpeter, 2017), but also because, as early as Mill (1965), social scientists have hypothesized that a diverse social environment fosters the exchange and integration of ideas.

Beyond coeducation, our paper also contributes to the literature seeking to understand the influence of exposure to diverse environments on idea generation and behaviors. Posch et al. (2023) study how the surname diversity of a society affects innovation. In terms of how past gender exposure affects future behaviors, Duchin, Simutin, and Sosyura (2021) finds that male CEOs who grew up in male-dominated families or attended all-male high schools allocate more investment capital to male division managers than to female ones, suggesting that limited interaction with the opposite sex during formative years may contribute to gender disparities in decision-making later in life. Calder-Wang and Gompers (2021) shows that venture capital partners are more likely to hire female partners when they have more daughters. Our finding complements this literature by showing that gender-inclusive educational environments foster startups in industries with greater female presence and employ more women.

Last, our work contributes to the ongoing literature seeking to understand the roles of education, peer environments, and gender in entrepreneurship. Levine and Rubinstein (2017) show that better-educated people with "smart and illicit" traits are more likely to run incorporated businesses. We focus on the gender diversity of the higher education environment rather than the level of education. Several papers also study the impact of one's surrounding environment, such as peer effects at universities and workplaces, on entrepreneurship (e.g. Nanda and Sørensen, 2010; Kacperczyk, 2013; Lerner and Malmendier, 2013; Wallskog, 2022). A growing literature seeks to understand the role of gender in entrepreneurial financing (e.g. Ewens and Townsend, 2020; Hebert, 2023; Gornall and Strebulaev, 2024; Hebert, Tookes, and Yimfor, 2024; Howell and Nanda, 2024. See Ewens (2022) for a literature review), and the entry of women in entrepreneurship (e.g. Zandberg, 2021; Gottlieb, Townsend, and Xu, 2022; Mertz, Ronchi, and Salvestrini, 2024).

2 Coeducation of Higher Education in the U.S.

Coeducation in higher education has a longstanding history in the United States. From 1835 to the 1950s, coeducational universities steadily opened, and single-sex universities transitioned to coeducation at a consistent pace. However, the process accelerated significantly between the 1960s and 1990s. This later period is particularly notable in the history of coeducation, as it saw the admission of women into several of the nation's most prestigious universities, including Columbia, Dartmouth, Princeton, and Yale.

Several factors influenced universities' decisions to adopt coeducation. First, for Catholic institutions, a key motivator was the reforms of the Second Vatican Council (1962–1965), which declared that women had the right "to acquire an education...equal to that recognized for men" (Poulson, 1995; Goldin and Katz, 2011; Calkins et al., 2023). Second, for universities following the wave of coeducation, the primary motivation was to remain relevant and competitive. Third, financial considerations often played a role. While some all-male universities established coordinate schools or separate campuses for women, transitioning to a coeducational campus was more cost-effective for institutions facing financial constraints. Allowing students of both genders to study and reside on the same campus proved more budget-friendly than maintaining separate facilities (Malkiel, 2016). Lastly, other factors such as alumni resistance, local student demand, and proximity to schools of the opposite gender also influenced the decision. Alumni often favored preserving single-sex traditions, and institutions with stronger alumni influence—typically older universities—tended to adopt coeducation later (Malkiel, 2016).

The decision-making process for coeducation typically involves complex negotiations among university presidents, trustees, and faculty members, which makes the exact time of transition more uncertain. Some all-male universities initially considered merging with women's colleges as coordinate schools, such as the proposed pairings of Yale-Vassar and Princeton-Sarah Lawrence. Institutions also conducted surveys and commissioned reports to evaluate the potential benefits of coeducation and to address financial requirements, such as the need for women's dormitories and

other facilities (Malkiel, 2016).

3 Data and Variable Construction

3.1 University Data

We construct a list of historically all-male universities and their dates of transition to coeducation from the Higher Education General Information Survey (HEGIS) and Integrated Postsecondary Education Data System (IPEDS) following Calkins et al. (2023). We collect all baccalaureate degree granting universities that transitioned to coeducation. We manually check the year of turning coeducational through Wikipedia and school official websites. For universities whose names cannot be uniquely identified, we drop them from the sample since we rely on university names to match to the resume database.[3] We also drop universities without a clear transition time and universities that were merged with others.

Our sample includes 99 previously all-male universities with transition dates spanning from 1942 to 1997, which include renowned institutions such as Ivy League universities (Columbia, Dartmouth, Princeton, and Yale) as well as prestigious liberal arts colleges such as Amherst, Wesleyan, and Williams. The distribution of university transition years is plotted in Figure 1. We focus on this time period as it marks a critical acceleration in the transition to coeducation in the U.S. with several prestigious universities admitting women for the first time, and provides a relatively long post-window to observe the career trajectories of affected students.

We obtain other university characteristics from HEGIS and IPEDS collected by the U.S. Department of Education, including the year of establishment, location, public/private status, religious affiliation, annual enrollment, and degrees-awarded data. To illustrate how transitioning to coeducation changed the composition of students at these universities, Figure 2 plots the number of bachelor's degrees awarded to male and female students relative to the year of coeducation adoption.

[3]For instance, some universities have similar names, such as "King's College", making it impossible to identify the specific institution of interest. We exclude such duplicate names from the sample.

Degrees awarded to females rise significantly around adoption dates, while male numbers remain steady. This also suggests that coeducation is associated with an overall expansion of student bodies.

3.2 Resume Data

For data on individuals' education and job experiences, we rely on a 2017 snapshot of LinkedIn profiles. LinkedIn users usually backfill their educational backgrounds on their profiles, which allows us to identify the cohorts affiliated with the universities around the transition years. When matching the resume data with the IPEDS data, we keep the cohorts entering the university at most 10 years before and 10 years after the year of transitioning to coeducation in the stacked sample. We also account for instances where universities changed their names during our sample period. We compare the LinkedIn data coverage to administrative enrollment statistics reported by IPEDS, and find that LinkedIn data cover approximately 42% of individuals who attended the university in our sample period.

Using an unsupervised clustering algorithm, we classify the majors based on information on educational degrees into twelve groups: art & architecture, humanity, social science, economics, business, nursing & health, science, computer related, engineering, education, law, and other liberal arts. Detailed steps on major classification are described in Online Appendix A. For tests related to majors, we drop individuals who do not provide such information on LinkedIn. The resume database also provides detailed career trajectories of each individual. Because high-skilled professionals are the primary users of LinkedIn, we believe the startups that can be observed in our sample are likely of higher impact, if compared with administrative data sources.

3.3 Other Data

We obtain local economic variables from the 1960 Public Use Microdata Sample (PUMS) and the 1969 Bureau of Economic Analysis (BEA) data. The 1960 PUMS is a 5% sample of the 1960 U.S. Census of Population and Housing, with anonymized individual-level data. It is the earliest

decennial census that releases a publicly available 5% sample.[4] We use variables such as gender, marital status, personal income, and educational background to construct county-level measures of employment and income gap between college-educated and non-college-educated married women. Specifically, we categorize married women into college-educated and non-college-educated groups based on whether they have finished a four-year college education. Then, we calculate the county-level average employment probability and personal income across the two groups, and take the difference. Since the 1960 PUMS data lack county identifiers in some areas for confidentiality reasons, we also obtain BEA's county-level income and population data in 1969, the earliest available year, to explore local economic factors driving the timing of coeducation adoption.

3.4 Constructing Variables

3.4.1 Identifying Gender

We determine the gender of the individuals by their first names using the Social Security Administration (SSA) Baby Names dataset.[5] The SSA Baby Names dataset includes the top 1000 popular baby names and covers around 70% of all names in the U.S. every year. We use data from 1930 to 2022 to calculate the number of times the name is given to a female or male and estimate the associated gender probability.[6] We define a name as strongly gendered if over 90% of the people that are given this name are exclusively male or exclusively female (Van Buskirk, Clauset, and Larremore, 2023). Testing the performance of this gender classification method for a randomly chosen set of 1,000 individuals in our sample, 8 percent have no name match in the SSA dataset, and 85.4 percent of the sample names are strongly gendered. We keep individuals classified as strongly gendered. We choose this gender classification method over others because this classification is not simply based on training data limited to recent generations, and the majority of the individuals in our sample went to college in and after the 1960s.

[4] For years before 1960, 1% sample is available instead.

[5] https://www.ssa.gov/oact/babynames/limits.html

[6] We choose the start year of 1930 because a) this is early enough to cover the birth years of the vast majority of individuals in our sample, and b) there are potential selection issues in the database due to reporting at retirement age prior to 1930.

3.4.2 Identifying Entrepreneurs

Our entrepreneur indicator comes from the self-reported job title on LinkedIn. We determine whether the individual has ever been a founder based on the following rules. First, we filter a subsample where the job title includes "founder", "owner", "president", "CEO", "chief executive officer". Second, we require that this is the individual's first job at this company, and the time of the individual joining the company is no later than one year after this company ever occurs in the dataset. This time limit will rule out the cases where the individual holds a management position in a corporation but is not a founder. However, this limit may exclude true entrepreneurs if two or more different firms share the same name and have overlapping periods of operation. We define an alternative measure below to address this potential bias. Third, we exclude the job experiences that are related with chain businesses or incumbent corporations. This type of job title usually contains the following words: "vice", "senior", "assistant", "promoted", "board, "department", "group", "advisor", "intern", "product", "process owner", "project", "representative", "program", "service", "lead", "consultant", "forum", "retired", "retire", "franchise", "agency", "operator", "agent", "insurance", "agent/owner", "owner/agent", "unfranchise", "store", "shop", "operator", "independant", "franchisee", "unfranchise", "allstate", "distributor", "franchise".

We also define a narrow version and a wide version of the founder variable for robustness checks. The narrow definition of founder indicator excludes the cases where the job title includes "president", "CEO", "chief executive officer" even if the individual joins the company at its inception in the LinkedIn database. This approach eliminates the possibility of mistakenly identifying corporate managers as founders. The wide definition of founder indicator takes the value of one when the job title includes "founder" or "owner" regardless of when the company's name string first appears on LinkedIn. This approach helps correct for cases where two or more firms share the same name and have overlapping periods of operation. We believe the narrow definition is likely too restrictive, and that the true entrepreneurship rate falls between the main version and the wide version.

A concern for the resume data is that earlier cohorts mechanically have longer observable career

histories. Since the resume data is a snapshot of career profiles in 2017, the observable career length is right-hand truncated, which may mechanically lead to a decreasing trend in entrepreneurship rate across cohorts over time. While we include cohort fixed effects in all regressions, we further alleviate this concern by constructing a founder measure while limiting the career history for every individual to 18 years.

3.4.3 Start-up Characteristics

Start-up characteristics include the gender of employees and industry distribution. To calculate the share of female employees, we first find all employees for the startup and determine their genders based on the methodology described in Section 3.4.1, and then calculate the percentage of females, conditional on having at least one employee other than the founder himself. We drop startups with more than five individuals in the founder team, which is usually a data error when the start-ups share the same names or when they are chain businesses. We classify the startups with two-digit NAICS industry labels based on the information contained in company names. We leverage an open large language model, ChatGPT, to classify each startup to the most related industry. Our prompt asks the model to classify company names into two-digit NAICS industry groups based on the textual similarity between the company names and keywords contained in NAICS categories and subcategories.[7]

4 Empirical Strategy

4.1 Stacked Sample Construction

Our main specification leverages variations in the timing of the coeducation adoption across universities in stacked DiD specifications, which is akin to a collection of event studies and is similar to the methodology used in Cengiz et al. (2019), Deshpande and Li (2019), Butters et al.

[7]The prompt says "You will be provided with company names. Please classify the company into one category from the following. The classification is based on the similarity between the company name and category key words". NAICS codes and related keywords can be found in Online Appendix B.

(2022), and Wing et al. (2024).[8] Specifically, we group universities by their years of turning coeducation into event groups, so that each event group has a unique treatment year, referred to as the focal treatment year. Within each event group, the treated universities include those that turned coeducation in the focal treatment year, and the event window includes 10 years before and after the focal treatment year. We show robustness when we use a shorter event window of five years around the focal year.

For each event, we next define a set of control universities, which include those that do not themselves turn coeducation during the entire event window around the focal treatment year. For example, in the year of 1978, Hanover College and Hamilton College turned coeducation. Therefore, the treated sample for this event group includes all cohorts entering Hanover College and Hamilton College in the 10 years before and after 1978, i.e. from 1968 to 1988. Then, the control sample for this event group includes cohorts attending universities that turned coeducation outside of the entire event window—either before 1968 or after 1988—during the same event window. We also conduct robustness tests while restricting the control group to universities turning coeducation only after the entire event window, as discussed in Section 4.2. We use this procedure to construct every event group and append all event groups together into a stacked sample.

