When Harry Met Sally: Coeducation and Startup Formation

Zhuo Chen *Boston College*

Hanyi Livia Yi* *Boston College*

November 17, 2025

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

We study how gender integration in higher education affects entrepreneurial entry and subsequent startup quality. Leveraging staggered coeducation reforms at 91 previously all-male U.S. universities between 1942 and 1995, and linking historical enrollment records with resume data, we find that exposure to female peers increases male students' likelihood of becoming entrepreneurs. Three mechanisms drive this effect: cross-gender interactions that expand human and social capital, improved household risk-sharing, and sorting into masculine-perceived fields. The resulting startups survive longer, but show similar employment sizes and grow more gradually, consistent with a rise in sustainable, necessity-driven entrepreneurship rather than high-growth, transformative ventures. Overall, our findings suggest that the gender environment of higher education has lasting economic consequences by altering both the quantity and nature of entrepreneurial activity.

JEL classification: I24, J16, L26
Keywords: Entrepreneurship, Coeducation, Higher Education, Gender Integration, Startup Outcomes

*We thank Andra Ghent, Edie Hotchkiss, Jonathan Karpoff, Xing Liu, Andrey Malenko, Zheng Sun, and participants at the Boston College Ph.D. Student Seminar, University of Washington Fostering Inclusion Workshop, AFA AFFECT Mentoring Workshop, and CICF for their thoughtful feedback and comments. Sofia Utoft provided excellent research assistance. Emails: Zhuo Chen: zhuo.chen@bc.edu. Hanyi Livia Yi: livia.yi@bc.edu.

1 Introduction

Scholars have long argued that interactions among diverse individuals can spur new ideas, innovation, and economic growth (e.g., Mill, 1965; Jacobs, 1969; Glaeser, Kallal, Scheinkman, and Shleifer, 1992; Posch, Schulz, and Henrich, 2023). Such interactions may facilitate knowledge exchange, collaboration, and network formation, creating fertile ground for entrepreneurship. By expanding social and marriage markets, they may also foster family formation, improving entrepreneurs' ability to bear downside risk. At the same time, these interactions may reinforce identity-driven sorting, potentially intensifying group differences and segregation (Akerlof and Kranton, 2005; Charles and Bradley, 2009; Getik and Meier, 2022). These mechanisms can operate simultaneously, jointly shaping economic outcomes.

In this paper, we study how exposure to gender-diverse peers during undergraduate education affects male entrepreneurship and the characteristics of the ventures they create. Higher education is an ideal setting to examine the various economic tensions underlying gender integration for two reasons. First, as a formative stage for both cognitive and identity development, higher education is likely to shape long-term behavior and outcomes. Second, the U.S. higher education system underwent a profound transformation over the 20th century, as many institutions shifted from single-sex to coeducational environments, significantly increasing gender diversity on campuses. Although the vast majority of U.S. undergraduate students today attend coeducational universities, no institutions admitted women before 1835, and some of the most prestigious universities remained single-sex until the 1960s. While this shift has been studied from various socioeconomic angles—such as its effects on women's educational 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).

Leveraging the staggered transition of formerly all-male universities to coeducation, we find that male students exposed to female peers are more likely to become entrepreneurs. This result is driven by three complementary channels: cross-gender interactions that expand professional

networks and shape founding teams; improved household risk-sharing and downside protection that enable greater entrepreneurial risk-taking; and gendered sorting into "masculine"-perceived majors and occupations. Furthermore, we show that coeducation affects the nature of new ventures: the resulting startups tend to survive longer, but show similar employment sizes and grow more gradually, consistent with an increase in sustainable, necessity-driven entrepreneurship rather than high-growth, transformative ventures (Schoar, 2010; Botelho, Fehder, and Hochberg, 2025).

We construct a novel dataset combining university enrollment and graduation records, their coeducation transition dates, and individual career histories from resume data. Our sample covers 91 previously all-male universities that shifted to coeducation between 1942 and 1995. To exploit variation in adoption timing, while addressing concerns in staggered settings, we implement a stacked difference-in-differences (DiD) design, which is akin to a collection of event studies (Cengiz, Dube, Lindner, and Zipperer, 2019; Deshpande and Li, 2019; Butters, Sacks, and Seo, 2022; Wing, Freedman, and Hollingsworth, 2024).[1] Our identification rests on the assumption that, absent coeducation, entrepreneurial outcomes at treated and control universities would have moved in parallel. This assumption is plausible if adoption timing was independent of factors that might otherwise affect students' entrepreneurial choices—a condition largely supported by the historical context and motivations for coeducation. These motivations include efforts to boost applications, improve cost efficiency, respond to evolving religious norms regarding women's education, and idiosyncratic factors such as alumni resistance or proximity to opposite-gender schools, all of which are less likely to directly influence entrepreneurial outcomes in the absence of coeducation. When empirically examining institutional and local economic factors that predict adoption timing, an institution's public versus private control and religious affiliation emerge as two determinants. To address these factors and other potential local economic confounders, we show robustness to controls of time fixed effects interacted with public vs. private status, religious affiliation, and local economic growth deciles, as well as to a matching procedure.

[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 et al., 2019; Deshpande and Li, 2019; Butters et al., 2022; Wing et al., 2024).

We first examine coeducation's impact on students' likelihood of becoming entrepreneurs. When all-male universities transition to coeducation, male students are 1.1 percentage points more likely to become entrepreneurs—an 8.5% increase relative to the sample's unconditional entrepreneurship probability of 13%. Our main specification includes university fixed effects, which control for institution-specific factors such as location and selectivity, and entry year fixed effects, which capture macroeconomic shocks—such as changes in unemployment or business cycles—that affect entrepreneurship rates across cohorts. This specification also includes event fixed effects, ensuring that the identifying variation stems from comparisons between treated and control universities within the same coeducation-turning event. Dynamic DiD coefficients show that treated and control universities followed similar trends in the decade preceding coeducation adoption. The rise in entrepreneurship begins approximately one to two years before official coeducation adoption, likely reflecting spillover effects from newly enrolled female students on pre-existing male students, and this rise persists over time. 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). They also withstand a range of additional tests addressing endogenous timing of adoption, confounding local economic trends, selection concerns, and alternative definitions of founders, sample windows, and control groups.

We then explore three complementary mechanisms behind the rise in entrepreneurship. First, coeducation may facilitate cross-gender interactions that drive knowledge exchange, network expansion, and collaboration. This mechanism implies larger effects in settings with greater potential for cross-gender interaction. Indeed, the effects are stronger in majors and cohorts with greater increases in female representation, which offer more exposure to female peers, and areas with lower gender inequality, which offer a more supportive social environment for such interactions. These heterogeneities also highlight that cross-gender interactions depend crucially on both institutional and cultural environments. The increased interactions also manifest in two tangible outcomes: the size of one's professional network, and entrepreneurial team formation. Coeducation increases the probability of having an extensive professional network (>500 connections on LinkedIn) by 1.7

percentage points, or 7% relative to the sample mean, consistent with a broader and more diverse peer group enhancing the social capital essential for entrepreneurship (Garfinkel, Mayer, Strebulaev, and Yimfor, 2021; Fluegge and Bailey, 2024). Coeducation also raises the probability of having peer female co-founders from the same university by 1.5–2.0 percentage points. This represents a 75–100% effect relative to the sample mean, and a 32–43% increase relative to the 4.66% benchmark rate observed in always-coeducational institutions. Meanwhile, the probability of having a female co-founder from outside the university remains unchanged, suggesting that the increase in cross-gender collaboration is fostered specifically within the higher education environment, rather than by broader societal shifts in gender norms.

Second, coeducation may influence entrepreneurship by affecting household formation, risk-sharing, and downside protection. Kirkebøen, Leuven, Mogstad, and Mountjoy (2025) argue that universities function as marriage markets, with stronger spousal matching effects when opposite-sex peers are more abundant.[2] By reducing marital search costs and increasing men's chances of marrying a college-educated spouse, coeducation could improve household income, financial stability, and risk-bearing capacity, thus encouraging entrepreneurial entry (Benham, 1974; Wong, 1986; Özcan, 2011; Hombert, Schoar, Sraer, and Thesmar, 2020; Barrios, Hochberg, and Yi, 2022, 2024; Denes, Lagaras, and Tsoutsoura, 2025). We test this mechanism by examining variation in local family educational structures and assessing whether the effect is larger where the marginal benefit of a college-educated spouse is greater. Consistent with the mechanism, we find stronger effects in regions where marrying a college-educated partner yields greater improvements in financial security—specifically, in counties where college-educated spouses are scarce, where the income premium for college-educated women is higher, and where average male income is lower.

Third, coeducation could induce sorting into more masculine-perceived academic majors and occupations. Identity economics posits that the introduction of a salient out-group can reinforce in-group identity, leading individuals to adhere more strongly to its norms and update their expectations

[2] Ford (2020) documents that approximately 33% of college graduates who entered between 1975 and 2005 married someone from the same university.

based on the outcomes of in-group peers (Akerlof and Kranton, 2005; Charles and Bradley, 2009; Akerlof and Kranton, 2010; Liqui Lung, 2022). This framework suggests that the influx of women into previously all-male universities likely makes gender identity more salient for men, encouraging them to assert a masculine identity. Consistent with this hypothesis, we find that male students shift toward traditionally male-dominated majors such as business and science, and away from female-dominated fields such as humanities and social sciences after coeducation. Moreover, this identity-driven sorting extends to occupational choices. Following coeducation, entry into entrepreneurship, an already male-dominated field, increases most in regions where it was even more male-dominated historically. Taken together, these findings suggest that, to some extent, coeducation reinforced gendered sorting and stereotypes in both education and occupational choices.

Finally, beyond entrepreneurial entry, we assess how coeducation affects the nature and quality of the new ventures. Coeducation leads to more enduring startups that survive longer. Although the average employment for these startups remains unchanged, they take more time to reach their peak employment. These patterns suggest that the increased entrepreneurship induced by coeducation is more likely to be subsistence-based and limited in scale, rather than transformative or growth-oriented (Schoar, 2010). These results are consistent with our mechanisms operating on two margins: interactions and network channels likely improve new firm resilience, while the downside protection channel primarily lowers the entry barrier for stable, necessity-driven businesses.

Related Literature Our paper contributes to several strands of literature. First, we add to the coeducation literature, which has examined impacts on women's educational attainment, faculty research direction, and student outcomes, including traits, perceptions, and major choices. In U.S. college settings, Goldin and Katz (2011) find that greater accessibility of coeducational institutions is associated with increased women's educational attainment. Truffa and Wong (2024) find that coeducation leads to an increase in gender-related research publications. Related to our study, Calkins et al. (2023) examine college major choices following the transition of women's colleges to coeducation and find that female students shift away from STEM fields and toward health and psychology majors, mirroring our results on male students' major choices when formerly all-male

universities transition to coeducation. In the setting of randomly assigned high schools in South Korea, researchers find that coeducation affects gender stereotypes, future educational attainment, earnings, and workplace relationships (Lee, 2021; Lee and Nakazawa, 2022; Park, 2023). In the U.K., Booth, Cardona-Sosa, and Nolen (2014) find that random exposure to single-sex college courses increases risk-taking among women, suggesting that the observed gender differences might reflect social learning rather than inherent traits.

While existing work largely focuses on effects for the newly admitted gender, we contribute to this literature by documenting meaningful consequences for existing male students when universities become coeducational. Moreover, we focus on a previously understudied aspect of students' career choices: their likelihood of becoming entrepreneurs, and the characteristics of the startups they found. This focus is important not only because entrepreneurship has long been recognized as a driver of economic development (Schumpeter, 2017), but also because, as early as Mill (1965), social scientists have posited that diverse social environments should foster the exchange and integration of ideas. More recently, Posch et al. (2023) provide empirical evidence that surname diversity within a society spurs innovation. In a different setting, Greenberg, Wasserman, and Weber (2024) examine the impact of admitting women into previously all-male U.S. military units and find no negative impact on men's behaviors such as misconduct and career outcomes, and a small positive impact for some outcomes. Our results are largely consistent with this view in showing that, in higher education, the integration of women increases the resilience of the startups men create.

Last, our findings highlight an interesting spillover effect: educational investments in women meaningfully affect their male peers' outcomes. In this aspect, our work contributes to the broad literature on the roles of peer environments, gender, and education in entrepreneurship. Several papers document entrepreneurial peer effects at universities and workplaces (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). Levine and Rubinstein (2017) show that better-educated people with "smart and illicit" traits are more likely to run incorporated businesses. We show that the gender composition of educational environments, distinct from educational attainment levels, meaningfully shapes entrepreneurial outcomes.

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. For Catholic universities, the motivation for coeducation was also affected by doctrinal reforms. The Second Vatican Council (1962–1965) affirmed women's right "to acquire an education...equal to that recognized for men," encouraging many Catholic-affiliated institutions to consider coeducation (Poulson, 1995; Goldin and Katz, 2011; Calkins et al., 2023). For institutions facing budgetary pressures, integrating male and female students on a single campus was significantly more cost-effective than maintaining separate facilities or establishing coordinate colleges. Coeducation allowed schools to consolidate resources for instruction, housing, and administration, offering an efficient and financially viable solution to growing constraints (Malkiel, 2016). Beyond religious and efficiency motivations, some universities adopted coeducation to remain competitive during a broader wave of institutional change. Finally, other idiosyncratic factors such as alumni resistance, local student demand, and proximity to schools of the opposite gender also influenced universities' decisions. 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 was often lengthy and complex, involving extensive discussions among university presidents, trustees, and faculty members. This made the exact timing of the transition more uncertain. Many institutions conducted internal surveys and reports to evaluate the academic and financial implications of coeducation, such as the costs of providing women's dormitories and updating other campus facilities. Some all-male universities also considered alternative proposals, including mergers with nearby women's colleges to form coordinate arrangements, such as the proposed pairings of Yale-Vassar and Princeton-Sarah Lawrence.[3] In Section 4.3, we empirically examine the determinants of a university's transition timing to coeducation and account for these factors in our empirical design.

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 transition to coeducation through Wikipedia and schools' official websites. For universities whose names cannot be uniquely identified, we drop them from the sample because we rely on university names to match with the resume database.[4] We also drop universities without a clear transition time and universities that were merged with others. We exclude military universities given their specialized curricula and different post-graduate career paths.

[3] We exclude schools that became coeducational through merging with other institutions, since such mergers could involve more confounding changes in resource, faculty, and governance beyond the decision to admit women.

[4] 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 duplicated names from the sample.

Our sample consists of 91 previously all-male universities with transition dates spanning from 1942 to 1995, which include renowned institutions such as Ivy League universities (Columbia, Dartmouth, Princeton, and Yale), prestigious liberal arts colleges such as Amherst, Wesleyan, and Williams, and many other public and private institutions across the country. 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 maintained by the U.S. Department of Education, including the year of establishment, location, public vs. private status, religious affiliation, annual enrollment, revenue, expenditure, and degrees-awarded data. We also collect the earliest available average Scholastic Assessment Test (SAT) scores of admitted students from College Scorecard, which have been reported since 2001.

To visualize 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. 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 career experiences, we rely on a 2017 snapshot of LinkedIn profiles. LinkedIn users typically backfill their educational histories, which allows us to identify the cohorts of students enrolled at each university around the transition years. When matching the resume data with IPEDS, we include the cohorts that enter the university within 10 years before and after the year of coeducation adoption in the stacked sample. We also manually account for instances where universities changed their names during our sample period. When we compare the resume data coverage to administrative enrollment statistics reported by IPEDS, we

find that the resume data cover approximately 42% of individuals who attended the universities in our sample during the sample period. This partial coverage should not bias our results as long as LinkedIn usage is not systematically correlated with the variation in coeducation that we exploit. Moreover, because high-skilled professionals are the primary users of LinkedIn, we believe the startups that can be observed in our sample are likely to have higher impact compared with administrative data sources.

