Colleges, Assortative Mating, and Intergenerational Mobility in the United States

Ran Abramitzky, Santiago Pérez, and Joseph Price*

March 23, 2026

Abstract

Do colleges shape marriage markets and, in doing so, reinforce intergenerational inequality? We study this question using newly digitized records from 63 coeducational U.S. colleges between 1915 and 1949, linked to population censuses. About 5% of men and 8% of women marry someone from their own college, rates over 22 times higher than what would occur under random mating to a college-educated spouse. Exploiting within-college variation in cohort gender composition and variation in the amount of time students overlap on campus, we show that this homogamy reflects a causal effect. Even within the same institution, same-college marriages are strongly assortative by parental income, particularly at elite colleges, implying that colleges amplify socioeconomic sorting. These marriage-market effects have intergenerational consequences: children of same-college couples are more likely to attend selective colleges themselves. Colleges thus reinforce inequality not only through education, but also through assortative mating.

*Abramitzky: Stanford University and NBER (email: ranabr@stanford.edu); Pérez: University of California - Davis and NBER (email: seperez@ucdavis.edu), Price: Brigham Young University and NBER (email: joe_price@byu.edu). Abramitzky, Pérez, and Price gratefully acknowledge funding from the National Science Foundation and the Russell Sage Foundation.

1 Introduction

A large body of research has examined the relationship between family background and college attendance, documenting that students from higher-income families tend to enroll in more selective institutions.[1] This work shows that when access to college is stratified by parental income, higher education may reinforce the persistence of inequality rather than promote upward mobility. This literature primarily emphasizes an individual-level mechanism: children from high-income families attend more selective colleges, become high-earning adults, and their children then gain access to similarly selective colleges, perpetuating advantage across generations.

At the same time, colleges may shape intergenerational mobility through a second, less studied channel: by serving as marriage markets (Kirkebøen, Leuven and Mogstad, 2021). By increasing the likelihood that students meet and marry partners from similar socioeconomic backgrounds, colleges may facilitate assortative mating, thereby reinforcing the persistence of economic status across generations (Chadwick and Solon, 2002).

The central question motivating this paper is whether colleges shape who marries whom and, in doing so, contribute to the persistence of inequality across generations. To answer this question, we examine three sub-questions. First, do colleges function as marriage markets—that is, does attending the same institution causally increase the likelihood that two individuals marry one another, and if so, by how much? Second, do they reinforce social stratification by facilitating matches between students from similar socioeconomic backgrounds, or do they instead operate as "marriage melting pots," fostering matches across class lines? Third, what are the consequences of same-college marriages for the human capital of the next generation?

Our analysis draws on newly assembled data on students at 63 co-educational U.S. colleges spanning 1915–1949. This historical setting has two key features that make it particularly well suited for studying the intergenerational consequences of same-college marriages:

The first feature is that we are able to observe students' socio-economic backgrounds prior to college, their own adult post-college outcomes (including whom they marry), and the outcomes of their *children*. Although certain datasets capture portions of this information (e.g., Chetty et al. (2020)), to our knowledge there is no U.S. dataset that simultaneously includes all of these components.

To assemble our data, we digitize 2 million student records from 63 U.S. colleges spanning 1915-1949. We chose these colleges to represent the selective four-year segment of U.S. higher education, combining elite private institutions with a broad set of flagship and land-grant public universities. The sample includes nine elite private institutions that were already coeducational by 1915—four Ivy Plus universities (Cornell, Penn, Chicago, and Stanford) and five Little Ivies (Bates, Colby, Tufts, Swarthmore, and Middlebury)—as well as 54 public universities representing most U.S. states through their flagship or land-grant institution.[2] We then link these students to decennial U.S. censuses before and after their college attendance. By doing so, we observe students' socioeconomic backgrounds prior to attending college as well as their post-college outcomes, including whether they are married to a spouse from the same college. For students who attended college in the earlier part of our sample, we can also observe if their *children* attended, prior to

[1] See, for instance, Bowen and Bok (1998); Bailey and Dynarski (2011); Chetty et al. (2020); Chetty, Deming and Friedman (2023); Abramitzky et al. (2024*a*); Bleemer and Quincy (2025); Hoxby and Avery (2012).

[2] Together, these colleges account for 16% of *all* college students in the 1897–1931 birth cohorts and about a third of those enrolled in public universities.

1966, one of the colleges in an extended sample of 95 colleges (those colleges in the current paper plus those in Abramitzky et al. (2024*a*)), allowing us to study the longer-run intergenerational consequences of same-college marriages.

The second convenient feature is that, unlike today, where many colleges in the U.S. maintain stable gender ratios, our period features substantial variation in the gender composition of incoming cohorts.[3] We exploit this within-college variation to examine how exposure to opposite-sex peers affects the likelihood of marrying someone from the same college and how this, in turn, shapes assortative mating by social class and human capital in the next generation.

We begin by documenting the prevalence of same-college marriages. Across all colleges in our data, about 5% of men and 8% of women marry someone from their own college. For men, these marriage rates are 22 times higher than what would occur under random mating to other college-educated women, and for women, they are 21 times higher. These findings imply that conventional measures (based on whether individuals marry others with similar educational attainment) substantially underestimate the true extent of educational homogamy. For example, college-educated men are "only" 4 times more likely to marry another college-educated person than someone from the general population. Finally, comparing our estimates to modern survey data (Rosenfeld, Thomas and Hausen, 2024), we find that the share of college-educated individuals who marry a spouse they met during college has remained broadly stable over the last century.

We then explore which mechanisms could account for this strong same-college homogamy, focusing on two main explanations. First, it may simply reflect *selection*: individuals who attend the same college may share characteristics, such as socioeconomic background or academic ability, that make them both more likely to attend that college and to marry each other. Second, it may reflect a *causal* effect, whereby attending the same institution causally increases the likelihood that two individuals marry.

We implement two empirical exercises that show that the observed homogamy reflects a causal effect of colleges. First, we exploit within-college across cohorts variation in gender ratios to assess whether students who are exposed to a larger fraction of opposite-sex students are more likely to marry someone from their own college. The identifying assumption in this strategy is that, conditional on gender, college, and entry-cohort fixed effects, students who start college in years with different proportions of opposite-sex classmates would have similar marriage market outcomes. Consistent with this assumption, we find no correlation between the share of opposite-sex students in the cohort and students' background characteristics such as parental income rank, parental occupations, or race.

We find that students exposed to a larger fraction of opposite-sex students are significantly more likely to marry within their college: Increasing exposure to opposite-sex students from 25% to 50% raises the probability of same-college marriages by 2.5 percentage points, a large increase relative to a mean of 5.5%. These effects are relatively similar by gender, in private or public colleges, in larger or smaller cohorts, and across the parental income distribution. This finding makes it unlikely that the observed homogamy is simply due to selection: If that were the case, greater exposure to potential partners should *not* increase the likelihood of marrying someone from the same college.

[3]Focusing on the same set of colleges as in our main sample but in the 2000-2022 period, the standard deviation of the female share by cohort is only 5%. By contrast, the standard deviation in our setting is 15% in the raw data and 10% after netting out college and entry cohort fixed effects.

While this evidence rules out a purely selection-based explanation for the observed college homogamy, it does not tell us *how much* of the observed homogamy reflects selection versus a causal effect of colleges. To quantify the causal channel, we implement a second research design comparing the marriage propensity of *pairs* of individuals who overlapped in college to those who narrowly missed overlapping. Specifically, we estimate dyadic regressions where the outcome is an indicator for whether a pair of individuals is married to each other, and the main explanatory variables capture whether the two individuals attended the same college at the same time. Because individuals who start college in the same year may be more likely to marry due to assortative mating on the basis of age, we construct dyads using all possible pairs of students (including pairs of students who attended different colleges). This strategy allows us to include fixed effects for each pair of entry cohorts, thus absorbing age-related marriage patterns.

We find that individuals who overlapped in college are significantly more likely to marry each other than those who attended the same college but did not overlap. If we assume that students who started college four years or more apart experience zero on-campus overlap, our estimates imply that 77% of the observed same-college homogamy can be attributed to causal on-campus exposure effects.

After establishing that college homogamy reflects a causal effect of attending the same institution, we next examine its implications for homogamy by *social class* (as captured by parental income quintiles). Within-college social class homogamy reflects two distinct forces. The first is a *compositional channel*: because colleges are segregated by parental income, same-college marriages would mechanically involve individuals from similar backgrounds even if mating *within* colleges were random. In the extreme, if all students at a college came from the same income group, social-class homogamy would arise purely from this composition. The second is a *sorting channel*: within colleges, students may be more likely to marry other students from similar backgrounds.

We find that same-college couples are strongly homogamous on the basis of parental income, and that this homogamy is not purely due to the compositional channel: Even within the same institution, students are more likely to marry students from similar social backgrounds than what would be predicted if mating were random. For instance, because the fraction of bottom 20% students at Ivy Plus colleges is only 3%, random mating would make it unlikely that such students marry someone who is also from the bottom 20%. Yet, 19% of bottom 20% students in the Ivy Plus who are married to a college classmate are married to someone who is also from the bottom 20%.

Homogamy rates are particularly high at Ivy Plus colleges, where 57% of same-college marriages are between individuals from the same parental income quintile (compared to 32% if mating were random). Homogamy rates are more comparable (39%) than those in the general population in the case of the non-top public colleges. This lower homogamy is partly driven by the more balanced social class composition of these colleges.

Within-college marriages are on average *more* homogamous than those in the general population and also than those marriages formed by *other* students in the same institution. For instance, while among all marriages 38% include partners who were both raised at the same quintile of the parental income distribution, that proportion is 47% among *same-college* marriages.

Yet, we observe notable heterogeneity by social background in the relative degree of social class homogamy: While individuals in the top 20% of the parental income distribution are more likely to marry another person from the top 20% when they marry within their college than when

they marry outside college, the opposite is true for individuals from the bottom 80% of the distribution.

One potential explanation for the stronger social class homogamy of same-college marriages is that the partners that students meet in college tend to be more similar to them in social class than those they would meet through other channels. An alternative explanation is that individuals who end up marrying a college peer may possess unobserved characteristics that independently increase their propensity to marry within their social class.

To help disentangle these two explanations, we again exploit within-college across cohorts variation in the gender ratio of incoming classes. If, for reasons outside students' control, they become more likely to marry within their college, does this in turn make them more likely to marry within their *social class*?

We find that, on average, higher exposure to opposite-sex students does not increase the probability of marrying someone from the same parental income quintile. However, this average masks important heterogeneity. In private colleges, a higher share of opposite-sex peers *increases* the probability of socially homogamous marriages for students in the top 20% of the distribution but *lowers* it for other students. The explanation for this pattern is straightforward: Because there is a high concentration of top 20% students at the private colleges in our sample, increasing the probability of same-college marriages increases the chances that top 20% students marry within their group but decreases the chances that other students do so. In contrast, the effects are weaker in public colleges. A likely explanation for the private-public contrast is that, in our sample, public colleges have a more balanced social-class mix than the private ones. Hence, increasing the probability of same-college marriages leads to marital matches that are not that dissimilar to those that would occur outside of college.

In the final section, we investigate the intergenerational consequences of same-college marriages. Throughout this section, we begin with descriptive comparisons between same-college and other marriages, and then use within-college variation in exposure to opposite-sex peers to assess how changes in likelihood of same-college marriages relate to children outcomes.

We first investigate whether same-college couples have *children* who are themselves more educated, focusing on whether the children of students themselves attend one of the 95 colleges in the data in Abramitzky et al. (2024*a*) and this paper. These colleges account for 19% of all individuals with at least one year of college in the 1915 to 1950 birth cohorts, and an even higher share of those attending the most prestigious colleges. We focus on this measure (rather than educational attainment from the census) because the 1950 census is the last census that is currently available and relatively few of the children of our 1915-1949 original students had reached young adulthood by 1950. Focusing on this measure enables us to observe children's educational outcomes up to 1966.[4]

We find that children of same-college couples are more likely to attend their parents' college than children for whom only one parent attended. Furthermore, they are also more likely to attend *other* colleges in our data, consistent with an overall increase in the likelihood of attending a selective college (rather than simply a change in which college they attend).

This advantage also arises if we consider the children whose parents had a higher chance of marrying within their college due to a higher fraction of opposite-sex students: Increasing the proportion of opposite-sex students from 25% to 50% would increase the probability that the child

[4]For a similar reason, we have small sample sizes for looking at children labor market outcomes.

attends the same college as their parent by 0.5 percentage points (relative to a mean of 5.0%). Increasing the proportion of opposite-sex students from 25% to 50% would also increase the chance of attending *other* selective colleges (by 0.2 percentage points, relative to a mean of 5.1%) again implying that our results likely capture an actual increase in college attainment.

What explains the higher human capital of children of same-college marriages? We consider two main explanations. First, same-college couples might have more educated children simply because they have fewer children and there is a quality-quantity tradeoff. We find little evidence in favor of this explanation. If anything, same-college couples have *more* children than those who marry outside their college. Second, same-college couples might have more educated children because those who marry within their college end up with more educated spouses than those who marry outside their college, which in turn increases human capital in the next generation. Consistent with this explanation, we show that students married to someone from their own college end up with partners that have, on average, 2.4 extra years of schooling than those who marry outside of college. This higher level of spousal education also arises when students are exogenously more likely to marry within their college: Increasing the share of opposite-sex students in the cohort from 25% to 50% increases spousal years of schooling by 0.2 years.

1.1 Related Literature

This paper contributes to four main strands of the literature. Our main contribution is to the literature investigating the role of colleges in intergenerational mobility (Chetty et al., 2020; Chetty, Deming and Friedman, 2023; Bleemer and Quincy, 2025). We do so by empirically documenting that colleges shape intergenerational mobility not only through their impact on individual-level labor market outcomes, but also by facilitating marital matches between individuals from similar socio-economic backgrounds.

Second, by focusing on the role of colleges in the marriage market, our research is also related to the literature investigating the marriage market consequences of college (Barrios Fernández, Neilson and Zimmerman, n.d.; Ge, Isaac and Miller, 2022; Kaufmann, Messner and Solis, 2013; Kirkebøen, Leuven and Mogstad, 2021). We contribute to this literature by investigating the extent to which colleges facilitate marriages between individuals from similar socio-economic backgrounds and how this, in turn, affects human capital in the next generation.

Third, we contribute to the literature on educational assortative mating (see, for instance, Schwartz (2013); Kalmijn (1998)). We offer some of the first evidence on assortative mating based on the *specific college* attended. This distinction is important because there is substantial heterogeneity in outcomes by the type of college attended.[5] We also provide causal evidence that such assortative mating arises from on-campus overlap.

Finally, we contribute to the literature on assortative mating by parental background (e.g., Olivetti et al., 2024; Schwartz, 2013; Goldman, Gracie and Porter, 2024; Eriksson et al., 2023; Bratsberg et al., 2023). First, we document the role of colleges in generating such assortative mating. Second, by showing that even conditional on attending the same institution, marital sorting by parental income remains strong, we shed light on the relative role of search frictions versus preferences for explaining assortative mating by social class (see, for instance, Abramitzky, Delavande

[5]The only exception in the U.S. context is Arum, Roksa and Budig (2008), who uses the PSID to show that a large share of college-educated married individuals have spouses who attended institutions with similar characteristics.

and Vasconcelos (2011)).[6]

2 Data

In this section, we begin by introducing our main dataset on students attending 63 co-educational colleges between 1915 and 1949. We then outline the procedure used to identify spousal links among these students. Finally, we explain how we recover information on the college outcomes of their *children*.

Main Sample: Students at U.S. Co-Educational Colleges (1915-1949). We compile a dataset including 2 million students who enrolled as degree-seeking undergraduates at one of 63 co-educational colleges between 1915 and 1949.[7] Table A1 shows the list of colleges included in our main sample, their corresponding school group, and the total number of students at each college.

To do so, we start from the data in Abramitzky et al. (2024*b*), which compiled a dataset of students at elite private and public colleges. Out of these colleges, we include the nine private colleges and 27 public colleges that were already co-educational by 1915. This list includes four of the Ivy Plus colleges (Cornell, Penn, Chicago, and Stanford), five of the Little Ivies (Bates, Colby, Tufts, Swarthmore, Middlebury), as well as the universities classified as "Public Ivies" in the 2001 book *"Greenes' Guides to Educational Planning"* (Greene and Greene, 2001).