4.2 Main Specification

We denote event group by e, university by s, entering-year cohort by c, and individual by i. We first estimate the following model:

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \varepsilon_{sci,e}, \quad (1)$$

[8] Recent advancement in the econometric literature has identified stacked DiD regressions as a solution to the potential estimation biases resulting from a two-way fixed-effects model (TWFE) in a staggered DiD setting as documented in previous literature (see Roth, Sant'Anna, Bilinski, and Poe (2023) for a review). To summarize, the main issues with the TWFE model in a staggered DiD setting are implicit weighting and estimation bias when dynamic treatment effects are present (Goodman-Bacon, 2021).

where $\mathbb{1}(\text{founder}_{sci,e})$ is a dummy variable indicating whether during event e, the individual i from university s, cohort c becomes a founder. Cohorts are assigned based on their college entry year. $\text{Treat}_{s,e}$ is a dummy variable that takes the value of one for universities that turned coeducation in event e. $\text{Post}_{sc,e}$ is a dummy variable that takes the value of one for all cohorts entering the university after the focal treatment year in event e. φ_c are entering cohort fixed effects, which control for the time-series trends in founder rates in the economy. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. We cluster standard errors by university, the level at which coeducation adoption varies. Our standard errors therefore account for the possibility that changes in entrepreneurship rate may be correlated within a university, and also address the concern that the same university can serve as a control in multiple events.

We also run specifications in which we include event-specific fixed effects, or the interaction of event-by-university and event-by-cohort fixed effects:

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,e}, \quad (2)$$

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \varphi_{c,e} + \theta_{s,e} + \varepsilon_{sci,e}, \quad (3)$$

where ρ_e are event fixed effects, which ensure that comparisons are made between individuals in treated and control universities within each coeducation turning event. $\varphi_{c,e}$ and $\theta_{s,e}$ are event-by-cohort and event-by-university fixed effects, respectively. These interactive fixed effects allow for heterogeneous estimates of the influence of university characteristics and macroeconomic shocks across different turning events (Butters et al., 2022). In the last specification, the estimated effects on the lower order interaction terms $\text{Treat}_{s,e}$ and $\text{Post}_{sc,e}$ are absorbed.

In our main specification, the control group includes both the later-treated universities (i.e. those that turned coeducation 10 years *after* the focal treatment year) and the early-treated universities (i.e. those already turned coeducation 10 years *before* the focal treatment year). However, if treatment effects are not complete in 10 year, the presence of dynamic treatment effects can potentially bias

our estimates downward. Dynamic DiD estimates suggest that effects are largely complete around the time of adoption, and remain flat several years out, alleviating this concern. Nevertheless, we also run specifications excluding early-treated universities as controls, and confirm that the results are quantitatively similar. In other robustness tests, we also assign sample-share weights when aggregating estimates across event groups following Wing et al. (2024), and show that our main results are robust to using alternative definitions of founders as described in Section 3.4.2, alternative sample windows, and alternative sample structures such as an individual-year sample.

4.3 Identification Assumption

The key identifying assumption underlying our empirical strategy is that individuals' likelihood of becoming entrepreneurs in treated and control universities would move in parallel in the absence of coeducation adoption. While this assumption is not directly testable, we explore the dynamic effects of coeducation on entrepreneurship and examine whether the trends between treated and control groups diverge in the decade prior to coeducation adoption. We estimate the following dynamic model:

$$\begin{aligned}\mathbb{1}(\text{founder}_{sci,e}) = \alpha &+ \sum_{\tau \geq -10}^{\tau \leq 10} \beta_0 \tau \, \mathbb{1}(\text{Year} - \text{Focal Treatment Year} = \tau)_{sc\tau,e} \times \text{Treat}_{s,e} \\ &+ \beta_1 \times \text{Treat}_{s,e} + \sum_{\tau \geq -10}^{\tau \leq 10} \beta_2 \tau \times \mathbb{1}(\text{Year} - \text{Focal Treatment Year} = \tau)_{sc\tau,e} \\ &+ \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,e}, \end{aligned} \tag{4}$$

where $\mathbb{1}(\text{Year} - \text{Focal Treatment Year} = \tau)_{sc\tau,e}$ are a series of event time indicators for years relative to the focal treatment year, and β_τ captures the dynamic treatment effects in each event year. The estimates suggest no systematic pre-trend divergence in the entrepreneurship rate between the treated and control cohorts, as shown in Figure 3.

The identification assumption is more likely to hold if universities' timing of transition to coeducation is independent of factors that might otherwise affect students' decisions to become

entrepreneurs. As described in Section 2, for individual universities, the exact timing of the coeducation decision can be related to factors such as alumni resistance, financial budget, and proximity to other opposite gender schools. Furthermore, an important driver for the coeducation wave after the 1960s is that catholic universities are liberated from the Second Vatican Council admitting women's equal right to education, and this drove other universities to turn coeducation to remain relevant and competitive. Therefore, the decision is less likely to be affected by the entrepreneurial potential of the student body.

Empirically, we examine the factors predicting the timing of a university's transition to coeducation in Appendix Table A1, incorporating both university characteristics and local economic conditions. Undergraduate enrollment, public/private status, local population, population growth, income, income growth, and religious affiliation do not appear to predict the timing of coeducation in a statistically significant way. However, later-established universities tend to transition to coeducation earlier, consistent with anecdotal evidence that influential alumni resistance can delay coeducation adoption. We control for the university's year of establishment through university fixed effects.

Although observable local economic characteristics do not drive adoption, to address any remaining concerns that unobserved local characteristics may simultaneously influence both the universities' decision to adopt coeducation and students' entrepreneurship rates, we use location-by-cohort fixed effects (Appendix Table A6) and find quantitatively similar results.

Another concern might be that coeducational universities attract different types of students through selection rather than influence their future entrepreneurship rates through the collegiate environment. To address this concern, we test our main specification on students who are less likely to base their college choices on coeducation status—those attending public universities.[9] Appendix Table A5 shows that students at public universities exposed to coeducation are 2.3% more likely to become entrepreneurs.

[9]We exclude public universities that are military.

4.4 Summary Statistics

Our sample includes individuals attending 99 coeducation-transitioning universities that were previously all-male, and the transition dates span from 1942 to 1997. The average year of turning coeducation is 1961 and the students' average year of entry is 1984. We present the summary statistics in Table 1, divided into individual-level variables (Panel A) and startup-level variables (Panel B). Panel A includes variables such as majors, connections, cohort, and founder indicators. Panel B includes variables such as female employee shares and industry classifications. In Online Appendix Table A2, we also tabulate the mean of the variables for the treated and control groups separately, in the year prior to the focal treatment year (Panel A) and all pre-treatment periods (Panel B). We also calculate the standardized differences in the means.

5 Results

5.1 Main Results

We begin by documenting that male students who enter the traditionally all-male universities after they turn coeducation have a higher probability of becoming entrepreneurs later in life. In Table 2, the outcome variable is an indicator of whether the individual ever becomes an entrepreneur as defined in Section 3.4.2. Columns (1) to (3) use the [-10, +10] event window around coeducation, and Columns (4) to (6) use the [-5, +5] event window. In Columns (1) and (4), we include school and entry year (cohort) fixed effects. In Columns (2) and (5), we add event fixed effects so that the comparisons are within an event. In Columns (3) and (6), we further control for event-entering cohort and event-university fixed effects to absorb any unobserved persistent characteristics on the level of universities or the year of entering college within each event.

All columns consistently show an increase in entrepreneurship rate of around 1 percentage point, which remains stable with varying fixed effects. For instance, Column (2) coefficients suggest that coeducation leads to a 1.2 percentage point increase in students' entrepreneurship rate, which is

economically meaningful, representing 9.2% of the mean and 3.5% of the standard deviation of the outcome variable.

The estimates remain stable in robustness tests using alternative definitions of the outcome variable, sample lengths, and control groups in Online Appendix Table A3. For each Panel, we show results for the narrow, medium (same as used in Table 2), and wide definitions of the entrepreneur measure. While we show all estimates to be comprehensive, as explained in Section 3.4.2, the narrow definition is likely too restrictive, and the true entrepreneurship rate likely falls between the medium version and the wide version of the variable. We first vary the control groups by only allowing later treated universities as controls in Panel B, and then add weights to the event group estimates following Wing et al. (2024) in Panel C.[10] Last, we vary the event time window by using [-5, +5] year event window in Panel D and Panel E, while also assigning weights to the estimates in Panel E.

One concern with the resume data is that earlier cohorts mechanically have longer observable career histories. Since the resume data is a snapshot of career profiles in 2017, the observable career length is right-hand-side truncated, which will mechanically lead to a decreasing trend in entrepreneurship rate across cohorts over time. To alleviate this concern, we construct a more balanced outcome variable by limiting the career experience to 18 years after graduation for every individual (Appendix Table A4). The number of observations decreases compared with Table 2 because cohorts that entered universities after 1995 are dropped to ensure an 18-year career history.

To examine the dynamics of the treatment effects, we estimate Eq. (5) and plot the event-time coefficients in Figure 3, where Panel A shows results from [-10, +10] year sample and Panel B for [-5, +5] year sample. There is no differential trend between treat and control groups until about a year before the universities officially turn coeducation, and then the effects appear and remain steady over time. The reason for a positive effect prior to the coeducation could be that previous

[10] When aggregating the event group estimates into a single coefficient estimate, Wing et al. (2024) assign weights to the estimates from each event by its share of the treated sample. Specifically, the treatment weight is one, and control weight is $\frac{N_a^D/N_D}{N_a^C/N_C}$, where N_a^D, N_a^C are the number of treatment and control observations in the event group a, and N^D, N^C are the number of total treatment and control observations.

cohort who entered the university before it turns coeducation is also impacted by the new female students from their interaction in the classroom or on campus. However, for more senior students, direct interactions in classrooms with newly entered students are less likely.

Last, we extend our analysis to individual-year level. The individual-year level founder indicator takes the value of one if individual becomes an entrepreneur in a given year. This allows us to estimate the effect on an individual's likelihood of becoming an entrepreneur in a given year of her career, rather than throughout her career. The individual-year level indicator of entrepreneurship takes the value of one if individual i from university s cohort c becomes an entrepreneur in year t. We present the results in Appendix Table A7 and further include year fixed effects to control for the influence of macroeconomic shocks in certain years, in addition to entering cohort, university and event fixed effects. The coefficient in Column (2) is interpreted as coeducation increases the probability of becoming an entrepreneur in a certain year by 0.3%, which represents a 4.3% effect on the mean. Column (5) shows that results using the [-5, +5] event window are of similar magnitude. The dynamic effects for the individual-year level regressions are shown in Appendix Figure A1. These results are similar to previous specifications.

5.2 Startup Industries and Female Employee Shares

Next, we study whether coeducation affects the characteristics of the startups. We begin by examining their industry distribution. We classify these startups based on the information contained in their names, and leverage ChatGPT, an open large language model, to label them with the most related two-digit NAICS industry category as described in Section 3.4.3.

Table 3 shows that post-coeducation, the startups founded by students in affected universities and cohorts are more likely to be in fields such as Arts & Recreation, Accommodation (including hospitality, food services, and more), Wholesale, and Retail. Meanwhile, these startups are less likely to be in fields such as Manufacturing, Information, and Finance, although the standard errors on these estimates are large. Interestingly, the industries that experience increases in entrepreneurship tend to have greater female representation historically than those that experience a decrease.

For example, in 1992 Census data (the earliest we could find that has industry-level female ownership rates), Retailer (31.32%) and Wholesaler (24.86%) have a higher share of female ownership compared to manufacturing (22.46%) and finance (23.98%).[11] Hebert (2023) also documents that start-ups in Arts & Recreation (36.3%), Accommodation (34.2%), and Wholesale and Retail (31.2%) are more likely to be founded by females compared to those in Manufacturing (25.9%), Information (17.4%), and Finance (24.7%).[12] This shift in industry distribution could be consistent with information exchange with female students on campus or after graduation.

Beyond industry distribution, we also find an increase in the share of female employees at the post-coeducation startups. The outcome variables in Table 4 Columns (1) and (2), (3) and (4), and (5) and (6) are the maximum, average, and median female employee share throughout the life of the startups, respectively. We limit the sample to startups with at least two employees. Column (3) shows that the startups founded by affected students after coeducation have 6.9% more female employees on average, which represents 16.2% (0.069/0.425) of a standard deviation and 17.8% (0.069/0.388) of the mean. We also tabulate the results when using the maximum and median female employee share for the startup over its life in Columns (1) and (5). After controlling for startup industries, the increase in the average female employee share is slightly smaller at 5.8% in Columns (2), suggesting that gender composition changes are not solely driven by industry distribution shifts.

These results are consistent with gender diversity in education also influencing gender diversity in the workplace, either through changes in the gender composition of founders' professional networks, or overall shifts in the gender perception and culture of the startup. The intuition is consistent with Duchin et al. (2021) and Calder-Wang and Gompers (2021), who show that the gender experience, such as all-male high school or having more daughters, will impact gender-

[11] These statistics are tabulated in Appendix Table A8. Industry classification standards are different between the 1992 census and the NAICS 2-digit categories. Food and drinking places were listed in Retail in the 1992 Census but changed into Accommodation in NAICS groups. Hotels was listed in Services in the 1992 census but changed into Accommodation in NAICS groups. Services include health, education, professional, management, art, accommodation, and others. See https://www2.census.gov/econ/sbo/92/wb92-1.pdf for reference.

[12] Hebert (2023) uses French administrative data of entrepreneurs from 2002 to 2018, and the female ownership shares are, on average, higher than those in the 1992 US Census data. However, the patterns across industries are relatively similar.

related conceptions and decisions later in life.

5.3 Mechanisms

After establishing the main effects, we explore the mechanisms behind the positive impact of coeducation on entrepreneurship rate. We find that coeducation affects students' major choices and expands their professional networks, which are likely important for the knowledge structure and social capital of the entrepreneur. Furthermore, our evidence is consistent with coeducation affecting household dynamics and risk-sharing, leading to increased entrepreneurial risk-taking.