Using an unsupervised, machine learning-based clustering algorithm, we classify students' college majors into twelve groups based on information provided in their educational records: art & architecture, humanities, social science, economics, business, nursing & health, science, computer-related, engineering and education. Detailed steps are described in Online Appendix A. For tests related to college majors, we drop individuals who do not report this information on LinkedIn. The resume database also contains detailed information on each individual's career trajectory that allows us to identify founders and characterize startups, as discussed in Section 3.4.2 and 3.4.3.

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.[5] We use variables such as gender, marital status, personal income, and educational background to construct county-level measures of the income gap between college-educated and non-college-educated married women, which we refer to as the female college income premium. For further heterogeneity analysis, we also calculate county-level average male income, as well as gender inequality in income and education, defined respectively as the difference between the average male and female incomes and between male and female bachelor's degree attainment rates. Additionally, we construct a county-level measure of

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

the family share in which only the male holds a bachelor's degree. 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 that may have influenced 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.[6] This 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 frequency with which each name is given to males and females, and estimate the corresponding gender probability.[7] We define a name as strongly gendered if over 90% of individuals with that name are either male or female (Van Buskirk, Clauset, and Larremore, 2023). We assess the performance of the gender classification method using a randomly selected sample of 1,000 individuals. In this sample, 8% have no matching name in the SSA dataset, while 85.4% have strongly gendered names. We keep individuals with strongly gendered names. We choose this gender classification method over newer alternatives because it is not limited to training data from recent generations, which is important given most individuals in our sample attended college in or before the 1990s.

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", or "chief ex-

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

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

ecutive officer". Second, we impose two timing-related requirements: (i) the position must be the individual's first job at this company, and (ii) the individual must have joined the company within one year of the company's first appearance in the dataset. These restrictions rule out cases where the individual holds a management position in a corporation but is not a founder. However, they may also omit true founders in cases where different firms share the same name and have overlapping periods of operation. Thus, we define an alternative measure below to address this potential downward bias. Third, we exclude the job experiences that are related to chain businesses or incumbent corporations.[8]

We also define a narrow and a wide version of the founder variable for robustness checks. The narrow definition excludes cases where the job title includes "president", "CEO", or "chief executive officer", even if the individual joins the company at its inception in the resume database. This approach effectively eliminates the risk of mistakenly identifying corporate managers as founders, but is likely overly restrictive. The wide definition assigns a value of one when the job title includes "founder" or "owner", regardless of when the company's name first appears in the resume database. This approach helps capture founders in cases where two or more firms share the same name and have overlapping periods of operation, though it may be too permissive. Overall, we expect the true entrepreneurship rate to lie between our main version and the wide definition, with the narrow version serving as a lower bound.

A concern with the resume data is that earlier cohorts naturally have longer observable career histories. Because the data are a 2017 snapshot, the observed career length is right-censored, which may mechanically create a decline in entrepreneurship rate across cohorts. While all regressions include cohort fixed effects, we further address this concern by measuring founders based on the first 18 years of each individual's career, ensuring comparability across cohorts.

[8]This type of job titles 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", "agency", "operator", "agent", "insurance", "agent/owner", "owner/agent", "enfranchise", "store", "shop", "independent", "franchisee", "allstate", "distributor", and "franchise".

3.4.3 Startup Characteristics

Startup characteristics include longevity, years to mature, employment size, the presence of female co-founders, and industry. Startup longevity is measured using years of survival, which captures the total number of years a startup remains active (having at least one employee) in the resume database. We also construct a series of dummy variables that equal one if the startup survives more than 10, 15, 20, and 25 years. Years to mature is defined as the time it takes for a startup to reach its peak employment, calculated as the difference between the founding year and the year of maximum employment. To measure startup size, we calculate both the average and maximum number of employees over its life, reflecting its typical and peak employment levels. Because we can only observe employees with LinkedIn profiles, both the startup longevity and employment measures likely understate the actual values. To identify the presence of female co-founders in a startup, we first identify all founders associated with a startup using the resume data, and then determine their genders using the methodology described in Section 3.4.1. We further categorize female co-founders into two groups: those who attended the same university as the male student and graduated in peer cohorts (defined as within three years before or after graduation), and those who did not.

We also classify startups into 2-digit NAICS industries, and use this industry label as a control variable in tests related to longevity and employment, as industry-specific factors may influence both the survival and growth of new ventures. This classification algorithm uses a two-step procedure that combines large language model (LLM) predictions with embedding-based correction. We use the Your-economy Time Series (YTS) dataset as our training data, which contains the names of U.S. establishments from years 1998 to 2022 and is similar in spirit to the establishment-level data provided by the U.S. Census Bureau. In the first step, we prompt GPT-4o with examples of company name–industry pairs from the training data and keywords for each of the 18 NAICS-based industries to generate initial classifications. We repeat the classification five times independently for each company to minimize random variation and improve consistency. In the second step, we address the challenge that many startup names contain insufficient information for reliable inference by applying

an embedding-based method. Specifically, we generate vector representations of startup names and use supervised learning to identify and retain only those GPT predictions that likely match the true industries. Validated against company names and NAICS codes from the YTS dataset, this procedure achieves an 84% classification accuracy. Detailed steps on industry classification are provided in Online Appendix B.

3.4.4 Other Measures

To measure heterogeneity in treatment effects across areas with differing occupational stereotypes of entrepreneurship (i.e., its perceived masculinity among college graduates), we calculate the ex ante share of male entrepreneurs among college graduates in each county. We match individuals in the resume database to data on U.S. baccalaureate degree–granting institutions. For each county, prior to the adoption of coeducation, we calculate the male share of entrepreneurs among (i) all historical graduates, (ii) graduates from the previous ten years, and (iii) graduates from the previous five years. A higher male share of entrepreneurs in a county is interpreted as reflecting a stronger local norm that codes entrepreneurship as a masculine occupation.

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. (2022), and Wing et al. (2024).[9] Specifically, we group universities by their years of transition to coeducation into event groups, so that each event group has a unique treatment year, referred to

[9]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).

as the focal treatment year. Within each event group, the treated universities include those that adopted coeducation in the focal treatment year, and the event window consists of ten years before and after the focal treatment year. The results are robust to using a shorter event window of five years.

For each event, we define a set of control universities, which include those that do not themselves turn coeducation during the entire event window. We illustrate the sample construction with a specific example below. In 1978, Hanover College and Hamilton College turned to coeducation. Therefore, the treated sample for this event group includes all cohorts entering Hanover College and Hamilton College within a 10-year window before and after 1978, i.e., from 1968 to 1988. Then, the corresponding control sample for this event group includes cohorts attending universities during the same time period (i.e., from 1968 to 1988), but only at universities that transitioned to coeducation outside of the event window (i.e., either before 1968 or after 1988). 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:

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

where $\mathbb{1}(\text{Founder}_{sci,e})$ is a dummy variable indicating whether in 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 turn 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-year (cohort) fixed effects,

which control for broad economic trends over time that may affect founder rates across university cohorts. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. We cluster standard errors by event, corresponding to the level at which treatment timing varies (Bertrand, Duflo, and Mullainathan, 2004; Deshpande and Li, 2019). Our standard errors, therefore, account for the possibility that error terms may be correlated across all universities that undergo the coeducation transition in the same year. This clustering structure is robust to both serial correlation within event groups and cross-sectional correlation across universities within the same event group.

We also run specifications in which we include event fixed effects, as well as interactive fixed effects such as event-by-university and event-by-cohort:

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

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

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, which 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 of 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 to coeducation 10 years *after* the focal treatment year) and the early-treated universities (i.e. those already turned to coeducation 10 years *before* the focal treatment year). If treatment effects are not fully realized in 10 years, the presence of dynamic treatment effects can potentially bias our estimates downward. However, the 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 estimate specifications omitting early-treated universities from the control

group, and confirm that the results remain quantitatively similar. In other robustness tests, we assign sample-share weights when aggregating estimates across event groups following Wing et al. (2024), use alternative definitions of founders described in Section 3.4.2, and alternative sample windows. These results are described in Section 5.1.

4.3 Identification Assumption

The key identifying assumption underlying our empirical strategy is that individuals' likelihood of becoming entrepreneurs across treated and control universities would have moved 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 denoting 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 years before coeducation, discussed in detail in Section 5.1.

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, the exact timing of the coeducation decision can be related to factors such as budget, alumni resistance, and proximity to other opposite gender schools. Furthermore, a key driver of the coeducation wave after the 1960s was the Second Vatican Council,

which recognized women's equal right to education and prompted many Catholic universities to admit women. These considerations suggest that universities' decisions are less likely to be affected by the entrepreneurial potential of the student body. To mitigate further concerns about endogeneity in the adoption timing of coeducation, unobserved local economic trends, and selection issues, we conduct a series of robustness exercises in Section 6.

4.4 Summary Statistics

Our sample includes individuals attending 91 coeducation-transitioning universities that were previously all-male, and the transition dates span from 1942 to 1995. The average year of turning coeducation is 1965, and the students' average year of entry is 1985. We present the summary statistics in Table 1, divided into individual-level variables (Panel A and B) and startup-level variables (Panel C). Panel A includes individual characteristics variables such as university entry year, major, connections, and founder indicators. Panel B presents university characteristics and local economic factors, as matched to individuals. Panel C includes startup characteristics such as years of survival, number of employees, and the presence of female co-founders. In Online Appendix Table A1, we also tabulate the means of the variables for the treated and control groups separately, in the year prior to the focal treatment year. We also calculate the standardized differences in the means.

5 Results

5.1 Main Results

We begin by documenting that male students who enter traditionally all-male universities after they adopt coeducation have a higher probability of becoming entrepreneurs later in life. In Table 2, the outcome variable is an indicator of whether the individual becomes an entrepreneur. 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-by-entry year and event-by-school fixed effects to absorb any unobserved characteristics of universities or cohorts that may vary across events.

All columns consistently show an increase in the entrepreneurship rate of around one percentage point, which remains stable with varying fixed effects. For instance, Column (2) coefficients suggest that coeducation leads to a 1.1 percentage points increase in students' entrepreneurship rate, which is economically meaningful, representing 8.5% of the mean and 3.3% of the standard deviation of the outcome variable.

To examine the dynamics of the treatment effects, we estimate Eq. (4) and plot the event-time coefficients in Figure 3, where Panel A presents results using the longer sample and Panel B using the shorter sample. There is no systematic difference in trends between the treat and control groups until about one to two years before universities officially adopt coeducation, after which the effects appear and remain steady over time. The positive effect observed prior to the official coeducation reform may reflect spillovers to earlier cohorts who were still on campus when female students arrived. These students—though admitted before the reform—could have interacted with women through shared classes or campus activities. However, such interactions are likely less frequent for more senior students, as they would have had fewer overlapping courses and activities with the newly admitted women. Additionally, some universities began admitting small numbers of female students in the years leading up to the formal adoption of coeducation. This is consistent with the trend in Figure 2, which shows a modest increase in female enrollment preceding the official reform date. In some instances, formerly all-male institutions had female coordinate schools whose students could attend classes. We retain these institutions in the sample but note that prior exposure to female peers makes the coeducation transition less pronounced and might attenuate our estimates.[10]

The results are robust to using alternative definitions of the outcome variable, sample windows,

[10]For example, Columbia University did not become fully coeducational until 1983, but women from Barnard College could take courses there prior to that year.

and control groups, as shown in Online Appendix Table A2. Each panel presents results for the narrow, medium (our main measure, 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 we expect the true entrepreneurship rate to fall between the medium and the wide versions of the variable, with the narrow version serving as an extremely conservative lower bound. 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.[11] Last, we vary the event time window by using the [-5, +5] event window in Panels D and E, while also assigning weights to the estimates in Panel E.

A potential 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-censored, leading to naturally shorter observable windows for more recent cohorts and thus lower observed entrepreneurship rates for them. We already control for cohort fixed effects, which helps address this. To further alleviate this concern, we construct a more balanced outcome variable by limiting the career span to the first 18 years after graduation for all individuals (Appendix Table A3). This restriction reduces the number of observations relative to Table 2, as we exclude cohorts entering university after 1995 to ensure a full 18-year career history.[12] The coefficients are smaller in magnitude compared to the main specification, consistent with individuals having lower probabilities of becoming entrepreneurs within an 18-year window rather than throughout their entire career span.

[11] When aggregating the event-group estimates into a single coefficient estimate, Wing et al. (2024) assign weights to each event by its share of the treated sample. Specifically, each treated observation is assigned a weight of one, while control observations are weighted proportionally to the relative sample shares across event groups, $\frac{N_a^D/N_D}{N_a^C/N_C}$. N_a^D, N_a^C are the number of treated and control observations in the event-group a, and N^D, N^C are the total number of treated and control observations.

[12] Our results are not sensitive to choosing a different length of career history. We face the tradeoff of losing more individuals if we choose a longer unified career length.

5.2 Mechanisms

Our baseline results establish that exposure to a more diverse student body following coeducation increases entrepreneurial activity. We next explore three non-mutually exclusive channels that may drive this effect. First, exposure to female peers could facilitate interaction across genders, expanding entrepreneurial social capital and fostering idea exchange. Second, coeducation might affect household formation and improve intra-household risk-sharing and downside protection to enable increased entrepreneurial risk-taking. Finally, coeducation may steer male students toward more gender-stereotypical educational tracks and occupations. Identity economics suggests that increased visibility of women in such contexts could paradoxically increase men's awareness of gendered expectations, making entrepreneurship seem more "appropriately male", thus fostering entry (Akerlof and Kranton, 2005).

5.2.1 Peer Interactions, Networks, and Team Formation

We first explore the mechanism that coeducation facilitates entrepreneurial idea, knowledge, and skill exchange through increased interaction with female peers. If this mechanism is at play, the effect should be stronger in settings where the potential for cross-gender interaction is greater. We test this hypothesis by examining heterogeneity based on the increase in female peer exposure post-coeducation. For each university, we calculate the increase in female student share by major and cohort, respectively, in the decade after coeducation relative to the decade before, and partition the sample based on whether this increase is above or below the median. Table 3 shows that the increase in entrepreneurship is larger in contexts with above-median increases in female representation: a 1.5 percentage point effect in above-median female share majors (Column 1) versus negligible change in below-median majors (Column 2). This pattern is stronger at the major level, where daily peer interactions and idea exchanges are most direct, than at the cohort level (Columns 3 and 4).[13]

While higher female representation provides the critical mass necessary for interactions to

[13]This result holds when controlling for the interaction of post and student body size, suggesting that the coeducation effect acts through gender composition changes rather than solely through overall enrollment changes.

materialize, the efficacy of these interactions likely depends on the surrounding social norms. We hypothesize that more egalitarian social and professional norms and environments provide a more fertile ground for cross-gender exchanges. We measure the degree of gender egalitarianism using county-level historical data on the gender inequality in income and educational attainment (Fortin, 2005), measured as the differences between male and female average income and bachelor rates, and find support for this hypothesis. As shown in Table 4, the coeducation effect is stronger in lower-inequality counties. Columns (1) and (2) show that in regions with low gender inequality in income, entrepreneurship rises by 1.2 percentage points, compared to 0.8 percentage points in high-inequality regions. Similarly, Columns (3) and (4) show that in regions with low gender inequality in education, entrepreneurship increases by 1.6 percentage points, whereas no effect is observed in high educational inequality regions.[14] A potential concern is that our measure of gender inequality might be artificially low in less-developed regions, where both men and women have similarly low levels of income and education. To address this concern, we scale gender gaps by local average income and bachelor's degree rates in Appendix Table A5. After this adjustment, the contrast between low- and high-inequality regions becomes even stronger, further supporting the view that egalitarian social environments foster cross-gender exchanges and strengthen the coeducation effect.