We complement these data by compiling, transcribing, and standardizing publicly available documents (e.g., registers, directories, catalogs, bulletins, yearbooks) that include lists of students at 27 *additional* public universities. Adding these universities enables to cover nearly every U.S. state through its flagship or land-grant institution. Overall, our sample covers about a third of students in public universities and 17.3% of *all* students in "Universities, colleges, and professional schools" as of 1949.[8]

Expanded Sample: Students at U.S. Colleges (1915-1966). Our main sample ends in 1949 as this is the last year for which we can link students to an adult census and observe their marriage outcomes (the 1950 census is currently the latest full-count census available). However, we also collect data on students who attended college in the 1950-1966 period so as to measure the college attainment of the *children* of students in our main sample. This extended sample includes students in 95 colleges (the ones in the current paper plus all non-coeducational colleges in Abramitzky et al. (2024*a*), see Table A2 for a full list of such colleges and the corresponding sample sizes). Using data from the 1970 census, we estimate that these colleges account for 19% of all individuals with

[6] Methodologically, our results are related to the literature investigating the marriage market consequences of skewed sex ratios (Angrist, 2002; Abramitzky, Delavande and Vasconcelos, 2011; Grosjean and Khattar, 2019). They are also related to a few papers that use variation in gender ratios within colleges to investigate impacts on students' educational outcomes Hill (2017); Bostwick and Weinberg (2022).

[7] We define a "degree-seeking undergraduate" as any student enrolled full-time in their first degree program. This excludes students with prior degrees, those pursuing certificates, etc. See Abramitzky et al. (2024*b*) for further details.

[8] According to the 1948-1950 Biennial Survey of Education (U.S. Office of Education, 1953), there were 161 publicly controlled institutions in the "Universities, colleges, and professional schools" category as of the 1949 academic year. Hence, our sample includes roughly a third of all public universities. The same source reports that in the 1949 academic year there were 171,874 newly enrolled students in publicly controlled institutions in the "Universities, colleges, and professional schools" category. Our sample includes 63,126 students in public universities for the same academic year.

at least one year of college in the 1915 to 1950 birth cohorts. These colleges, however, represent a much higher share of students attending the most prestigious colleges.[9]

Measuring Students' Socio-Economic Backgrounds and Post-College Outcomes. After compiling the student-level data, we match these records to census records before and after their college attendance. This procedure enables us to obtain information on students' socioeconomic backgrounds prior to attending college as well as their post-college outcomes, including whether they are married to a spouse from the same college. Figure A1 provides a flowchart of the procedure. Our linking procedure has two steps.

In the first step, we link students to the *two closest* censuses *prior* to their college enrollment, allowing us to observe them in their childhood household and recover information on their socioeconomic background. This matching process, in the spirit of Abramitzky, Boustan and Eriksson (2014), relies on students' full names, hometowns (available for 86% of records) and approximate birth years (assuming they would have been between the ages of 17 and 28 when they started college). In this step, we take advantage of the fact that most female students were unmarried during college, allowing us to use their maiden names to link them to earlier census records.[10]

In the second step, we use census-to-census crosswalks from the Census Tree Project (Buckles et al., 2023) and the IPUMS MLP project (Ruggles et al., 2019) to link each student *forward* to all available post-college censuses.[11] We use these links because they enable us to follow women, including those who marry and adopt their husband's last name. The Census Tree is able to follow women who marry by using genealogical links created by users of the FamilySearch platform (Buckles et al., 2023). MLP is able to follow women who marry by matching census records to the Social Security's NUMIDENT file, which includes both birth and married surnames (Ruggles et al., 2019). This step allows us to observe students as adults living with their spouse and thus identify spousal relationships among students.

The example in Figure 1 illustrates our linking procedure. This figure shows George Hicks, a student at Berkeley in 1915, and his spouse Beatrice Beardslee, another student at Berkeley in 1915. We establish their spousal relationship through the following steps:

1. We first link Berkeley records to population censuses *prior* to Berkeley enrollment, which allows us to observe both George and Beatrice in the 1910 census.
2. We then link both students *forward* to the 1920 census, where we observe them living together as husband and wife. The census lists George as household head and Beatrice as his spouse. From this step, we establish the spousal relationship between the students.

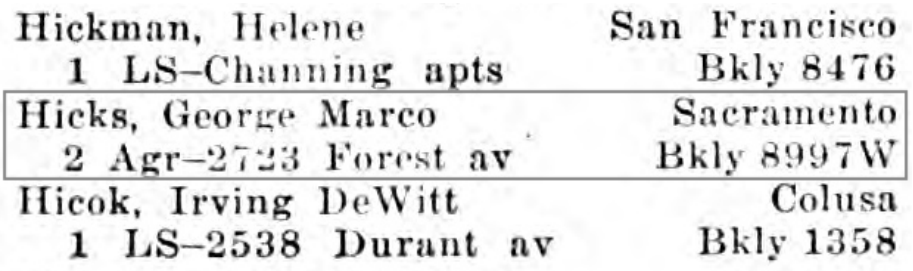
[9]For instance, our sample includes 12 out of the 14 universities that the first edition of U.S. News College Ranking ranked as the best Universities in the country (Reiter, 2024).

[10]Whenever we match someone to more than one census, we keep the observation corresponding to the census that is closest in time to the student's first year of college (so that parental background is measured as close as possible to when students' college enrollment decisions occurred).

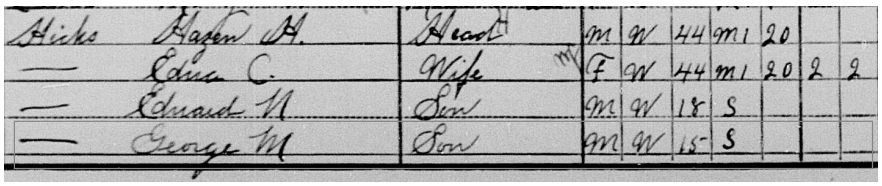
[11]We use the Census Tree crosswalks whenever available and IPUMS MLP crosswalks for census pairs involving the 1950 census (which are currently not available through the Census Tree). Note that the Census Tree aggregates links from various sources including the Census Linking Project (Abramitzky et al., 2020) and MLP (Ruggles et al., 2019).

Figure 1: Illustration of Data Digitization and Linking Strategy

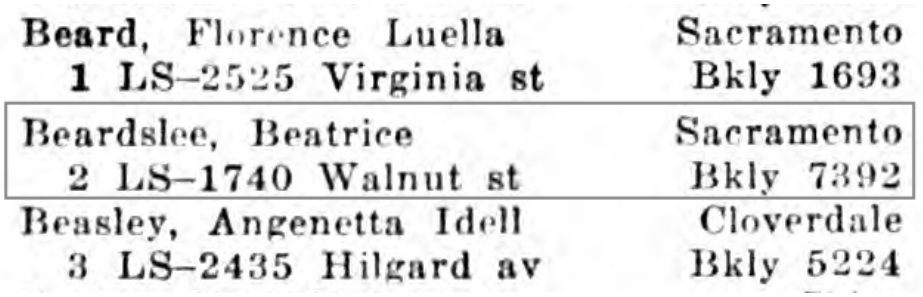
(a) George Hicks in Register of the University of California, Berkeley, 1915



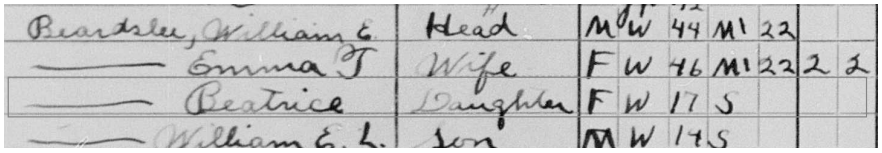
(b) George Hicks as a child, 1910 Census



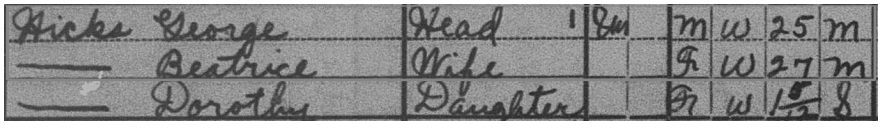
(c) Beatrice Beardslee in Register of the University of California, Berkeley, 1915



(d) Beatrice Beardslee as a child, 1910 Census



(e) George and Beatrice as husband and wife, 1920 Census



Notes: Panels (a) and (c) show excerpts from the 1915 Register of the University of California, Berkeley. Panel (a) highlights student George Hicks, whereas Panel (c) highlights student Beatrice Beardslee. Panels (b) and (d) show these students in their childhood households in the 1910 U.S. Population Census. Panel (e) shows George and Beatrice as husband and wife in the 1920 census.

Matching Rates and Sample Sizes. Panel (a) of Figure A3 reports the number of students we match to at least one childhood and one adult census, while Panel (b) presents the corresponding match rates (by gender and entry decade). In total, our sample includes 465,712 students who are matched to both a childhood and an adult census. We consider an individual to be matched to an adult census if we observe them at ages 25 or older in at least one post-college census (so that

we can potentially observe their spouse). This number represents 28% of all students in our main sample. The matching rates to a *childhood* census are in general similar by gender and by entry decade, averaging 54% across all students in our sample. The matching rates to an *adult* census are lower for more recent entry decades (as we have fewer adult censuses to which to match the person) and are lower for women (particularly in the later censuses). We match 68% of the students who started college before 1940 to an adult census, but only 29% of those who started after 1940. Similarly, we match 61% of the men we match to a childhood census to at least one post-college adult census, whereas we only match 37% of the women. Although women have lower matching rates than men, we show below that the resulting sample of women is similarly representative of the population of students at these colleges than the sample of men.

Identifying Colleges Attended by Children of the Original Students. Our procedure for identifying the colleges attended by the children of the original students follows Abramitzky et al. (2025). This approach takes advantage of the fact that, when we match the original students to adult censuses, we can observe them with children of their own and establish whether such children themselves are matched to any of the colleges in our data. Figure A2 provides a flowchart of the process.

Representativeness of Linked Sample. Since we are only able to match a subset of students, a concern is that our sample may not be representative. There are two main sources of selection in our data. First, there might be selection of who matches from the college registers to a childhood census (step 1 of our linking process). Second, there might be selection of who matches from a childhood to an adult census (step 2 of our linking process).

Because the college registers have little information about students' backgrounds (as described above, they only include students' names and hometowns), it is challenging to directly compare the backgrounds of matched students to the backgrounds of those who do not match. Hence, to address the first source of selection, we show that the background characteristics of individuals that we match to a childhood census closely align with those of individuals in a nationally representative survey that includes alumni from the same group of colleges. This comparison sample comprises 489 *men* who attended the same set of colleges between the mid-1920s and 1950, as recorded in the Survey of Occupational Change in a Generation (OCG). The OCG was administered as a supplement to the 1962 and 1973 CPS, with the 1973 wave including information on the college attended.

Figure A5 shows that men in our sample closely resemble those in the OCG data with respect to parental occupations, regional origins, and racial makeup. This result confirms the findings in Abramitzky et al. (2024*a*), which showed a similar pattern for some of the colleges in our sample.

To address the second source of selection (from childhood censuses to an adult census), we can directly compare the childhood characteristics of individuals who are matched to an adult census to the background characteristics of those who are not matched. We start by noting that Buckles et al. (2023) show that individuals in the Census Tree (our main approach for matching individuals forward) are mostly comparable to the full U.S. population, although matched individuals are slightly more likely to be of higher socio-economic status. We expect selection to be even milder in our setting, because the individuals in our sample all attended the same set of colleges and hence are a more homogeneous group than the broader U.S. population.

To assess the link between socio-economic status and selection into the sample, Figure A4 shows the relationship between students' parental income ranks (based on the childhood census) and the likelihood of observing them in at least one census at ages 25 or more. Panel (a) focuses

on men whereas panel (b) focuses on women.

The figure yields two main findings. First, the matching rates are higher for men, as women are harder to match to future censuses due to name changes upon marriage. This is reflected in the higher intercept for men. Second, for both men and women, there is a weak relationship between the likelihood of observing them as adults and parental income ranks. The slope for men implies that the matching probability for a student at percentile 100 is only 4 percentage points higher than the probability for a student at percentile 0. For women, the slope implies that the matching probability for a student at percentile 100 is in fact lower (by -3 percentage points) than the probability for a student at percentile 0.

Table A3 shows the relationship between a broader set of background characteristics and the likelihood of matching to an adult census. The top row in the table shows the fraction of individuals who match to at least one adult census, by gender. The remaining rows show the coefficient associated to each variable in a regression in which the outcome is equal to 100 if the observation matches to at least one adult census (each row is a separate regression).

The table shows that differences in the matching probability by parental occupations are generally small. The largest difference is that male children of unskilled workers are 2.8 percentage points less likely to match to an adult census than children of other workers, relative to an average matching probability of 61%. The differences for other parental occupations are smaller (at most one percentage point).

The largest difference in matching probabilities is that White children are much more likely (28.7 percentage points for men and 18.7 percentage points for women) to match than non-White children. But note that our sample is overwhelmingly White (98%), so we will show below that our linked sample is very similar to the full sample despite this difference in matching probabilities.

Figures A6 and A7 replicate the comparison in Figure A5, but contrasting students observed in both a childhood and an adult census with all students observed as children. Across parental occupations, regional origins, and racial composition, differences between the full and matched samples are small for both genders.

Overall, we conclude that our sample is mostly comparable to the target population: neither of the matching steps (from registers to childhood census and from childhood census to adult census) appears to introduce quantitatively meaningful selection into the sample.

3 Prevalence of Same-College Marriages

We first show that there is a high fraction of individuals who are married to someone from their own college (college homogamy). The challenge in calculating this share is that we only observe whether a student married to someone from their own college if we: (1) match them and their spouse to a childhood census, (2) match them and their spouse from a childhood to an adult census where they are observed co-residing as spouses. In Appendix Section A.1, we show how we can use Bayes rule together with observed probabilities in the data (namely, the fraction of individuals who are observed with a same-college spouse and the fraction of married individuals who are matched to at least one post-college census) to estimate the true proportion of individuals married to someone from their own college.

Panel (a) of Column 1 in Table 1 shows the observed levels of homogamy for men and women, after adjusting for the fact that we only observe a subset of same-college marriages. Across all

Table 1: Educational Homogamy and Assortativity

(a) Same-College Homogamy

	Homogamy			Assortativity		
	Observed	Random	Maximal	Ratio	Absolute	Rescaled
	(1)	(2)	(3)	(1)/(2)	(1)-(2)	$\frac{(1)-(2)}{(3)-(2)}$
Men	0.0497	0.0023	0.5141	21.61	0.0474	0.09
Women	0.0775	0.0036	0.9682	21.53	0.0739	0.08

(b) Homogamy Based on College Attainment

	Homogamy			Assortativity		
	Observed	Random	Maximal	Ratio	Absolute	Rescaled
	(1)	(2)	(3)	(1)/(2)	(1)-(2)	$\frac{(1)-(2)}{(3)-(2)}$
Men	0.5816	0.1588	0.8048	3.66	0.42	0.65
Women	0.6980	0.1974	1.0000	3.54	0.50	0.62

Notes: Observed homogamy is the share of individuals who married someone from their own college. Random is the share that would be observed if individuals had the same chance of marrying within their college than with any college-educated opposite-sex person in their birth cohort. Maximal is the highest share of students that could marry within their college given the observed proportion of men and women in the college.

colleges in our data, we find that 5% of men and 8% of women are married to someone from their own college.

To interpret these magnitudes, the remaining columns of Table 1 compare the observed level of homogamy to two benchmarks. First, we compare it to the probability of observing a marriage between two individuals of the same college by chance, conditional on both individuals being *college educated*. This is a useful benchmark as it enables us to compare the degree of *same-college* assortative mating to a more standard measure of educational assortative mating (namely, whether college-educated individuals marry other other college-educated individuals). For men, this rate can be calculated as:

$$\sum_c \alpha_c^m s_c^w \tag{1}$$

where α_c^m is the fraction of all men (m) in our college sample who attended college c, and s_c^w is the fraction of *all* college-educated women (w) in the 1897-1931 birth cohorts who attended college c. The rate for women can be calculated analogously.

Column 2 shows that the probability of observing a same-college marriage by chance is very low. This is not surprising because students from any particular college represent a very small fraction of all college-educated individuals. As a result, male students are 22 times more likely to marry someone from their own college than what would be predicted based purely on chance and female students are 21 times as likely to do so.