5.3.1 Major Choice

We find that male students are more likely to major in business and science fields, and less likely to major in humanity after coeducation. We classify the individuals' college majors into twelve categories as shown in Table 5, and the outcome variables are indicators that are equal to one if the individual majors in a given field. Table 5 Columns (3), (6), and (8) show that the probability of majoring in Humanity decreases by 1.9%, and the probability of majoring in Business and Science increases by 3.2% and 1.3%. These findings are consistent with anecdotal evidence from the records of the University of Rochester and Lehigh University. For instance, after the University of Rochester turned coeducation for the first time in 1900, the number of men who obtained a bachelor's degree in Arts and Philosophy dropped from 82 to 45 percent, while the percentage of men who graduated with a Bachelor of Science degree increased from 18 to 55 percent between 1905 and 1909. After Lehigh University turned coeducation in 1971, the share of male undergraduates majoring in Arts & Science dropped from 35% to 25% and while the share in Engineering rose from 51% to 56% over the next ten years (Miller-Bernal and Poulson, 2004, p.57, p.171). These findings also mirror previous evidence that female students can be crowded out of STEM majors after female-exclusive universities turn coeducation (Calkins et al., 2023).

The increased probability of becoming Business and Science majors might positively affect students' entrepreneurship rate. As shown in a survey of 652 U.S.-born tech company founders

with companies established from 1995 through 2005, 33% of them hold a terminal degree in Business, 47% in STEM fields, and 3% in Arts, Humanities, Social Sciences (Wadhwa et al., 2008). The Business- and Tech-related majors might offer students the knowledge and skills in assessing market opportunities, tackling management and financing issues, technical problem-solving and innovation, which enhances entrepreneurship rate (Wadhwa et al., 2008).

5.3.2 Professional Network

We also find that the size of students' professional networks increases after the universities adopt coeducation. We measure the students' network scale by the number of connections shown in their online professional profiles, which are no longer displayed at specific numbers above 500. Thus, we construct dummy variables that equal to one if the number of connections is within $[0, 100)$, $[100, 200)$, $[200, 300)$, $[300, 400)$, $[400, 500)$, and $[500, +\infty)$, respectively. Table 6 shows that coeducation generally increases the likelihood of having more than 500 connections and decreases the likelihood of having connections between 100 and 200. The probability of of having $[500, +\infty)$ connections increases by 1.2%, which represents a 5.4% effect on the mean. When universities become coeducational, student enrollment typically increases, potentially expanding social networks and raising the likelihood of entrepreneurship. This expanded professional network could, in turn, increase the probability of becoming entrepreneurs. While this mechanism may not be unique to coeducation, it coincides with previous evidence on the importance of alumni networks in affecting entrepreneurship (Garfinkel et al., 2021).

5.3.3 Household Risk-Sharing

Finally, coeducation may affect household dynamics and risk sharing. If coeducation increases the likelihood of marrying a college-educated spouse, this could then improve household risk-sharing and downside protection, potentially encouraging greater entrepreneurial risk-taking.[13]

[13]Coeducation might also increase the marriage rate among a university's students. For instance, anecdotal evidence suggests that after Notre Dame adopted coeducation, the percentage of Notre Dame male graduates marrying Saint Mary's female graduates declined, implying that Notre Dame men increasingly married their female schoolmates instead. https://magazine.nd.edu/stories/nd-saint-mary-s-marriages-dwindling/

This channel would suggest that the effect should be stronger in regions where labor force participation and income disparities between college-educated and non-college-educated married women are larger, where the improvements in income and downside protection are likely to be more substantial.

To measure such gaps, we obtain data on employment and income differences between college-educated and non-college-educated married women at the county level from the 1960 PUMS. Table 7 Columns (1) and (2) show that in regions with above-median income gaps, entrepreneurship increases by 1.9%, compared to 0.3% in regions with below-median gaps. Similarly, Columns (3) and (4) show that in regions with an above-median employment gap, entrepreneurship increases by 1.5%, compared to 0.8% in other regions.[14] Figure 4 further shows the heterogeneous dynamic effects of coeducation on entrepreneurship for these groups.[15]

6 Conclusion

Leveraging the historical context of U.S. universities' transition to coeducation, we examine how exposure to a gender-diverse higher education environment influences an individual's decision to become an entrepreneur. We find that the shift to coeducation results in a statistically significant and economically meaningful increase in the likelihood of entrepreneurship for affected students. This rise in entrepreneurship is partly explained by changes in major choices, the expansion of professional networks, and improved risk-sharing within households. Furthermore, gender-inclusive educational environments not only raise entrepreneurship rates, but also shape the nature of entrepreneurial ventures—post-coeducation, male entrepreneurs' startups are more likely to be in industries with larger female presence historically, and have a higher proportion of female employees. These findings suggest that gender diversity in higher education has a lasting impact on startup creation and entrepreneurial dynamics.

[14]The sample size decreases compared to previous tables due to missing county identifiers in the 1960 PUMS data.

[15]To address concerns that some early-turning universities may have influenced local income and employment gaps by 1960, we further restrict the analysis to transitions occurring after 1960. The results, presented in Table A9 and Figure A2, remain similar.

References

Barrios, J. M., Hochberg, Y. V., Yi, H., 2022. Launching with a parachute: The gig economy and new business formation. Journal of Financial Economics 144, 22–43.

Becker, G. S., 1974. A theory of marriage. Economics of the Family: Marriage, Children, and Human Capital/Post Secondary of Chicago Press .

Benham, L., 1974. Benefits of women's education within marriage. Journal of political Economy 82, S57–S71.

Booth, A., Cardona-Sosa, L., Nolen, P., 2014. Gender differences in risk aversion: Do single-sex environments affect their development? Journal of economic behavior & organization 99, 126–154.

Butters, R. A., Sacks, D. W., Seo, B., 2022. How do national firms respond to local cost shocks? American Economic Review 112, 1737–1772.

Calder-Wang, S., Gompers, P. A., 2021. And the children shall lead: Gender diversity and performance in venture capital. Journal of Financial Economics 142, 1–22.

Calkins, A., Binder, A. J., Shaat, D., Timpe, B., 2023. When sarah meets lawrence: The effects of coeducation on women's college major choices. American Economic Journal: Applied Economics 15, 1–34.

Cengiz, D., Dube, A., Lindner, A., Zipperer, B., 2019. The effect of minimum wages on low-wage jobs. The Quarterly Journal of Economics 134, 1405–1454.

Denes, M. R., Lagaras, S., Tsoutsoura, M., 2025. Entrepreneurship and the gig economy: Evidence from us tax returns. Tech. rep., National Bureau of Economic Research.

Deshpande, M., Li, Y., 2019. Who is screened out? application costs and the targeting of disability programs. American Economic Journal: Economic Policy 11, 213–248.

Duchin, R., Simutin, M., Sosyura, D., 2021. The origins and real effects of the gender gap: Evidence from ceos' formative years. The Review of Financial Studies 34, 700–762.

Ewens, M., 2022. Gender and race in entrepreneurial finance .

Ewens, M., Townsend, R. R., 2020. Are early stage investors biased against women? Journal of Financial Economics 135, 653–677.

Garfinkel, J. A., Mayer, E. J., Strebulaev, I. A., Yimfor, E., 2021. Alumni networks in venture capital financing. SMU Cox School of Business Research Paper .

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., Shleifer, A., 1992. Growth in cities. Journal of political economy 100, 1126–1152.

Goldin, C., Katz, L. F., 2011. Putting the "co" in education: Timing, reasons, and consequences of college coeducation from 1835 to the present. Journal of Human Capital 5, 377–417.

Goodman-Bacon, A., 2021. Difference-in-differences with variation in treatment timing. Journal of econometrics 225, 254–277.

Gornall, W., Strebulaev, I. A., 2024. Gender, race, and entrepreneurship: A randomized field experiment on venture capitalists and angels. Management Science .

Gottlieb, J. D., Townsend, R. R., Xu, T., 2022. Does career risk deter potential entrepreneurs? The Review of Financial Studies 35, 3973–4015.

Hebert, C., 2023. Gender stereotypes and entrepreneur financing. In: *10th Miami behavioral finance conference*.

Hebert, C., Tookes, H., Yimfor, E., 2024. Financing the next vc-backed startup: The role of gender. Available at SSRN 4784414 .

Hombert, J., Schoar, A., Sraer, D., Thesmar, D., 2020. Can unemployment insurance spur entrepreneurial activity? evidence from france. The Journal of Finance 75, 1247–1285.

Howell, S. T., Nanda, R., 2024. Networking frictions in venture capital, and the gender gap in entrepreneurship. Journal of Financial and Quantitative Analysis 59, 2733–2761.

Jacobs, J., 1969. The economy of. Cities. New York: Penguin Books .

Jones, C. I., 2023. Recipes and economic growth: A combinatorial march down an exponential tail. Journal of Political Economy 131, 1994–2031.

Kacperczyk, A. J., 2013. Social influence and entrepreneurship: The effect of university peers on entrepreneurial entry. Organization Science 24, 664–683.

Lee, K. H., 2021. Causal effects of single-sex schools on long-term outcomes: Evidence from random high school assignment in south korea. Available at SSRN 3957630 .

Lee, Y., Nakazawa, N., 2022. Does single-sex schooling help or hurt labor market outcomes? evidence from a natural experiment in south korea. Journal of Public Economics 214, 104729.

Lerner, J., Malmendier, U., 2013. With a little help from my (random) friends: Success and failure in post-business school entrepreneurship. The Review of Financial Studies 26, 2411–2452.

Levine, R., Rubinstein, Y., 2017. Smart and illicit: who becomes an entrepreneur and do they earn more? The Quarterly Journal of Economics 132, 963–1018.

Malkiel, N. W., 2016. Keep the damned women out: The struggle for coeducation .

Mertz, M., Ronchi, M., Salvestrini, V., 2024. Female representation and talent allocation in entrepreneurship: the role of early exposure to entrepreneurs. Available at SSRN 4920176 .

Mill, J. S., 1965. Principles of political economy. U of Toronto P .

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

Nanda, R., Sørensen, J. B., 2010. Workplace peers and entrepreneurship. Management science 56, 1116–1126.

Özcan, B., 2011. Only the lonely? the influence of the spouse on the transition to self-employment. Small Business Economics 37, 465–492.

Park, S. Y., 2023. Do gender enclaves help girls' empowerment? evidence from a natural experiment in south korea. Evidence from a Natural Experiment in South Korea (November 22, 2023) .

Posch, M., Schulz, J., Henrich, J., 2023. Surname diversity, social ties and innovation. Social Ties and Innovation (August 3, 2023) .

Poulson, S. L., 1995. From single-sex to coeducation: the advent of coeducation at georgetown, 1965-1975. US Catholic Historian 13, 117–137.

Roth, J., Sant'Anna, P. H., Bilinski, A., Poe, J., 2023. What's trending in difference-in-differences? a synthesis of the recent econometrics literature. Journal of Econometrics 235, 2218–2244.

Schumpeter, J., 2017. Theorie der wirtschaftlichen Entwicklung. Duncker und Humblot.

Truffa, F., Wong, A., 2024. Undergraduate gender diversity and the direction of scientific research .

Van Buskirk, I., Clauset, A., Larremore, D. B., 2023. An open-source cultural consensus approach to name-based gender classification. In: *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 17, pp. 866–877.

Wadhwa, V., Freeman, R. B., Rissing, B. A., 2008. Education and tech entrepreneurship. Available at SSRN 1127248 .

Wallskog, M., 2022. Entrepreneurial spillovers across coworkers. Tech. rep., Working paper, Duke University.

Weitzman, M. L., 1998. Recombinant growth. The Quarterly Journal of Economics 113, 331–360.

Wing, C., Freedman, S. M., Hollingsworth, A., 2024. Stacked difference-in-differences. Tech. rep., National Bureau of Economic Research.

Wong, Y.-C., 1986. Entrepreneurship, marriage, and earnings. The Review of Economics and Statistics pp. 693–699.

Zandberg, J., 2021. Family comes first: Reproductive health and the gender gap in entrepreneurship. Journal of Financial Economics 140, 838–864.

Figure 1: Distribution of Coeducation Transition Dates

This figure displays the distribution of years in which traditionally all-male universities transitioned to coeducation. In our sample, 99 all-male universities adopted coeducation between 1942 and 1997. The figure below shows the number of traditionally all-male universities making the shift to coeducation each year.