So far, the evidence suggests that the baseline effects are stronger in settings with more female representation and gender-friendly social environments, where cross-gender interactions are more likely to occur. We next examine whether these interactions manifest through professional networking and entrepreneurial team formation, two natural channels for peer effects to materialize. We begin by exploring whether coeducation expands students' professional networks later on. The predicted effect is not obvious ex ante, as previous evidence in other contexts suggests that, despite the presence of the opposite gender, students might interact only with the same-gender peers due to homophily (Getik and Meier, 2022, 2025). We measure network scale using the number of

[14]Sample size is smaller due to missing county identifiers in the PUMS. Because our measures are constructed using the 1960 PUMS data, we show that results are robust to restricting the sample to post-1960 transition events (Appendix Table A4), alleviating concerns that early transitions influenced the 1960 inequality measures.

connections listed on individuals' online professional profiles, which are reported as exact counts up to 500, but displayed as "500+" beyond that point. Thus, we construct indicator variables that take the value of one if the number of connections falls within the following ranges, respectively: $[0, 100)$, $[100, 200)$, $[200, 300)$, $[300, 400)$, $[400, 500)$, and $[500, +\infty)$. Table 5 shows that coeducation shifts the distribution of professional connections rightward, increasing the probability of an individual having 500 or more connections by 1.7 percentage points (a 7% increase relative to the mean), and generally decreasing the probability of having connections below 400. This result is consistent with a larger, more diverse peer pool providing a broader foundation for the social capital essential to entrepreneurship (Garfinkel et al., 2021; Fluegge and Bailey, 2024).[15]

In addition to professional networks, we analyze the gender composition of entrepreneurial teams to provide concrete, direct evidence of cross-gender interaction and collaboration. Despite a well-documented tendency for gender homophily in entrepreneurial team formation—with Gompers, Huang, and Wang (2017) documenting that students are 25% more likely to work with same-gender peers—we find that coeducation leads to a 1.5-2 percentage point increase in the probability that a male founder has a female co-founder from the same university and peer cohorts (Table 6, Panel A). We interpret this magnitude using two benchmarks. First, relative to the sample mean, it represents a 75-100% increase. Second, relative to the 4.66% baseline rate of having a female co-founder in always-coeducational institutions, it corresponds to a 32–43% increase, which is also substantial.[16] By contrast, the probability of having a female co-founder from a different university or outside their peer cohorts remains unchanged (Panel B). These divergent effects suggest that the rise in cross-gender collaboration is more likely to stem from interactions within the higher

[15] Getik and Meier (2022) and Getik and Meier (2025) find that greater female presence in elementary school classes reduces boys' future social network scale. A plausible reason our results differ is that, in our setting, coeducation expands the overall cohort size by adding women without displacing men, whereas in their setting, a higher female share implies fewer male peers. Another difference between our setting and theirs is the students' age (college vs. elementary school), which might imply different peer interaction dynamics.

[16] By showing that a gender-diverse educational environment fosters collaboration with female cohorts in the formation of founder teams, we add to the broad literature seeking to understand the role of diverse environments in economic behavior. Duchin, Simutin, and Sosyura (2021) find that male CEOs who grew up in male-dominated environments 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 future decision-making. Calder-Wang and Gompers (2021) find that venture capital partners are more likely to hire female partners when they have more daughters.

education environment itself, rather than from broader societal trends or a general shift in gender perceptions. While idea exchange is generally unobservable, the shift in co-founding relationships provides tangible evidence that the interactions facilitated by coeducation are productive enough to culminate in new venture creation.

5.2.2 Household Formation and Downside Protection

A second channel through which coeducation may encourage entrepreneurship is by improving intra-household risk-sharing and downside protection. By increasing the likelihood or speeding up the process of finding a college-educated spouse, coeducation can improve a household's ability to absorb the financial risk of starting a business (Love, 2010; Christiansen, Joensen, and Rangvid, 2015; Doepke and Tertilt, 2016).[17]

We test this mechanism by examining variation in local family and educational structures and assessing whether the effect is larger in contexts where the marginal benefit of a college-educated spouse is greater. First, we hypothesize that the effect should be stronger in regions where college-educated partners were previously scarce. To this end, we measure the pre-treatment share of households in a county where only the male holds a bachelor's degree or above. Table 7 shows that the rise in entrepreneurship is larger in counties with an above-median share of such households: a 1.2 percentage point increase (Column 1) versus a 0.7 point increase in counties with a below-median share (Column 2), consistent with a greater benefit of marrying a college-educated spouse in regions where households would otherwise rely primarily on the male's education and earnings.[18]

Second, the value of this risk-sharing depends on the income potential of a college-educated spouse. In other words, the potential gains from marrying a college-educated spouse should be larger in regions with greater income disparities between college-educated and non-college-educated women—we refer to this gap as the female college income premium. Consistent with the mechanism, the coeducation effect on entrepreneurship is twice as large in counties with an

[17] Kirkebøen et al. (2025) find that college graduates tend to marry someone from the same institution.

[18] The sample size decreases compared to previous tables due to missing county identifiers for confidentiality reasons in the 1960 PUMS data. All results in this section are robust to restricting to post-1960 transitions (Appendix Table A6), alleviating concerns that early transitions influenced 1960 county characteristics.

above-median female college income premium (1.4 pp, Table 7 Column 3) than in those with a below-median premium (0.7 pp, Column 4).

Third, we posit that this downside protection is most valuable in regions where the expected male income tends to be lower, and therefore the marginal value of additional household income is higher for household financial security. Consistent with this hypothesis, Columns (5) and (6) of Table 7 suggest that entrepreneurship increases by 1.8 percentage points in below-median male income counties (Column 6), compared to a statistically insignificant change in wealthier counties (Column 5). A potential concern is that lower male income counties also have lower female incomes, which would dampen the marginal benefit of a college-educated spouse. However, the correlation between male and female income across counties is very small (0.0486), suggesting that regional variation in male earnings does not systematically align with variation in female earnings during this historical sample period.

Taken together, we find stronger effects where educated partners were scarce, their income potential was high, and local average male income was low ex ante. These patterns of heterogeneity provide support for the downside protection channel. The "economic insurance" provided by a college-educated spouse's earnings appears to lower the barrier to entrepreneurial entry. This finding may also imply a nuance: the resulting ventures may be more likely to represent subsistence-oriented or necessity-driven entrepreneurship, enabled by improved downside protection, rather than transformative, high-growth ventures (Schoar, 2010). In Section 5.3, we further explore this possibility by analyzing startup quality along the dimensions of employment size and longevity.

5.2.3 Major and Occupational Sorting

A third channel through which coeducation may increase male entrepreneurship is by reinforcing gendered sorting into educational tracks and occupations perceived as masculine. A growing literature shows that the early gender environment can have persistent effects on education and occupational choice by affecting entry into gender-stereotypical fields. For instance, exposure to more female peers in primary school leads women to select into less female-stereotypical

occupations, thereby raising their earnings (Getik and Meier, 2022). Similarly, Calkins et al. (2023) find that when women's colleges transition to coeducation, female students shift away from STEM majors.

We posit that the exposure to women at the higher education level can trigger an analogous but opposite response for men. Identity economics suggests that the introduction of a salient out-group can reinforce in-group identity and the desire to adhere to its norms (Akerlof and Kranton, 2005; Charles and Bradley, 2009; Akerlof and Kranton, 2010; Bertrand et al., 2011; Olivetti and Petrongolo, 2016). The influx of women into previously all-male universities may have made gender identity more salient, prompting male students to assert a masculine identity by specializing in traditionally male-dominated fields like Business and Science and entering male-dominated pursuits like entrepreneurship. This framework predicts that coeducation would not dampen but rather amplify existing gender stereotypes in occupational choice, driving entry into entrepreneurship.

To test this hypothesis, we begin by examining college major choices. We classify majors into twelve broad categories and estimate the impact of coeducation on the probability of majoring in each field. As shown in Table 8, the probability of majoring in business and science increases by 4.3 and 1.7 percentage points, respectively (Columns 1 and 2). Conversely, the probability of majoring in humanities, social science, nursing & health, and computer-related fields decreases by 3.7, 1.5, 0.3, and 0.7 percentage points (Columns 6, 7, 8, and 10).[19] The results are largely consistent with a re-specialization pattern, where coeducation steers male students toward traditionally male-dominated majors and away from female-dominated ones.[20]

[19]Though perceived as masculine today, computer programming was once seen as "women's work" until the end of the 1960s (Hicks, 2017). Historically, the term "computer" referred to people who performed laborious and low-paid scientific calculations by hand, a role dominated by women (Grier, 2013). In the early 1940s, women still made up the majority of early computer operators, programmers, and software developers, since programming was considered clerical and low-skilled compared with hardware making. The gender composition of the industry shifted in the 1960s when programming began to be viewed as more technical (Hicks, 2017). Therefore, it is plausible that computing remained relatively female-dominated during the earlier period of our sample, but transitioned to a more male-dominated field in later periods. Given the ambiguity in the historical gender composition, we interpret results for computer-related majors with caution.

[20]Altonji, Humphries, Yuksel, and Zhong (2025) provide evidence on the historical gender composition across majors, and the role of the gender education gap in explaining the gender earnings gap.

These effects are corroborated by historical anecdotal evidence from university records. For instance, after the University of Rochester turned coeducation in 1900, the share of male students obtaining bachelor's degrees in arts and philosophy dropped from 82 to 45 percent, while the share of male students earning bachelor's degrees in science rose from 18 to 55 percent between 1905 and 1909. Similarly, at Lehigh University, the share of male undergraduates in arts & science dropped from 35% to 25%, while the share in engineering rose from 51% to 56% in the decade following coeducation (Miller-Bernal and Poulson, 2004, p.57, p.171). The gendered sorting into business and science—fields that provide relevant skills for venture creation—could positively affect entrepreneurship rates.[21] These majors might offer students the knowledge and skills in assessing market opportunities, tackling management and financing issues, technical problem-solving, and innovation, thereby increasing entrepreneurship ideas and opportunities (Wadhwa et al., 2008).

More importantly, we further ask whether this sorting extends to occupational choice, especially given that entrepreneurship is often perceived as a masculine pursuit characterized by risk-taking, independence, and assertiveness (Lewis, 2006; Gupta, Turban, and Bhawe, 2008; Gupta, Turban, Wasti, and Sikdar, 2009). Men may also update their beliefs about success in entrepreneurship by drawing on the past success of other men (Liqui Lung, 2022). If this mechanism is at play, then the effect on entrepreneurship should be stronger in regions where entrepreneurs have been predominantly male, reflecting a strong local norm that codes entrepreneurship as a masculine profession. To capture this, we measure the male share of entrepreneurs among university graduates in a county prior to coeducation. We calculate this share across all historical graduates, as well as among graduates within the past 10 or 5 years, separately, to capture more recent regional norms.

The results are consistent across the three measures. Table 9 shows that in regions with an above-median historical male share of entrepreneurship, the effects are around 1.4-1.5 percentage points (Columns 1, 3, and 5). In stark contrast, the effect is negligible in regions with below-median share and thus lower perceived masculinity (Columns 2, 4, and 6). This dichotomy suggests that

[21] 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, Freeman, and Rissing, 2008).

coeducation catalyzes entrepreneurial entry partially by reinforcing male sorting into domains—either majors or occupations—that are culturally coded as masculine.

5.3 Startup Quality

Having established that coeducation increases entrepreneurial entry, we now assess whether the nature of these startups differs. Specifically, we examine two dimensions of startup quality: longevity and size. We find a nuanced pattern: coeducation leads to startups that survive longer, but do not grow larger. This is consistent with our proposed mechanisms operating on two margins: interactions and network channels (e.g., idea exchange, knowledge spillover) likely help improve new firm resilience, while household risk-sharing and downside protection primarily lower the entry barrier for stable, necessity-driven businesses rather than high-growth firms.

5.3.1 Startup Longevity

We first analyze startup longevity, defined as the total number of years a startup remains active. We estimate Poisson models to account for the count nature of the outcome variable. Coefficients from the Poisson model are interpreted as percentage changes in the expected years of survival. To account for potential confounding macroeconomic trends and industry-specific factors, all specifications include startup founding year and industry fixed effects. Industry classifications are inferred from startup names using an LLM-based method as described in Section 3.4.3 and Online Appendix B. Founding year fixed effects also help mitigate concerns arising from the right-censoring in our data, which ends in 2017.

The results, presented in Table 10 Panel A, show a statistically significant positive effect on startup longevity. Columns (1) to (3) suggest that coeducation increases average startup longevity by 2.3% to 3.7%. To further assess long-term survival, we examine the probability of surviving beyond specific thresholds. Columns (4) through (7) show that startups founded after coeducation are 1.8 percentage points more likely to survive beyond 15 years, 2.2 percentage points more likely to survive beyond 20 years, and 1.4 percentage points more likely to survive beyond 25 years. This

pattern confirms that the marginal startups enabled by coeducation are more enduring.

5.3.2 Startup Size and Growth Trajectory

We next examine startup scale, measured by the average and maximum number of employees over a firm's lifetime, reflecting its typical and peak employment sizes. The predicted impact of coeducation on startup size is ambiguous ex ante. On the one hand, the knowledge and network benefits could foster larger ventures. On the other hand, the household risk-sharing and downside protection channel may primarily facilitate smaller, subsistence-oriented entrepreneurship. The net effect of these two mechanisms might suggest that the overall impact is muted. Consistent with this, we find no statistically significant effect of coeducation on either average or maximum employment size (Table 10 Panel B). These regressions also control for industry fixed effects to address concerns about comparability across sectors (e.g., a 5-person software firm could be high-growth, whereas a 20-person retail firm may be subsistence-oriented).

We investigate this interpretation further by analyzing startup growth dynamics. Although the startup sizes remain unchanged, these startups may differ in their speed of growth. If the marginal startups are primarily necessity-driven ventures, they might exhibit *slower* upscaling. To test this, we define a third metric of quality: the number of years a startup takes to reach its peak employment. As shown in Table 10 Panel C, startups founded after coeducation take 10–12 percent longer to mature (i.e. reach its peak size). This slower growth trajectory provides further evidence that the coeducation-induced ventures are more likely to be necessity-based, prioritizing stability and longevity over rapid expansion.

6 Robustness

In this section, we conduct a series of robustness checks to address three potential concerns: the estimated impact of coeducation is driven by (i) endogenous adoption of coeducation by universities, (ii) confounding local economic trends, or (iii) student self-selection. Across these exercises, the

results remain consistent, strengthening the causal interpretation.

6.1 Endogeneity in Adoption Timing

A second concern is that university characteristics may simultaneously influence the timing of coeducation adoption and the entrepreneurship outcomes of their students. To investigate this, we directly examine whether university characteristics predict the timing of a university's transition to coeducation in Appendix Table A7 Panel A. Prior evidence using even longer historical periods of coeducation transition shows that private, Catholic, and earlier-established institutions transitioned later, while those in the Midwest and states with fewer coeducation enrollments adopted earlier (Goldin and Katz, 2011). Our results are largely similar. In particular, we find that public universities transitioned earlier than private institutions, reflecting pressure from public opinion supporting education equality and weaker reliance on alumni donations, as alumni often resisted coeducation for fear it would undermine traditions. Meanwhile, Catholic affiliated institutions transitioned later, likely because they faced doctrinal restrictions and largely remained single-sex until the reform of the Second Vatican Council in the 1960s. Founding year does not significantly predict adoption timing in our sample, though the coefficient directionally aligns with the prior finding that earlier-established universities transitioned later likely due to stronger alumni resistance. Geographical region, undergraduate enrollment, and state coeducational enrollment share (Panel B Column 1) show no statistically significant correlation with the timing of coeducation in our sample time period.