Even if individuals only married within their college, the observed homogamy rate would be below 100% if there are differences between the number of men and women in a college (as it is the case in our setting). Our second benchmark is then the *maximal* homogamy which can be

calculated for men as:

$$\sum_c s_c^m \frac{min(n_c^m, n_c^w)}{n_c^m} \tag{2}$$

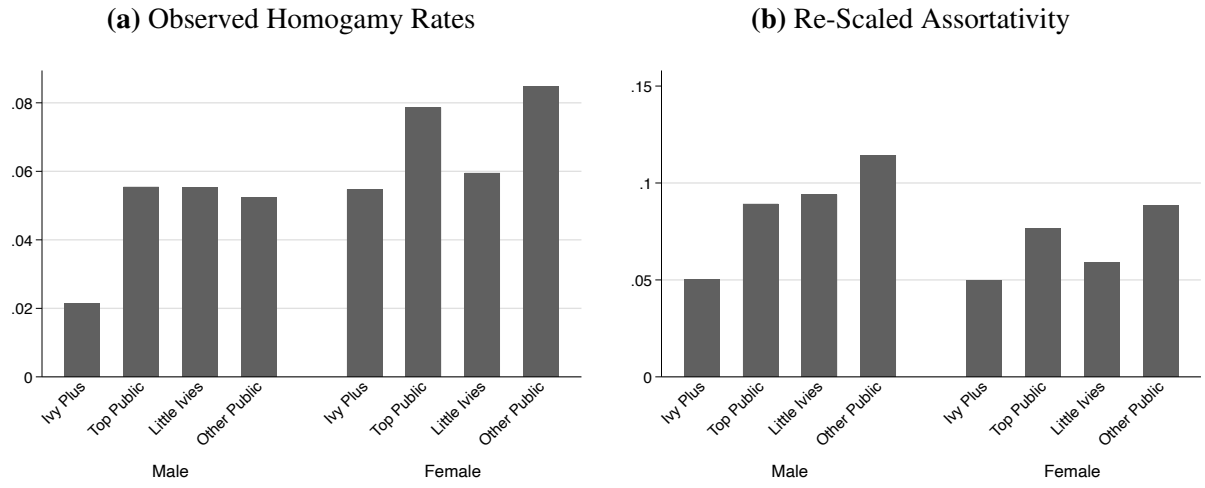
where α_c^m is the fraction of all men in our college sample who attended college c, n_c^m is the number of men in college c, and n_c^w is the number of women in college c (the calculation for women is analogous). Maximal homogamy is only 51% for men but nearly 100% for women given that men constituted the majority of students in our setting.

In the last column, we combine these statistics into a single measure following Liu and Lu (2006) and Kirkebøen, Leuven and Mogstad (2021). This measure ranges from zero to one: it is equal to zero in the absence of assortativity (purely random mating), and equal to one when assortativity reaches the maximum feasible value. This column shows that, once adjusting for unbalanced sex ratios (which imply that the maximal homogamy is lower for men than for women), men and women exhibit similar levels of assortativity: men achieve 9% and women achieve 8% of the maximal assortativity.

Heterogeneity by College Group. So far, we have calculated homogamy and assortativity rates by pooling all colleges in our data. We next perform a similar set of calculations but splitting the analysis by group of colleges: Ivy Plus, Little Ivies, Top Public Colleges, and Other Public Colleges.

Figure 2 shows the observed levels of homogamy as well as the rescaled homogamy, by gender and college group. Homogamy levels among men are the lowest in Ivy Plus colleges but relatively comparable in the remaining group of colleges. Among women, homogamy levels are the lowest in the Ivy Plus and the Little Ivies.

Figure 2: Homogamy and Assortativity by Group of Colleges



Notes: This figure shows observed homogamy rates (panel (a)) and rescaled assortativity (panel (b)), separately by group of colleges.

Same-College Homogamy versus Educational Homogamy. To benchmark these results, we compare them to the degree of homogamy on the basis of simply having attended college. To do so, we use the 1960 population census (Ruggles et al., 2020) to calculate analogous measures to the ones above but focusing on the probability of having a college-educated spouse rather than a

same-college spouse. To make these findings as comparable to those in the previous section, we focus on the same birth cohorts that are included in our main sample, namely individuals who likely started college between 1915 and 1950 (1897 to 1931 birth cohorts).

Panel (b) Table 1 shows that there is strong homogamy on the basis of college attendance: 58% of college-educated men are married to a college-educated spouse and 70% of college-educcated women are married to a college-educated spouse. Yet, assortativity on the basis of college attainment is substantially weaker than that based on specific colleges: College-educated men are "only" 4 times more likely to marry someone with a college education than what would be expected at random (compared to 22 times for someone who attended their same college). College-educated women are 4 times more likely to marry someone with a college education than what would be expected at random (compared to 21 times for someone who attended their same college).

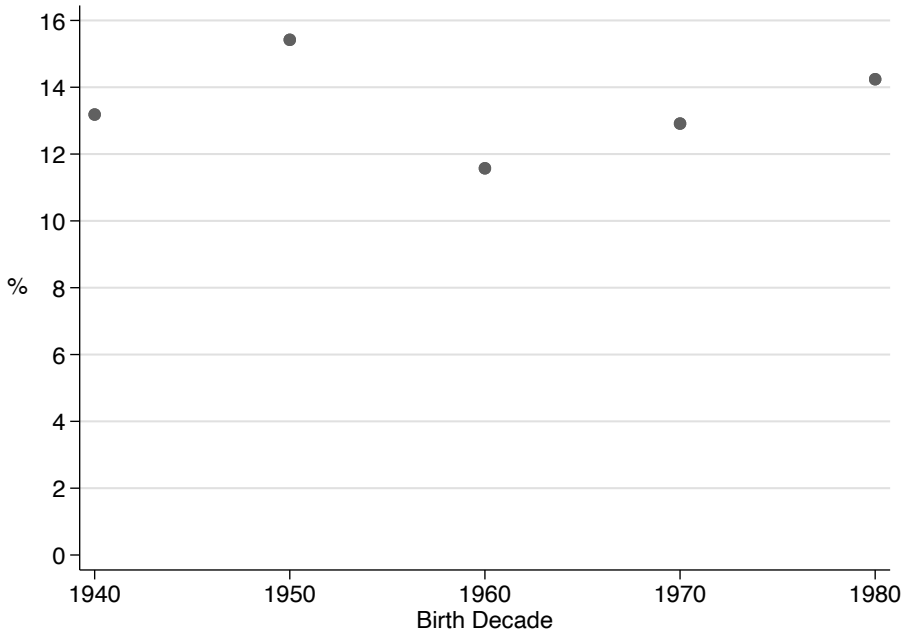
Comparison to Modern Data. The analysis above documents the prevalence of same-college homogamy among individuals who began college between 1915 and 1949. A natural follow-up question is whether similar patterns persist today.

To answer this question, we draw on data from the How Couples Meet and Stay Together (HCMST) study (Rosenfeld, Thomas and Hausen, 2024), a nationally representative survey designed to collect detailed information on how Americans form and maintain romantic relationships. We focus on respondents in the 2017 survey wave, including subsequent follow-up interviews conducted in 2020 and 2022.

To construct a sample comparable to our historical cohort, we restrict the data to individuals who completed at least one year of college and who were age 35 or older at their last interview. We then measure the share of respondents who, in any of the three survey waves, report being married to someone who attended their same college.

Figure 3 shows that the share of college-educated adults who married someone they met in college has remained relatively stable (around 12 percent) for birth cohorts from the 1940s through the 1980s. We focus on these birth cohorts so that as to give each cohort enough time to "complete" their marriage decisions as of 2022, the last year of the survey. These levels are broadly similar to those observed in our historical college-based data, suggesting that colleges remain an important avenue through which college-educated individuals meet their future spouses.

Figure 3: Percentage of College-Educated Individuals who Married Someone who Attended their Same College, by Birth Decade (HCMST Study)



Notes: This figure shows the fraction of individuals who married someone who attended their own college, by birth cohort. The figure uses data from the *How Couples Meet and Stay Together* (HCMST) study (Rosenfeld, Thomas and Hausen, 2024). We restrict the sample to individuals with at least one year of college and who were at least 35 years old by the time they were last surveyed.

4 Explaining Same-College Marriages: Selection versus Exposure

The previous section shows that there is strong college homogamy: individuals are dramatically more likely to marry someone from their same college than what would be expected by chance. This homogamy might be the result of two distinct forces. First, it might simply reflect selection. That is, individuals who attend the same college might share attributes (such as socioeconomic background or academic ability) that make them both more likely to attend the same college and to marry each other. Second, it might reflect a causal effect, whereby attending the same institution increases the chances that individuals marry each other. Below, we implement two empirical strategies, both of which indicate that the observed homogamy reflects a causal effect of colleges.

4.1 Exposure to Opposite-Sex Students and Likelihood of Same-College Marriages

As a first step to assess whether the observed homogamy is the result of selection or a causal effect, we test whether homogamy rises when students are exposed to a larger share of opposite-sex peers. If homogamy were solely the result of selection, differences in exposure to potential

partners within the same college should *not* increase the likelihood of marrying someone from that college.

Panel (a) in Figure 4 shows that there is a strong positive relationship between the share of opposite-sex students in a student's cohort and the likelihood of marrying a spouse from the same college. When the fraction of students from the opposite sex is below 20%, only 1% of students are observed with a spouse from the same college. That proportion rises to 7% when the fraction of opposite-sex students is above 80%.

While suggestive, the relationship above might reflect sorting of students rather than a causal effect. For instance, colleges with more skewed sex ratios might tend to attract students with a higher propensity to marry each other. Another possibility is that women that started college in periods in which there were fewer other women studying (and hence were exposed to a higher proportion of opposite-sex students) might have characteristics that increase their propensity of marrying someone from their college.

To address this possibility, we estimate:

$$SameCollegeSpouse_{ict} = \alpha_c + \alpha_t + \alpha Male_{ict} + \beta \%OppositeSex_{ict} + \varepsilon_{ict} \quad (3)$$

where $SameCollegeSpouse_{ict}$ is a dummy that is one if student i at college c from entry cohort t is observed in the census with a spouse who attended the same college, α_c are college fixed effects, α_t are entry cohort fixed effects, and $Male_{ict}$ is a dummy that is one for male students. We include a male dummy as, in this setting, women are systematically exposed to a higher share of opposite-sex students than are men. Our coefficient of interest is β, which measures the increase in the probability of observing a same-college marriage for a 1 percentage points increase in the proportion of students from the opposite gender. Our sample includes all students who can be matched to at least one adult census, regardless of their marital status in the adult census (below, we investigate changes in the extensive margin of marriage). In this section of the analysis and throughout the paper, we cluster standard errors at the college level.[12]

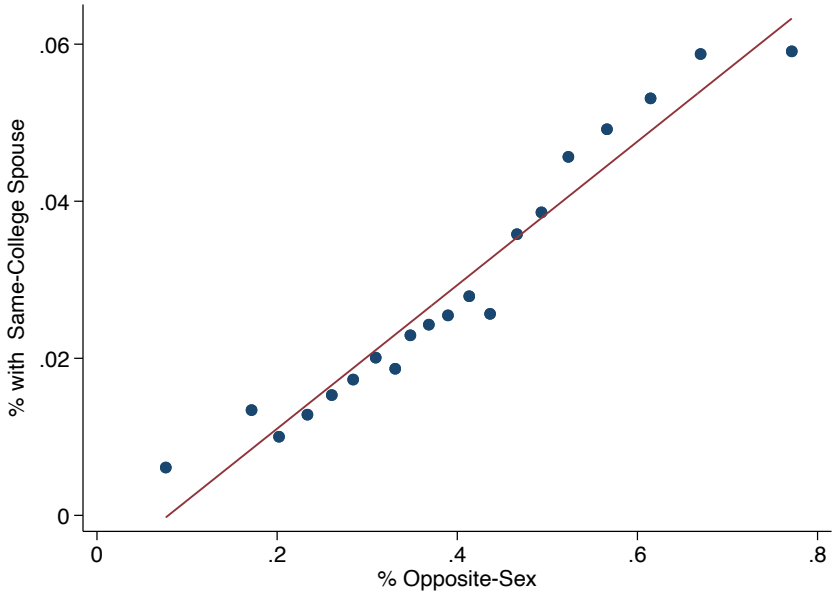
The identifying assumption in this strategy is that, conditional on college, entry cohort, and gender fixed effects, individuals who start college in a year with a higher proportion of opposite-sex students have similar potential outcomes than those who start in years with a lower proportion of such students.

Evidence on the Validity of the Identification Strategy. There are two main concerns with this strategy. The first concern is that there might not be enough within-college variation in gender ratios. Figure 5 shows that this is not the case. Panel (a) in this figure shows the distribution of the share of female students by entry cohort. Panel (b) shows that this variation is not solely explained by which college students attended or when they enrolled: The standard deviation of the residualized share of women is roughly two-thirds of the raw standard deviation, indicating that a large fraction of the observed variation reflects within-college changes in gender composition across cohorts. Panel (c), which excludes the 1941-1949 cohorts from the sample, shows that the variation in gender composition across cohorts is not driven either by anomalous gender ratios during World War II and the subsequent years.

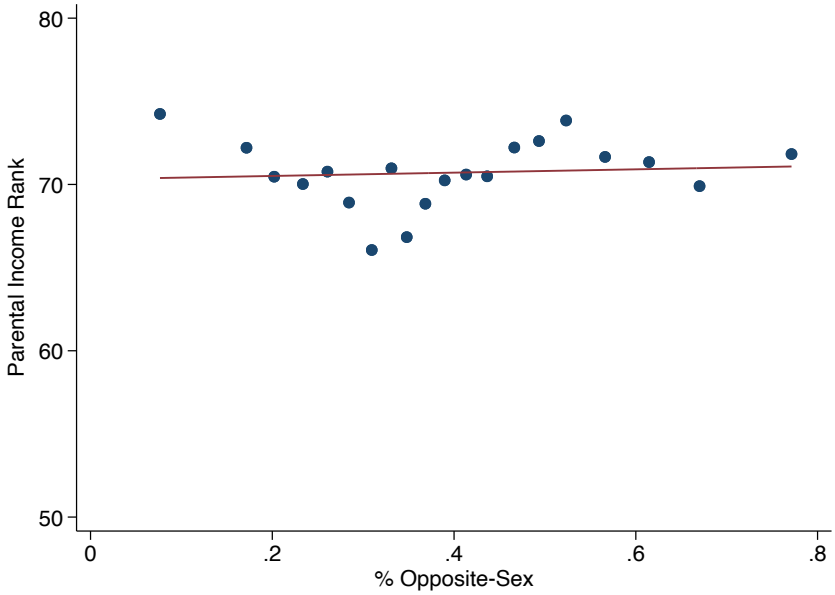
[12]This strategy is conceptually similar to designs used in the peer-effects literature, which exploit within-school variation in cohort composition to estimate causal effects. In particular, our approach parallels studies that measure the impact of a higher share of female classmates on student outcomes (e.g., Hoxby (2000); Lavy and Schlosser (2011); Black, Devereux and Salvanes (2013)).

Figure 4: Relationship between % of Opposite-Sex Students in Cohort and Likelihood of Same-College Spouse

(a) Same-College Spouse



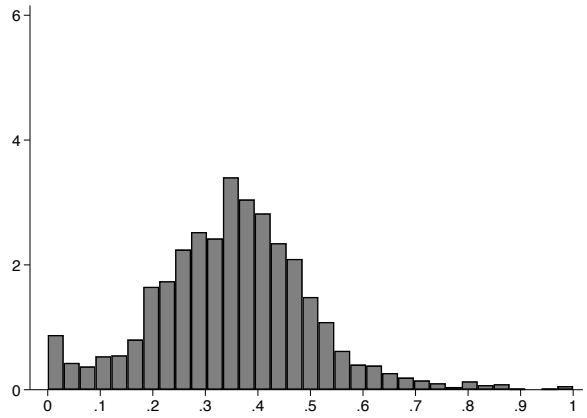
(b) Parental Income Rank



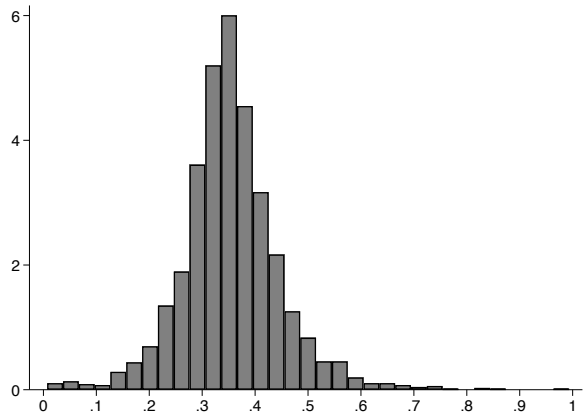
Notes: *% Opposite-Sex* ranges from 0 to 1 and measures the proportion of individuals in a student college cohort (defined as those starting in the same college and in the same year) who are of the opposite sex. . Panel (a) shows a binned scatterplot of the relationship between such variable and the likelihood that the student is observed in the census with a spouse from the same college. Panel (b) shows the relationship between such variable and students' parental income ranks.