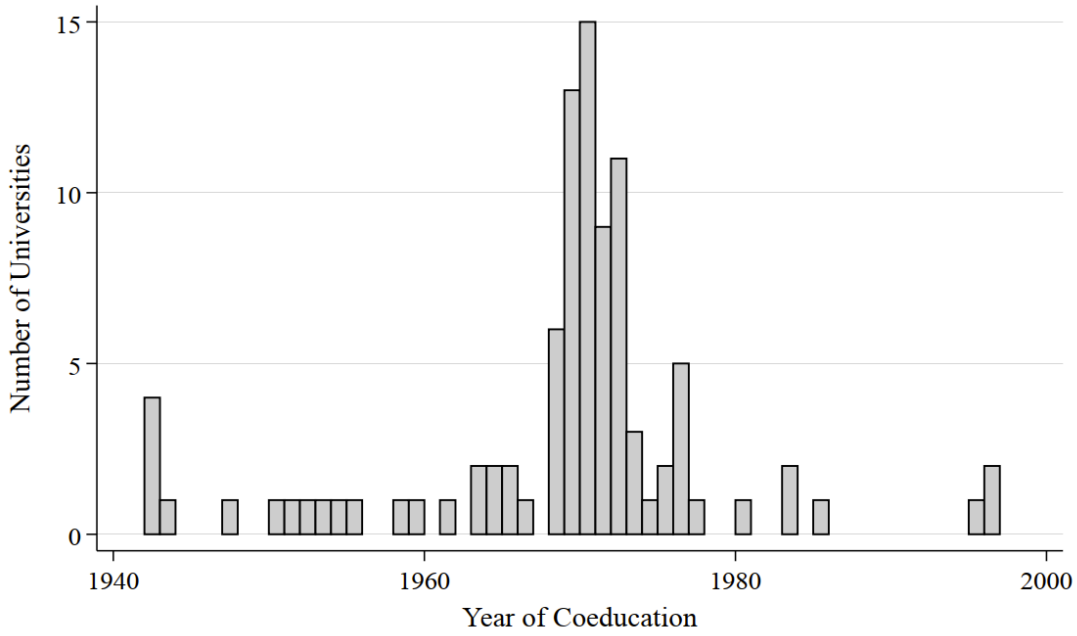


Figure 2: Coeducation and Student Gender Composition

This graph illustrates the gender composition of students in traditionally all-male universities transitioning to coeducation. The y-axis represents the natural logarithm of the number of degrees awarded to male and female students, respectively, while the x-axis shows the number of years relative to the year of coeducation adoption.

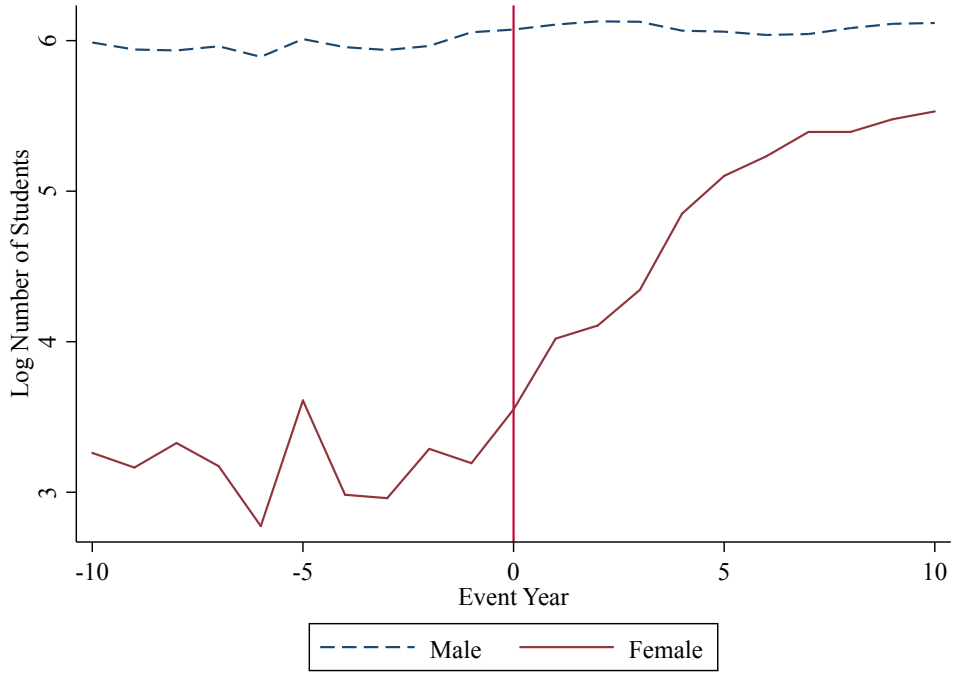


Figure 3: Dynamic Effects of Coeducation on Entrepreneurship

We estimate the following model:

$$\begin{aligned}\mathbb{1}(\text{founder}_{sci,e}) = \alpha &+ \sum_{\tau \geq -10}^{\tau \leq 10} \beta_{0\tau}\, \mathbb{1}(\text{Year} - \text{Focal Treatment Year} = \tau)_{sc\tau,e} \times \text{Treat}_{s,e} \\ &+ \beta_1 \times \text{Treat}_{s,e} + \sum_{\tau \geq -10}^{\tau \leq 10} \beta_{2\tau} \times \mathbb{1}(\text{Year} - \text{Focal Treatment Year} = \tau)_{sc\tau,e} \\ &+ \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,e}, \end{aligned} \tag{5}$$

The dependent variable is an indicator that takes the value of one if individual i from cohort c in university s becomes a founder. $\text{Treat}_{s,e}$ equals one if in event e, university s turns coeducation. $\mathbb{1}(\text{Year} - \text{Focal Treatment Year} = \tau)_{sc\tau,e}$ is a dummy variable that takes the value of one if cohort c enters university s in the τth year relative to the year of transition in event e. φ_c are entering cohort fixed effects, which control for the time-series trends in founder rates in the economy. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. ρ_e are event fixed effects, which ensure that comparisons are made between individuals in treated and control universities within each coeducation turning event. β_τ captures the dynamic effects for each event year. Panel A uses a [-10, +10] year event window relative to the year of transition, and Panel B uses a [-5, +5] year event window. Both panels allow early- and later-treated groups in the control group. Estimates at $t = -5$ are omitted. Standard errors are clustered by university.

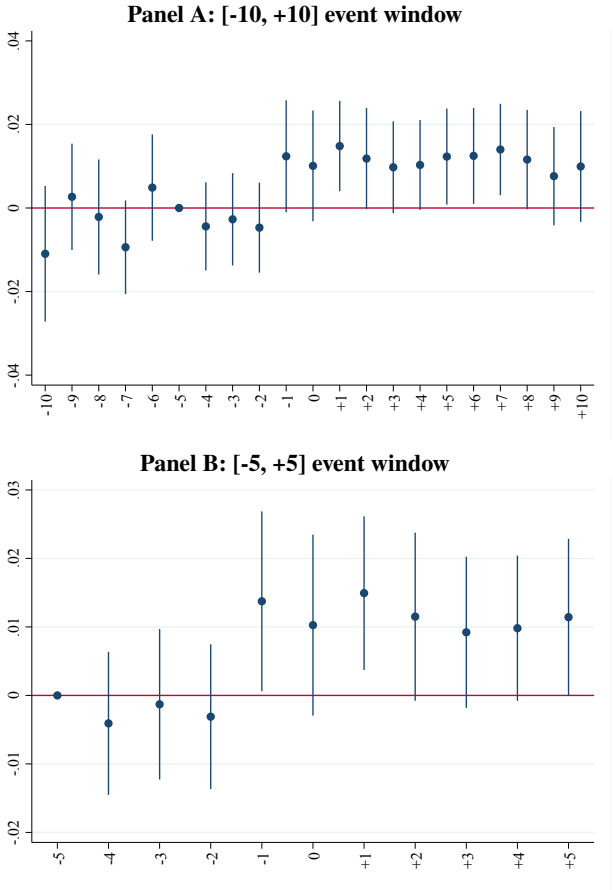


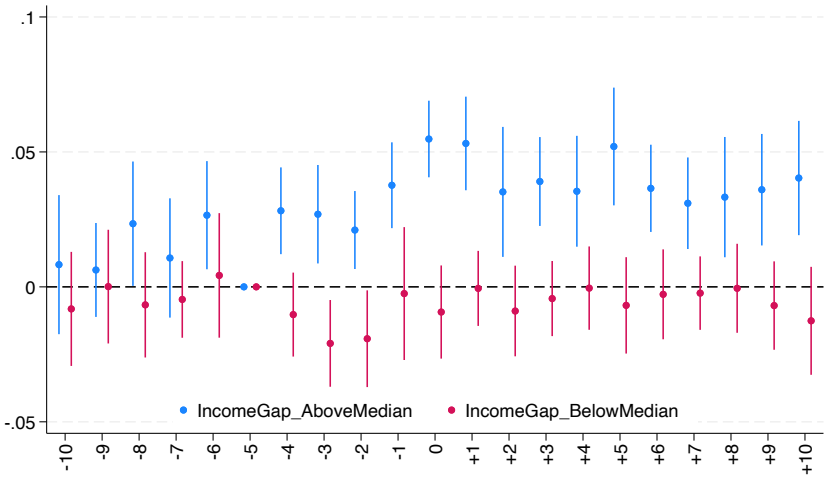
Figure 4: Coeducation, Spousal Employment and Income, and Entrepreneurship

We estimate the following model, dividing the sample based on the median values of the income and employment gap between college-educated and non-college-educated married women:

$$\begin{aligned}\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \sum_{\tau \geq -10}^{\tau \leq 10} \beta_{0\tau}\, \mathbb{1}(\text{Year} - \text{Focal Treatment Year} = \tau)_{sc\tau,e} \times \text{Treat}_{s,e} \\ + \beta_1 \times \text{Treat}_{s,e} + \sum_{\tau \geq -10}^{\tau \leq 10} \beta_{2\tau} \times \mathbb{1}(\text{Year} - \text{Focal Treatment Year} = \tau)_{sc\tau,e} \\ + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,e}, \end{aligned} \tag{5}$$

The dependent variable is an indicator that takes the value of one if individual i from cohort c in university s becomes a founder. $\text{Treat}_{s,e}$ equals one if in event e, university s turns coeducation. $\mathbb{1}(\text{Year} - \text{Focal Treatment Year} = \tau)_{sc\tau,e}$ is a dummy variable that takes the value of one if cohort c enters university s in the τth year relative to the year of transition in event e. φ_c are entering cohort fixed effects, which control for the time-series trends in founder rates in the economy. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. ρ_e are event fixed effects, which ensure that comparisons are made between individuals in treated and control universities within each coeducation turning event. β_τ captures the dynamic effects for each event year. Panel A divides the sample based on the median value of income gap between college-educated and non-college-educated married women at county level, and Panel B divides the sample based on the median value of the employment gap. We calculate the income gap and employment gap between college-educated and non-college-educated married women using 1960 PUMS data at county level. Both panels use a [-10, +10] year event window and allow early- and later-treated groups in the control group. Estimates at $t = -5$ are omitted. Standard errors are clustered by university.

Panel A: Subsample by Income Gap between College-Educated and Non-College-Educated Women



Panel B: Subsample by Employment Gap between College-Educated and Non-College-Educated Women

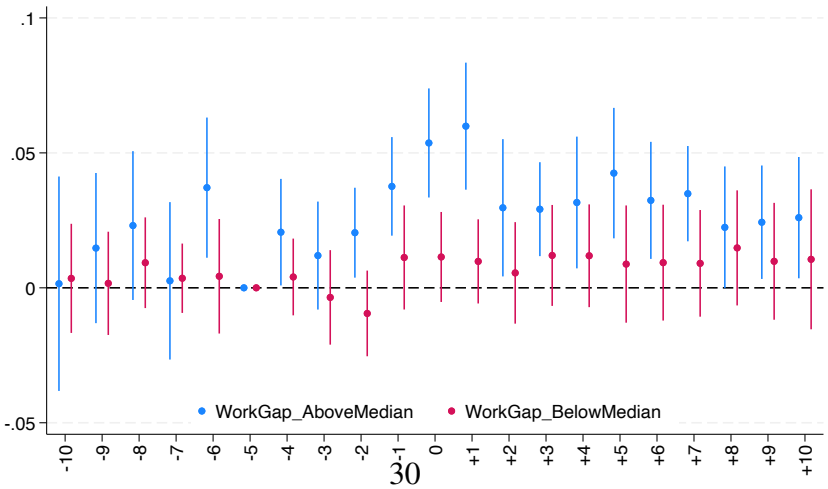


Table 1: Summary Statistics

This table presents the summary statistics of individual- and startup-level variables. Panel A shows individual characteristics related to cohort, founder indicators, connections, and majors. Panel B shows startup characteristics related to female employee shares and industry classifications. *Year of Entry* and *Year of Graduation* describe the entering and graduation cohorts for each individual. *Founder_Narrow*, *Founder_Medium*, and *Founder_Wide* are dummy variables that takes the value of one if this individual becomes an entrepreneur according to the narrow, medium and wide definitions of entrepreneurship, as defined in Section 3.4.2. *Founder_Narrow_18yr*, *Founder_Medium_18yr*, and *Founder_Wide_18yr* are dummy variables that equal one if this individual becomes an entrepreneur within 18 years after college graduation. *Connection* variables are dummy variables that takes the value of one if an individual's number of connections falls within the range $[0, 100)$, $[100, 200)$, $[200, 300)$, $[300, 400)$, $[400, 500)$, and $[500, +\infty)$. *Major* variables are dummy variables that takes the value of one if an individual major in certain field. *Female Employee Share* refers to the proportion of female employees in a startup, calculated for firms with at least two employees. We present the maximum, average and median share of female employees throughout the life of startups. *Industry* variables are indicators that denote whether a startup belongs to a specific two-digit NAICS industry code..