To account for these differences across institution types, we saturate our model with public/private university-by-entry year, religious affiliation-by-entry year, and Catholic affiliation-by-entry year fixed effects. These additional fixed effects absorb any time-varying, unobserved trends correlated with the institutional characteristics predictive of coeducation adoption. As shown in Appendix Table A8, the estimated effects remain quantitatively similar, alleviating concerns that our results are confounded by the endogenous timing of adoption.

6.2 Confounding Local Economic Trends

Moreover, we address the possibility that local economic shocks simultaneously influence coeducation decisions and entrepreneurship. In addition to prior evidence that coeducation is not correlated with broad macroeconomic conditions such as recessions (Goldin and Katz, 2011), we examine more geographically granular measures, including local population, income levels, and their growth rates, and find that they also do not systematically predict the timing of adoption (Appendix Table A7 Panel B).

We further address this concern with two strategies. First, we incorporate a rich set of interacted fixed effects, including state-by-entry year, income growth decile-by-entry year, and population growth decile-by-entry year fixed effects to account for time-varying, unobserved trends in local economic conditions. These specifications effectively compare individuals entering treated and control schools located in the same state and/or counties with similar economic development trajectories in the same year. The results remain quantitatively similar under these stringent controls (Appendix Table A9).

Second, we implement a matching strategy to further address concerns that unobserved institutional and local factors may bias our estimates. Specifically, for each treated university, we identify the closest-matched controls from all U.S. baccalaureate degree–granting institutions (excluding women's colleges, since our analysis requires counterfactuals of male student outcomes). We define the closest match by standardizing all pre-transition characteristics and selecting universities that minimize the sum of squared differences across this set of variables. These variables include both institutional factors, such as total revenue, average SAT score, and the pre-treatment student entrepreneurship rate, and local economic characteristics measured at the county level, including income, population, per capita income, and their growth rates. To minimize the influence of treatment, all matching variables are measured in the year prior to a university's transition, or use the earliest available observation if data are missing. For robustness, we employ a variety of specifications: we allow the same control university to be matched to multiple treated institutions, and present

results using samples with the five, three, and one closest matches. We further construct alternative matched samples by selecting the geographically nearest three and one universities from the top five closest matches to ensure comparability in local economic conditions and unobserved regional culture and norms. Appendix Table A10 presents the balance statistics for treated and control groups, confirming that the matched samples are comparable across a wide range of institutional and local economic dimensions, including variables not directly used in matching. Appendix Table A11 shows that the positive effect of coeducation on entrepreneurship holds across all matched specifications, including those based on the geographically nearest neighbors.

6.3 Student Self-Selection

Finally, a potential concern is that universities adopting coeducation may attract students of different types, in particular those with a pre-existing propensity for entrepreneurship. It is worth noting that college markets were far less nationalized during our sample period. Interstate infrastructure and air travel were underdeveloped, and state universities did not yet aggressively recruit out-of-state students for revenue.[22] Therefore, a university's coed status likely played only a limited role in students' choices. Nevertheless, to alleviate this concern, we further estimate our main specification on a subsample of students who are less likely to base their college choices on coeducation status: those attending public universities. These students' college choices are likely influenced more by geographic proximity and in-state tuition than by coeducation status, making them less susceptible to this form of selection.

As shown in Appendix Table A12, the effect of coeducation on entrepreneurship remains positive and statistically significant, increasing to approximately 1.9 percentage points. This result confirms that the effect is driven by coeducation's influence on male students' career choice, rather than by the selection of entrepreneurial students into coeducational institutions.

Taken together, these robustness tests alleviate concerns that our results are driven by endogenous adoption, local economic confounders, or selection, bolstering the causal interpretation that

[22] According to Cooperative Institutional Research Program (CIRP) Freshman Survey, the majority of students attended universities within 50 miles from home.

coeducation increases male entrepreneurship.

7 Conclusion

Leveraging the historical transition of all-male universities to coeducation in the U.S., we examine how gender integration in undergraduate education shapes entrepreneurial dynamics. We find that coeducation leads to a statistically and economically meaningful increase in male students' likelihood of becoming entrepreneurs. This effect operates through three complementary channels. First, greater female representation facilitates cross-gender interactions and idea exchanges, which is reflected in expanded professional networks and a higher probability of forming mixed-gender founding teams. Second, by potentially increasing the probability of marrying a college-educated spouse, coeducation improves household risk-sharing and downside protection. This encourages entrepreneurial risk-taking, especially in regions where the marginal value of improved financial insurance is greater. Third, the changed gender environment appears to reinforce the salience of male identity, encouraging men to pursue majors and occupations perceived as masculine, and to enter entrepreneurship.

Beyond increasing entrepreneurial entry, coeducation also affects the nature and quality of startups. While not larger in size, these marginal startups are more resilient, surviving longer despite growing more gradually—a pattern consistent with a rise in stable, necessity-driven entrepreneurship. Overall, our findings imply that the gender composition of college peer groups has lasting economic consequences, shaping career trajectories, entrepreneurial activity, and the nature of new business formation in the economy.

References

Akerlof, G. A., Kranton, R. E., 2005. Identity and the economics of organizations. Journal of Economic perspectives 19, 9–32.

Akerlof, G. A., Kranton, R. E., 2010. Identity economics: How our identities shape our work, wages, and well-being. In: *Identity Economics*, Princeton University Press.

Altonji, J. G., Humphries, J. E., Yuksel, Y., Zhong, L., 2025. Decomposing trends in the gender gap for highly educated workers. Tech. rep., National Bureau of Economic Research.

Barrios, J. M., Hochberg, Y., Yi, H. L., 2024. Hustling from home? work from home flexibility and entrepreneurial entry. Tech. rep., National Bureau of Economic Research.

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.

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

Bertrand, M., Duflo, E., Mullainathan, S., 2004. How much should we trust differences-in-differences estimates? The Quarterly journal of economics 119, 249–275.

Bertrand, M., et al., 2011. New perspectives on gender. Handbook of labor economics 4, 1543–1590.

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.

Botelho, T. L., Fehder, D., Hochberg, Y., 2025. Innovation-driven entrepreneurship. Tech. rep., Journal of Economic Literature.

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.

Charles, M., Bradley, K., 2009. Indulging our gendered selves? sex segregation by field of study in 44 countries. American journal of sociology 114, 924–976.

Christiansen, C., Joensen, J. S., Rangvid, J., 2015. Understanding the effects of marriage and divorce on financial investments: The role of background risk sharing. Economic Inquiry 53, 431–447.

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.

Doepke, M., Tertilt, M., 2016. Families in macroeconomics. In: *Handbook of macroeconomics*, Elsevier, vol. 2, pp. 1789–1891.

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.

Fluegge, R. B., Bailey, M., 2024. Where do friends matter for founders? the role of social connections in us entrepreneurship. The Role of Social Connections in US Entrepreneurship*(July 30, 2024) .

Ford, K. S., 2020. Marrying within the alma mater: Understanding the role of same-university marriages in educational homogamy. Sociological Research Online 25, 254–272.

Fortin, N. M., 2005. Gender role attitudes and the labour-market outcomes of women across oecd countries. oxford review of Economic Policy 21, 416–438.

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 .

Getik, D., Meier, A. N., 2022. Peer gender and mental health. Journal of Economic Behavior & Organization 197, 643–659.

Getik, D., Meier, A. N., 2025. The long-run effects of peer gender on occupational sorting and the wage gap. American Economic Journal: Economic Policy 17, 35–70.

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.

Gompers, P. A., Huang, K., Wang, S. Q., 2017. Homophily in entrepreneurial team formation. Tech. rep., National Bureau of Economic Research.

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.

Greenberg, K., Wasserman, M., Weber, E. A., 2024. The effects of gender integration on men: Evidence from the us military. Tech. rep., National Bureau of Economic Research.

Grier, D. A., 2013. When computers were human .

Gupta, V. K., Turban, D. B., Bhawe, N. M., 2008. The effect of gender stereotype activation on entrepreneurial intentions. Journal of applied psychology 93, 1053.

Gupta, V. K., Turban, D. B., Wasti, S. A., Sikdar, A., 2009. The role of gender stereotypes in perceptions of entrepreneurs and intentions to become an entrepreneur. Entrepreneurship theory and practice 33, 397–417.

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 .

Hicks, M., 2017. Programmed inequality: How Britain discarded women technologists and lost its edge in computing. MIT press.

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 .

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

Kirkebøen, L., Leuven, E., Mogstad, M., Mountjoy, J., 2025. College as a marriage market .

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.

Lewis, P., 2006. The quest for invisibility: Female entrepreneurs and the masculine norm of entrepreneurship. Gender, work & organization 13, 453–469.

Liqui Lung, C. W., 2022. Optimal self-screening and the persistence of identity-driven choices .

Love, D. A., 2010. The effects of marital status and children on savings and portfolio choice. The Review of Financial Studies 23, 385–432.

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.

Olivetti, C., Petrongolo, B., 2016. The evolution of gender gaps in industrialized countries. Annual review of Economics 8, 405–434.

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.

Schoar, A., 2010. The divide between subsistence and transformational entrepreneurship. Innovation policy and the economy 10, 57–81.

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.

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, 91 all-male universities adopted coeducation between 1942 and 1995. The figure below shows the number of traditionally all-male universities making the transition to coeducation each year.

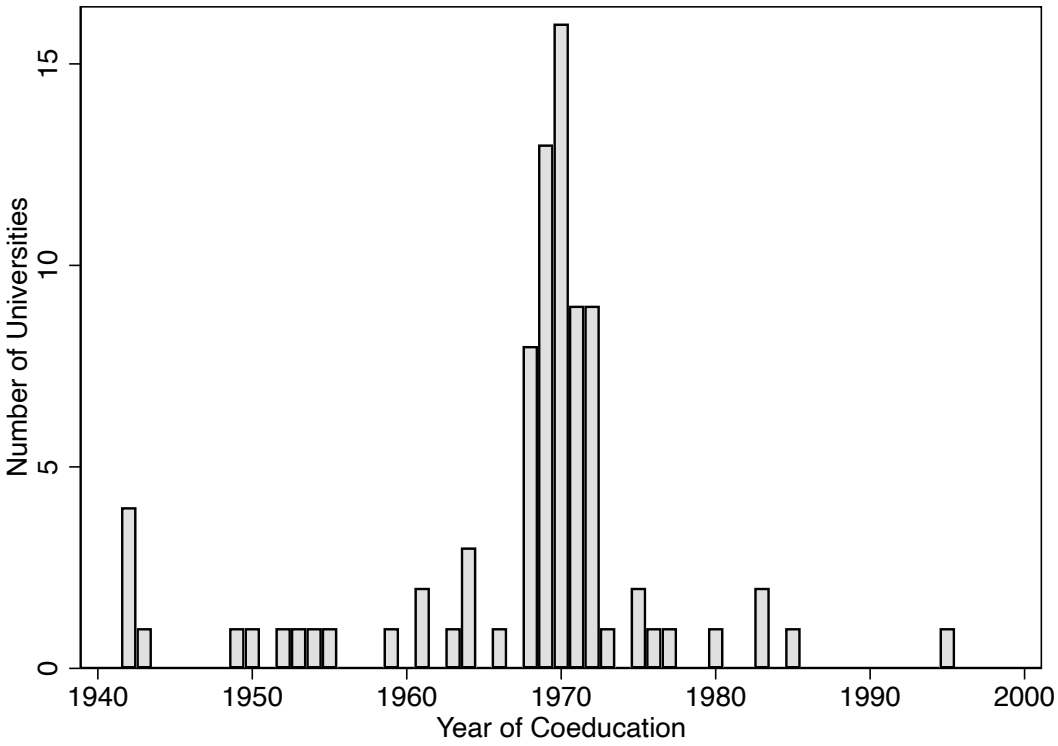


Figure 2: Coeducation and Student Gender Composition

This graph illustrates the gender composition of students in our sample of 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.

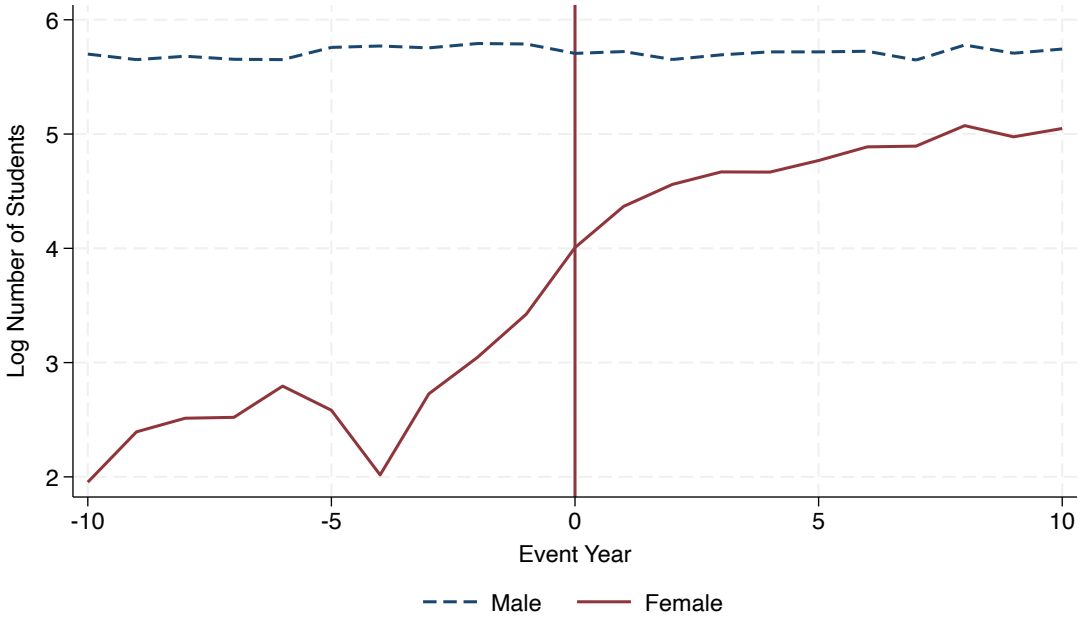


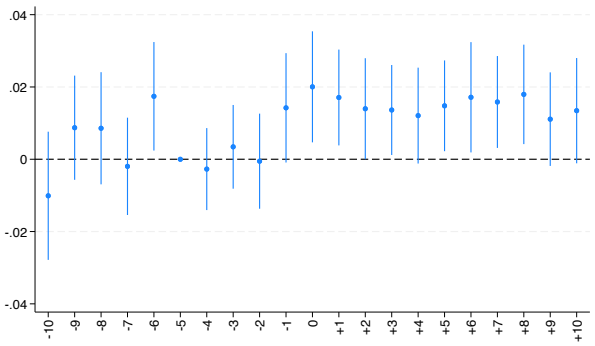
Figure 3: Dynamic Effects of Coeducation on Entrepreneurship

We estimate the following model:

$$\begin{aligned}\mathbb{1}(\text{Founder}_{sci,e}) = \alpha + \sum_{\tau \geq -10}^{\tau \leq 10} \beta_{0\tau}\, \mathbb{1}(\text{Year} - \text{Focal Treatment Year} = \tau)_{sc\tau,e} \times \text{Treat}_{s,e} \\ + \beta_1 \times \text{Treat}_{s,e} + \sum_{\tau \geq -10}^{\tau \leq 10} \beta_{2\tau} \times \mathbb{1}(\text{Year} - \text{Focal Treatment Year} = \tau)_{sc\tau,e} \\ + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,e} \end{aligned} \tag{4}$$

The dependent variable is an indicator that takes the value of one if individual i from cohort c in university s becomes a founder. $Treat_{s,e}$ equals one if in event e, university s adopts coeducation. $\mathbb{1}(Year - Focal\ Treatment\ Year = \tau)_{sc\tau,e}$ 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 entry-year fixed effects, capturing overall time trends in founder rates across university entering cohorts. 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. Estimates at $t = -5$ are omitted and used as the baseline. Standard errors are clustered by event.