Figure 5: Female Share by Entry Cohort

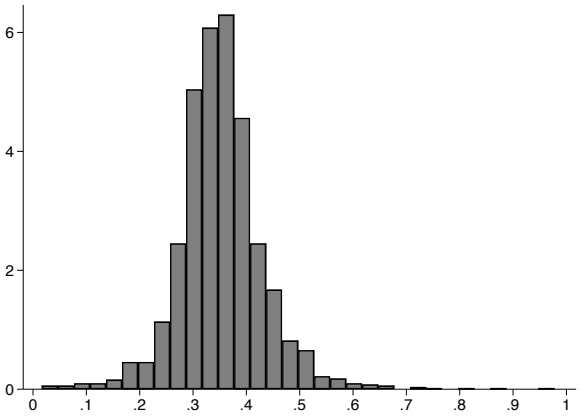
(a) Unadjusted



(b) College and Entry Cohort Fixed Effects



(c) Excluding World War II Cohorts



Notes: Panel (a) shows the empirical distribution of the share of women by college cohort (each observation is a college-entry-cohort combination). Panel (b) shows the distribution of the residuals of a regression of the share of women on entry cohort and college fixed effects. The residuals are re-centered around the mean share of women. Panel (c) repeats the analysis in panel (b) but excluding cohorts that coincided with World War II or during which G.I. Bill benefits for returning veterans were available (1941-1949).

The second concern is that changes in the gender ratio might be correlated with other determinants of students' marital outcomes. This possibility is unlikely: while students might know the average gender mix of a college, they are unlikely to be able to anticipate the gender mix of their actual class, particularly in the less-information rich setting of the early and mid-20th century.

Nevertheless, to address this concern, we show that there is little relationship between students' background characteristics and the proportion of opposite-sex students in their cohort. Panel (b) in Figure 4 is analogous to panel (a) of the same figure but shows parental income ranks in the y-axis (measured using the approach detailed in Abramitzky et al. (2024*a*)). The figure reveals that there is no correlation between the share of opposite-sex students in a students' cohort and such variable.

Table B1 provides further evidence consistent with our identification assumption. In this table, we estimate equation 3 using as outcome variables a set of pre-determined student characteristics: parental income ranks, indicators for a student's father occupational categories (professional, white-collar non-professional, farmer, skilled blue-collar, unskilled), and an indicator that is one if the student is white. Overall, there is little relationship between the share of opposite-sex students and students' background characteristics. The point estimates are in most cases small, particularly after the inclusion of gender, college, and year fixed effects. The only exception is that students who start college in years with a higher proportion of opposite-sex students appear more likely to have a professional father. This significant relationship is likely due to multiple testing given the large set of pre-determined characteristics that we consider in this table. We nevertheless show that our results are nearly identical when controlling for this full set of background characteristics.

Second, we show that our results are robust to excluding cohorts that started college during World War II and during the years during which WWII G.I. Bill benefits were available (1941-1951). Gender ratios varied substantially over this period: Figure B1 shows that the proportion of women in our sample increased during WWII given women's lower likelihood of serving in the war, and then subsequently declined as returning WWII veterans entered college. A concern is that students who attended college during the War or under G.I. Bill eligibility might have attributes that make them more likely to marry within their college. However, our results are nearly identical when we exclude such cohorts.

Third, we show that our results are robust to the inclusion of college-specific linear trends. This specification enables us to account for college-specific trends in the gender mix that might be correlated with the likelihood of same-college marries.

Fourth, we show that our results are robust to exploiting variation in the share of opposite-sex students across only *nearby* cohorts within a college. The goal of this strategy is to deal with potential time varying factors correlated with both the share of opposite-sex students and the likelihood that students within a cohort marry each other. To implement this strategy, we estimate equation 3 after including college$\times$ five-year bin entry cohorts. This strategy compares students who attended the same college, in the same 5-year period (1915-1919, 1920-1924, etc.), but who were exposed to different fractions of opposite-sex students.[13]

Main Results. Having established the validity of this research strategy, we turn to estimate the impacts of a higher share of opposite-sex students on students' marital outcomes. Table 2 shows that the proportion of opposite-sex students is a strong determinant of the likelihood of

[13]We chose this specification rather than a specification just comparing adjacent cohorts as in Hoxby (2000) and Lavy and Schlosser (2011) because, as we show below, we see meaningful across-cohort marriages for individuals starting college in adjacent years.

Table 2: Share of Opposite-Sex Students in Cohort and Likelihood of Same-College Spouse

	(1)	(2)	(3)
% Opposite-Sex	0.0600***	0.0570***	0.0569***
	(0.00499)	(0.00746)	(0.00729)
Gender FEs	Yes	Yes	Yes
College FEs	No	Yes	Yes
Entry Cohort FEs	No	Yes	Yes
Individual Controls	No	No	Yes
Observations	479284	479284	479284
Mean of Dep. Var.	0.0290	0.0290	0.0290

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. The outcome variable is an indicator that is one if the student is ever observed in an adult census with a spouse who attended the same college. *% Opposite-Sex* ranges from 0 to 1 and measures the proportion of individuals in a student college cohort (defined as those starting in the same college and in the same year) who are of the opposite sex. Column 1 shows the unconditional relationship between the two variables. Column 2 adds college, entry cohort, and gender fixed effects. In Column 3, we control for father's income rank, indicators for father's occupation (professional, white-collar non-professional, farmer, skilled blue collar, with unskilled as the omitted category), and an indicator for being White. Standard errors are clustered at the college level.

marrying someone from the same college. Column 1 shows the relationship between the two variables without any control variables. In column 2, we include gender, college, and entry cohort fixed effects. In column 3, we add the full set of individual controls. All columns yield very similar point estimates, consistent with the balancing in students' pre-determined characteristics documented above.

Our estimates in column 3 imply that increasing exposure to opposite-sex students from 25% to 50% would increase the probability of being observed with a same-college spouse in the census by 1.4 percentage points. This is a large impact given that the fraction of students who are observed with a same-college spouse is only 2.9%.

We apply an analogous re-scaling procedure to the one outlined in the previous section to adjust this coefficient for the fact that we observe only a subset of same-college marriages. When doing so, we find that increasing exposure to opposite-sex students from 25% to 50% would increase the probability of a same-college spouse by 2.5 percentage points, relative to a mean of 5.5% (Table B2).

Table B3 shows the robustness of this main result to the set of alternative specifications and sample restrictions outlined above.

Heterogeneity. We next explore the heterogeneity of this main result by gender, whether the college is private or public, whether the students' cohort is below or above the median with respect to size, and whether the student grew up in a family in the bottom 80% or top 20% of the parental income distribution. Overall, there is relatively little heterogeneity: a higher share of opposite-sex students increases the chances of same-college marriages across all sub-samples we consider and by a similar magnitude.

In Columns 1 and 2 of Table 3 we split the sample by gender. A number of studies suggest that

shifts in the gender composition of a cohort may influence broader gender norms, which could in turn affect marriage outcomes (Grosjean and Khattar, 2019). Thus, our results could in principle reflect changing gender norms within a cohort rather than reduced search frictions in the marriage market. Similarly, a number of studies document a relationship between a class gender's mix and academic performance, which could in turn also affect marriage outcomes (Hoxby, 2000; Lavy and Schlosser, 2011). However, we find similar effects for both men and women. This symmetric pattern appears inconsistent with a pure gender-norms channel or with the effects operating through changes in academic performance.

Table 3: Share of Opposite-Sex Students in Cohort and Likelihood of Same-College Spouse, Heterogeneity

	Gender		Public/Private		Cohort Size		Parental Income Rank	
	(1) Women	(2) Men	(3) Public	(4) Private	(5) Below Median	(6) Above Median	(7) Bottom 80%	(8) Top 20%
% Opposite-Sex	0.0575***	0.0552***	0.0594***	0.0537***	0.0476***	0.0598***	0.0474***	0.0678***
	(0.0195)	(0.0107)	(0.00864)	(0.0137)	(0.0105)	(0.00888)	(0.00647)	(0.00895)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	122909	356375	404844	74440	81561	397723	240761	238523
Mean of Dep. Var.	0.0545	0.0202	0.0313	0.0163	0.0329	0.0282	0.0243	0.0337

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. Each column in this table corresponds to a different sample, as indicated by the column headings. The outcome variable is an indicator that is one if the student is ever observed in an adult census with a spouse who attended the same college. All columns control for father's income rank, indicators for father's occupation (professional, white-collar non-professional, farmer, skilled blue collar, with unskilled as the omitted category), and an indicator for being White. Columns 1-2 split the sample by students' gender, columns 3-4 split the sample based on whether the student attended a public or private college, and columns 5-6 split the sample based on whether the students' cohort was below or above the median cohort size in our data. Standard errors are clustered at the college level.

Table 4: Share of Opposite-Sex Students in Cohort and Likelihood of Marriage

	(1)	(2)	(3)
% Opposite-Sex	0.0273	-0.0467*	-0.0463*
	(0.0466)	(0.0242)	(0.0240)
Gender FEs	Yes	Yes	Yes
College FEs	No	Yes	Yes
Entry Cohort FEs	No	Yes	Yes
Individual Controls	No	No	Yes
Observations	479284	479284	479284
Mean of Dep. Var.	0.717	0.717	0.717

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. The outcome variable is an indicator that is one if the student is ever observed in an adult census with a spouse. *% Opposite-Sex* ranges from 0 to 1 and measures the proportion of individuals in a student college cohort (defined as those starting in the same college and in the same year) who are of the opposite sex. Column 1 shows the unconditional relationship between the two variables. Column 2 adds college, entry cohort, and gender fixed effects. In Column 3, we control for father's income rank, indicators for father's occupation (professional, white-collar non-professional, farmer, skilled blue collar, with unskilled as the omitted category), and an indicator for being White. Standard errors are clustered at the college level.

In Columns 3 and 4, we re-estimate these results separately for private and public colleges. We find similar point estimates for both private and public colleges, although the impact is larger in private colleges when compared to the mean of the dependent variable. Indeed, Figure B2 shows that we obtain similar results if we exclude all colleges within any given group of schools. This finding implies that our results are not driven by any particular college or group of colleges.

Columns 5 and 6 consider heterogeneity on the basis of cohort size. Here, we split the sample based on whether a cohort was below or above the median cohort size in our sample (441 students). A higher share of opposite-sex students increases the probability of a same-college marriage in both the below and the above median size cohorts, and does so by a similar magnitude.

Finally, Columns 7 and 8 split the sample based on whether the student grew up in a family in the bottom 80% or top 20% of the parental income distribution. We find that a higher share of opposite-sex students increases the probability of same-college marriages for both groups, although the point estimates are larger for students from top 20% families.

Impacts on the Extensive Margin of Marriage. A higher share of opposite-sex students might increase the probability of same-college marriages simply by increasing marriage propensities overall. To assess this possibility, we next investigate whether a higher share of opposite-sex students increases the likelihood of marriage overall. We re-estimate equation 3 but using as the outcome variable an indicator that is one if the person is ever observed with a spouse in an adult census. Table 4 shows that the effects we document on college homogamy are not driven by changes in the marriage extensive margin: We find no evidence that a higher share of opposite-sex students leads to higher marriage rates in general. If anything, students exposed to a higher share of opposite-sex students are *less* likely to marry as adults (but this result is noisier than our main result and less robust to alternative specifications, see Table B4).

4.2 On-Campus Overlap and Likelihood of Same-College Marriages

The previous subsection shows that students exposed to a higher proportion of opposite-sex peers are more likely to marry someone from their own college. This pattern implies that the homogamy we document cannot be fully explained by selection alone. Yet, this evidence does not tell us *how much* of the observed homogamy reflects selection versus exposure. In this subsection, we develop an empirical strategy to separate the contribution of each mechanism.

We start from the observation that attending the same college can increase the probability that two individuals marry each other through two conceptually distinct causal channels:

(i) Non-exposure (indirect) effects. This first channel is that, even without overlapping on campus, alumni of the same institution may be more likely to meet and marry later if attending that college steers them into similar post-college environments (occupations, industries, cities, social networks).

(ii) On-Campus Exposure (direct meeting opportunities). The second channel is that students who overlap on campus (same institution and nearby entry cohorts) have more direct opportunities to meet.

To quantify the *on-campus exposure* channel, we start from all the students that we observe as adults married to another student in the data. This includes students married to someone from their same college as well as those married to someone outside their college. We then create a dyadic sample, where each observation is a *pair* of students. We limit the sample to pairs where the difference in entry cohorts is at most ten years (in absolute value). Using these data, we estimate the following dyadic regression:

$$P_{ict\,jc't'} = \alpha_{cc'} + \alpha_{tt'} + \sum_{d=-9}^{10} \beta_d \, \mathbb{1}\{\, c = c',\, t - t' = d \,\} + \varepsilon_{ict\,jc't'}. \tag{4}$$

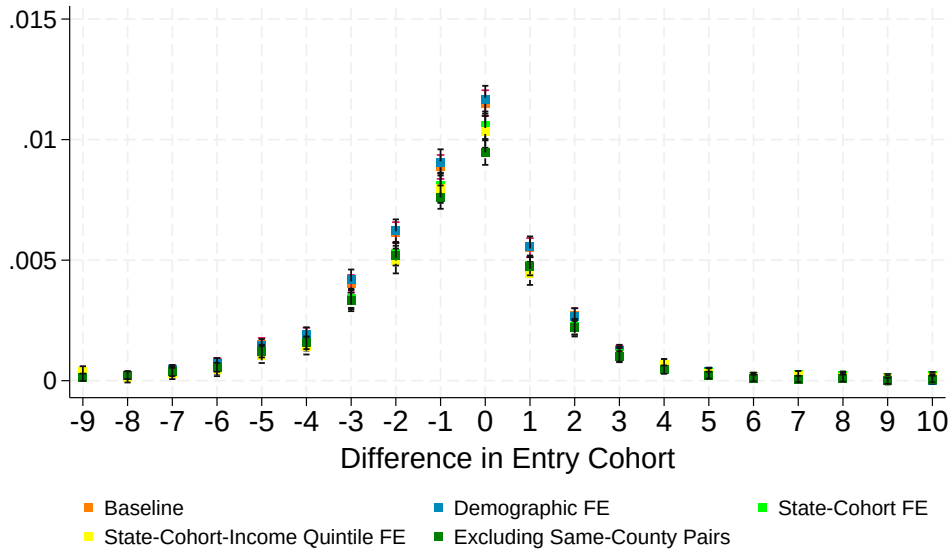
where $P_{ict\,jc't'}$ is an indicator that is one if male student i who attended college c with entry cohort t is married to female student i' who attended college c' with entry cohort t', $\alpha_{cc'}$ are *college-pair* fixed effects, and $\alpha_{tt'}$ are *entry cohort-pair* fixed effects. The college-pair fixed effects capture selection into similar colleges by similar types. The entry cohort-pair fixed effects account for the fact that individuals are more likely to marry others of similar ages.

Our parameters of interest are the β_d coefficients (the omitted category is all pairs whose entry-cohort gap satisfies $d = -10$). These coefficients trace the *exposure profile*: β_0 is the incremental marriage probability for exact classmates (same-college, same entry cohort), while $\beta_{\pm 1}, \beta_{\pm 2}, \ldots$ show how the effect changes with less overlap. If students spend exactly four years in college and college homogamy is mostly due to *on-campus exposure*, we expect the β_d coefficients to be close to zero when $|d| >= 4$. Note that, when $d > 0$, this corresponds to cases in which the male student i started later than the female student j (male in pair is likely younger than female in pair), whereas $d < 0$ corresponds to cases in which the male student started earlier (male in pair is likely older). In general, $\beta_d \neq \beta_{-d}$ as the propensity of marrying younger individuals might differ by gender.[14]

Figure 6 shows the results of estimating this equation. The figure shows a spike in the likelihood of marriage for individuals who start college in the exact same year and hence have maximum on-campus overlap. The probability declines rapidly as the difference in entry cohort grows from

[14]This strategy is similar to that of Bayer, Ross and Topa (2008), who estimate dyadic regressions comparing the likelihood that individuals work for the same employer when they live on the same block to the likelihood when they live in the same neighborhood but on different blocks.

Figure 6: Marriage Probability by Difference in Entry Cohort: Alternative Specifications



Notes: This figure shows the estimated values of $\{\beta_d\}$ from equation 6. The baseline specification includes only entry-cohort pair and college-pair fixed effects. The demographic controls specification includes parental income quintile and race pair fixed effects. The state-cohort specification includes entry-cohort $\times$ state of residence pair fixed effects. The state-cohort-income-quintile specification includes entry-cohort $\times$ state of residence $\times$ parental income quintile pair fixed effects.