Panel A: Individual Characteristics			
Variable	Mean	Standard Deviation	Median
Year of Entry	1984	13.25	1984
Year of Graduation	1988	13.24	1988
Founder_Narrow	0.08	0.27	0
Founder_Medium	0.13	0.34	0
Founder_Wide	0.15	0.35	0
Founder_Narrow_18yr	0.05	0.22	0
Founder_Medium_18yr	0.08	0.27	0
Founder_Wide_18yr	0.09	0.29	0
Connection_0_100	0.32	0.47	0
Connection_100_200	0.18	0.38	0
Connection_200_300	0.13	0.33	0
Connection_300_400	0.09	0.29	0
Connection_400_500	0.06	0.24	0
Connection_500	0.22	0.42	0
Art & Architecture	0.03	0.16	0
Other Liberal Art	0.12	0.33	0
Humanity	0.10	0.30	0
Social Science	0.11	0.31	0
Economics	0.05	0.22	0
Business	0.23	0.42	0
Nursing & Health	0.01	0.10	0
Science	0.09	0.29	0
Computer	0.05	0.21	0
Engineering	0.21	0.41	0
Education	0.01	0.07	0
Law	0.01	0.08	0

Table 1: Summary Statistics (continued)

Panel B: Startup Characteristics			
Variable	Mean	Standard Deviation	Median
Max(Female Employee Share)	0.559	0.511	0.5
Avg(Female Employee Share)	0.388	0.425	0.248
Med(Female Employee Share)	0.381	0.447	0.200
Agriculture	0.014	0.116	0
Mining	0.009	0.097	0
Construction	0.017	0.129	0
Manufacturing	0.049	0.217	0
Wholesaler	0.011	0.103	0
Retailer	0.037	0.189	0
Transportation	0.028	0.166	0
Information	0.140	0.347	0
Finance	0.106	0.308	0
Real Estate	0.036	0.187	0
Professional Service	0.200	0.400	0
Management	0.074	0.261	0
Administrative Service	0.028	0.164	0
Education	0.033	0.180	0
Healthcare	0.079	0.270	0
Arts & Recreation	0.078	0.268	0
Accommodation	0.034	0.181	0
Other Services	0.026	0.158	0

Table 2: Coeducation and the Probability of Becoming an Entrepreneur

To estimate the impact of coeducation on students' probability of becoming an entrepreneur, we estimate the following specifications in OLS with varying fixed effects:

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \varepsilon_{sci,e}, \quad (1)$$

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,e}, \quad (2)$$

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \varphi_{c,e} + \theta_{s,e} + \varepsilon_{sci,e}. \quad (3)$$

The dependent variable is a dummy variable indicating whether in event e, the individual i from university s, cohort c becomes a founder. $\text{Treat}_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $\text{Post}_{sc,e}$ is a dummy variable that takes the value of one for all cohorts entering the university after the focal treatment year in event e. φ_c are entering cohort fixed effects, which control for overall trends in founder rates related with cohorts. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. ρ_s are event fixed effects that forces comparisons to be across individuals in treated and control universities within each coeducation turning event. Further, $\varphi_{c,e}$ and $\theta_{s,e}$ are event-by-entering cohort and event-by-university fixed effects, respectively. These interactive fixed effects allow for heterogeneous estimates of the influence of university characteristics and macroeconomic shocks across events. β estimates the impact of entering the university after its transition to coeducation on the individual rate of entrepreneurship. Columns (1) and (4) show results from Eq.(1). Columns (2) and (5) show results from Eq.(2). Columns (3) and (6) show results from Eq.(3). Standard errors clustered by university are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	𝟙(Founder)					
Event Window:	[-10, +10]			[-5, +5]		
	(1)	(2)	(3)	(4)	(5)	(6)
Treat × Post	0.011***	0.012***	0.012***	0.009***	0.010***	0.011***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Treat	-0.006**	-0.007**		-0.004	-0.005	
	(0.003)	(0.003)		(0.003)	(0.003)	
Post	-0.002**	-0.001*		-0.001**	-0.000	
	(0.001)	(0.001)		(0.001)	(0.001)	
Observations	3,454,538	3,454,538	3,454,490	2,225,874	2,225,874	2,225,815
R-squared	0.025	0.025	0.026	0.022	0.022	0.024
School FE	Yes	Yes		Yes	Yes	
Entry Year FE	Yes	Yes		Yes	Yes	
Event FE		Yes			Yes	
Event × School FE			Yes			Yes
Event × Entry Year FE			Yes			Yes

Table 3: Coeducation and Startup Industry Distribution

To estimate the impact of coeducation on startups' industry, we estimate the following model in OLS:

$$\text{industry}_{sci,f,e} = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,f,e}, \quad (6)$$

The dependent variables are dummy variables that take the value of one if startup f, founded by individual i from cohort c university s in event e, is categorized under certain two-digit NAICS industry code. We ask ChatGPT to classify startups based the similarity between startup names and NAICS category keywords (described in Section 3.4.3). $\text{Treat}_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $\text{Post}_{sc,e}$ is a dummy variable that takes the value of one for all cohorts entering the university after the focal treatment year in event e. φ_c are entering cohort fixed effects, which control for the time-series trends in founder rates in the economy. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. ρ_e are event fixed effects, which ensure that comparisons are made between individuals in treated and control universities within each coeducation turning event. β estimates the impact of entering the university after its transition to coeducation on the startup's industry. Standard errors clustered by university are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Real Estate	Professional Service	Management	Administrative Service	Education	Healthcare	Arts & Recreation	Accmodation	Other Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treat × Post	-0.004	-0.013	0.015	-0.014	0.014	0.008	0.024*	0.023*	0.012
	(0.015)	(0.027)	(0.013)	(0.012)	(0.009)	(0.021)	(0.014)	(0.013)	(0.009)
Treat	0.003	0.009	-0.006	0.006	-0.008	-0.006	-0.006	-0.010	-0.005
	(0.007)	(0.015)	(0.007)	(0.006)	(0.005)	(0.011)	(0.008)	(0.006)	(0.005)
Post	0.006	0.013	-0.002	0.006	-0.014*	-0.018	-0.027**	-0.026**	-0.009
	(0.013)	(0.024)	(0.010)	(0.011)	(0.008)	(0.019)	(0.011)	(0.011)	(0.007)
Observations	111,025	111,025	111,025	111,025	111,025	111,025	111,025	111,025	111,025
R-squared	0.009	0.012	0.010	0.008	0.010	0.017	0.012	0.009	0.010
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Event FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

	Agriculture	Mining	Construction	Manufacturing	Wholesaler	Retailer	Transportation	Information	Finance
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treat × Post	0.003	0.003	0.002	-0.023	0.014**	0.024**	0.004	-0.048	-0.044
	(0.006)	(0.008)	(0.007)	(0.023)	(0.005)	(0.010)	(0.006)	(0.031)	(0.039)
Treat	-0.001	-0.003	-0.000	0.008	-0.007**	-0.013**	-0.002	0.023	0.020
	(0.003)	(0.004)	(0.004)	(0.012)	(0.003)	(0.005)	(0.004)	(0.015)	(0.021)
Post	-0.004	0.001	-0.007	0.026	-0.009*	-0.014*	-0.004	0.037	0.046
	(0.006)	(0.007)	(0.006)	(0.021)	(0.005)	(0.008)	(0.005)	(0.028)	(0.036)
Observations	111,025	111,025	111,025	111,025	111,025	111,025	111,025	111,025	111,025
R-squared	0.011	0.009	0.013	0.015	0.005	0.012	0.020	0.019	0.014
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Event FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 4: Coeducation and Startups' Employee Gender Composition

To estimate the impact of coeducation on startups' employee gender composition, we estimate the following model in OLS:

$$\text{femaleshare}_{sci,f,e} = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,f,e}, \quad (7)$$

The dependent variable is the female employee share in startup f, founded by individual i from cohort c university s in event e. For each startup, we calculate its maximum, average, median share of female employees throughout the life of startups. $\text{Treat}_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $\text{Post}_{sc,e}$ is a dummy variable that takes the value of one for all cohorts entering the university after the focal treatment year in event e. φ_c are entering cohort fixed effects, which control for the time-series trends in founder rates in the economy. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. ρ_e are event fixed effects, which ensure that comparisons are made between individuals in treated and control universities within each coeducation turning event. β estimates the impact of entering the university after its transition to coeducation on the startup's female employee share. Standard errors clustered by university are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Max(Female Employee Pct)		Avg(Female Employee Pct)		Median(Female Employee Pct)	
	(1)	(2)	(3)	(4)	(5)	(6)
Treat × Post	0.131***	0.120***	0.069***	0.058**	0.061**	0.049*
	(0.035)	(0.036)	(0.025)	(0.027)	(0.026)	(0.027)
Treat	-0.067***	-0.063***	-0.031**	-0.027*	-0.023	-0.019
	(0.020)	(0.020)	(0.014)	(0.016)	(0.015)	(0.016)
Post	-0.130***	-0.121***	-0.088***	-0.079***	-0.083***	-0.073***
	(0.031)	(0.032)	(0.022)	(0.024)	(0.023)	(0.024)
Observations	41,077	41,077	41,077	41,077	41,077	41,077
R-squared	0.032	0.061	0.027	0.062	0.026	0.060
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	No	Yes	No	Yes
Event FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 5: Coeducation and College Major Choice

To explore the mechanisms of how coeducation affects entrepreneurship, we test the impact of coeducation on students' major choices by the following model in OLS:

$$\text{major}_{sci,e} = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,e}, \quad (8)$$

The dependent variables are dummy variables indicating whether individual i from cohort c university s in event e takes their undergraduate degree in certain major. $\text{Treat}_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $\text{Post}_{sc,e}$ is a dummy variable that takes the value of one for all cohorts entering the university after the focal treatment year in event e. φ_c are entering cohort fixed effects, which control for the time-series trends in founder rates in the economy. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. ρ_e are event fixed effects, which ensure that comparisons are made between individuals in treated and control universities within each coeducation turning event. β estimates the impact of entering the university after its transition to coeducation on the major choice. Standard errors clustered by university are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	The Probability of Choosing the Major of ...					
	Art & Architecture	Other Liberal Art	Humanity	Social Science	Economics	Business
	(1)	(2)	(3)	(4)	(5)	(6)
Treat × Post	0.001	-0.005	-0.019**	-0.002	-0.003	0.032*
	(0.003)	(0.006)	(0.008)	(0.006)	(0.006)	(0.018)
Treat	-0.001	-0.006	0.015*	0.009*	0.006	-0.028*
	(0.004)	(0.005)	(0.008)	(0.005)	(0.006)	(0.016)
Post	0.000	0.001	0.000	-0.000	-0.001*	-0.001
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Observations	2,954,162	2,954,162	2,954,162	2,954,162	2,954,162	2,954,162
R-squared	0.023	0.022	0.101	0.046	0.041	0.148
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Event FE	Yes	Yes	Yes	Yes	Yes	Yes
	The Probability of Choosing the Major of ...					
	Nursing & Health	Science	Computer	Engineering	Education	Law
	(7)	(8)	(9)	(10)	(11)	(12)
Treat × Post	-0.001	0.013*	-0.008	-0.009	0.001	-0.001
	(0.002)	(0.007)	(0.006)	(0.020)	(0.001)	(0.004)
Treat	0.004*	-0.012*	0.016**	-0.003	-0.001	-0.000
	(0.002)	(0.007)	(0.006)	(0.016)	(0.001)	(0.003)
Post	-0.000	-0.000	0.001	-0.000	0.000	-0.000
	(0.000)	(0.001)	(0.001)	(0.003)	(0.000)	(0.000)
Observations	2,954,162	2,954,162	2,954,162	2,954,162	2,954,162	2,954,162
R-squared	0.010	0.035	0.037	0.244	0.009	0.041
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Event FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 6: Coeducation and Professional Networks

To explore the mechanisms of how coeducation affects entrepreneurship, we test the impact of coeducation on individuals' professional networks by the following model in OLS:

$$\text{connections}_{sci,e} = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,e}, \quad (9)$$

The dependent variables measures the number of connections for individual i from cohort c university s in event e. The original variable is right-hand-side truncated at 500 and is transformed into dummy variables, each equal to one if the value falls within the ranges $[0, 100)$, $[100, 200)$, $[200, 300)$, $[300, 400)$, $[400, 500)$, and $[500, +\infty)$, presented in columns (1) through (6), respectively. $\text{Treat}_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $\text{Post}_{sc,e}$ is a dummy variable that takes the value of one for all cohorts entering the university after the focal treatment year in event e. φ_c are entering cohort fixed effects, which control for the time-series trends in founder rates in the economy. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. ρ_e are event fixed effects, which ensure that comparisons are made between individuals in treated and control universities within each coeducation turning event. β estimates the impact of entering the university after its transition to coeducation on the networking scale. Standard errors clustered by university are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Connections 0-100	Connections 100-200	Connections 200-300	Connections 300-400	Connections 400-500	Connections 500+
	(1)	(2)	(3)	(4)	(5)	(6)
Treat × Post	-0.002	-0.006*	-0.003	-0.002	0.002	0.012***
	(0.005)	(0.004)	(0.003)	(0.003)	(0.002)	(0.005)
Treat	0.003	0.007*	0.002	0.000	-0.001	-0.012*
	(0.005)	(0.004)	(0.003)	(0.002)	(0.001)	(0.007)
Post	-0.000	0.000	0.001*	0.000	0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Observations	3,454,538	3,454,538	3,454,538	3,454,538	3,454,538	3,454,538
R-squared	0.063	0.003	0.003	0.005	0.005	0.037
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Event FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 7: Coeducation, Spousal Employment and Income, and Entrepreneurship