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



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

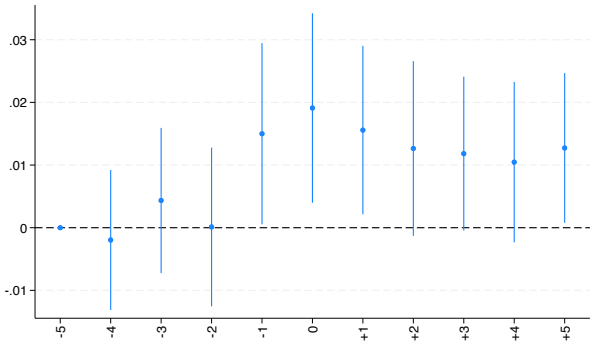


Table 1: Summary Statistics

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 presents university characteristics and local economic conditions, as matched to individuals. Panel C shows startup characteristics related to startup longevity, years to mature, size, and founder team composition. In Panel A, $\mathbb{1}$*(Founder)* is a dummy variable that takes the value of one if an individual becomes an entrepreneur in one's career, as defined in Section 3.4.2. *Year of Entry* and *Year of Graduation* describe the entering and graduating cohorts for each individual. We also report major-related variables, which are dummy variables that take the value of one if an individual majors in a certain field, and connection-related variables, which are dummy variables that take the value of one if an individual's number of LinkedIn connections falls within the range $[0, 100)$, $[100, 200)$, $[200, 300)$, $[300, 400)$, $[400, 500)$, and $[500, +\infty)$, repsectively. In Panel B, we first report university characteristics from HEGIS/IPEDS. *Num(Undergraduates)* and *Num(Faculties)* refer to the number of undergraduates and faculties in each university. *% Female in Major* and *% Female in Cohort* measure the change in exposure to females at the major and cohort levels, computed as the difference between the average female share in the ten years after and the ten years before coeducation adoption for each major and cohort at a given university. *Revenue from Tuition*, *Total Revenue*, *Expenditure on Instruction*, and *Total Expenditure* describe the level of revenue, tuition, expenditure, and expenditure on instruction for each university. *Income*, *Population*, *Income Per Capita* are measured in 1969, and *Income Growth*, and *Population Growth* are measured in 1970, all sourced from the Bureau of Economic Analysis. *Gender Inequality in Income (Education)* measures the county-level difference between average male and female income (bachelor rates), calculated using PUMS Data. *Female College Income Premium* measures the county-level income gap between college-educated and non-college-educated married women, calculated using PUMS Data. *Male Income* refers to the county average male income from PUMS. *% Male-Only BA+ Households* measures the county-level share of families in which only the male holds a bachelor's degree or above, calculated using PUMS Data. *Historical % Male Entrepreneurs* measures the regional share of entrepreneurs who are male among college graduates from the county. *Rolling % Male Entrepreneurs (10 Yrs)* and *Rolling % Male Entrepreneurs (5 Yrs)* are defined analogously, but restricted to college graduates within the past 10 and 5 years, respectively. In Panel C, $\mathbb{1}$*(Survive 10/15/20/25)* are indicator variables that equal one if a startup survives at least 10, 15, 20, or 25 years, respectively. *Years to Mature* measures the number of years from the founding year until the startup reaches its maximum employment size. *Avg Num(Employees)* measures the average number of employees across all observed years of a startup's life. *Max Num(Employees)* measures the maximum number of employees throughout a startup's life. $\mathbb{1}$*(Female Co-founder in Cohort)* takes the value of one if a startup has at least one female co-founder who attended the same university and graduated in peer cohorts as the male founder. *1(Female Co-founder Outside Cohort)* takes the value of one if a startup has at least one female co-founder who did not attend the same university or graduate in peer cohorts as the male founder. Revenue from tuition, total revenues, expenditure on instruction, total expenditure, and population are reported in millions and income is reported in billions.

Variable	Mean	Standard Deviation	Median
Panel A: Individual Characteristics			
$\mathbb{1}$(Founder)	0.130	0.330	0
Year of Entry	1985	12.96	1985
Year of Graduation	1989	12.95	1989
Art & Architecture	0.030	0.160	0
Humanity	0.100	0.300	0
Social Science	0.110	0.310	0
Economics	0.060	0.230	0
Business	0.230	0.420	0
Nursing & Health	0.010	0.100	0
Science	0.090	0.290	0
Computer	0.050	0.220	0
Engineering	0.210	0.410	0
Education	0.010	0.070	0
Connections 0-100	0.280	0.450	0
Connections 100-200	0.190	0.390	0
Connections 200-300	0.130	0.340	0
Connections 300-400	0.100	0.290	0
Connections 400-500	0.070	0.250	0
Connections 500+	0.240	0.430	0

Variable	Mean	Standard Deviation	Median
Panel B: University and Local Economic Characteristics			
Num(Undergraduates)	9093.0	8159.0	5500.0
Num(Faculties)	568.7	457.0	457.0
% Female in Major	0.079	0.094	0.068
% Female in Cohort	0.066	0.062	0.052
Revenue from Tuition	108.3	128.2	65.1
Total Revenue	381.7	493.1	188.4
Expenditure on Instruction	88.6	143.6	34.3
Total Expenditure	370.5	429.7	200.4
Income	17.6	33.6	7.4
Population	0.767	1.079	0.613
Income Per Capita	19527.0	16085.0	15809.0
Income Growth	0.070	0.030	0.070
Population Growth	0.005	0.010	0.004
Gender Inequality in Income	3058.0	686.9	2826.0
Gender Inequality in Education	0.050	0.020	0.050
Female College Income Premium	1127.0	438.9	1019.0
Male Income	4374.0	626.2	4331.0
%Male-Only BA+ Households	0.080	0.020	0.080
Historical %Male Entrepreneurs	0.640	0.200	0.680
Rolling % Male Entrepreneurs (10 Yrs)	0.610	0.200	0.630
Rolling % Male Entrepreneurs (5 Yrs)	0.600	0.210	0.620
Panel C: Startup Characteristics			
Years of Survival	13.5	10.9	10.0
$\mathbb{1}$(Survive 10)	0.520	0.500	1.000
$\mathbb{1}$(Survive 15)	0.350	0.480	0.000
$\mathbb{1}$(Survive 20)	0.240	0.420	0.000
$\mathbb{1}$(Survive 25)	0.160	0.370	0.000
Years to Mature	3.1	7.1	0.0
Avg Num(Employees)	2.6	17.5	1.0
Max Num(Employees)	4.5	52.4	1.0
$\mathbb{1}$(Female Co-founder in Cohort)	0.020	0.150	0.000
$\mathbb{1}$(Female Co-founder outside Cohort)	0.161	0.367	0.000

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. $Treat_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $Post_{sc,e}$ 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 entry-year fixed effects, capturing overall time trends in founder rates across university entering 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 force 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-entry year 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. The first (last) three columns use an event window that includes the ten (five) years before and after the coeducation adoption year. 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 event 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	0.010***	0.011***	0.012***	0.009**	0.009***	0.012***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Treat	-0.003	-0.003		-0.003	-0.003	
	(0.003)	(0.003)		(0.004)	(0.004)	
Post	-0.002***	-0.002**		-0.001**	0.000	
	(0.000)	(0.001)		(0.000)	(0.000)	
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
Observations	3,074,393	3,074,393	3,074,367	1,944,144	1,944,144	1,944,106
R^2	0.023	0.023	0.024	0.019	0.019	0.021

Table 3: Coeducation, Female Peer Share, and Entrepreneurship

To examine the heterogeneous impact of coeducation on entrepreneurship across female peer share, we estimate the following model:

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

The dependent variable is whether individual i from cohort c, university s, in event e has ever become an entrepreneur. $Treat_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $Post_{sc,e}$ 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 entry-year fixed effects, capturing overall time trends in founder rates across university entering cohorts. 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 entrepreneurship rate. *% Female in Major* and *% Female in Cohort* measure the change in exposure to females at the major and cohort levels, computed as the difference between the average female share in the ten years after and the ten years before coeducation adoption for each major and cohort at a given university. Columns (1) and (2) report results for sub-samples where *% Female in Major* is above and below the median, respectively. Similarly, Columns (3) and (4) report results for sub-samples where *% Female in Cohort* is above and below the median. The sample size is smaller than in Table 2 due to missing major information for some individuals. Lower-order interaction terms are omitted from the table for brevity. Standard errors clustered by event are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	𝟙(Founder)			
	by % Female in Major		by % Female in Cohort	
	Above Median	Below Median	Above Median	Below Median
	(1)	(2)	(3)	(4)
Treat × Post	0.015***	-0.000	0.012***	0.007**
	(0.004)	(0.009)	(0.003)	(0.003)
School FE	Yes	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes	Yes
Event FE	Yes	Yes	Yes	Yes
Observations	1,026,520	1,025,939	1,338,089	1,338,314
R^2	0.014	0.021	0.014	0.017

Table 4: Coeducation, Gender Inequality, and Entrepreneurship

To examine the heterogeneous impact of coeducation on entrepreneurship across measures of gender equality, we estimate the following model:

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

The dependent variable is whether individual i from cohort c, university s, in event e has ever become an entrepreneur. $Treat_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $Post_{sc,e}$ 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 entry-year fixed effects, capturing overall time trends in founder rates across university entering cohorts. 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 entrepreneurship rate. *Gender Inequality in Income (Education)* measures the gap between male and female income and bachelor rates, respectively, at the county level, calculated using PUMS Data. Columns (1) and (2) report results for sub-samples where *Gender Inequality in Income* is above and below the median, respectively. Similarly, Columns (3) and (4) report results for sub-samples where *Gender Inequality in Education* is above and below the median. The sample size decreases compared to Table 2 because some areas have missing county identifiers due to confidentiality requirements in the PUMS data. Lower-order interaction terms are omitted from the table for brevity. Standard errors clustered by event are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	𝟙(Founder)			
	by Gender Inequality in Income		by Gender Inequality in Education	
	Above Median	Below Median	Above Median	Below Median
	(1)	(2)	(3)	(4)
Treat × Post	0.008*	0.012**	0.001	0.016***
	(0.004)	(0.005)	(0.005)	(0.004)
School FE	Yes	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes	Yes
Event FE	Yes	Yes	Yes	Yes
Observations	794,379	966,631	826,317	934,694
R^2	0.026	0.020	0.025	0.022

Table 5: Coeducation and Professional Networks

We test the impact of coeducation on individuals' professional networks using the following model:

$$\text{Connections}_{sci,e} = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,e}, \quad (5)$$

The dependent variables measure the number of connections for individual i from cohort c, university s in event e. Because LinkedIn connection counts are censored at 500, it is transformed into a set of indicator variables, each taking the value of one if the count falls within the range of $[0, 100)$, $[100, 200)$, $[200, 300)$, $[300, 400)$, $[400, 500)$, and $[500, +\infty)$, respectively. These variables are used as outcome variables in Columns (1) through (6). $Treat_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $Post_{sc,e}$ 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 entry-year fixed effects, capturing overall time trends in founder rates across university entering cohorts. 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 one's networking scale. Lower-order interaction terms are omitted from the table for brevity. Standard errors clustered by event 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	-0.003	-0.012***	-0.002	-0.002	0.001	0.017***
	(0.005)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
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
Observations	3,074,393	3,074,393	3,074,393	3,074,393	3,074,393	3,074,393
R^2	0.057	0.004	0.003	0.004	0.004	0.036

Table 6: Coeducation and Female Co-founders

To study the impact of coeducation on gender-collaborative founding, we estimate the following four specifications across the four columns, respectively:

$$\mathbb{1}(\text{Female Co-founder})_{sci,e} = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \varepsilon_{sci,e}, \quad (6)$$

$$\mathbb{1}(\text{Female Co-founder})_{sci,e} = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,e}, \quad (7)$$

$$\mathbb{1}(\text{Female Co-founder})_{sci,e} = \alpha + \beta \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \varphi_{c,e} + \theta_{s,e} + \varepsilon_{sci,e}. \quad (8)$$

$$\mathbb{1}(\text{Female Co-founder})_{sci,e} = \alpha + \beta \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \varphi_{c,e} + \theta_{s,e} + \gamma_t + \varepsilon_{sci,e}. \quad (9)$$

We classify female co-founders into two groups: those who attended the same university and graduate in peer cohorts (defined as within three years before or after graduation) as the male student and those who did not. The dependent variable in Panel A is whether individual i from cohort c university s in event e has at least one female co-founder from the same university and peer cohorts. The dependent variable in Panel B is whether individual i from cohort c university s in event e has female co-founders who are from a different university or outside their peer cohorts. $Treat_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $Post_{sc,e}$ 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 entry-year fixed effects, capturing overall time trends in founder rates across university entering cohorts. 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-entry year 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. γ_t are startup founding year fixed effects, which control for macroeconomic conditions at the time of founding that may influence startup characteristics. β estimates the impact of entering the university after its transition to coeducation on the individual's probability to collaborate with female co-founders. Columns (1) to (4) report results from Eqs. (6) to (9). Lower-order interaction terms are omitted from the table for brevity or absorbed by fixed effects. Standard errors clustered by event are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A.	**𝟙(Female Co-founder in Cohort)**			
	(1)	(2)	(3)	(4)
Treat × Post	0.015*	0.016*	0.020**	0.019**
	(0.008)	(0.009)	(0.008)	(0.007)
School FE	Yes	Yes		
Entry Year FE	Yes	Yes		
Event FE		Yes		
School × Event FE			Yes	Yes
Entry Year × Event FE			Yes	Yes
Founding Year FE				Yes
Observations	27,356	27,356	27,213	27,213
R^2	0.027	0.027	0.044	0.060
Panel B.	**𝟙(Female Co-founder outside Cohort)**			
	(1)	(2)	(3)	(4)
Treat × Post	-0.045	-0.045	-0.037	-0.044
	(0.032)	(0.033)	(0.038)	(0.038)
School FE	Yes	Yes		
Entry Year FE	Yes	Yes		
Event FE		Yes		
School × Event FE			Yes	Yes
Entry Year × Event FE			Yes	Yes
Founding Year FE				Yes
Observations	27,356	27,356	27,213	27,213
R^2	0.046	0.046	0.071	0.090

Table 7: Coeducation, Downside Protection, and Entrepreneurship

To examine the heterogeneous impact of coeducation on entrepreneurship across variations in downside protection, we estimate the following specification in sub-samples split at the median of three measures: (i) *% Male-Only BA+ Households*, (ii) *Female College Income Premium*, and (iii) *Male Income*:

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

% Male-Only BA+ Households measures the county-level share of families in which only the male holds a bachelor's degree or above, calculated using PUMS Data. *Female College Income Premium* measures the county-level income gap between college-educated and non-college-educated married women, calculated using PUMS Data. *Male Income* refers to the county average male income from PUMS. The dependent variable is whether individual i from cohort c, university s, in event e has ever become an entrepreneur. $Treat_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $Post_{sc,e}$ 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 entry-year fixed effects, capturing overall time trends in founder rates across university entering cohorts. 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 entrepreneurship rate. The sample size decreases compared to Table 2 because some areas have missing county identifiers due to confidentiality requirements in the PUMS data. Lower-order interaction terms are omitted from the table for brevity. Standard errors clustered by event are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	𝟙(Founder)					
	by % Male-Only BA+ Households		by Female College Income Premium		by Male Income	
	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median
	(1)	(2)	(3)	(4)	(5)	(6)
Treat × Post	0.012**	0.007*	0.014***	0.007	0.002	0.018***
	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)
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
Observations	994,459	766,552	859,809	901,201	878,621	882,390
R^2	0.023	0.024	0.024	0.023	0.026	0.021

Table 8: Coeducation and College Major Choice

We test the impact of coeducation on students' major choices using the following model:

$$\text{Major}_{sci,e} = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,e}, \quad (10)$$

The dependent variables are dummy variables indicating whether individual i from cohort c, university s in event e takes their undergraduate degree in a certain major. $Treat_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $Post_{sc,e}$ 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 entry-year fixed effects, capturing overall time trends in founder rates across university entering cohorts. 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. The sample size decreases compared to Table 2 because individuals have missing major information. Lower-order interaction terms are omitted from the table for brevity. Standard errors clustered by event are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	The Probability of Choosing a Major in ...				
	Business	Science	Engineering	Economics	Art & Architecture
	(1)	(2)	(3)	(4)	(5)
Treat × Post	0.043***	0.017**	-0.002	0.004	-0.001
	(0.008)	(0.007)	(0.005)	(0.004)	(0.001)
School FE	Yes	Yes	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes	Yes	Yes
Event FE	Yes	Yes	Yes	Yes	Yes
Observations	2,730,461	2,730,461	2,730,461	2,730,461	2,730,461
R^2	0.150	0.041	0.259	0.045	0.021
	The Probability of Choosing a Major in ...				
	Humanity	Social Science	Nursing & Health	Education	Computer
	(6)	(7)	(8)	(9)	(10)
Treat × Post	-0.037***	-0.015**	-0.003***	0.000	-0.007*
	(0.005)	(0.006)	(0.001)	(0.001)	(0.004)
School FE	Yes	Yes	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes	Yes	Yes
Event FE	Yes	Yes	Yes	Yes	Yes
Observations	2,730,461	2,730,461	2,730,461	2,730,461	2,730,461
R^2	0.105	0.051	0.010	0.007	0.036

Table 9: Coeducation, Perceived Masculinity, and Entrepreneurship

To examine the heterogeneous impact of coeducation on entrepreneurship across regions with differing perceived masculinity of entrepreneurship, we estimate the following model in sub-samples split at the median of three measures: *Historical % Male Entrepreneurs*, *Rolling % Male Entrepreneurs (10 Yrs)*, and*Rolling % Male Entrepreneurs (5 Yrs)*:

$$\mathbb{1}(\text{Founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \rho_e + \varepsilon_{sci,e}, \tag{11}$$

Historical % Male Entrepreneurs measures the regional share of entrepreneurs who are male among college graduates from the county. *Rolling % Male Entrepreneurs (10 Yrs)* and *Rolling % Male Entrepreneurs (5 Yrs)* are defined analogously, but restricted to college graduates within the past 10 and 5 years, respectively. The dependent variable is whether individual i from cohort c, university s, in event e has ever become an entrepreneur. $Treat_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $Post_{sc,e}$ 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 entry-year fixed effects, capturing overall time trends in founder rates across university entering cohorts. 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 entrepreneurship rate. The sample size is smaller than in Table 2 because some counties have no observed entrepreneurs in certain years. Lower-order interaction terms are omitted from the table for brevity. Standard errors clustered by event are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	𝟙(Founder)					
	by Historical % Male Entrepreneurs		by Rolling % Male Entrepreneurs (10 Yrs)		by Rolling % Male Entrepreneurs (5 Yrs)	
	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median
	(1)	(2)	(3)	(4)	(5)	(6)
Treat × Post	0.014**	-0.000	0.014***	-0.005	0.015***	0.000
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.003)
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
Observations	1,100,300	1,100,727	1,098,781	1,102,246	1,100,070	1,100,957
R^2	0.018	0.022	0.015	0.024	0.016	0.024

Table 10: Coeducation, Startup Longevity, and Size

To examine the quality of startups founded by students exposed to coeducation, we estimate the following model:

$$\text{Startup Characteristics}_{sci,f,e} = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \varphi_c + \theta_s + \gamma_t + \delta_{ind} + \rho_e + \varepsilon_{sci,f,e}, \quad (12)$$

The dependent variables measure the quality of startup f, founded by individual i from cohort c university s in event e, including the years of survival, indicators of surviving beyond 10, 15, 20, and 25 years (Panel A), the average and maximum number of employees throughout its life (Panel B), and years to mature(Panel C), defined as the number of years it takes the startup to reach its peak employment. Detailed variable definitions can be found in the legends to Table 1. When the outcome variables are count-based (i.e. years of survival, employee counts, and years to mature), we estimate the models using Poisson. $Treat_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $Post_{sc,e}$ 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 entry-year fixed effects, capturing overall time trends in founder rates across university entering cohorts. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. γ_t are startup founding year fixed effects, which control for macroeconomic conditions at the time of founding that may influence startup characteristics. δ_{ind} are industry fixed effects, which control for systematic differences across industries that may affect startup outcomes. We infer industry from startup names using a combination of large language model classification and embedding-based correction, retaining only high-confidence inferences. This procedure reduces the number of observations, with further details provided in Appendix B. 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 startup quality. Lower-order interaction terms are omitted from the table for brevity or absorbed by fixed effects. Standard errors clustered by event are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Startup Longevity

	Years of Survival			𝟙(Survive 10)	𝟙(Survive 15)	𝟙(Survive 20)	𝟙(Survive 25)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treat × Post	0.037**	0.023**	0.023**	0.003	0.018**	0.022**	0.014*
	(0.016)	(0.011)	(0.011)	(0.007)	(0.008)	(0.011)	(0.007)
School FE	Yes	Yes					
Entry Year FE	Yes	Yes					
Event FE		Yes					
School × Event FE			Yes	Yes	Yes	Yes	Yes
Entry Year × Event FE			Yes	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	218,247	218,247	218,175	218,175	218,175	218,175	218,175
Model	Poisson	Poisson	Poisson	OLS	OLS	OLS	OLS
R^2				0.598	0.646	0.671	0.684

Panel B. Startup Size

	Avg Num(Employees)			Max Num(Employees)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treat × Post	0.088	0.041	-0.022	0.038	-0.037	-0.128
	(0.104)	(0.064)	(0.056)	(0.154)	(0.099)	(0.086)
Treat	-0.179	-0.112		-0.183	-0.058	
	(0.235)	(0.119)		(0.326)	(0.143)	
Post	-0.068*	-0.001		-0.088*	0.023	
	(0.037)	(0.017)		(0.047)	(0.034)	
School FE	Yes	Yes		Yes	Yes	
Entry Year FE	Yes	Yes		Yes	Yes	
Event FE		Yes			Yes	
School × Event FE			Yes			Yes
Entry Year × Event FE			Yes			Yes
Founding Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	218,247	218,247	218,247	218,247	218,247	218,247
Model	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson

Panel C. Startup Growth Speed

	Years to Mature		
	(1)	(2)	(3)
Treat × Post	0.124***	0.105***	0.122***
	(0.029)	(0.024)	(0.027)
Treat	-0.051**	-0.036	
	(0.022)	(0.024)	
Post	-0.028***	-0.006	
	(0.006)	(0.013)	
School FE	Yes	Yes	
Entry Year FE	Yes	Yes	
Event FE		Yes	
School × Event FE			Yes
Entry Year × Event FE			Yes
Founding Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Observations	218,207	213,603	213,307
Model	Poisson	Poisson	Poisson

Online Appendix

A College 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, humanities, social science, economics, business, nursing & health, science, computer-related, engineering, and education.

The specific fields within each major group are as follows:

1. Art & architecture includes design, dance, music, theater, architecture, etc.
2. Humanities includes philosophy, languages, literature, history, etc.
3. Social science includes sociology, political science, international studies, media, anthropology, human resources, criminal studies, etc.
4. Business includes finance, accounting, marketing, business communication, entrepreneurial studies, etc.
5. Nursing & health includes nursing, exercise science, physical science, and pre-meds.
6. Science includes physics, mathematics, chemistry, biology, geology, etc.
7. Computer-related includes computer science, software engineering, information system, etc.
8. Engineering includes civil engineering, electric engineering, mechanical engineering, etc.

B Industry Classification using an LLM-Based Approach

We classify startups into 2-digit NAICS industries by combining LLM predictions with embedding-based correction procedures. Our test data consists of startup names obtained from individuals' LinkedIn profiles, which lack standardized industry information. For our training data, we use the YTS dataset, which records company names and NAICS codes for all U.S. establishments. Compared to public company datasets (e.g., Compustat), YTS is more comparable to our LinkedIn startup sample in terms of the coverage of smaller, privately held companies.

B.1 Name Cleaning

We first clean and standardize company names to improve consistency across datasets. We convert all characters to lowercase and drop any text enclosed in parentheses. We remove common corporate suffixes and abbreviations such as inc., ltd., corp., corporation, incorporated, limited, co., lc, llc, and other variations.

B.2 GPT-Based Industry Classification

We use GPT-4o to generate initial industry classifications from startup names. When providing input to the model, we consider the tradeoff that too little information can make industry definitions ambiguous, while too much can overwhelm the algorithm. To balance this, our prompt provides examples of company name–industry pairs, as well as short keywords for each of the 18 NAICS-based categories. The company name–industry pairs are randomly drawn from the YTS training data, ensuring at least one name–industry pair from each industry, which helps the model learn how actual names map to industries. Furthermore, the manually selected keywords provide additional semantic cues without introducing excessive text. The full set of industry categories and associated keywords is listed below:

1. Agriculture, Forestry, Fishing, and Hunting: agriculture, forestry, logging, fishing, hunting, crops, farming, ranching
2. Mining, Quarrying, and Oil and Gas Extraction: mining, quarrying, oil, gas extraction
3. Construction: construction, contractors, exteriors
4. Manufacturing: manufacturing
5. Wholesale Trade: wholesaler

6. Retail Trade: retailers, dealers, stores, gasoline stations, groceries, florists
7. Transportation and Warehousing: transportation, logistics, freight trucking, taxi, bus, delivery, warehousing, loading/unloading, storage
8. Information: publishing, motion picture, video, recording, broadcasting, internet, telecommunications, data, software, radio, television, media networks, computing infrastructure
9. Finance and Insurance: finance, insurance, banking, credit, investments, securities, funds, trusts, mortgages, brokerage
10. Real Estate and Rental and Leasing: real estate, rental, leasing, car rental, home inspection, appraisal
11. Professional, Scientific, and Technical Services: lawyers, legal services, accounting, tax preparation, engineering services, design services, advertising, public relations, photography, consulting, CPA, veterinary
12. Management of Companies and Enterprises: management of companies, company offices
13. Administrative and Support, Waste Management Services: administrative support, waste management, travel agencies, security guards, investigation, pest control, janitorial services, carpet cleaning, damage cleaning, gardening, employment placement agencies
14. Educational Services: colleges, schools, driving schools, taekwondo, yoga
15. Health Care and Social Assistance: health care, hospitals, nursing, residential care, social assistance, dental care, child development, chiropractic, elder care
16. Arts, Entertainment, and Recreation: arts, entertainment, recreation, spectator sports, museums, historical sites, amusement, gambling, parks
17. Accommodation and Food Services: hospitality, accommodation, food services, drinking places, hotels, restaurants, cafeterias
18. Other Services (Repair, Personal, and Religious): repair and maintenance, laundry services, religious organizations, churches, auto collision, hair, nail, skin care, barber, beauty services

To reduce randomness in GPT responses, we repeat the classification process independently five times for each company. The final industry classification is assigned by majority vote across the five runs, and the associated confidence score is calculated as the average confidence across runs. Inferences with a confidence score below 0.75 are excluded from the sample.

Applying this procedure to classify 30,000 YTS companies yields an accuracy of 54% against the YTS industry labels. This relatively low accuracy reflects the task's inherent difficulty: many startup names do not contain sufficient information to determine their industry.

B.3 Embedding-Based Correction

To address this inherent difficulty mentioned above, we introduce an embedding-based correction. Specifically, we use vector representations of company names and supervised learning to identify and retain only GPT classifications likely to be correct.

We generate vector representations of company names using a sentence-transformer, which maps sentences to a 384-dimensional dense vector space. This model captures semantic similarity and contextual meaning, rather than treating texts as an unordered collection of tokens. Therefore, it generally outperforms bag-of-words or TF-IDF approaches in short-text classification tasks.

We compute a measure of within-cluster cohesion for each industry. Specifically, we calculate the average pairwise cosine distance between company-name embeddings within each industry cluster. Lower values indicate that company names within the industry cluster are more semantically similar in the embedding space, implying higher cohesion.

Then, we train a Support Vector Machine (SVM) classifier on the YTS sample. Inputs include company-name embeddings, the initial GPT classifications, and each industry's within-cluster cohesion. The outcome is a binary variable equal to one if the initial GPT classification matches the true industry.

We choose SVM because it handles high-dimensional data, such as embeddings, efficiently, and can capture non-linear relationships via kernel functions. Among available kernels, the radial basis function (RBF) kernel performs best because it measures similarity based on distance and allows a curved, flexible boundary instead of a rigid, linear boundary in the space. We train the model using 5-fold cross-validation, and select hyperparameters to maximize the out-of-sample F1-score.[23]

B.4 Filtering by Predicted Confidence

For each LinkedIn company, we apply the embedding-based SVM to assess whether the initial GPT-based industry classification is likely to be correct. We retain only the classifications that the SVM identifies as correct. This filtering procedure retains approximately 58% of the LinkedIn

[23]The F1-score is the harmonic mean of precision (i.e. the share of predicted positives that are truly correct) and recall (i.e. the share of true positives that are successfully identified).

sample, and within this subset, the classification accuracy increases substantially to 84%. This is the final sample used in regressions with industry controls.

C Additional Tables and Figures

Table A1: 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, measured in the year prior to transition to coeducation, and calculates the standardized differences in the means. Panel A shows individual characteristics related to cohort, founder indicators, connections, and majors. Panel B presents university characteristics and local economic conditions, as matched to individuals. Panel C shows startup characteristics related to startup longevity, years to mature, size, and founder team composition. Variable definitions can be found in Table 1 legend. Revenue from tuition, total revenues, expenditure on instruction, total expenditure, and population are reported in millions and income is reported in billions.