0 to 4, and it then levels of at a value close to zero: if student i started college 4 years after student j, they are no more likely to marry each other than students starting college with a larger entry cohort gap. There is also evidence that the on-campus exposure effects are asymmetric: the estimated coefficients are larger when the differences in entry cohorts are negative (men in the pair starts college earlier than women in the pair) than when they are positive. As a result, the same total overlap in years leads to a higher marriage propensity when the man started earlier than the woman.

Under the assumption that on-campus exposure is effectively zero when $|d| >= 4$, we can use the estimated values of β_d to calculate the exposure-induced increase in the probability that the *average student* marries a same-college spouse.

To do so, note that the probability of marrying someone from your own college depends on: (1) on how many potential partners you are exposed to, and (2) the probability that you marry each of them.

Assume that a male student in cohort (c,t) is only exposed on campus to students in the same college with relative entry-year difference $|d| < 4$, and let $N^{\text{opp}}_{c,t+d}$ denote the number of opposite-sex students in college c who entered in year $t+d$. Under these assumptions, the exposure-induced increase in the probability of a same-college marriage $(e_{c,t})$ for male students in cohort (c,t) can be written as:

$$e_{c,t} \equiv \sum_{d=-3}^{3} \beta_d N^{\text{opp}}_{c,t+d}.$$

Aggregating this expression over colleges and cohorts yields the *average student-level exposure effect*,

$$h^{\text{exposure}} = \sum_c \sum_t s_{c,t} e_{c,t} \tag{5}$$

where $s_{c,t}$ is the fraction of male students in our sample who are from college c and cohort t. Let h^{observed} denote the observed marriage probability for students from the same college with relative entry-year difference $|d| < 4$. The share of observed same-college homogamy attributable to *direct exposure* is then:

$$\frac{h^{\text{exposure}}}{h^{\text{observed}}}$$

Our estimates imply that 77% of the observed homogamy is due to exposure effects.

Robustness. A challenge for identification is that students who enroll in the same college and in the same year may share more attributes (and hence be more likely to marry) than those who enroll in the same college but different years (Manski, 1993). For instance, if a college admissions criteria or outreach policies change over time, students who start in the same year might be more demographically similar to each other than to students in adjacent cohorts. To address this concern, we estimate a version of the model that includes *pair-level* demographic controls, thereby accounting for observable demographic differences across cohorts. We estimate:

$$P_{ict\,jc't'} \;=\; \alpha_{cc'} + \alpha_{tt'} + \sum_{d=-9}^{10} \beta_d \, \mathbb{1}\{\, c = c',\, t - t' = d \,\} + \gamma X_{ict\,jc't'} + \varepsilon_{ict\,jc't'} \,. \tag{6}$$

where $X_{ict\,jc't'}$ are indicators representing different combinations of parental income quintiles (both students from quintile 1, student i is from quintile 1 and student j is from quintile 2, etc.) and race (both students are white, etc). Figure 6 shows that the results are very similar to those in our baseline specification after the inclusion of these controls.

A related concern is reverse causality, whereby students that were already romantically involved prior to attending college (for instance, because they met during high school) decide to enroll in college together. To address this concern, Figure 6 shows that our results are robust to excluding from the sample pairs of students who are from the same hometown.

Our main specification includes entry-cohort–pair fixed effects. The goal of including these controls is to account for the fact that individuals who enter college in the same year may be more likely to marry each other simply because of assortative mating by age. A potential concern, however, is that individuals who attend the same college and enter in the same year may also have a higher ex-ante propensity to marry than our "control" pairs—those who start college in the same year but at a different college. For example, students who attend the same college in the same year may be more likely to come from the same state or to be from similar socioeconomic backgrounds. To address this concern, we estimate two additional models. First, we include entry-cohort $\times$ state-of-residence-before-college fixed effects. Second, we further refine the specification by including entry-cohort $\times$ state-of-residence-before-college $\times$ parental-income-quintile fixed effects. Figure 6 shows that in all cases the results are nearly identical than those in our baseline specification.

5 Do Colleges Foster Social Mixing or Reinforce Social-Class Homogamy?

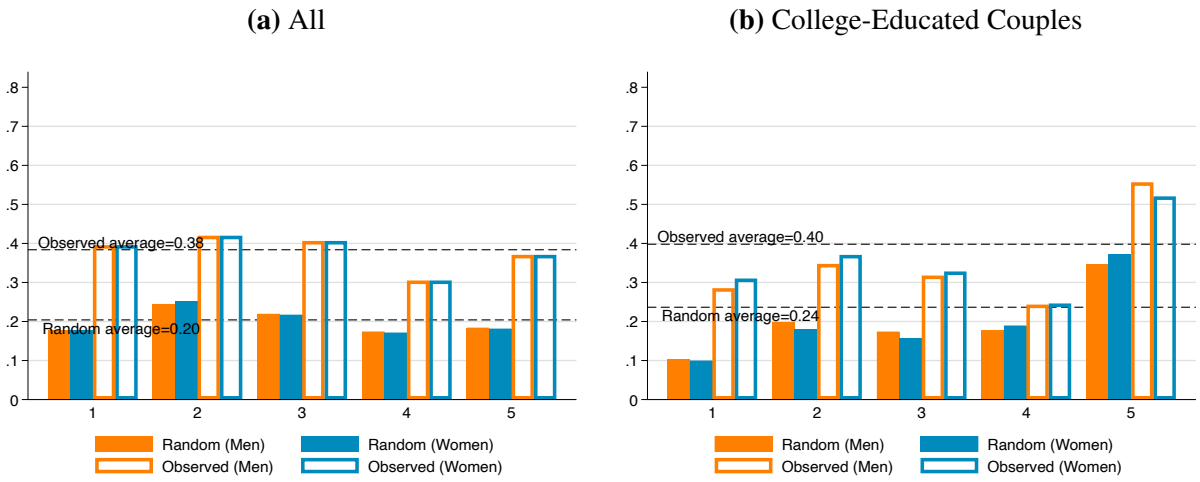
The results above show that colleges serve as a marriage market: attending the same college at the same time has a causal impact on the likelihood that two individuals marry. In this section, we investigate whether colleges serve as a marriage "melting pot" in which individuals from different social classes meet and marry, or whether they mostly provide an opportunity for people from similar socioeconomic backgrounds to meet. To do so, we first show patterns of social class homogamy, both in the general population and among same-college couples. We then compare same-college couples to couples formed outside of college.

5.1 Patterns of Social Class Homogamy

All Marriages. To provide context to our college-specific results, we start by documenting the extent of social class homogamy in the *general* population as well as among all college-educated couples (regardless of whether they attended one of the colleges in our sample). To do so, we use a sample matching 5.7 million married adults ages 30 to 50 in the 1940 full count U.S. census to their 1910 childhood households using links from the Census Tree Project (Price et al., 2021). The sample includes 492,686 college-educated individuals who are married to another college-educated individual. Our focus throughout this section is on parental income quintiles.

Panel (a) of Figure 7 shows the fraction of *all* married individuals who married someone raised at their same parental income quintile, separately based on a person's own parental income quintile and gender. Panel (a) also plots the predicted fraction under random mating, which in this case is close to 20% for all parental income quintiles both for men and women (as the marriage probability is relatively uniform across parental income quintiles in this setting). Among all marriages, about 38% are comprised of individuals who were both raised at the same quintile. This average varies little across quintiles and is about twice as large as what would be expected under random mating.

Figure 7: Social Class Homogamy by Parental Income Quintile in the General Population



Notes: This figure shows the observed proportion of couples in the 1940 census in which both members of the couple grew up in the same parental income quintile, separately by parental income quintile. The figure also shows the predicted such proportion if individuals had the same chance of marriage regardless of parental background. Panel (a) focuses on all married couples, whereas Panel (b) restricts the analysis to couples where both members have at least one year of college.

Panel (b) of Figure 7 instead focuses on couples where both spouses have at least one year of college. Again, we report the fraction of individuals married to someone from their same parental income quintile, as well as the predicted share under random mating. In this case, the predicted share under random mating takes into account the fact that individuals who grew up higher in the income distribution are overrepresented among college-educated individuals. Hence, even if mating was random *within* the college-educated population, men (results are similar for women) from the top 20% would have a higher chance of mating with someone from their same quintile (a 35% chance) than men from the bottom 20% (a 10.4% chance).

Across the parental income distribution, we find that college-educated individuals are more likely to marry someone from their own parental income quintile than what random mating would predict. This finding implies that there is social class homogamy even *within* the college-educated population. For instance, the proportion of college-educated individuals from the top 20% who are married to someone from the top 20% is 54% (versus a 35% random chance). The proportion of college-educated individuals from the bottom 20% who are married to someone from the bottom 20% is also much higher (30%) than what would be predicted under random mating (10.4%).

Within-Colleges Social Class Homogamy. We next investigate patterns of social class homogamy within specific colleges. The probability that a male student from quintile q at college c marries someone from their same quintile can be written as:

$$p_{cq} = s^{w}_{cq} + \Delta^{(m)}_{cq}$$

where s^{w}_{cq} is the share of women in college c who are from quintile q and $\Delta^{(m)}_{cq}$ is:

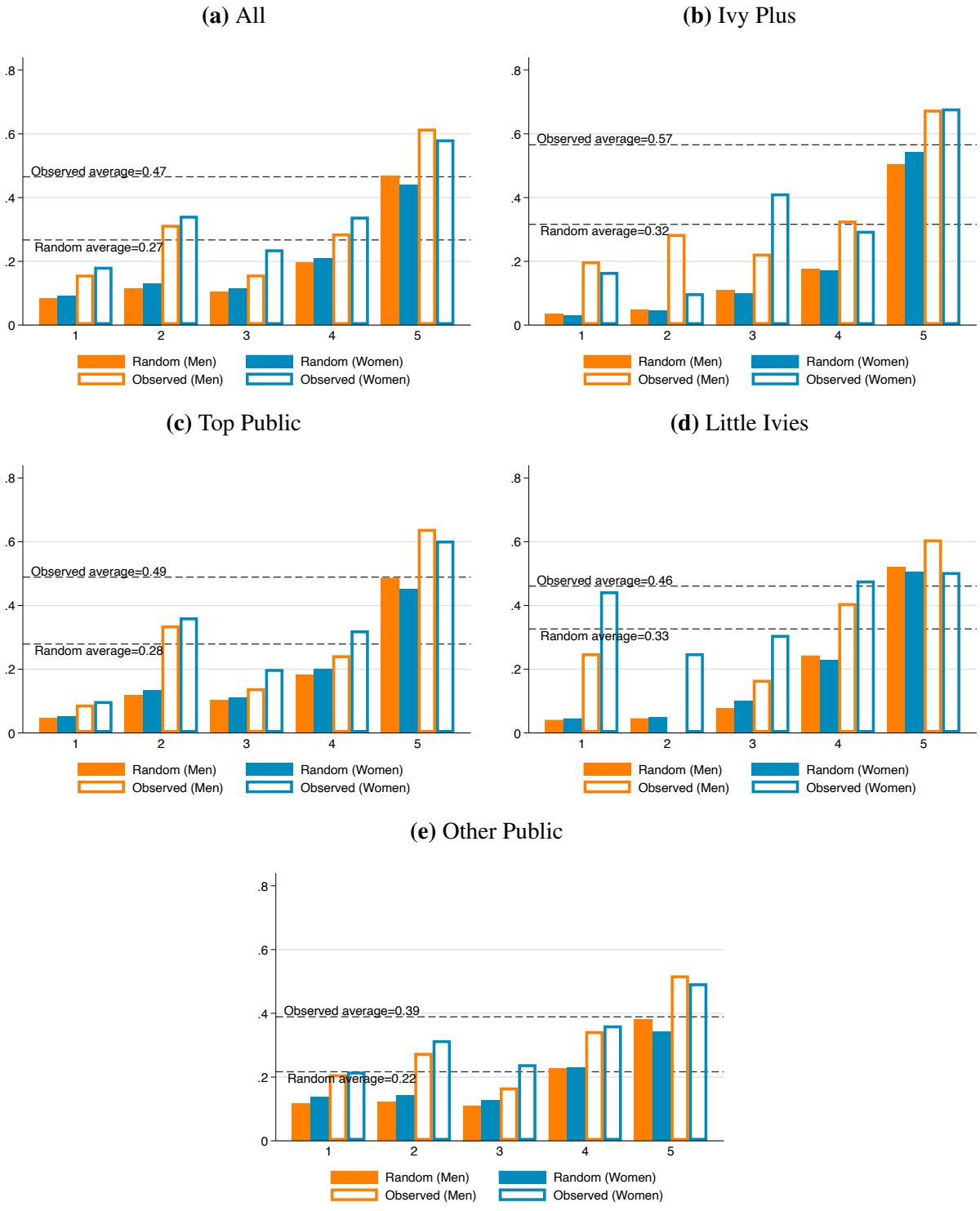
$$\Delta^{(m)}_{cq} = p_{cq} - s^{w}_{cq}$$

The first term s^{w}_{cq} ("compositional effect") captures the fact that colleges are stratified by parental income (Chetty et al., 2020; Abramitzky et al., 2024*a*). Hence, even if mating within a college was random, same-college marriages would tend to be more homogamous than those in the general population. In the most extreme case, if all students in a college were from the same parental income quintile, all same-college marriages would be homogamous even under random mating.

The second term $\Delta^{(m)}_{cq}$ ("within-college sorting") captures that individuals from similar social backgrounds who attend the same college might have a preference toward marrying each other or might be more exposed to each other (for instance, if there is within-college stratification in majors). This term is equal to the difference between the observed probability of a same-quintile marriage and the expected probability if mating were random within a college.

Figure 8 shows the fraction of individuals who married someone from their same parental income quintile, by parental income quintile. The figure also plots the predicted proportions under random within-college mating. Panel (a) pools all colleges together and the remaining panels split the analysis by group of colleges. The sample in this section of the analysis includes the 122,518 individuals who we observe living with a spouse in a later census *and* whose spouse can be matched to a childhood census.

Figure 8: Observed and Predicted Social Class Homogamy, by Parental Income Quintile and Group of Colleges



Notes: This figure shows the observed proportion of same-college in which both members of the couple grew up in the same parental income quintile, separately by parental income quintile. The figure also shows the predicted such proportion if individuals had the same chance of marriage regardless of parental background. Panel (a) focuses on all couples in our college sample. The remaining panels split the analysis by group of colleges.

The figure shows that there is strong *within* college homogamy, with students at each parental income quintile and in each group of colleges more likely to marry within their income quintile than what would be predicted under random mating.

Homogamy rates are particularly high at Ivy Plus colleges, where 57% of same-college marriages are between individuals from the same parental income quintile (compared to 32% if mating were random). For instance, in these colleges only 3% of students are from families in the bottom 20%. Hence, if mating occurred at random, bottom 20% students who marry within their college would only have a 3% chance of marrying another bottom 20% student. Yet, the observed proportion is instead 19%. The chances that students higher in the parental income distribution marry within their income group are also higher than what would be expected at random: 68% of students from top 20% families at the Ivy Plus who are married within their college marry someone from the same income group (compared to a predicted 50% under random mating).

Homogamy rates are the lowest (39%) in the case of of the non-top public colleges. This lower homogamy rate is partly driven by the more balanced social class composition of these colleges: just 22% of couples would be homogamous if mating were random within these colleges.

5.2 Comparing Same-College Marriages to Marriages Formed Outside of College

The previous subsection shows that marriages formed in college are socially homogamous, with students more likely to marry partners from similar socio-economic backgrounds. Yet it remains unclear whether such marriages are *more* socially homogamous than other marriages. In this subsection, we compare same-college marriages to two benchmarks. First, we compare them to marriages in the general population. Second, we compare them to marriages formed by students who attended the same college but married partners outside their college.

Same-College Marriages vs Marriages in the General Population. On average, same-college marriages are more homogamous than marriages in the general population: While among all marriages, 38% are between individuals from the same quintile (40% among all college-educated couples), that proportion is 47% among *same-college* marriages. The higher homogamy rates among same-college marriages are not simply driven by the compositional channel: if mating were random within colleges, only 27% of marriages would be among individuals from the same parental income quintile.

However, we observe notable heterogeneity by social background in the relative degree of social class homogamy. For students in the bottom 80% of the parental income distribution, within-college marriages are generally *less* homogamous than marriages outside of college. For instance, while nearly 40% of individuals from the bottom 20% marry within their group in the general population (30% among the college educated), only 17% do so when considering within-college marriages. By contrast, for students at the top of the income distribution, within-college marriages are more homogamous than other marriages.