To examine the heterogeneous impact of coeducation on entrepreneurship, we estimate the following OLS model in sub-samples split at the median of income and work probability gaps between college-educated and non-college-educated married women:

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,f,e}, \quad (2)$$

The dependent variable is whether individual i from cohort c university s in event e has ever become an entrepreneur. $\text{Treat}_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $\text{Post}_{sc,e}$ is a dummy variable that takes the value of one for all cohorts entering the university after the focal treatment year in event e. φ_c are entering cohort fixed effects, which control for the time-series trends in founder rates in the economy. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. ρ_e are event fixed effects, which ensure that comparisons are made between individuals in treated and control universities within each coeducation turning event. β estimates the impact of entering the university after its transition to coeducation on individual entrepreneurship rate. We compute the differences in average income and employment between college-educated and non-college-educated women using the 1960 PUMS data at the county level. Columns (1) and (2) present results for sub-samples where the income gap variable is above and below the median, respectively. Similarly, Columns (3) and (4) display results for sub-samples where the employment gap variable is above and below the median. The sample size decreases compared to Table 2 because some areas have missing county identifiers due to confidentiality reasons in the PUMS data. Standard errors clustered by university are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	𝟙(Founder)			
	by Income Gap		by Employment Gap	
	Above Median	Below Median	Above Median	Below Median
	(1)	(2)	(3)	(4)
Treat × Post	0.019***	0.003	0.015**	0.008
	(0.006)	(0.004)	(0.006)	(0.007)
Treat	-0.012	-0.000	-0.004	-0.012
	(0.008)	(0.004)	(0.007)	(0.008)
Post	-0.001	0.001	0.000	0.000
	(0.001)	(0.002)	(0.001)	(0.002)
Observations	1,094,152	1,094,896	1,052,718	1,136,330
R-squared	0.027	0.025	0.024	0.028
School FE	Yes	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes	Yes
Event FE	Yes	Yes	Yes	Yes

Online Appendix

A Major Classification

We apply an unsupervised clustering algorithm to classify the majors from self-reported educational degrees. We first split the multiple educational degrees by "minor" and ";". Next, we clean and tokenize the strings by removing symbols, numbers and stopwords. Stopwords include common English stopwords and in our case: degree abbreviations, "degree", "magna", "cum", "laude", "phi", "beta", "kappa", "honors", "distinction", "summa", "laud", "valedictorian", "school", "year", "class", "classman", "freshmen", "freshman", "alumni", "sophomore", "junior", "senior", "major", "minor", "study", "studies", "deans list", "dean list", "applied", "apply", "concentration", "science", "art", "arts", "gpa", "major gpa", "general", "fine arts", "sciences".

Then, we vectorize the text using Term Frequency-Inverse Document Frequency (TF-IDF), which assigns weights that reflect the significance of a word within the document relative to general. We use K-means clustering and then manually sort the clusters into twelve groups: art & architecture, humanity, social science, economics, business, nursing & health, science, computer related, engineering, education, law and other liberal arts.

The specific fields within each major group are as follows:

Art & architecture includes design, dance, music, theater, architecture, etc.

Humanity includes philosophy, languages, literature, history, etc.

Social science includes sociology, political science, international studies, media, anthropology, human resource, etc.

Business includes finance, accounting, marketing, business communication, entrepreneurial studies, etc.

Nursing & health includes nursing, exercise science, physical science, and pre-meds.

Science includes physics, mathematics, chemistry, biology, geology, etc.

Computer related include computer science, software engineering, information system, etc.

Engineering includes civil engineering, electric engineering, mechanical engineering, etc.

Law includes justice, criminal studies, etc.

Other liberal arts include strings related with "liberal art" but has no further specification.

B Key Words by NAICS Industry Classification

NAICS codes and related keywords are as follows:

11 Agriculture: agriculture, forestry, logging, fishing, hunting, crop, farming, ranching.

21 Mining: mining, quarrying, oil and gas extraction.

23 Construction: construction, contractors.

31-33 Manufacturing: manufacturing.

42 Wholesaler: merchant wholesaler.

44-45 Retailer: retailers, dealers, stores, gasoline stations, grocery, florists.

48-49 Transportation: transportation, logistics, freight trucking, taxi, bus, delivery, warehousing storage.

51 Information: information, publishing, motion picture, video, recording, broadcasting, internet, telecommunications, data, software, radio, television, media networks, computing infrastructure.

52 Finance: finance, insurance, bank, credit, investments, securities, funds, trusts, mortgage, brokerage.

53 Real Estate: real estate, rental, leasing, car rental.

54 Professional Service: professional, scientific, and technical services, lawyers, legal services, accounting, tax preparation, engineering service, design services, advertising, public relations, photographic.

55 Management: management of companies and enterprises, company offices.

56 Administrative Support: administrative support, waste management services, travel agencies, security guard, investigation, pest control, janitorial, carpet cleaning.

61 Education: college, schools, driving schools.

62 Healthcare: health care, hospitals, nursing, residential care, social assistance, dental.

71 Art & Recreation: arts, entertainment, recreation, spectator sports, museums, historical sites, amusement, gambling, parks.

72 Accommodation: hospitality, accommodation, food services, drinking places, hotels, eating, restaurant, cafeterias.

81 Other Services: repair and maintenance, laundry services, religious organizations, hair, nail, and skin care, barber.

C Additional Tables and Figures

Table A1: What Predicts the Timing of Turning to Coeducation

This table examines what predicts the year of coeducation by estimating the following OLS regressions:

$$\text{Year of Coeducation}_s = \alpha + \beta_0 \times \text{University Characteristics}_s + \varepsilon_s, \tag{10}$$

$$\text{Year of Coeducation}_s = \alpha + \beta_0 \times \text{Local Economic Factors}_s + \varepsilon_s. \tag{11}$$

The dependent variable, Year of Coeducation$_s$, represents the year in which a university transitioned to coeducation. The explanatory variables include University Characteristics$_s$, such as the year of establishment, undergraduate enrollment, and indicators for public or private (benchmark) status and religious affiliation. Religious affiliation is categorized in two ways: religious and non-religious (benchmark), Catholic and non-Catholic (benchmark). Local Economic Factors$_s$ include county-level measures such as income, population, income growth, population growth and income per capita. Columns (1) to (5) present results focusing on university characteristics, while Columns (6) to (11) examine local economic factors. We use the county-level income, population, and income per capita for 1969, along with income and population growth for 1970, sourced from the Bureau of Economic Analysis. We also use the county-level average personal income from the 1960 PUMS data. Notably, Columns (5) to (8) include only universities that transitioned to coeducation no earlier than 1969, as these variables rely on 1969 data. Similarly, Columns (9) to (11) restrict the sample to universities that transitioned no earlier than 1970, 1970 and 1960, respectively. The sample sizes in Columns (9) and (10) decrease due to this restriction, while the sample size in Column (11) is smaller because some areas lack county identifiers for confidentiality reasons in the PUMS data.

	Year of Coeducation										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Year of Establishment	-0.038*										
	(0.022)										
𝟙(Public University)		-0.283									
		(3.239)									
𝟙(Religious)			0.660								
			(2.502)								
𝟙(Catholic)				2.270							
				(2.617)							
Num(Undergraduates)					-0.159						
					(1.026)						
ln(Income) from BEA						0.073					
						(0.501)					
ln(Population) from BEA							0.030				
							(0.563)				
ln(Income Per Capita) from BEA								1.884			
								(3.082)			
Income Growth from BEA									45.697		
									(29.164)		
Population Growth from BEA										6.035	
										(59.117)	
ln(Income) from PUMS											1.147
											(3.847)
Observations	94	99	93	93	66	62	62	62	49	48	49
R-squared	0.032	0.000	0.001	0.008	0.000	0.000	0.000	0.006	0.050	0.000	0.002

Table A2: Summary Statistics for Individuals in Treated and Control Universities Pre-Coeducation

This table shows means of the variables for the treated and control groups separately. Panel A shows means of the variables in the year prior to the focal treatment year and Panel B shows means of the variables in all pre-treatment periods. Variable definitions can be found in Table 1 legend. Standard errors are clustered by university.

Panel A: τ = -1						
	Treat = 0		Treat = 1			
	Mean	SD	Mean	SD	P-value	Std. Diff.
	Individual Characteristics					
Number of Male Freshmen	1,141.25	(1,183.77)	1,007.49	(671.46)	(0.62)	-0.10
Art & Architecture	0.03	(0.16)	0.01	(0.12)	(0.02)	-0.06
Other Liberal Art	0.12	(0.33)	0.13	(0.34)	(0.23)	0.02
Humanity	0.09	(0.28)	0.16	(0.37)	(0.00)	0.16
Social Science	0.11	(0.31)	0.13	(0.33)	(0.10)	0.05
Economics	0.05	(0.22)	0.07	(0.25)	(0.07)	0.05
Business	0.24	(0.43)	0.16	(0.36)	(0.00)	-0.16
Nursing & Health	0.01	(0.10)	0.02	(0.13)	(0.08)	0.04
Science	0.09	(0.28)	0.11	(0.32)	(0.01)	0.06
Computer	0.06	(0.23)	0.02	(0.13)	(0.00)	-0.15
Engineering	0.20	(0.40)	0.18	(0.39)	(0.70)	-0.03
Education	0.01	(0.08)	0.01	(0.09)	(0.29)	0.02
Law	0.01	(0.09)	0.00	(0.05)	(0.15)	-0.06
Founder_Narrow	0.08	(0.27)	0.09	(0.29)	(0.00)	0.03
Founder_Medium	0.13	(0.33)	0.17	(0.38)	(0.00)	0.09
Founder_Wide	0.14	(0.35)	0.18	(0.39)	(0.00)	0.08
Observations	124,468		8,405			
	Startup Characteristics					
Max(Female Employee Share)	0.55	(0.51)	0.53	(0.49)	(0.67)	-0.02
Avg(Female Employee Share)	0.38	(0.41)	0.38	(0.42)	(0.97)	-0.00
Med(Female Employee Share)	0.37	(0.43)	0.37	(0.44)	(0.85)	0.01
Agriculture	0.02	(0.12)	0.01	(0.12)	(0.90)	-0.00
Mining	0.01	(0.09)	0.01	(0.09)	(0.99)	-0.00
Construction	0.02	(0.16)	0.02	(0.13)	(0.24)	-0.03
Manufacturing	0.06	(0.23)	0.05	(0.22)	(0.54)	-0.02
Wholesaler	0.00	(0.07)	0.01	(0.08)	(0.65)	0.01
Retailer	0.03	(0.17)	0.03	(0.17)	(0.92)	-0.00
Transportation	0.02	(0.15)	0.02	(0.15)	(0.95)	0.00
Information	0.16	(0.36)	0.15	(0.36)	(0.89)	-0.01
Finance	0.11	(0.31)	0.10	(0.30)	(0.78)	-0.01
Real Estate	0.04	(0.19)	0.04	(0.19)	(0.81)	0.01
Professional Service	0.17	(0.37)	0.20	(0.40)	(0.07)	0.07
Management	0.07	(0.25)	0.07	(0.26)	(0.51)	0.02
Administrative Service	0.03	(0.18)	0.03	(0.16)	(0.41)	-0.02
Education	0.03	(0.17)	0.03	(0.16)	(0.77)	-0.01
Healthcare	0.11	(0.31)	0.08	(0.27)	(0.14)	-0.06
Arts & Recreation	0.07	(0.26)	0.08	(0.27)	(0.64)	0.02
Accommodation	0.04	(0.19)	0.04	(0.20)	(0.55)	0.02
Other Services	0.03	(0.17)	0.03	(0.16)	(0.65)	-0.01
Observations	936		1,748			

Table A2: Summary Statistics for Individuals in Treated and Control Universities Pre-Coeducation (continued)