Panel A: Individual Characteristics						
	Treat = 0		Treat = 1			
	Mean	SD	Mean	SD	P-value	Std. Diff.
𝟙(Founder)	0.130	(0.336)	0.181	(0.385)	(0.000)	0.100
Art & Architecture	0.027	(0.163)	0.016	(0.125)	(0.000)	-0.055
Humanity	0.089	(0.285)	0.197	(0.398)	(0.000)	0.220
Social Science	0.112	(0.315)	0.146	(0.354)	(0.000)	0.073
Economics	0.054	(0.226)	0.083	(0.276)	(0.000)	0.080
Business	0.230	(0.421)	0.162	(0.369)	(0.000)	-0.121
Nursing & Health	0.011	(0.105)	0.013	(0.115)	(0.107)	0.015
Science	0.095	(0.293)	0.124	(0.330)	(0.000)	0.067
Computer	0.056	(0.229)	0.016	(0.124)	(0.000)	-0.154
Engineering	0.208	(0.406)	0.108	(0.311)	(0.000)	-0.195
Education	0.006	(0.076)	0.011	(0.102)	(0.000)	0.038
Connections 0-100	0.277	(0.447)	0.346	(0.476)	(0.000)	0.106
Connections 100-200	0.185	(0.388)	0.204	(0.403)	(0.000)	0.034
Connections 200-300	0.135	(0.341)	0.119	(0.324)	(0.000)	-0.032
Connections 300-400	0.094	(0.292)	0.078	(0.268)	(0.000)	-0.040
Connections 400-500	0.066	(0.248)	0.054	(0.226)	(0.000)	-0.036
Connections 500+	0.243	(0.429)	0.197	(0.398)	(0.000)	-0.079
Observations	126,855		6,690			

Panel B: University and Local Economic Characteristics

	Treat = 0		Treat = 1			
	Mean	SD	Mean	SD	P-value	Std. Diff.
Num(Undergraduates)	8,963.1	(8,068.0)	5,716.1	(5,756.0)	(0.000)	-0.328
Num(Faculties)	567.5	(464.6)	410.9	(382.7)	(0.000)	-0.260
Revenue from Tuition	94.4	(102.2)	38.2	(44.9)	(0.000)	-0.503
Total Revenue	333.7	(378.1)	160.2	(199.7)	(0.000)	-0.406
Expenditure on Instruction	70.9	(101.0)	21.4	(43.1)	(0.000)	-0.451
Total Expenditure	341.9	(383.0)	162.6	(203.0)	(0.000)	-0.414
Income	13.4	(21.5)	4.9	(8.7)	(0.000)	-0.368
Population	0.689	(0.766)	0.707	(0.974)	(0.094)	0.014
Income Per Capita	16,873.2	(11,951.8)	5,884.4	(4,645.4)	(0.000)	-0.857
Income Growth	0.071	(0.029)	0.076	(0.026)	(0.000)	0.121
Population Growth	0.005	(0.012)	0.005	(0.010)	(0.000)	-0.043
Gender Inequality in Income	3,047.8	(687.4)	3,244.9	(669.9)	(0.000)	0.205
Gender Inequality in Education	0.054	(0.019)	0.055	(0.017)	(0.001)	0.039
% Male-Only BA+ Households	0.082	(0.024)	0.082	(0.022)	(0.477)	0.008
Female College Income Premium	1,123.9	(437.6)	1,133.5	(465.4)	(0.172)	0.015
Male Income	4,366.5	(623.0)	4,527.1	(605.6)	(0.000)	0.185
Historical % Male Entrepreneurs	0.645	(0.190)	0.733	(0.203)	(0.000)	0.316
Rolling % Male Entrepreneurs (10 Yrs)	0.616	(0.188)	0.718	(0.206)	(0.000)	0.363
Rolling % Male Entrepreneurs (5 Yrs)	0.608	(0.201)	0.702	(0.223)	(0.000)	0.312
Observations	126,855		6,690			

Panel C: Startup Characteristics

	Treat = 0		Treat = 1			
	Mean	SD	Mean	SD	P-value	Std. Diff.
Years of Survival	13.6	(10.8)	17.2	(12.3)	(0.000)	0.221
𝟙(Survived 10 yrs)	0.528	(0.499)	0.639	(0.481)	(0.000)	0.159
𝟙(Survived 15 yrs)	0.349	(0.477)	0.496	(0.500)	(0.000)	0.213
𝟙(Survived 20 yrs)	0.233	(0.423)	0.362	(0.481)	(0.000)	0.202
𝟙(Survived 25 yrs)	0.163	(0.369)	0.270	(0.444)	(0.000)	0.186
Years to Mature	3.157	(7.076)	4.316	(8.788)	(0.000)	0.103
Avg Num(Employees)	2.493	(6.721)	2.502	(9.797)	(0.962)	0.001
Max Num(Employees)	4.192	(16.653)	4.314	(21.057)	(0.789)	0.005
𝟙(Have Female Cohort Cofounder)	0.021	(0.144)	0.015	(0.121)	(0.723)	-0.034
𝟙(Have Other Female Cofounder)	0.162	(0.368)	0.191	(0.396)	(0.525)	0.054
Observations	20,132		1,505			

Table A2: Coeducation and Entrepreneurship, with Alternative Event Windows, Control Groups, and Weighting

We present robustness checks for estimates from the following model:

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

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. $Treat_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $Post_{sc,e}$ 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 entry-year fixed effects, capturing overall time trends in founder rates across university entering cohorts. 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 the 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 11). 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. Lower-order interaction terms are omitted from the table for brevity. Standard errors clustered by event are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: [-10,+10], early- and lately-treated groups as control, no weights			
	Founder_Narrow	Founder_Medium	Founder_Wide
	(1)	(2)	(3)
Treat × Post	0.005**	0.011***	0.013***
	(0.002)	(0.003)	(0.004)
School FE	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes
Event FE	Yes	Yes	Yes
Observations	3,074,393	3,074,393	3,074,393
R^2	0.007	0.023	0.023
Panel B: [-10,+10], lately-treated group as control, no weights			
	Founder_Narrow	Founder_Medium	Founder_Wide
	(1)	(2)	(3)
Treat × Post	0.013***	0.013***	0.014***
	(0.002)	(0.003)	(0.004)
School FE	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes
Event FE	Yes	Yes	Yes
Observations	550,896	550,896	550,896
R^2	0.008	0.016	0.016

Panel C: [-10,+10], early- and lately-treated groups as control, with weights

	Founder_Narrow	Founder_Medium	Founder_Wide
	(1)	(2)	(3)
Treat × Post	0.005**	0.010***	0.013***
	(0.002)	(0.003)	(0.004)
School FE	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes
Event FE	Yes	Yes	Yes
Observations	3,074,393	3,074,393	3,074,393
R^2	0.007	0.019	0.019

Panel D: [-5,+5], early- and lately-treated groups as control, no weights

	Founder_Narrow	Founder_Medium	Founder_Wide
	(1)	(2)	(3)
Treat × Post	0.003	0.009***	0.009**
	(0.002)	(0.003)	(0.004)
School FE	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes
Event FE	Yes	Yes	Yes
Observations	1,944,144	1,944,144	1,944,144
R^2	0.006	0.019	0.019

Panel E: [-5,+5], early- and lately-treated groups as control, with weights

	Founder_Narrow	Founder_Medium	Founder_Wide
	(1)	(2)	(3)
Treat × Post	0.004*	0.011***	0.011**
	(0.002)	(0.003)	(0.004)
School FE	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes
Event FE	Yes	Yes	Yes
Observations	1,944,144	1,944,144	1,944,144
R^2	0.006	0.017	0.016

Table A3: 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 following specifications 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. $Treat_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $Post_{sc,e}$ 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 entry-year fixed effects, capturing overall time trends in founder rates across university entering cohorts. 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-entry year 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. Column (1) shows results from Eq.(1). Column (2) shows results from Eq.(2). Column (3) shows results from Eq.(3). Lower-order interaction terms are omitted from the table for brevity or absorbed by fixed effects. Standard errors clustered by event are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	$\mathbb{1}$(Founder within 18 Yrs of Graduation)		
	(1)	(2)	(3)
Treat × Post	0.006**	0.007***	0.007**
	(0.002)	(0.002)	(0.003)
School FE	Yes	Yes	
Entry Year FE	Yes	Yes	
Event FE		Yes	
Event × School FE			Yes
Event × Entry Year FE			Yes
Observations	2,293,660	2,293,660	2,293,634
R^2	0.005	0.005	0.006

Table A4: Coeducation, Gender Inequality, and Entrepreneurship, for Events after 1960

To examine the heterogeneous impact of coeducation on entrepreneurship across measures of gender equality, we estimate the following model:

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

The dependent variable is whether individual i from cohort c, university s, in event e has ever become an entrepreneur. $Treat_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $Post_{sc,e}$ 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 entry-year fixed effects, capturing overall time trends in founder rates across university entering cohorts. θ_s are university fixed effects that control for time-invariant university characteristics such as 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 entrepreneurship rate. *Gender Inequality in Income (Education)* measures the gap between male and female income and bachelor rates, respectively, at the county level, calculated using PUMS Data. Columns (1) and (2) report results for sub-samples where *Gender Inequality in Income* is above and below the median, respectively. Similarly, Columns (3) and (4) report results for sub-samples where *Gender Inequality in Education* 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. In addition, the sample size decreases compared to Table 2 because some areas have missing county identifiers due to confidentiality requirements in the PUMS data. Lower-order interaction terms are omitted from the table for brevity. Standard errors clustered by event are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	𝟙(Founder)			
	by Gender Inequality in Income		by Gender Inequality in Education	
	Above Median	Below Median	Above Median	Below Median
	(1)	(2)	(3)	(4)
Treat × Post	0.009**	0.012**	0.002	0.016***
	(0.004)	(0.005)	(0.005)	(0.005)
School FE	Yes	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes	Yes
Event FE	Yes	Yes	Yes	Yes
Observations	738,816	928,775	777,730	889,861
R^2	0.021	0.019	0.023	0.019

Table A5: Coeducation, Gender Inequality, and Entrepreneurship, with Scaled Gender Inequality Measures

To examine the heterogeneous impact of coeducation on entrepreneurship across measures of gender equality, we estimate the following model:

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

The dependent variable is whether individual i from cohort c, university s, in event e has ever become an entrepreneur. $Treat_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $Post_{sc,e}$ 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 entry-year fixed effects, capturing overall time trends in founder rates across university entering cohorts. θ_s are university fixed effects that control for time-invariant university characteristics such as 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 entrepreneurship rate. *Gender Inequality in Income (Education)* measures the county-level gap between male and female income (bachelor rates), scaled by the average income (bachelor rate), calculated using PUMS Data. Columns (1) and (2) report results for sub-samples where *Gender Inequality in Income* is above and below the median, respectively. Similarly, Columns (3) and (4) report results for sub-samples where *Gender Inequality in Education* is above and below the median. In addition, the sample size decreases compared to Table 2 because some areas have missing county identifiers due to confidentiality requirements in the PUMS data. Lower-order interaction terms are omitted from the table for brevity. Standard errors clustered by event are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	𝟙(Founder)			
	by Gender Inequality in Income (Scaled)		by Gender Inequality in Education (Scaled)	
	Above Median	Below Median	Above Median	Below Median
	(1)	(2)	(3)	(4)
Treat × Post	0.008**	0.016***	0.000	0.020***
	(0.004)	(0.006)	(0.004)	(0.006)
School FE	Yes	Yes	Yes	Yes
Entry Year FE	Yes	Yes	Yes	Yes
Event FE	Yes	Yes	Yes	Yes
Observations	794,029	966,981	864,308	896,702
R^2	0.027	0.019	0.024	0.022

Table A6: Coeducation, Downside Protection, and Entrepreneurship, for Events after 1960

To examine the heterogeneous impact of coeducation on entrepreneurship across variations in downside protection, we estimate the following specification in sub-samples split at the median of three measures: (i) *% Male-Only BA+ Households*, (ii) *Female College Income Premium*, and (iii) *Male Income*:

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

% Male-Only BA+ Households measures the county-level share of families in which only the male holds a bachelor's degree or above, calculated using PUMS Data. *Female College Income Premium* measures the county-level income gap between college-educated and non-college-educated married women, calculated using PUMS Data. *Male Income* refers to the county average male income from PUMS. The dependent variable is whether individual i from cohort c, university s, in event e has ever become an entrepreneur. $Treat_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $Post_{sc,e}$ 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 entry-year fixed effects, capturing overall time trends in founder rates across university entering cohorts. 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 entrepreneurship rate. 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. In addition, the sample size further reduces compared to Table 2 because some areas have missing county identifiers due to confidentiality requirements in the PUMS data. Lower-order interaction terms are omitted from the table for brevity. Standard errors clustered by event are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	𝟙(Founder)					
	by %Male-Only BA+ Households		by Female College Income Premium		by Male Income	
	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median
	(1)	(2)	(3)	(4)	(5)	(6)
Treat × Post	0.012**	0.008*	0.013***	0.008	0.003	0.018***
	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)
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
Observations	949,286	718,305	817,748	849,843	829,265	838,326
R^2	0.021	0.019	0.021	0.021	0.023	0.018

Table A7: What Predicts the Timing of Turning to Coeducation

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

$$\text{Year of Coeducation}_s = \alpha + \beta_0 \times \text{University Characteristics}_s + \varepsilon_s, \tag{14}$$

$$\text{Year of Coeducation}_s = \alpha + \beta_0 \times \text{Local Economic Characteristics}_s + \varepsilon_s. \tag{15}$$

The dependent variable, *Year of Coeducation*$_s$, represents the year in which a university transitions to coeducation. The explanatory variables include *University Characteristics*$_s$, such as the year of establishment, indicators for public or private (benchmark) status and religious affiliation, geographical location, and undergraduate enrollment. Religious affiliation is categorized in two ways: religious and non-religious (benchmark), Catholic and non-Catholic (benchmark). Undergraduate enrollment comes from the earliest HEGIS records in 1969. *Local Economic Characteristics*$_s$ include state-level coeducational enrollment share, and county-level measures such as income, population, income growth, population growth, and income per capita. The coeducational enrollment share is defined as the share of students enrolled in coeducational institutions within the state in 1969, the earliest year available from HEGIS. 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. Panel A reports estimates from Eq. (14) using university characteristics, and Panel B reports estimates from Eq. (15) using local characteristics. Notably, Panel A Column (6) and Panel B Columns (1) to (4) include only universities that transitioned to coeducation no earlier than 1969, as these variables rely on 1969 data. Similarly, Panel B Columns (5) and (6) restrict the sample to universities that transitioned no earlier than 1970, and the sample sizes decrease due to this time restriction. Panel B Column (7) restricts the sample to universities that transitioned no earlier than 1960, and the sample size further shrinks because some areas lack county identifiers for confidentiality requirements in the PUMS data. Undergraduate enrollment, income, population, income growth, population growth, and income per capita are standardized.