Same-College Marriages vs. Other Marriages Among Students. We next compare same-colleges to other marriages among students. We begin with descriptive comparisons between same-college and other marriages, and then use within-college variation in exposure to opposite-sex peers to assess how changes in likelihood of same-college marriages relate to the probability of marrying

a same-quintile spouse. We first estimate:

$$SameQuintileSpouse_{ict} = \alpha_c + \alpha_t + \alpha Male_{ict} + \beta SameCollegeSpouse_{ict} + \gamma X_{ict} + \varepsilon_{ict} \quad (7)$$

where $SameQuintileSpouse_{ict}$ is a dummy that is one if student i married someone from their same parental income quintile, α_c are college fixed effects, and α_t are entry cohort fixed effects. X_{ict} is a vector of students' background characteristics (parental income ranks, parental occupations, and race).

Columns 1 and 2 in Panel (a) of Table 5 show that, on average, same-college marriages are more homogamous than other marriages among students. However, similar to our results above, this average difference masks important heterogeneity. Children from the top 20% (Columns 3 and 4) who marry a same-college spouse are 11 percentage points *more* likely to end up with a spouse of the same parental quintile than those who marry outside their college. By contrast, children from the rest of the parental income distribution (Columns 5 and 6) are 3 percentage points *less* likely to do so.

One possible interpretation of these results is that college introduces students, particularly those from higher-income backgrounds, to a pool of potential partners who are more alike in social background. Another possibility is that students who ultimately marry a college peer differ along unobserved dimensions that also make them more likely to marry within their social class.

To help disentangle these two explanations, we investigate the consequences of greater exposure to opposite-sex peers (which we have already shown increases the likelihood of marrying within the same college). We estimate:

$$SameQuintileSpouse_{ict} = \alpha_c + \alpha_t + \alpha Male_{ict} + \beta \%OppositeSex_{ict} + \gamma X_{ict} + \varepsilon_{ict} \quad (8)$$

Table B5 shows that, after including gender, college, and entry cohort fixed effects, there is still no correlation between students' pre-determined characteristics and the share of opposite-sex students in the cohort in this sample (the sample is limited to married individuals whose spouses can be matched to a childhood census).

Panel (b) of Table 5 shows the results of estimating equation 8 in the full sample of colleges. Columns 1 and 2 pool all students together, whereas the remaining columns split the sample by parental income quintile (students from the top 20% of the parental income distribution, which represent roughly half of the sample, and the rest). On average, there is little evidence that higher exposure to opposite-sex students shapes the probability of same-quintile marriages.

Table 6 repeats the analysis but now splitting the sample between students at private (panel (a)) and students at public colleges (panel (b)). This table reveals that the results pooling all colleges mask important heterogeneity. Panel (a) shows that, in private colleges, a higher proportion of opposite-sex students in the cohort *increases* the chances of a same-quintile marriage among top 20% students: increasing exposure to opposite-sex students from 25% to 50% would increase the probability of a same-quintile spouse by 1 percentage points (relative to a mean of 66%). By contrast, among students below the top 20%, a similar increase actually *decreases* the probability of a same quintile spouse by 1.7 percentage points (relative to a mean of 31.4%). The explanation for this pattern is straightforward: Because there is a high concentration of top 20% students at private colleges, increasing the probability of same-college marriages increases the chances that top 20% students marry within their group but decreases the chances that other students do so.

Table 5: Same-College Couples and Likelihood of Same-Quintile Spouse

(a) Same-College Spouse

	All		Top 20%		Rest	
	(1)	(2)	(3)	(4)	(5)	(6)
Same-College Spouse	0.0720***	0.0495***	0.113***	0.109***	-0.0349***	-0.0337***
	(0.0114)	(0.00927)	(0.0113)	(0.0109)	(0.00973)	(0.00886)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	109834	109834	53998	53998	55836	55836
Mean of Dep. Var.	0.454	0.454	0.579	0.579	0.333	0.333

(b) Share of Opposite-Sex Students in Cohort

	All		Top 20%		Rest	
	(1)	(2)	(3)	(4)	(5)	(6)
% Opposite-Sex	-0.0212	-0.0112	-0.00534	-0.00383	-0.0128	-0.00629
	(0.0218)	(0.0165)	(0.0216)	(0.0202)	(0.0163)	(0.0163)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	109834	109834	53998	53998	55836	55836
Mean of Dep. Var.	0.454	0.454	0.579	0.579	0.333	0.333

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. The outcome variable is an indicator equal to one if the student is married to a spouse in the same quintile of the parental income distribution. *% Opposite-Sex* ranges from 0 to 1 and measures the proportion of individuals in a student college cohort (defined as those starting in the same college and in the same year) who are of the opposite sex. All columns include gender, college, and entry cohort fixed effects. The odd columns further control for father's income rank, indicators for father's occupation (professional, white-collar non-professional, farmer, skilled blue collar, with unskilled as the omitted category), and an indicator for being White. Standard errors are clustered at the college level.

Table 6: Share of Opposite-Sex Students and Likelihood of Same-Quintile Spouse, Private vs Public Colleges

(a) Private

	All		Top 20%		Rest	
	(1)	(2)	(3)	(4)	(5)	(6)
% Opposite-Sex	-0.0971***	-0.0304	0.0308	0.0388*	-0.0778**	-0.0689**
	(0.0271)	(0.0241)	(0.0250)	(0.0198)	(0.0268)	(0.0284)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	15324	15324	8974	8974	6350	6350
Mean of Dep. Var.	0.514	0.514	0.656	0.656	0.314	0.314

(b) Public

	All		Top 20%		Rest	
	(1)	(2)	(3)	(4)	(5)	(6)
% Opposite-Sex	0.00874	0.00873	0.0104	0.00959	-0.00573	0.00327
	(0.0175)	(0.0150)	(0.0216)	(0.0213)	(0.0193)	(0.0189)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	94510	94510	45024	45024	49486	49486
Mean of Dep. Var.	0.444	0.444	0.563	0.563	0.336	0.336

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. The outcome variable is an indicator equal to one if the student is married to a spouse in the same quintile of the parental income distribution. *% Opposite-Sex* ranges from 0 to 1 and measures the proportion of individuals in a student college cohort (defined as those starting in the same college and in the same year) who are of the opposite sex. Even columns control for father's income rank, indicators for father's occupation (professional, white-collar non-professional, farmer, skilled blue collar, with unskilled as the omitted category), and an indicator for being White. Panel (a) restricts the sample to students in private colleges, whereas Panel (b) restricts the sample to students in public colleges. Standard errors are clustered at the college level.

In contrast, the effects are weaker in public colleges. A likely explanation for the private-public contrast is that, in our sample, public colleges have a more balanced social-class mix than the private ones (Figure B3). Hence, increasing the probability of same-college marriages leads to matches that are not that dissimilar to those that would occur outside of college.

Robustness. Our finding that, in private colleges, a higher share of opposite-sex students decreases the chances of same-income quintile marriages for lower income students but increases it for those at the top is robust to the same set of robustness exercises as those we implemented in section 4.1. In particular, Table B6 shows that we obtain similar results if we (1) exclude the 1941-1949 cohorts, (2) include college-specific linear trends, and (3) include college $\times$ five-year entry cohort bins fixed effects.

6 Intergenerational Consequences: Human Capital in the Next Generation

In the last section of the paper, we investigate the intergenerational consequences of same-college marriages. Our main result in this section is that the *children* of same-college couples are more likely to attend a selective college than the children of couples formed outside of college. We then explore potential mechanisms behind this main result, finding evidence that individuals with a same-college spouse marry higher human capital spouses than those who marry outside of their college. Throughout this section, we begin with descriptive comparisons between same-college and other marriages, and then use within-college variation in exposure to opposite-sex peers to assess how changes in likelihood of same-college marriages relate to intergenerational outcomes.

6.1 Children Educational Outcomes

To investigate whether same-college couples produce children with higher human capital, we first estimate:

$$ChildEducation_{jict} = \alpha_c + \alpha_t + \alpha Male_{ict} + \beta SameCollegeSpouse_{ict} + \gamma X_{ict} + \delta Z_{jict} + \varepsilon_{jict} \quad (9)$$

where $ChildEducation_{jict}$ is the educational attainment of child j of student i who attended college c starting in year t. Here, both the male dummy $Male_{ict}$ and the vector of background characteristics (X_{ict}) refer to the *original* student. The vector Z_{jict} includes gender and birth year fixed effects for child j.

Our measure of children's education is whether the children themselves are matched to one of the 95 colleges in our extended sample (either the same college of their parents or a different one). We estimate that these colleges account for 19% of all individuals with at least one year of college in the 1915 to 1950 birth cohorts, and an even higher proportion of those attending a selective college.[15] We focus on this measure (rather than educational attainment from the census) because the 1950 census is the last census that is currently available and our sample starts in 1915. Hence,

[15] Pooling all data in the current paper and Abramitzky et al. (2024*a*), our sample includes 3.6 million students who attended college between 1933 and 1966. Using the 1970 Census, we estimate that 20 million US-born Whites in the 1915 to 1948 birth cohorts obtained at least one year of college.

relatively few of the *children* of the original students had reached young adulthood by 1950.[16] By contrast, focusing on attendance to one of the colleges in our sample enables us to observe children's educational outcomes up to 1966. To be consistent with our main sample restrictions, we restrict the analysis to children who would have been at least 27 by 1966 (so as to allow enough time for children to start college). Because we require the *children* of the original students to be at least 27 by 1966, our sample is mostly comprised of the descendants of students who attended college before 1940 (Figure A8). In total, our sample includes 315,979 children corresponding to 153,572 parents.

Panel (a) of Table 7 shows that children of same-college couples are 5.9 percentage points more likely to attend the same college as their parents, a large difference relative to a mean of 5.0%. Children of same-college couples are also 1.1 percentage points more likely (relative to a mean of 5.1%) to be observed attending *other* colleges in our data. Hence, our findings do not simply reflect a higher preference for attending the same college as their parents but rather an increase in the probability of attending a selective college overall.

In interpreting these magnitudes, it is important to bear in mind that *not all* children are matched to a childhood census. Specifically, given that we match 54% of students to a childhood census, the dependent variable should be re-scaled by a factor of about two.

We next examine how exposure to opposite-sex peers among *parents* affects the human capital of their *children*.[17] To do so, we estimate:

$$ChildEducation_{jict} = \alpha_c + \alpha_t + \alpha Male_{ict} + \beta \%OppositeSex_{ict} + \gamma X_{ict} + \delta Z_{jict} + \varepsilon_{jict} \qquad (10)$$

Panel (b) of Table 7 shows that greater exposure to opposite-sex peers increases the likelihood that students have children who later attend the same college as them. Increasing the proportion of opposite-sex students from 25% to 50% would increase the probability that the child attends the same college as their parent by 0.5 percentage points (relative to a mean of 5.0%). Increasing the proportion of opposite-sex students from 25% to 50% also increases the chance of attending *other* colleges in our data (by 0.2 percentage points, relative to a mean of 5.1%), again implying that our results likely capture an actual increase in the probability of attending a selective college.

Table B8 investigates the heterogeneity of these results on the basis of: (1) child gender, (2) parent gender, (3) whether the parent attended a public or a private college, and (4) whether the parent grew up in a family in the bottom 80% or the top 20%. The outcome variable in this table is whether the child is observed attending any of the colleges in the data (either the same as their parent or a different one).

Columns 1 and 2 show that these effects are concentrated among the *daughters* of the original students. Raising the proportion of opposite-sex peers in the parents' cohort from 25% to 50% increases the probability that their daughters attend one of the colleges in our sample by 1.1 percentage points, relative to a mean of 8.4

The remaining columns split the sample by the gender of the original student (columns 3 and 4), whether the student attended a private versus a public college (columns 5 and 6), and whether the student grew up in a family in the bottom 80% or the top 20% of the parental income distribution

[16] Furthermore, the education question was a sample line question in 1950 so it is available for only 20% of respondents, further reducing sample sizes.

[17] Table B7 shows that in this intergenerational sample there is also little correlation between the background characteristics of the original students and the share of opposite-sex students in their cohort.

Table 7: Same-College Couples and Likelihood of Having a College-Educated Child

(a) Same-College Spouse

	Same College as Parents		Other College		Any College	
	(1)	(2)	(3)	(4)	(5)	(6)
Same-College Spouse	0.0609***	0.0590***	0.0142***	0.0113***	0.0751***	0.0703***
	(0.00445)	(0.00421)	(0.00269)	(0.00260)	(0.00491)	(0.00459)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	315979	315979	315979	315979	315979	315979
Mean of Dep. Var.	0.0499	0.0499	0.0513	0.0513	0.101	0.101

(b) Share of Opposite-Sex Students in Cohort

	Same College as Parents		Other College		Any College	
	(1)	(2)	(3)	(4)	(5)	(6)
% Opposite-Sex	0.0195***	0.0197***	0.00829**	0.00847**	0.0278***	0.0281***
	(0.00670)	(0.00683)	(0.00408)	(0.00394)	(0.00930)	(0.00944)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	315979	315979	315979	315979	315979	315979
Mean of Dep. Var.	0.0499	0.0499	0.0513	0.0513	0.101	0.101

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. The outcome variable in columns 1 to 2 is an indicator that is one if the child attended the same college as their parent. The outcome variable in columns 3 to 4 is an indicator that is one if the child attended one of the 95 colleges in our extended sample other than the one attended by their parents. The outcome variable in columns 5 to 6 is an indicator that is one if the child attended any of the 95 colleges in our extended sample. *Same-College Spouse* is an indicator that is one if the student married another student from their same college. *% Opposite-Sex* ranges from 0 to 1 and measures the proportion of individuals in a student college cohort (defined as those starting in the same college and in the same year) who are of the opposite sex. All columns control for children's gender and birth year fixed effects. Even columns control for father's income rank, indicators for father's occupation (professional, white-collar non-professional, farmer, skilled blue collar, with unskilled as the omitted category), and an indicator for being White. Standard errors are clustered at the college level.

(Columns 7 and 8). The estimates in columns 1 to 6 are less precise and show modest differences across groups. Columns 7 and 8 indicate that the increase in children college attainment is present both for those who grew up at the bottom and those who grew up at the top of the parental income distribution.

Finally, Table B9 shows that our main results on children's educational attainment are robust to a similar set of sample selection and specification checks as the ones conducted in previous subsections.

6.2 Why Are Children of Same-College Couples More Educated?

We next investigate the potential mechanisms behind the higher probability of attending a selective college of children of same-college couples, focusing on two main explanations. First, we assess whether this is simply because these couples have fewer children, consistent with a quantity–quality tradeoff. Second, we examine whether it is because marrying within the college allows individuals to match with more educated spouses, whose human capital is transmitted to the next generation. Our findings are more consistent with the latter mechanism.

Fertility Outcomes. To explore the quality-quantity tradeoff hypothesis, we estimate:

$$Fertility_{ict} = \alpha_c + \alpha_t + \alpha Male_{ict} + \beta SameCollegeSpouse_{ict} + \gamma X_{ict} + \varepsilon_{ict} \tag{11}$$

where $Fertility_{ict}$ are different measures of fertility (the number of children, whether the student had at least one child) for student i from college c and entry cohort t. $SameCollegeSpouse_{ict}$ is an indicator that is one if the student married someone who also attended college c. X_{ict} is a vector of students' background characteristics (parental income ranks, parental occupations, and race). We restrict the sample to those individuals that we can match to an adult census (regardless of their marital status in adulthood).

Panel (a) of Table 8 shows that individuals married to a spouse from the same college have *more* children and are *more* likely to have at least one children than other students at the same college. Most of this difference is mechanical, as those married to someone from their same college have a higher likelihood of marrying *any* spouse: Table B10 shows much smaller differences once we restrict the sample to married individuals.

Panel (b) of Table 8 examines how exposure to opposite-sex peers affects fertility outcomes. We find little evidence that a higher share of opposite-sex students increases either the number of children or the probability of having at least one child. Thus, while greater exposure to opposite-sex students raises the likelihood of marrying within the college, its effects on fertility appear much more limited. This evidence suggests that the higher educational attainment in the next generation is not due to a quality-quantity tradeoff.

Spousal Human Capital. To investigate the role of spousal human capital, we estimate:

$$SpousalEducation_{ict} = \alpha_c + \alpha_t + \alpha Male_{ict} + \beta SameCollegeSpouse_{ict} + \gamma X_{ict} + \varepsilon_{ict} \tag{12}$$

where $SpousalEducation_{ict}$ are different measures (years of schooling, whether the person completed at least one year of college, etc.) of the education of the *spouse* of student i from college c with entry cohort t. We restrict the sample to individuals who are married and whose spouses can be matched to either the 1940 or the 1950 censuses (so that we can observe their educational attainment).