Panel B: Post = 0						
	Treat = 0		Treat = 1			
	Mean	SD	Mean	SD	P-value	Std. Diff.
	Individual Characteristics					
Number of Male Freshmen	1,113.73	(1,118.90)	869.51	(554.66)	(0.30)	-0.20
Art & Architecture	0.03	(0.16)	0.01	(0.12)	(0.02)	-0.06
Other Liberal Art	0.12	(0.33)	0.13	(0.34)	(0.27)	0.02
Humanity	0.10	(0.30)	0.15	(0.36)	(0.00)	0.12
Social Science	0.11	(0.31)	0.13	(0.33)	(0.22)	0.04
Economics	0.05	(0.22)	0.07	(0.25)	(0.18)	0.04
Business	0.24	(0.43)	0.14	(0.35)	(0.00)	-0.19
Nursing & Health	0.01	(0.10)	0.02	(0.12)	(0.04)	0.04
Science	0.09	(0.29)	0.11	(0.31)	(0.06)	0.05
Computer	0.04	(0.20)	0.02	(0.12)	(0.00)	-0.11
Engineering	0.20	(0.40)	0.23	(0.42)	(0.64)	0.04
Education	0.01	(0.08)	0.01	(0.08)	(0.92)	-0.00
Law	0.01	(0.09)	0.00	(0.04)	(0.07)	-0.07
Founder_Narrow	0.08	(0.27)	0.09	(0.28)	(0.34)	0.01
Founder_Medium	0.13	(0.34)	0.16	(0.37)	(0.00)	0.05
Founder_Wide	0.15	(0.36)	0.18	(0.38)	(0.00)	0.05
Observations	1,030,800		44,084			
	Startup Characteristics					
Max(Female Employee Share)	0.57	(0.51)	0.55	(0.51)	(0.38)	-0.02
Avg(Female Employee Share)	0.39	(0.42)	0.39	(0.43)	(0.59)	-0.01
Med(Female Employee Share)	0.39	(0.45)	0.38	(0.45)	(0.50)	-0.01
Agriculture	0.01	(0.12)	0.01	(0.11)	(0.46)	-0.01
Mining	0.01	(0.09)	0.01	(0.09)	(1.00)	-0.00
Construction	0.02	(0.13)	0.02	(0.13)	(0.64)	0.01
Manufacturing	0.05	(0.21)	0.05	(0.22)	(0.63)	0.01
Wholesaler	0.01	(0.11)	0.01	(0.10)	(0.10)	-0.02
Retailer	0.04	(0.18)	0.03	(0.18)	(0.83)	-0.00
Transportation	0.03	(0.16)	0.03	(0.17)	(0.15)	0.02
Information	0.15	(0.36)	0.14	(0.35)	(0.25)	-0.02
Finance	0.12	(0.32)	0.10	(0.30)	(0.04)	-0.04
Real Estate	0.03	(0.17)	0.04	(0.19)	(0.09)	0.02
Professional Service	0.19	(0.39)	0.20	(0.40)	(0.03)	0.03
Management	0.07	(0.25)	0.08	(0.27)	(0.03)	0.04
Administrative Service	0.03	(0.16)	0.03	(0.17)	(0.42)	0.01
Education	0.03	(0.18)	0.03	(0.17)	(0.29)	-0.01
Healthcare	0.09	(0.28)	0.08	(0.27)	(0.12)	-0.02
Arts & Recreation	0.08	(0.27)	0.08	(0.28)	(0.60)	0.01
Accommodation	0.04	(0.19)	0.04	(0.18)	(0.59)	-0.01
Other Services	0.03	(0.17)	0.03	(0.16)	(0.55)	-0.01
Observations	31,522		8,523			

Table A3: Robustness: Alternative Event Windows, Control Groups, and Weighting

We present robustness checks for estimates from following model in OLS:

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,e}, \quad (12)$$

The dependent variable is a dummy variable indicating whether in event e, the individual i from university s, cohort c, ever becomes a founder, as defined in Section 3.4.2. $\text{Treat}_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $\text{Post}_{sc,e}$ is a dummy variable that takes the value of one for all cohorts entering the university after the focal treatment year in event e. φ_c are entering cohort fixed effects, which control for the time-series trends in founder rates in the economy. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. ρ_e are event fixed effects, which ensure that comparisons are made between individuals in treated and control universities within each coeducation turning event. β estimates the impact of entering the university after its transition to coeducation on the individual rate of entrepreneurship. Panel A allows early- and later-treated groups in the control group and Panel B only allows later-treated group in the control group, both using [-10, +10] year event window. Panel C uses [-10, +10] year event window, allows early- and later-treated groups in the control group and aggregates event estimates with sample-share weights following Wing et al. (2024) (described in footnote 10). Panel D and Panel E uses [-5, +5] year event window relative to the coeducation year, allowing early- and later-treated groups in the control group. Panel E aggregates event estimates with sample-share weights. Standard errors clustered by university are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: [-10, +10], early- and later-treated group as control, no weights

	Founder_Narrow	Founder_Medium	Founder_Wide
	(1)	(2)	(3)
Treat × Post	0.005	0.012***	0.013***
	(0.003)	(0.003)	(0.004)
Treat	-0.005*	-0.007**	-0.008**
	(0.003)	(0.003)	(0.004)
Post	-0.001	-0.001*	-0.002**
	(0.001)	(0.001)	(0.001)
Observations	3,454,538	3,454,538	3,454,538
R-squared	0.007	0.025	0.025
School FE	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes
Event FE	Yes	Yes	Yes

Panel B: [-10, +10], later-treated group as control, no weights

	Founder_Narrow	Founder_Medium	Founder_Wide
	(1)	(2)	(3)
Treat × Post	0.011**	0.012**	0.012*
	(0.005)	(0.006)	(0.006)
Treat	0.003	0.005	0.004
	(0.004)	(0.003)	(0.003)
Post	-0.005***	-0.004	-0.004
	(0.002)	(0.003)	(0.003)
Observations	653,973	653,973	653,973
R-squared	0.007	0.017	0.017
School FE	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes
Event FE	Yes	Yes	Yes

Panel C: [-10, +10], early and lately treated group as control, with weights			
	Founder_Narrow	Founder_Medium	Founder_Wide
	(1)	(2)	(3)
Treat × Post	0.005	0.012***	0.012***
	(0.003)	(0.004)	(0.004)
Treat	-0.004	-0.005	-0.006*
	(0.002)	(0.003)	(0.003)
Post	-0.001	-0.002	-0.002*
	(0.001)	(0.001)	(0.001)
Observations	3,443,452	3,443,452	3,443,452
R-squared	0.004	0.012	0.012
School FE	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes
Event FE	Yes	Yes	Yes
Panel D: [-5, +5], early- and later-treated group as control, no weights			
	Founder_Narrow	Founder_Medium	Founder_Wide
	(1)	(2)	(3)
Treat × Post	0.002	0.010***	0.009**
	(0.002)	(0.003)	(0.003)
Treat	-0.003	-0.005	-0.005
	(0.003)	(0.003)	(0.004)
Post	0.000	-0.000	-0.000
	(0.000)	(0.001)	(0.000)
Observations	2,225,874	2,225,874	2,225,874
R-squared	0.006	0.022	0.021
School FE	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes
Event FE	Yes	Yes	Yes
Panel E: [-5, +5], early- and later-treated group as control, with weights			
	Founder_Narrow	Founder_Medium	Founder_Wide
	(1)	(2)	(3)
Treat × Post	0.002	0.010***	0.009**
	(0.002)	(0.003)	(0.004)
Treat	-0.003	-0.005	-0.005
	(0.003)	(0.004)	(0.004)
Post	0.002*	0.001	0.002
	(0.001)	(0.001)	(0.001)
Observations	2,198,068	2,198,068	2,198,068
R-squared	0.004	0.010	0.009
School FE	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes
Event FE	Yes	Yes	Yes

Table A4: Coeducation and Entrepreneurship, within 18 Years of College Graduation

We analyze 18-year career histories following graduation for each individual, excluding cohorts entering colleges after 1995, and estimate the same model in OLS with varying fixed effects:

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \varepsilon_{sci,e}, \quad (1)$$

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,e}, \quad (2)$$

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \varphi_{c,e} + \theta_{s,e} + \varepsilon_{sci,e}. \quad (3)$$

The dependent variable is a dummy variable indicating whether in event e, the individual i from university s, cohort c, becomes a founder within 18 years after college graduation, as defined in Section 3.4.2. $\text{Treat}_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $\text{Post}_{sc,e}$ is a dummy variable that takes the value of one for all cohorts entering the university after the focal treatment year in event e. φ_c are entering cohort fixed effects, which control for the time-series trends in founder rates in the economy. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. ρ_e are event fixed effects, which ensure that comparisons are made between individuals in treated and control universities within each coeducation turning event. Further, $\varphi_{c,e}$ and $\theta_{s,e}$ are event-by-entering cohort and event-by-university fixed effects, respectively. These interactive fixed effects allow for heterogeneous estimates of the influence of university characteristics and macroeconomic shocks across events. β estimates the impact of entering the university after its transition to coeducation on the individual rate of entrepreneurship within 18 years after college graduation. Columns (1) shows results from Eq.(1). Columns (2) shows results from Eq.(2). Columns (3) shows results from Eq.(3). Standard errors clustered by university are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Within 18 yrs after graduation...		
	$\mathbb{1}$(Founder)		
	(1)	(2)	(3)
Treat × Post	0.008***	0.009***	0.009***
	(0.002)	(0.002)	(0.002)
Treat	-0.006***	-0.007***	
	(0.002)	(0.002)	
Post	-0.002**	-0.001*	
	(0.001)	(0.001)	
Observations	2,621,812	2,621,812	2,621,812
R-squared	0.006	0.006	0.007
School FE	Yes	Yes	
Entry Year FE	Yes	Yes	
Event FE		Yes	
Event FE × School FE			Yes
Event FE × Entry Year FE			Yes

Table A5: Coeducation and Entrepreneurship, at Public Universities

We examine the impact of turning coeducation on students' entrepreneurship rate in OLS with varying fixed effects for public universities, excluding the public universities that are military universities:

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \varepsilon_{sci,e}, \quad (1)$$

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,e}, \quad (2)$$

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \varphi_{c,e} + \theta_{s,e} + \varepsilon_{sci,e}. \quad (3)$$

The dependent variable is a dummy variable indicating whether in event e, the individual i from university s, cohort c, becomes a founder. $\text{Treat}_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $\text{Post}_{sc,e}$ is a dummy variable that takes the value of one for all cohorts entering the university after the focal treatment year in event e. φ_c are entering cohort fixed effects, which control for the time-series trends in founder rates in the economy. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. ρ_e are event fixed effects, which ensure that comparisons are made between individuals in treated and control universities within each coeducation turning event. Further, $\varphi_{c,e}$, $\theta_{s,e}$ are event-by-entering cohort and event-by-university fixed effects, respectively. These interactive fixed effects allow for heterogeneous estimates of the influence of university characteristics and macroeconomic shocks related with student cohorts across events. β estimates the impact of entering the university after its transition to coeducation on the individual rate of entrepreneurship. Columns (1) shows results from Eq.(1). Columns (2) shows results from Eq.(2). Columns (3) shows results from Eq.(3). Standard errors clustered by university are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	$\mathbb{1}$(Founder)		
	(1)	(2)	(3)
Treat×Post	0.021*	0.023**	0.023**
	(0.009)	(0.008)	(0.007)
Treat	-0.020*	-0.023**	
	(0.008)	(0.008)	
Post	-0.001	-0.001	
	(0.001)	(0.001)	
Observations	812,340	812,340	812,334
R-squared	0.022	0.022	0.023
School FE	Yes	Yes	
Entry Year FE	Yes	Yes	
Event FE		Yes	
Event FE × School FE			Yes
Event FE × Entry Year FE			Yes

Table A6: Coeducation and Entrepreneurship, with Location-Entry Year Fixed Effects

To control for time-varying local economic factors, we examine the impact of turning coeducation on students' entrepreneurship rate in OLS, further including state-by-entering cohort and county-by-entering cohort fixed effects:

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \lambda_{location,c} + \varepsilon_{sci,e}, \quad (13)$$

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \lambda_{location,c} + \rho_e + \varepsilon_{sci,e}, \quad (14)$$

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \varphi_{c,e} + \theta_{s,e} + \lambda_{location,c,e} + \varepsilon_{sci,e}. \quad (15)$$

The dependent variable is a dummy variable indicating whether in event e, the individual i from university s, cohort c, ever becomes a founder. $\text{Treat}_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $\text{Post}_{sc,e}$ is a dummy variable that takes the value of one for all cohorts entering the university after the focal treatment year in event e. φ_c are entering cohort fixed effects, which control for the time-series trends in founder rates in the economy. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. $\lambda_{location,c}$ are location-by-cohort fixed effects, which control for the influences of state-level or county-level economic conditions that affect certain cohorts. ρ_e are event fixed effects, which ensure that comparisons are made between individuals in treated and control universities within each coeducation turning event. Further, $\varphi_{c,e}$, $\theta_{s,e}$, and $\lambda_{location,c,e}$ are event-by-entry year, event-by-university, event-by-location-by-entry year fixed effects, respectively. These interactive fixed effects allow for heterogeneous estimates of the influence of university characteristics, and macroeconomic shocks, state or county-level economic conditions related with student entering cohorts across events. β estimates the impact of entering the university after its transition to coeducation on the individual rate of entrepreneurship. Standard errors clustered by university are reported in parentheses. Panel A controls state-by-entering cohort fixed effects and Panel B controls county-by-entering cohort fixed effects. Columns (1) shows results from Eq.(13). Columns (2) shows results from Eq.(14). Columns (3) shows results from Eq.(15). ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: State-Entry Year FEs

	𝟙(Founder)		
	(1)	(2)	(3)
Treat×Post	0.010***	0.011***	0.015***
	(0.003)	(0.003)	(0.004)
Treat	-0.001	-0.002	
	(0.002)	(0.003)	
Post	-0.002***	-0.000	
	(0.001)	(0.000)	
Observations	3,399,668	3,399,668	3,399,079
R-squared	0.030	0.030	0.031
Entry Year FE	Yes	Yes	
School FE	Yes	Yes	
Event FE		Yes	
Entry Year FE × Event FE			Yes
School FE × Event FE			Yes
State FE × Entry Year FE	Yes	Yes	
State FE × Entry Year FE × Event FE			Yes