Panel A. University Characteristics

	Year of Coeducation					
	(1)	(2)	(3)	(4)	(5)	(6)
Year of Establishment	-0.028					
	(0.032)					
𝟙(Public University)		-21.111***				
		(5.782)				
𝟙(Religious)			7.130**			
			(3.479)			
𝟙(Catholic)				7.660**		
				(3.679)		
𝟙(Midwest)					-0.347	
					(0.717)	
Num(Undergraduates)						5.223
						(5.158)
Observations	91	91	91	91	91	53
R^2	0.009	0.130	0.045	0.046	0.011	0.005

Panel B. Local Economic Characteristics

	Year of Coeducation						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coed Enroll Share	-0.743						
	(6.907)						
ln(Income) from BEA		0.270					
		(0.614)					
ln(Population) from BEA			0.235				
			(0.611)				
ln(Income Per Capita) from BEA				0.392			
				(0.626)			
Income Growth from BEA					-0.482		
					(0.869)		
Population Growth from BEA						-1.131	
						(0.903)	
ln(Income) from PUMS							-0.098
							(0.698)
Observations	53	53	53	53	41	41	43
R^2	0.000	0.004	0.003	0.008	0.008	0.039	0.000

Table A8: Coeducation and Entrepreneurship, with Additional Time-Varying Fixed Effects for Public vs. Private University and Religious or Catholic Affiliation

To estimate the impact of coeducation on students' probability of becoming an entrepreneur, we estimate the following specifications with varying fixed effects: (i) public university-by-entry year, (ii) religious affiliation–by–entry year, and (iii) Catholic affiliation–by–entry year fixed effects (private, non-religious, and non-Catholic universities serving as benchmarks, respectively).

$$\mathbb{1}(\text{Founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \theta_s + \tau_{type,c} + \rho_e + \varepsilon_{sci,e}, \tag{16}$$

The dependent variable is a dummy variable indicating whether in event e, the individual i from university s, cohort c, becomes a founder. $Treat_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $Post_{sc,e}$ is a dummy variable that takes the value of one for all cohorts entering the university after the focal treatment year in event e. 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. We additionally control for time-varying institutional heterogeneity by including fixed effects interacted with entry year: $\tau_{\text{public},c}$ for public universities, $\tau_{\text{religious},c}$ for religious affiliation, and $\tau_{\text{catholic},c}$ for Catholic affiliation. These specifications allow for differential trends across university types that may confound treatment effects. β estimates the impact of entering the university after its transition to coeducation on the individual rate of entrepreneurship. Columns (1) adds public-by-entry year fixed effects; Columns (2) includes religious affiliation-by-entry year fixed effects; and Columns (3) adds Catholic affiliation-by-entry year fixed effects. Lower-order interaction terms are omitted from the table for brevity. Standard errors clustered by event are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	$\mathbb{1}$(Founder)		
	(1)	(2)	(3)
Treat × Post	0.009***	0.011***	0.011***
	(0.003)	(0.003)	(0.003)
School FE	Yes	Yes	Yes
Event FE	Yes	Yes	Yes
Public × Entry Year FE	Yes		
Religious × Entry Year FE		Yes	
Catholic × Entry Year FE			Yes
Observations	3,074,393	3,074,393	3,074,393
R^2	0.023	0.023	0.023

Table A9: Coeducation and Entrepreneurship, with Additional Time-Varying Fixed Effects for State and Economic Growth Deciles

To estimate the impact of coeducation on students' probability of becoming an entrepreneur, we estimate the following specifications with varying fixed effects: (i) state-by-entry year, (ii) income growth decile-by-entry year, (iii) population growth decile-by-entry year fixed effects:

$$\mathbb{1}(\text{Founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \theta_s + \rho_e + \tau_{state,c} + \varepsilon_{sci,e}, \quad (17)$$

$$\mathbb{1}(\text{Founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \theta_s + \rho_e + \tau_{state,c} + \zeta_{incg_{dec},c} + \varepsilon_{sci,e}, \quad (18)$$

$$\mathbb{1}(\text{Founder}_{sci,e}) = \alpha + \beta_0 \times \text{Treat}_{s,e} \times \text{Post}_{sc,e} + \beta_1 \times \text{Treat}_{s,e} + \beta_2 \times \text{Post}_{sc,e} + \theta_s + \rho_e + \tau_{state,c} + \kappa_{popg_{dec},c} + \varepsilon_{sci,e}, \quad (19)$$

The dependent variable is a dummy variable indicating whether in event e, the individual i from university s, cohort c, ever becomes a founder. $Treat_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $Post_{sc,e}$ is a dummy variable that takes the value of one for all cohorts entering the university after the focal treatment year in event e. 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. $\tau_{state,c}$ denotes state-by-entry year fixed effects, which absorb differences in economic trends across states. In addition, $\zeta_{incg_{dec},c}$ and $\kappa_{popg_{dec},c}$ denote income growth decile-by-entry year and population growth decile-by-entry year fixed effects, respectively. We group the counties into ten deciles based on the 1970 income growths and population growths from the Bureau of Economic Analysis. These interactive fixed effects control for potential time trends that may be driven by local income or population dynamics. 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.(17). Columns (2) shows results from Eq.(18). Columns (3) shows results from Eq.(19). Lower-order interaction terms are omitted from the table for brevity. Standard errors clustered by event are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	$\mathbb{1}$(Founder)		
	(1)	(2)	(3)
Treat × Post	0.009***	0.011**	0.009**
	(0.003)	(0.004)	(0.004)
School FE	Yes	Yes	Yes
Event FE	Yes	Yes	Yes
State × Entry Year FE	Yes	Yes	Yes
Income Growth Decile × Entry Year FE		Yes	
Pop Growth Decile × Entry Year FE			Yes
Observations	3,074,393	3,074,393	3,074,367
R^2	0.023	0.023	0.024

Table A10: Matched Sample Balance Table

We report balance statistics for university characteristics between treated and control groups in the year prior to the focal treatment. For each treated university, we identify the closest-matched controls from all U.S. baccalaureate degree–granting institutions (excluding women's colleges, since our analysis requires counterfactuals of male student outcomes). We find the closest-matched universities based on two sets of variables: (i) university characteristics, including total revenue, average SAT score, and the student entrepreneurship rate; and (ii) local economic conditions, including income, population, income per capita, income growth, and population growth, as described in Section 6.2. To minimize the influence of treatment, all matching variables are measured in the year prior to a university's transition, or use the earliest available observation if data are missing. We further report balance statistic for variables not directly used in matching, including number of undergraduates, faculties, tuition, expenditure on instruction and total expenditure. The number of treated schools decreases due to missing values in the variables required for matching. Panels A, B, and C report for samples with five, three, and one closest-matched universities, respectively. Within the top five matches, Panels D and E report for samples with the geographically nearest one and three universities within the top five matches, respectively. Revenue from tuition, total revenues, expenditure on instruction, total expenditure, and population are reported in millions and income is reported in billions.

Panel A: 5 Closest Match

	Treat = 0		Treat = 1			
	Mean	SD	Mean	SD	P-value	Std. Diff.
Total Revenue	69.1	(122.8)	63.7	(120.9)	(0.756)	-0.032
SAT Score	1,164.9	(125.8)	1,185.6	(147.2)	(0.268)	0.107
Founder Medium	0.171	(0.087)	0.173	(0.088)	(0.849)	0.019
Income Growth	0.073	(0.024)	0.073	(0.025)	(0.886)	-0.014
Population Growth	0.008	(0.011)	0.008	(0.011)	(0.620)	-0.049
Income Per Capita	4,502.3	(2,454.9)	4,557.9	(2,185.3)	(0.873)	0.017
Population	0.625	(1.217)	0.713	(1.285)	(0.619)	0.050
Income	3.299	(7.719)	3.620	(7.902)	(0.773)	0.029
Num(Undergraduates)	3,684.3	(3,719.0)	2,680.9	(3,588.1)	(0.150)	-0.194
Num(Faculties)	229.8	(243.8)	185.7	(200.2)	(0.203)	-0.140
Revenue from Tuition	21.3	(27.2)	16.0	(14.1)	(0.271)	-0.173
Expenditure on Instruction	6.6	(9.7)	4.5	(6.2)	(0.215)	-0.187
Total Expenditure	78.7	(129.4)	76.0	(147.6)	(0.913)	-0.014
Observations	290		58			

Panel B: 3 Closest Match

	Treat = 0		Treat = 1			
	Mean	SD	Mean	SD	P-value	Std. Diff.
Total Revenue	75.1	(140.2)	63.6	(120.6)	(0.575)	-0.062
SAT Score	1,169.2	(126.3)	1,185.6	(147.2)	(0.411)	0.085
Founder Medium	0.171	(0.088)	0.173	(0.088)	(0.889)	0.015
Income Growth	0.075	(0.025)	0.073	(0.025)	(0.618)	-0.054
Population Growth	0.008	(0.011)	0.008	(0.011)	(0.771)	-0.031
Income Per Capita	4,475.8	(2,674.4)	4,557.9	(2,185.3)	(0.833)	0.024
Population	0.610	(1.250)	0.713	(1.285)	(0.591)	0.057
Income	3.217	(7.826)	3.620	(7.902)	(0.735)	0.036
Num(Undergraduates)	3,690.9	(3,735.8)	2,680.9	(3,588.1)	(0.170)	-0.195
Num(Faculties)	239.9	(268.6)	185.7	(200.2)	(0.166)	-0.162
Revenue from Tuition	22.0	(29.3)	16.0	(14.1)	(0.259)	-0.182
Expenditure on Instruction	6.8	(10.3)	4.5	(6.2)	(0.213)	-0.194
Total Expenditure	86.9	(151.9)	76.0	(147.6)	(0.716)	-0.052
Observations	174		58			

Panel C: 1 Closest Match

	Treat = 0		Treat = 1			
	Mean	SD	Mean	SD	P-value	Std. Diff.
Total Revenue	84.8	(150.0)	63.6	(120.6)	(0.404)	-0.110
SAT Score	1,172.6	(128.0)	1,185.6	(147.2)	(0.611)	0.067
Founder Medium	0.173	(0.087)	0.173	(0.088)	(0.970)	0.005
Income Growth	0.074	(0.027)	0.073	(0.025)	(0.727)	-0.046
Population Growth	0.008	(0.012)	0.008	(0.011)	(0.879)	-0.020
Income Per Capita	4,341.6	(2,260.0)	4,558.0	(2,185.3)	(0.601)	0.069
Population	0.641	(1.300)	0.713	(1.285)	(0.764)	0.039
Income	3.213	(7.950)	3.620	(7.902)	(0.782)	0.036
Num(Undergraduates)	4,036.0	(4,070.1)	2,680.9	(3,588.1)	(0.150)	-0.250
Num(Faculties)	248.1	(279.8)	185.7	(200.2)	(0.177)	-0.182
Revenue from Tuition	23.5	(35.8)	16.0	(14.1)	(0.264)	-0.193
Expenditure on Instruction	8.1	(13.4)	4.5	(6.2)	(0.163)	-0.242
Total Expenditure	89.7	(148.3)	76.0	(147.6)	(0.703)	-0.066
Observations	58		58			

Panel D: 3 Geographically Closest Match

	Treat = 0		Treat = 1			
	Mean	SD	Mean	SD	P-value	Std. Diff.
Total Revenue	64.6	(119.1)	63.6	(120.6)	(0.959)	-0.006
SAT Score	1,165.6	(129.1)	1,185.6	(147.2)	(0.324)	0.102
Founder Medium	0.171	(0.088)	0.173	(0.088)	(0.887)	0.015
Income Growth	0.073	(0.025)	0.073	(0.025)	(0.922)	-0.011
Population Growth	0.008	(0.011)	0.008	(0.011)	(0.676)	-0.044
Income Per Capita	4,524.9	(2,312.5)	4,557.9	(2,185.3)	(0.924)	0.010
Population	0.661	(1.261)	0.713	(1.285)	(0.788)	0.029
Income	3.450	(7.860)	3.620	(7.902)	(0.886)	0.015
Num(Undergraduates)	3,515.4	(3,744.1)	2,680.9	(3,588.1)	(0.258)	-0.161
Num(Faculties)	222.3	(240.9)	185.7	(200.2)	(0.304)	-0.117
Revenue from Tuition	18.7	(23.8)	16.0	(14.1)	(0.540)	-0.096
Expenditure on Instruction	5.7	(7.9)	4.5	(6.2)	(0.439)	-0.115
Total Expenditure	65.3	(107.6)	76.0	(147.6)	(0.652)	0.058
Observations	174		58			

Panel E: 1 Geographically Closest Match

	Treat = 0		Treat = 1			
	Mean	SD	Mean	SD	P-value	Std. Diff.
Total Revenue	58.4	(119.2)	63.6	(120.6)	(0.816)	0.031
SAT Score	1,172.3	(142.9)	1,185.6	(147.2)	(0.623)	0.065
Founder Medium	0.176	(0.090)	0.173	(0.088)	(0.859)	-0.023
Income Growth	0.071	(0.022)	0.073	(0.025)	(0.742)	0.043
Population Growth	0.007	(0.010)	0.008	(0.011)	(0.770)	0.038
Income Per Capita	4,548.8	(2,186.9)	4,557.9	(2,185.3)	(0.982)	0.003
Population	0.767	(1.307)	0.713	(1.285)	(0.822)	-0.030
Income	3.867	(8.012)	3.620	(7.902)	(0.868)	-0.022
Num(Undergraduates)	2,995.6	(4,400.2)	2,680.9	(3,588.1)	(0.748)	-0.055
Num(Faculties)	192.6	(235.6)	185.7	(200.2)	(0.867)	-0.022
Revenue from Tuition	18.9	(30.2)	16.0	(14.1)	(0.615)	-0.087
Expenditure on Instruction	4.8	(8.0)	4.5	(6.2)	(0.849)	-0.033
Total Expenditure	61.137	(118.7)	76.0	(147.6)	(0.649)	0.078
Observations	58		58			

Table A11: Coeducation and Entrepreneurship, Matched Sample Analysis

To address concerns that unobserved institutional and local economic factors may bias our estimates, we identify the closest-matched control for each treated university as described in Section 6.2, and estimate the following specification:

$$\mathbb{1}(\text{Founder}_{sci}) = \alpha + \beta_0 \times \text{Treat}_s \times \text{Post}_{sc,group} + \beta_1 \times \text{Treat}_s + \beta_2 \times \text{Post}_{sc,group} + \theta_s + \tau_{group,c} + \varepsilon_{sci,e} \quad (20)$$

The dependent variable is a dummy variable indicating whether the individual i from university s, cohort c, ever becomes a founder. $Treat_s$ is a dummy variable that takes the value of one for universities in the treatment group. $Post_{sc,group}$ is a dummy variable that takes the value of one for all cohorts entering the university after the focal treatment year in the same matched group, which consists of the treated university and its paired control universities. θ_s are university fixed effects, which control for time-invariant university characteristics such as its location and selectivity. $\tau_{group,c}$ denotes matched group-by-entry year fixed effects, which absorb potential time trends in each matched group. β estimates the impact of entering the university after its transition to coeducation on the individual rate of entrepreneurship. We report results using samples based on the five, three, and one closest matches, allowing controls to be used multiple times. To ensure local comparability, we also construct matched samples by selecting the geographically nearest three and one universities from the top five closest matches. Balance statistics confirming comparability across a wide range of characteristics are provided in Appendix Table A10. Columns (1) to (3) shows results for matched samples with fice, three, one closest universities. Columns (4) and (5) show results for matched samples with the geographically nearest three, one universities within the top five matches. Lower-order interaction terms are absorbed by fixed effects. Standard errors clustered by university are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	$\mathbb{1}$(Founder)				
	5 Closest Match	3 Closest Match	1 Closest Match	3 Geographically Closest	1 Geographically Closest
	(1)	(2)	(3)	(4)	(5)
Treat × Post	0.011***	0.012***	0.014***	0.011***	0.013***
	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)
School FE	Yes	Yes	Yes	Yes	Yes
Matched Group × Entry Year FE	Yes	Yes	Yes	Yes	Yes
Observations	965,071	587,879	217,848	586,717	194,787
R^2	0.016	0.016	0.018	0.016	0.017

Table A12: Coeducation and Entrepreneurship, at Public Universities

To estimate the impact of turning coeducation on students' probability of becoming an entrepreneur in public universities, we estimate the following specifications 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. $Treat_{s,e}$ is a dummy variable that takes the value of one for universities in the treatment group for event e. $Post_{sc,e}$ 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 entry-year fixed effects, capturing overall time trends in founder rates across university entering cohorts. 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-entry year 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 to student cohorts across events. β estimates the impact of entering the university after its transition to coeducation on the individual rate of entrepreneurship. Column (1) shows results from Eq.(1). Column (2) shows results from Eq.(2). Column (3) shows results from Eq.(3). Lower-order interaction terms are omitted from the table for brevity or absorbed by fixed effects. Standard errors clustered by event are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	$\mathbb{1}$(Founder)		
	(1)	(2)	(3)
Treat × Post	0.018***	0.019***	0.020***
	(0.003)	(0.003)	(0.003)
School FE	Yes	Yes	
Entry Year FE	Yes	Yes	
Event FE		Yes	
School × Event FE			Yes
Entry Year × Event FE			Yes
Observations	834,224	834,224	834,215
R^2	0.021	0.021	0.021