Table 8: Fertility Outcomes of Same-College Couples

(a) Same-College Spouse

	Number of children			1+ Children		
	(1)	(2)	(3)	(4)	(5)	(6)
Same-College Spouse	0.684***	0.521***	0.533***	0.327***	0.268***	0.269***
	(0.0217)	(0.0198)	(0.0195)	(0.00680)	(0.00590)	(0.00591)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	No	Yes	Yes	No	Yes	Yes
Entry Cohort FEs	No	Yes	Yes	No	Yes	Yes
Individual Level Controls	No	No	Yes	No	No	Yes
Observations	479284	479284	479284	479284	479284	479284
Mean of Dep. Var.	1.196	1.196	1.196	0.573	0.573	0.573

(b) Share of Opposite-Sex Students in Cohort

	Number of children			1+ Children		
	(1)	(2)	(3)	(4)	(5)	(6)
% Opposite-Sex	-0.0249	-0.0128	-0.0130	0.0266	-0.0186	-0.0182
	(0.130)	(0.0434)	(0.0429)	(0.0488)	(0.0196)	(0.0195)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	No	Yes	Yes	No	Yes	Yes
Entry Cohort FEs	No	Yes	Yes	No	Yes	Yes
Individual Level Controls	No	No	Yes	No	No	Yes
Observations	479284	479284	479284	479284	479284	479284
Mean of Dep. Var.	1.196	1.196	1.196	0.573	0.573	0.573

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. The outcome variable in columns 1 to 3 is the number of children of the student. The outcome variable in columns 4 to 6 is an indicator that is one if the student has at least one child. *Same-College Spouse* is an indicator that is one if the student married another student from their same college. *% Opposite-Sex* ranges from 0 to 1 and measures the proportion of individuals in a student college cohort (defined as those starting in the same college and in the same year) who are of the opposite sex. Columns 3 and 6 control for father's income rank, indicators for father's occupation (professional, white-collar non-professional, farmer, skilled blue collar, with unskilled as the omitted category), and an indicator for being White. Standard errors are clustered at the college level.

Table 9: Same-College Spouses and Spousal Education

(a) Same-College Spouse

	Years of Schooling		1+ Years of College		4+ Years of College	
	(1)	(2)	(3)	(4)	(5)	(6)
Same-College Spouse	2.551***	2.415***	0.418***	0.398***	0.462***	0.449***
	(0.0536)	(0.0533)	(0.00742)	(0.00726)	(0.0157)	(0.0163)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	216761	216761	216761	216761	216761	216761
Mean of Dep. Var.	12.61	12.61	0.519	0.519	0.288	0.288

(b) Share of Opposite-Sex Students in Cohort

	Years of Schooling		1+ Years of College		4+ Years of College	
	(1)	(2)	(3)	(4)	(5)	(6)
% Opposite-Sex	0.684***	0.687***	0.127***	0.126***	0.163***	0.162***
	(0.156)	(0.137)	(0.0206)	(0.0189)	(0.0234)	(0.0212)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	216761	216761	216761	216761	216761	216761
Mean of Dep. Var.	12.61	12.61	0.519	0.519	0.288	0.288

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. The outcome variable in columns 1 to 2 is the years of schooling of the student's spouse. The outcome variable in columns 3 to 4 is an indicator that is one if the student's spouse has at least one year of college. The outcome variable in columns 5 to 6 is an indicator that is one if the student's spouse has four or more years of college. *Same-College Spouse* is an indicator that is one if the student married another student from their same college. *% Opposite-Sex* ranges from 0 to 1 and measures the proportion of individuals in a student college cohort (defined as those starting in the same college and in the same year) who are of the opposite sex. Columns 3 and 6 control for father's income rank, indicators for father's occupation (professional, white-collar non-professional, farmer, skilled blue collar, with unskilled as the omitted category), and an indicator for being White. Standard errors are clustered at the college level.

Panel (a) in Table 9 shows that individuals married to another student from the same college have spouses with more years of schooling, a higher likelihood of having at least one year of college, and a higher likelihood of having at least four years of college than individuals who married outside of their college. The difference in years of schooling is substantial: individuals with same-college spouses have partners with 2.4 more years of schooling than those married outside their college. The differences in spousal educational attainment are of similar magnitude when we control for students' background characteristics.

One explanation for these patterns is that marrying within one's college allows individuals to match with more educated partners than they otherwise would. An alternative interpretation is selection: Individuals who marry a spouse from their same college might have unobserved attributes that increase the likelihood that they marry another highly educated individual. To help disentangle these two explanations, we again exploit within-college across-cohort variation in exposure to opposite-sex students. We estimate:

$$SpousalEducation_{ict} = \alpha_c + \alpha_t + \alpha Male_{ict} + \beta \%OppositeSex_{ict} + \gamma X_{ict} + \varepsilon_{ict} \qquad (13)$$

Panel (b) in Table 9 shows that a higher share of opposite-sex students in the cohort increases spousal education. Increasing the share of opposite-sex students from 25% to 50% increases spousal years of schooling by 0.2 years, the likelihood of having a college-educated spouse by 3.2 percentage points, and the likelihood of having a spouse with four or more years of schooling by 4.1 percentage points. These are all substantial effects relative to the dependent variable means.

Table 10 shows the heterogeneity of this main result by by students' gender (Columns 1 and 2), whether the student attends a public or a private college (Columns 3 and 4), and whether the students belongs to a family in the bottom 80% or top 20% of the parental income distribution (Columns 5 and 6). Panel (a) focuses on years of schooling whereas panel (b) focuses on the likelihood of having four or more years of college.

The table shows that the increase in spousal education is especially pronounced among women. Increasing the share of opposite-sex students from 25% to 50% increases spousal years of schooling by .3 years, the likelihood of having a spouse with at least one year of college by 4.1 percentage points, and the likelihood of having a spouse with at least four years of college by 4.1 percentage points. For men, a similar increase in the share of opposite-sex students increases the probability of having a spouse with at least four years of college by 2.2 percentage points. The increase in spousal education are also stronger for students from top 20% families compared to students from the bottom 80%. By contrast, the differences between private and public colleges are in general more modest. Table B11 shows that these results are robust to (1) excluding students who started college in 1941 or after, (2) including college-specific linear trends, and (3) including college$\times$5-year entry cohort fixed effects.

Overall, this evidence is consistent with same-college marriages increasing human capital in the next generation by allowing students to match with more educated spouses.

7 Conclusions

In this paper, we studied the role of colleges in the marriage market and its implications for intergenerational inequality. To do so, we used newly digitized student records spanning 1915 to 1949 from 63 U.S. co-educational colleges matched to population censuses.

Table 10: Share of Opposite-Sex Students in Cohort and Spousal Education, Heterogeneity

(a) Years of Schooling

	Gender		Public/Private		Parental Income Rank	
	(1) Women	(2) Men	(3) Public	(4) Private	(5) Bottom 80%	(6) Top 20%
% Opposite-Sex	1.232***	0.349	0.659***	0.797**	0.494***	1.026***
	(0.425)	(0.310)	(0.169)	(0.270)	(0.157)	(0.159)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	47280	169481	182886	33875	109229	107532
Mean of Dep. Var.	13.21	12.44	12.62	12.51	12.08	13.14

(b) 4+Years of College

	Gender		Public/Private		Parental Income Rank	
	(1) Women	(2) Men	(3) Public	(4) Private	(5) Bottom 80%	(6) Top 20%
% Opposite-Sex	0.163***	0.0871***	0.170***	0.112***	0.130***	0.207***
	(0.0484)	(0.0280)	(0.0247)	(0.0329)	(0.0220)	(0.0257)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	47280	169481	182886	33875	109229	107532
Mean of Dep. Var.	0.427	0.249	0.289	0.280	0.236	0.340

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. Each column in this table corresponds to a different sample, as indicated by the column headings. *Same-College Spouse* is an indicator that is one if the student married another student from their same college. *% Opposite-Sex* ranges from 0 to 1 and measures the proportion of individuals in a student college cohort (defined as those starting in the same college and in the same year) who are of the opposite sex. All columns control for father's income rank, indicators for father's occupation (professional, white-collar non-professional, farmer, skilled blue collar, with unskilled as the omitted category), and an indicator for being White. Standard errors are clustered at the college level.

We first documented a striking degree of same-college homogamy, with about 5 percent of men and 8 percent of women marrying someone from their own college. This homogamy is far stronger than homogamy based merely on whether one attended college at all.

We then show that the observed college homogamy reflects a causal effect of colleges. First, exploiting within-college variation in the share of opposite-sex students across cohorts, we show that exposure to a higher proportion of opposite-sex partners during college increases the likelihood of marrying a same-college spouse. Second, comparing the marriage probabilities of students who overlap in college to the marriage probabilities of individuals who (barely) did not overlap, we show that most of the observed college homogamy reflects a causal on-campus exposure effect.

We next explore how same-college marriages shape assortative mating on the basis of social class. We find that same-college couples are strongly assortative on the basis of social class: Even within the same institution, students are more likely to marry students from similar social backgrounds. Same-college marriages are, on average, more socially homogamous than marriages formed outside of college. In private colleges, a higher share of opposite-sex peers *increases* the probability of socially homogamous marriages for students in the top 20% of the distribution but *lowers* it for other students.

In the last section of the paper we investigate the intergenerational consequences of same-college marriages. Children of same-college couples are more likely to attend a selective college than children of couples formed outside of college. A potential explanation for this finding is that students who marry within their college might end up with more educated spouses which, in turn, increases human capital in the next generation. Consistent with this explanation, we find that individuals who marry within their college indeed have higher human capital spouses.

Overall, our results show that colleges meaningfully shape who marries whom and how social status is transmitted across generations. Same-college marriages are substantially more socially homogamous, particularly among students from high-income families. These marriages have important intergenerational consequences: the children of same-college couples exhibit higher levels of human capital than children whose parents do not marry within their institution.

References

Abramitzky, Ran, Adeline Delavande, and Luis Vasconcelos. 2011. "Marrying up: the role of sex ratio in assortative matching." *American Economic Journal: Applied Economics*, 3(3): 124–157.

Abramitzky, Ran, Jennifer K Kowalski, Santiago Pérez, and Joseph Price. 2024*a*. "The GI Bill, Standardized Testing, and Socioeconomic Origins of the US Educational Elite Over a Century." National Bureau of Economic Research.

Abramitzky, Ran, Jennifer K Kowalski, Santiago Pérez, and Joseph Price. 2025. "Beyond Parents: College Legacy in the U.S. Over a Century." National Bureau of Economic Research.

Abramitzky, Ran, Jennifer Kowalski, Santiago Pérez, and Joseph Price. 2024*b*. "Socioeconomic Origins of the US Educational Elite: 1915-2013." Work in progress.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2014. "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration." *Journal of Political Economy*, 122(3): 467–506.

Abramitzky, Ran, Leah Platt Boustan, K Eriksson, S Pérez, and M Rashid. 2020. "Census linking project."

Angrist, Josh. 2002. "How do sex ratios affect marriage and labor markets? Evidence from America's second generation." *The Quarterly Journal of Economics*, 117(3): 997–1038.

Arum, Richard, Josipa Roksa, and Michelle J Budig. 2008. "The romance of college attendance: Higher education stratification and mate selection." *Research in Social Stratification and Mobility*, 26(2): 107–121.

Bailey, Martha J, and Susan M Dynarski. 2011. "Gains and gaps: Changing inequality in US college entry and completion." National Bureau of Economic Research.

Barrios Fernández, Andrés, Christopher Neilson, and Seth D Zimmerman. n.d.. "Elite universities and the intergenerational transmission of human and social capital."

Bayer, Patrick, Stephen L Ross, and Giorgio Topa. 2008. "Place of work and place of residence: Informal hiring networks and labor market outcomes." *Journal of Political Economy*, 116(6): 1150–1196.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2013. "Under pressure? The effect of peers on outcomes of young adults." *Journal of Labor Economics*, 31(1): 119–153.

Bleemer, Zachary, and Sarah Quincy. 2025. "Changes in the college mobility pipeline since 1900." National Bureau of Economic Research.

Bostwick, Valerie K, and Bruce A Weinberg. 2022. "Nevertheless she persisted? Gender peer effects in doctoral STEM programs." *Journal of Labor Economics*, 40(2): 397–436.

Bowen, William G., and Derek Bok. 1998. *The Shape of the River: Long-Term Consequences of Considering Race in College and University Admissions.* Princeton University Press.

Bratsberg, Bernt, Simen Markussen, Oddbjørn Raaum, Knut Røed, and Ole Røgeberg. 2023. "Trends in assortative mating and offspring outcomes." *The Economic Journal*, 133(651): 928–950.

Buckles, Kasey, Adrian Haws, Joseph Price, and Haley EB Wilbert. 2023. "Breakthroughs in Historical Record Linking Using Genealogy Data: The Census Tree Project." National Bureau of Economic Research.

Chadwick, Laura, and Gary Solon. 2002. "Intergenerational income mobility among daughters." *American Economic Review*, 92(1): 335–344.

Chetty, Raj, David Deming, and John Friedman. 2023. "Diversifying Society's Leaders? The Determinants and Causal Effects of Admission to Highly Selective Private Colleges." National Bureau of Economic Research, Inc NBER Working Papers 31492.

Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2020. "Income Segregation and Intergenerational Mobility Across Colleges in the United States." *The Quarterly Journal of Economics*, 135(3): 1567–1633.

Eriksson, Katherine, Gregory T Niemesh, Myera Rashid, and Jacqueline Craig. 2023. "Marriage and the Intergenerational Mobility of Women: Evidence from Marriage Certificates 1850-1920." Working Paper. https://niemesgt. github. io/files/WomenMobilityMarriage. pdf.

Ge, Suqin, Elliott Isaac, and Amalia Miller. 2022. "Elite schools and opting in: Effects of college selectivity on career and family outcomes." *Journal of Labor Economics*, 40(S1): S383–S427.

Goldman, Benjamin, Jamie Gracie, and Sonya R Porter. 2024. "Who Marries Whom? The Role of Segregation by Race and Class."

Greene, H., and M.W. Greene. 2001. *Greenes' Guide to Educational Planning:The Public Ivies. Greenes' Guides to Educational Planning*, HarperCollins.

Grosjean, Pauline, and Rose Khattar. 2019. "It's raining men! Hallelujah? The long-run consequences of male-biased sex ratios." *The Review of Economic Studies*, 86(2): 723–754.

Hill, Andrew J. 2017. "The positive influence of female college students on their male peers." *Labour Economics*, 44: 151–160.

Hoxby, Caroline M. 2000. "Peer effects in the classroom: Learning from gender and race variation."

Hoxby, Caroline M, and Christopher Avery. 2012. "The missing" one-offs": The hidden supply of high-achieving, low income students." National Bureau of Economic Research.

Kalmijn, Matthijs. 1998. "Intermarriage and homogamy: Causes, patterns, trends." *Annual review of sociology*, 24(1): 395–421.

Kaufmann, Katja Maria, Matthias Messner, and Alex Solis. 2013. "Returns to elite higher education in the marriage market: Evidence from Chile." *Available at SSRN 2313369*.

Kirkebøen, Lars, Edwin Leuven, and Magne Mogstad. 2021. "College as a marriage market." National Bureau of Economic Research.

Lavy, Victor, and Analia Schlosser. 2011. "Mechanisms and impacts of gender peer effects at school." *American Economic Journal: Applied Economics*, 3(2): 1–33.

Liu, Haoming, and Jingfeng Lu. 2006. "Measuring the degree of assortative mating." *Economics Letters*, 92(3): 317–322.

Manski, Charles F. 1993. "Identification of endogenous social effects: The reflection problem." *The review of economic studies*, 60(3): 531–542.

Olivetti, Claudia, M Daniele Paserman, Laura Salisbury, and E Anna Weber. 2024. "The emergence of the socioeconomic gradient in women's marriage outcomes." *Review of Economics and Statistics*, 1–29.

Price, Joseph, Kasey Buckles, Jacob Van Leeuwen, and Isaac Riley. 2021. "Combining family history and machine learning to link historical records: The Census Tree data set." *Explorations in Economic History*, 80: 101391.

Reiter, Andrew G. 2024. "Historical U.S. News & World Report College Rankings Dataset." `https://www.andyreiter.com/datasets/`, Accessed: 2025-01-09. Contains historical USNWR rankings for National Universities (1984–2026) and Liberal Arts Colleges (1985–2026).

Rosenfeld, Michael J., Reuben J. Thomas, and Sonia Hausen. 2024. "How Couples Meet and Stay Together (HCMST) 2017, 2020, 2022, United States." Version released 2024-02-07.

Ruggles, Steven, Catherine Fitch, Ron Goeken, J David Hacker, Jonas Helgertz, Evan Roberts, Matt Sobek, Kelly Thompson, John Robert Warren, and Jacob Wellington. 2019. "IPUMS Multigenerational Longitudinal Panel."."

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2020. "Ipums usa: Version 10.0." *(No Title)*.

Schwartz, Christine R. 2013. "Trends and variation in assortative mating: Causes and consequences." *Annual Review of Sociology*, 39(1): 451–470.