Panel B: County-Entry Year FEs

	𝟙(Founder)		
	(1)	(2)	(3)
Treat×Post	0.005***	0.007***	0.017***
	(0.002)	(0.002)	(0.005)
Treat	-0.001	-0.003*	
	(0.001)	(0.001)	
Post	-0.001**	-0.000	
	(0.000)	(0.000)	
Observations	3,347,987	3,347,987	3,346,281
R-squared	0.037	0.037	0.037
Entry Year FE	Yes	Yes	
School FE	Yes	Yes	
Event FE		Yes	
Entry Year FE × Event FE			Yes
School FE × Event FE			Yes
County FE × Entry Year FE	Yes	Yes	
County FE × Entry Year FE × Event FE			Yes

Table A7: Coeducation and Entrepreneurship, Individual-Year Level Panel

We construct an individual-year dataset and estimate the following model in OLS with varying fixed effects:

$$\mathbb{1}(\text{founder}_{sci,t,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \gamma_t + \varepsilon_{sci,t,e}, \quad (16)$$

$$\mathbb{1}(\text{founder}_{sci,t,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \gamma_t + \rho_e + \varepsilon_{sci,t,e}, \quad (17)$$

$$\mathbb{1}(\text{founder}_{sci,t,e}) = \alpha + \beta \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \varphi_{c,e} + \theta_{s,e} + \gamma_{t,e} + \varepsilon_{sci,t,e}. \quad (18)$$

The dependent variable is a dummy variable indicating whether in event e, the individual i from university s, cohort c, becomes a founder in year t. $\text{Treat}_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $\text{Post}_{sc,e}$ is a dummy variable that takes the value of one for all cohorts entering the university after the focal treatment year in event e. φ_c are entering cohort fixed effects, which control for the time-series trends in founder rates in the economy. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. γ_t are year fixed effects, which control for the influences of macroeconomic shocks in certain years. ρ_e are event fixed effects, which ensure that comparisons are made between individuals in treated and control universities within each coeducation turning event. Further, $\varphi_{c,e}$, $\theta_{s,e}$ and $\gamma_{t,e}$ are event-by-entering cohort, event-by-university, and event-by-year fixed effects, respectively. These interactive fixed effects allow for heterogeneous estimates of the influence of university characteristics and macroeconomic shocks related with student cohorts and years across events. β estimates the impact of entering the university after its transition to coeducation on the individual rate of entrepreneurship a given year t. Columns (1) and (4) show results from Eq.(16). Columns (2) and (5) show results from Eq.(17). Columns (3) and (6) show results from Eq.(18). Standard errors clustered by university are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	𝟙(Founder)					
Event Window:	[-10, +10]			[-5, +5]		
	(1)	(2)	(3)	(4)	(5)	(6)
Treat × Post	0.007***	0.007***	0.007***	0.006***	0.007***	0.007***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Treat	-0.004***	-0.004***		-0.003**	-0.003**	
	(0.001)	(0.002)		(0.001)	(0.001)	
Post	-0.002***	-0.001		-0.001***	-0.001	
	(0.001)	(0.001)		(0.000)	(0.000)	
Observations	103,052,586	103,052,586	103,052,572	72,631,113	72,631,113	72,631,108
R-squared	0.032	0.032	0.033	0.032	0.032	0.034
School FE	Yes	Yes		Yes	Yes	
Entry Year FE	Yes	Yes		Yes	Yes	
Year FE	Yes	Yes		Yes	Yes	
Event FE		Yes			Yes	
Event FE × School FE			Yes			Yes
Event FE × Entry Year FE			Yes			Yes
Event FE × Year FE			Yes			Yes

Table A8: Industry Share of Women-Owned Businesses

This table presents the number of women-owned businesses and the share of them by major industry groups. Data source: 1992 Economic Census.

	Number of Companies			Receipts ($1,000,000)			Number of Employees		
	Women-Owned	Total	% of Women-Owned	Women-Owned	Total	% of Women-Owned	Women-Owned	Total	% of Women-Owned
Agriculture	6006	22726	26.43%	3122	15161	20.59%	37129	219646	16.90%
Construction	49433	199584	24.77%	44169	296765	14.88%	374803	2340748	16.01%
Wholesaler	58517	235411	24.86%	179358	1380466	12.99%	669207	3319356	20.16%
Transportation and Utilities	25303	87321	28.98%	74490	765800	9.73%	653193	4645445	14.06%
Retailer	114624	365988	31.32%	117201	1342496	8.73%	1129049	11975846	9.43%
Finance, Insurance, and Real Estate	65130	271620	23.98%	155144	1860793	8.34%	666279	8603829	7.74%
Services	150521	632232	23.81%	73238	885895	8.27%	1827291	13243899	13.80%
Manufacturing	39517	175915	22.46%	274539	3957859	6.94%	1502969	16477305	9.12%
Mining	3636	16319	22.28%	5335	140375	3.80%	36449	518071	7.04%

Table A9: Coeducation, Spousal Employment and Income, and Entrepreneurship, for Events after 1960

To examine the heterogeneous impact of coeducation on entrepreneurship, we estimate the following OLS model in sub-samples split at the median of income and employment gaps between college-educated and non-college-educated married women:

$$\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,f,e}, \quad (2)$$

The dependent variable is whether individual i from cohort c university s in event e has ever become an entrepreneur. $\text{Treat}_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $\text{Post}_{sc,e}$ is a dummy variable that takes the value of one for all cohorts entering the university after the focal treatment year in event e. φ_c are entering cohort fixed effects, which control for the time-series trends in founder rates in the economy. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. ρ_e are event fixed effects, which ensure that comparisons are made between individuals in treated and control universities within each coeducation turning event. β estimates the impact of entering the university after its transition to coeducation on individual entrepreneurship rate. We compute the differences in average income and employment between college-educated and non-college-educated women using the 1960 PUMS data at the county level. Columns (1) and (2) present results for sub-samples where the income gap variable is above and below the median, respectively. Similarly, Columns (3) and (4) display results for sub-samples where the employment gap variable is above and below the median. To mitigate concerns that the 1960 variables may already be influenced by coeducation, we restrict the analysis to events with a focal treatment year later than 1960. The sample size decreases compared to Table 2 because some areas have missing county identifiers due to confidentiality reasons in the PUMS data. Standard errors clustered by university are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	𝟙(Founder)			
	by Income Gap		by Employment Gap	
	Above Median	Below Median	Above Median	Below Median
	(1)	(2)	(3)	(4)
Treat × Post	0.018***	0.003	0.015**	0.008
	(0.006)	(0.004)	(0.006)	(0.007)
Treat	-0.013	0.001	-0.005	-0.011
	(0.009)	(0.004)	(0.009)	(0.008)
Post	-0.000	-0.000	0.000	-0.000
	(0.001)	(0.002)	(0.001)	(0.002)
Observations	1,020,405	990,680	991,225	1,019,860
R-squared	0.025	0.021	0.022	0.024
School FE	Yes	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes	Yes
Event FE	Yes	Yes	Yes	Yes

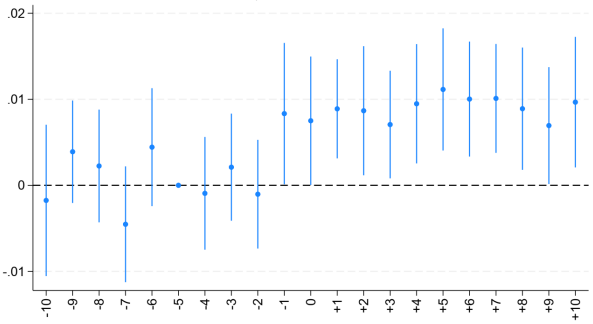
Figure A1: Dynamic Effects of Coeducation on Entrepreneurship, Individual-Year Level

We estimate the following model at individual-year level:

$$\begin{aligned}\mathbb{1}(\text{founder}_{sci,t,e}) = \alpha + \sum_{\tau \geq -10}^{\tau \leq 10} \beta_{0\tau}\, \mathbb{1}(\text{Year} - \text{Focal Treatment Year} = \tau)_{sc\tau,e} \times \text{Treat}_{s,e} \\ + \beta_1 \times \text{Treat}_{s,e} + \sum_{\tau \geq -10}^{\tau \leq 10} \beta_{2\tau} \times \mathbb{1}(\text{Year} - \text{Focal Treatment Year} = \tau)_{sc\tau,e} \\ + \varphi_c + \theta_s + \gamma_t + \rho_e + \varepsilon_{sci,t,e}, \end{aligned} \tag{19}$$

The dependent variable is an indicator that takes the value of one if individual i from cohort c in university s becomes a founder in year t. $\text{Treat}_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $\mathbb{1}(\text{Year} - \text{Focal Treatment Year} = \tau)_{sc\tau,e}$ is a dummy variable that takes the value of one if cohort c enters university s in the τth year relative to the transition year in event e. φ_c are entering cohort fixed effects, which control for the time-series trends in founder rates in the economy. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. γ_t are year fixed effects, which control for the influences of macroeconomic shocks in certain years. ρ_e are event fixed effects, which ensure that comparisons are made between individuals in treated and control universities within each coeducation turning event. β_τ captures the dynamic treatment effects of coeducation on entrepreneurship in each event year. Panel A uses a [-10, +10] year event window relative to coeducation year, and Panel B uses a [-5, +5] year window event. Both panels allow early- and later-treated groups in the control group. Estimates at $t = -5$ are omitted. Standard errors are clustered by university.

Panel A: [-10, +10] event window



Panel B: [-5, +5] event window

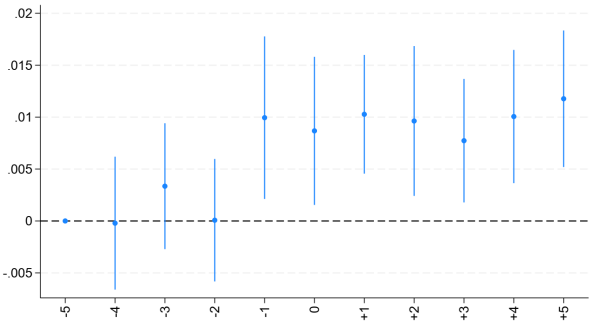


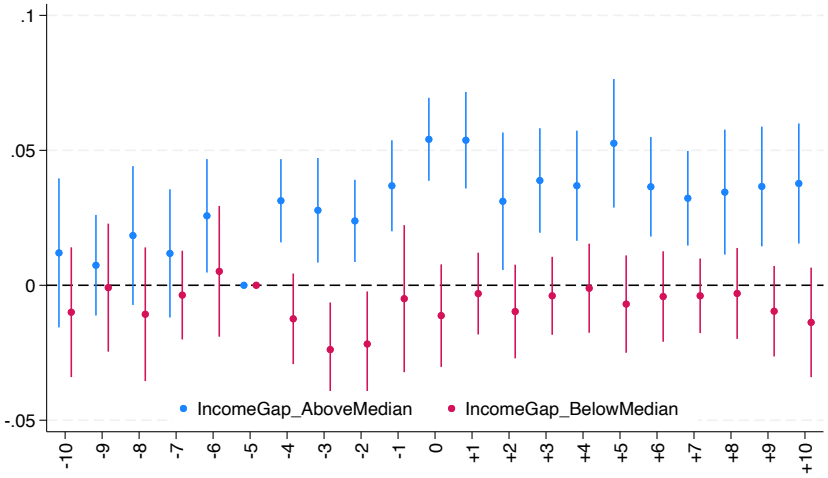
Figure A2: Coeducation, Spousal Employment and Income, and Entrepreneurship for Events after 1960

To examine the heterogeneous impact of coeducation on entrepreneurship, we estimate the following OLS model in sub-samples split at the median of income and employment gaps between college-educated and non-college-educated married women and restricting to events with a focal treatment year later than 1960:

$$\begin{aligned}\mathbb{1}(\text{founder}_{sci,e}) = \alpha + \sum_{\tau\geq -10}^{\tau\leq 10} \beta_{0\tau}\, \mathbb{1}(\text{Year} - \text{Focal Treatment Year} = \tau)_{sc\tau,e} \times \text{Treat}_{s,e} \\ + \beta_1 \times \text{Treat}_{s,e} + \sum_{\tau\geq -10}^{\tau\leq 10} \beta_{2\tau} \times \mathbb{1}(\text{Year} - \text{Focal Treatment Year} = \tau)_{sc\tau,e} \\ + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,e}, \end{aligned} \tag{5}$$

The dependent variable is an indicator that takes the value of one if individual i from cohort c in university s becomes a founder. $\text{Treat}_{s,e}$ equals one if in event e, university s turns coeducation. $\mathbb{1}(\text{Year} - \text{Focal Treatment Year} = \tau)_{sc\tau,e}$ is a dummy variable that takes the value of one if cohort c enters university s in the τth year relative to the year of transition in event e. φ_c are entering cohort fixed effects, which control for the time-series trends in founder rates in the economy. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. ρ_e are event fixed effects, which ensure that comparisons are made between individuals in treated and control universities within each coeducation turning event. β_τ captures the dynamic effects for each event year. Panel A divides the sample based on the median value of income gap between college-educated and non-college-educated married women at county level, and Panel B divides the sample based on the median value of the employment gap. Both panels uses a [-10, +10] year event window and allow early- and later-treated groups in the control group. To mitigate concerns that the 1960 variables may already be influenced by coeducation, we restrict the analysis to events with a focal treatment year later than 1960. Estimates at $t = -5$ are omitted. Standard errors are clustered by university.

Panel A: Subsample by Income Gap between College-educated and Not College-educated Married Women



Panel B: Subsample by Employment Gap between College-educated and Not College-educated Married Women