U.S. Office of Education. 1953. "Biennial Survey of Education, 1948–1950." U.S. Department of the Interior, Office of Education, Washington, D.C. Government Printing Office.

Online Appendix - Not for Publication

A Data Appendix

A.1 Calculating Fraction of Same-College Couples

The challenge in calculating the true fraction of same-college couples is that we only observe whether someone is married to someone from their own college if we: (1) match them and their spouse to a childhood census, (2) match them and their spouse from a childhood to an adult census where they are observed co-residing.

To address this issue, first note that in the data we can directly measure:

$$Pr(\text{Observed with Same-College Spouse})$$

that is, the fraction of individuals who are *observed* living with a spouse who attended the same college in the census. Using Bayes rule, this observed probability can be written as:

$$Pr(\text{Same-College Spouse}) \times Pr(\text{Observed with Same-College Spouse}/\text{Same-College Spouse})$$

Hence, the share of individuals married to someone from their own college can be written as:

$$Pr(\text{Same-College Spouse}) = \frac{Pr(\text{Observed with Same-College Spouse})}{Pr(\text{Observed with Same-College Spouse}/\text{Same-College Spouse})}$$

Because we cannot observe whether someone's spouse attended their same college unless we match them to the census, the denominator in this expression is not directly observed in the data. We can approximate this denominator using:

$$Pr(\text{Observing Spouse}/\text{Married})$$

This expression represents the probability of observing someone's spouse in a future census, conditional on this person being married. For men, such probability is equal to the probability of matching *married* women to at least one post-college census. For women, it is equal to the probability of matching *married* men to at least one post-college census.

Our estimate of the proportion of individuals married to someone from their own college is then:

$$\frac{\%\text{ Observed with Same-College Spouse}}{\%\text{ Opposite-Sex Married Matched to 1+ Post-College Census}}$$

A.2 Data Collection

We describe the collection of historical student registers in Abramitzky et al. (2024*b*), which details (1) the selection of colleges, (2) acquisition and compilation of PDFs with student names, (3) transcription, (4) identification of undergraduates, (5) data standardization, and (6) linkage to U.S. population censuses.

Table A1: List of Co-Educational Colleges (1915-1949 Sample Size)

College	College Group	Sample Size
Stanford University	Ivy Plus	43,760
University of Chicago	Ivy Plus	27,483
Cornell University	Ivy Plus	60,955
University of Pennsylvania	Ivy Plus	87,626
University of Arizona	Top Public	13,436
University of California, Berkeley	Top Public	123,350
University of California, Davis	Top Public	3,416
University of California, Los Angeles	Top Public	78,440
University of California, Santa Barbara	Top Public	2,800
University of Colorado Boulder	Top Public	51,040
University of Connecticut	Top Public	9,186
University of Delaware	Top Public	7,229
University of Georgia	Top Public	34,236
University of Illinois Urbana-Champaign	Top Public	49,063
Indiana University Bloomington	Top Public	44,474
University of Iowa	Top Public	21,800
University of Maryland, College Park	Top Public	11,876
Michigan State University	Top Public	26,887
University of Michigan	Top Public	63,437
State University of New York (SUNY)	Top Public	994
University of Miami	Top Public	27,625
The Ohio State University	Top Public	34,243
Pennsylvania State University	Top Public	36,378
University of Pittsburgh	Top Public	33,038
The University of Texas at Austin	Top Public	37,476
University of Vermont	Top Public	15,697
William & Mary	Top Public	17,070
University of Virginia	Top Public	23,949
University of Washington	Top Public	27,800
University of Wisconsin–Madison	Top Public	42,104
University of Minnesota, Twin Cities	Top Public	26,268
Bates College	Little Ivies	7,826
Colby College	Little Ivies	8,308
Tufts University	Little Ivies	9,372
Swarthmore College	Little Ivies	7,638
Middlebury College	Little Ivies	8,102
University of Alabama	Other Public	49,482
University of Alaska	Other Public	1,532
University of Arkansas	Other Public	24,904
Colorado State University	Other Public	23,442
University of Hawaii	Other Public	10,061
University of Idaho	Other Public	19,511
Iowa State University	Other Public	21,893
University of Kansas	Other Public	59,884
University of Kentucky	Other Public	41,985
Louisiana State University	Other Public	16,826
Louisiana$_{state}$	Other Public	16,814
University of Maine	Other Public	21,309
University of Mississippi	Other Public	18,164
University of Missouri	Other Public	25,126
University of Montana	Other Public	12,681
University of Nevada, Reno	Other Public	4,998
New Mexico State University	Other Public	7,307
University of New Mexico	Other Public	12,744
University at Buffalo (SUNY)	Other Public	15,640
North Dakota State University	Other Public	6,988
Ohio University	Other Public	15,042
University of Oregon	Other Public	13,969
University of South Carolina	Other Public	19,570
University of South Dakota	Other Public	5,865
University of Tennessee, Knoxville	Other Public	18,646
West Virginia University	Other Public	23,551
University of Wyoming	Other Public	13,699
Total		2,003,229

Notes: This table shows the list of co-educational colleges we include in our analysis and the corresponding sample size.

Table A2: Full List of Colleges for Intergenerational Analysis (1933-1966 Sample Size)

College	College Group	Sample Size	College	College Group	Sample Size
Stanford University	Ivy Plus	40,516	Bryn Mawr College	Elite Women	5,287
Yale University	Ivy Plus	43,312	Pembroke College in Brown University	Elite Women	8,877
University of Chicago	Ivy Plus	17,897	Connecticut College	Little Ivies	7,645
Harvard University	Ivy Plus	45,810	Trinity College	Little Ivies	8,940
Massachusetts Institute of Technology	Ivy Plus	20,410	Wesleyan University	Little Ivies	6,031
Dartmouth College	Ivy Plus	25,258	Bates College	Little Ivies	6,270
Princeton University	Ivy Plus	22,193	Bowdoin College	Little Ivies	7,347
Columbia University	Ivy Plus	45,641	Colby College	Little Ivies	9,265
Cornell University	Ivy Plus	69,033	Amherst College	Little Ivies	9,343
Duke University	Ivy Plus	32,461	Tufts University	Little Ivies	14,425
University of Pennsylvania	Ivy Plus	84,882	Williams College	Little Ivies	8,543
Brown University	Ivy Plus	12,745	Hamilton College	Little Ivies	3,667
University of Arizona	Top Public	16,038	Haverford College	Little Ivies	2,952
University of California, Berkeley	Top Public	110,816	Lafayette College	Little Ivies	241
University of California, Davis	Top Public	5,561	Swarthmore College	Little Ivies	6,753
University of California, Los Angeles	Top Public	70,189	Middlebury College	Little Ivies	9,042
University of California, Santa Barbara	Top Public	9,637	University of Alabama	Other Public	60,697
University of Colorado Boulder	Top Public	93,921	University of Alaska	Other Public	2,483
University of Connecticut	Top Public	26,228	University of Arkansas	Other Public	43,386
University of Delaware	Top Public	10,047	Colorado State University	Other Public	51,042
University of Florida	Top Public	36,429	University of Hawaii	Other Public	14,962
Georgia Institute of Technology	Top Public	30,272	University of Idaho	Other Public	28,162
University of Georgia	Top Public	54,912	Iowa State University	Other Public	28,826
University of Illinois Urbana-Champaign	Top Public	52,750	University of Kansas	Other Public	94,628
Indiana University Bloomington	Top Public	39,716	University of Kentucky	Other Public	38,734
University of Iowa	Top Public	26,924	Louisiana State University	Other Public	49,674
University of Maryland, College Park	Top Public	25,084	University of Louisiana at Lafayette	Other Public	39,758
Michigan State University	Top Public	63,943	University of Maine	Other Public	28,463
University of Michigan	Top Public	49,032	University of Massachusetts Amherst	Other Public	31,340
State University of New York (SUNY)	Top Public	3,235	University of Mississippi	Other Public	30,106
University of North Carolina at Chapel Hill	Top Public	31,642	University of Missouri	Other Public	34,928
University of Miami	Top Public	38,605	University of Montana	Other Public	22,111
The Ohio State University	Top Public	43,048	University of Nebraska–Lincoln	Other Public	28,726
Pennsylvania State University	Top Public	59,967	University of Nevada, Reno	Other Public	10,955
University of Pittsburgh	Top Public	28,719	University of New Hampshire	Other Public	20,088
The University of Texas at Austin	Top Public	68,229	New Mexico State University	Other Public	11,740
University of Vermont	Top Public	15,907	University of New Mexico	Other Public	24,707
William & Mary	Top Public	13,591	University at Buffalo (SUNY)	Other Public	15,311
University of Virginia	Top Public	22,549	North Dakota State University	Other Public	10,944
University of Washington	Top Public	31,271	Ohio University	Other Public	19,710
University of Wisconsin–Madison	Top Public	47,336	University of Oregon	Other Public	15,189
University of Minnesota, Twin Cities	Top Public	28,429	University of Rhode Island	Other Public	14,169
Radcliffe College	Elite Women	11,466	University of South Carolina	Other Public	30,746
Mount Holyoke College	Elite Women	12,810	University of South Dakota	Other Public	13,375
Smith College	Elite Women	18,500	University of Tennessee, Knoxville	Other Public	23,852
Wellesley College	Elite Women	10,507	West Virginia University	Other Public	34,405
Barnard College	Elite Women	14,296	University of Wyoming	Other Public	29,351
Vassar College	Elite Women	6,897			

Total sample size: 2,705,857

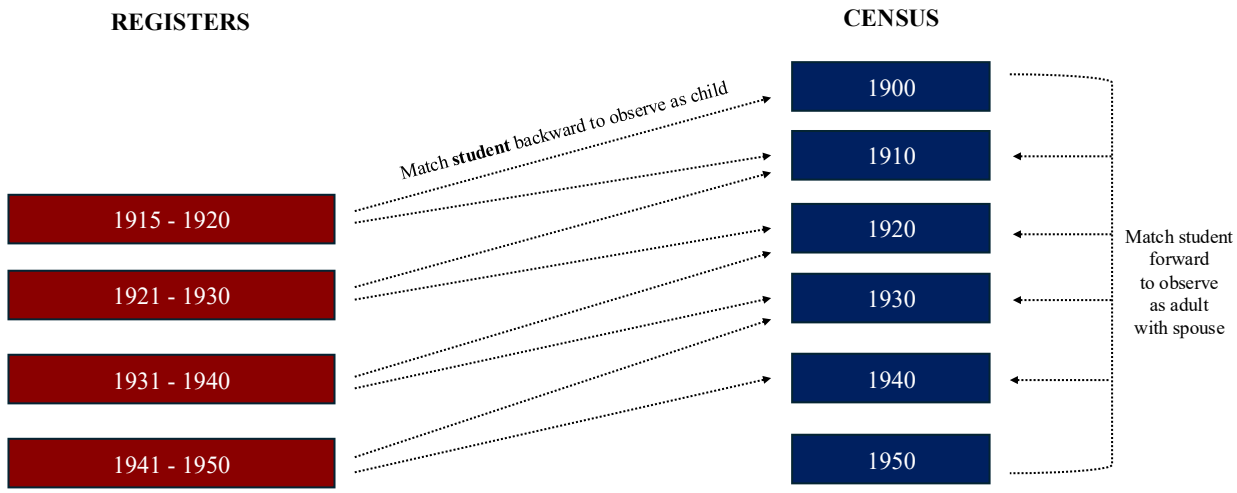
Notes: This table shows the full list of colleges that we use for our intergenerational analysis and the corresponding sample size.

Table A3: Relationship Between Pre-Determined Characteristics and Probability of Matching to an Adult Census

	Men	Women
Matching Rate (%)	62.2	38.3
i. Parental Occupations		
Non-Prof White-Collar	1.0	0.0
Professional	0.8	1.1
Farmer	0.2	0.7
Skilled Blue-Collar	-0.9	-0.4
Unskilled	-2.8	-3.0
ii. Parental Income Rank	0.04	0.03
iii. Race		
White	28.66	18.70

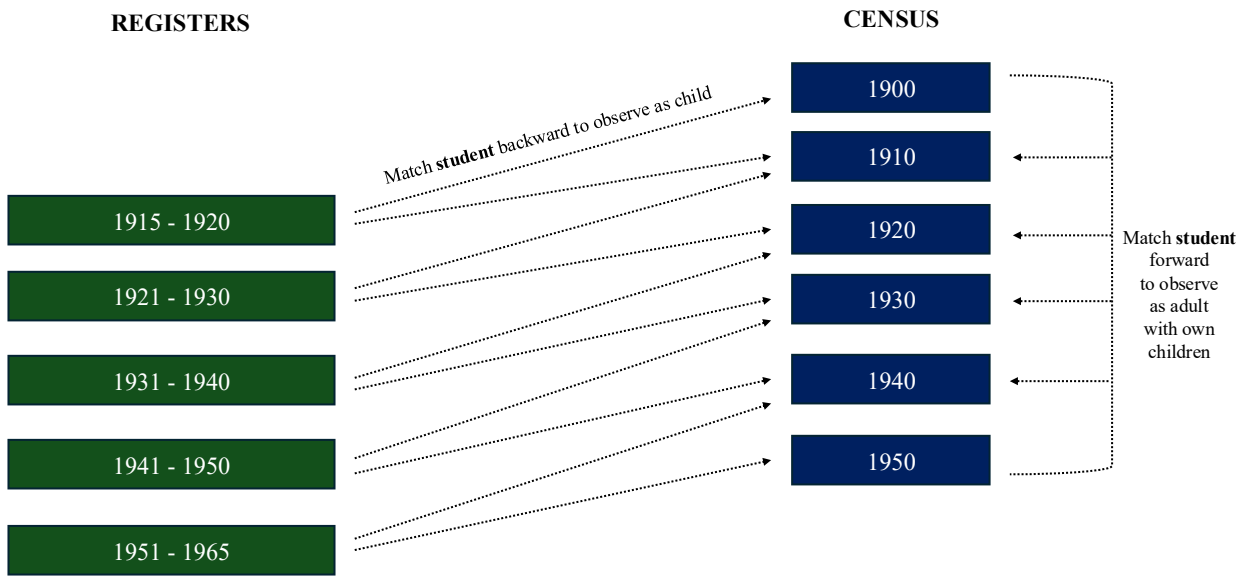
Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. . The top row of the table shows the proportion of individuals in our sample that we match to an adult census in which they are at least 25 years old. The remaining rows show the coefficient associated to each variable in a regression in which the outcome is equal to 100 if the observation matches to an adult census (each row is a separate regression).

Figure A1: Identifying Colleges Attended by Spouses



Notes: This figure illustrates the procedure that we use to identify the college attended by spouses. Our procedure first links college registers *backwards* to the two closest population censuses, where we observe students in their childhood households. We then link students' census records *forward* where we observe them as adults potentially living together as spouses.

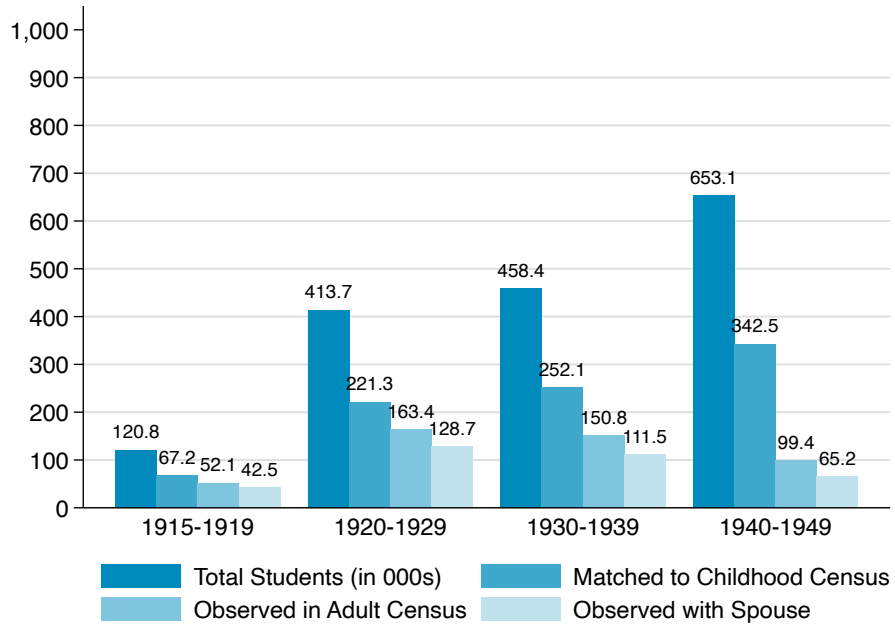
Figure A2: Identifying Colleges Attended by Child-Parent Pairs



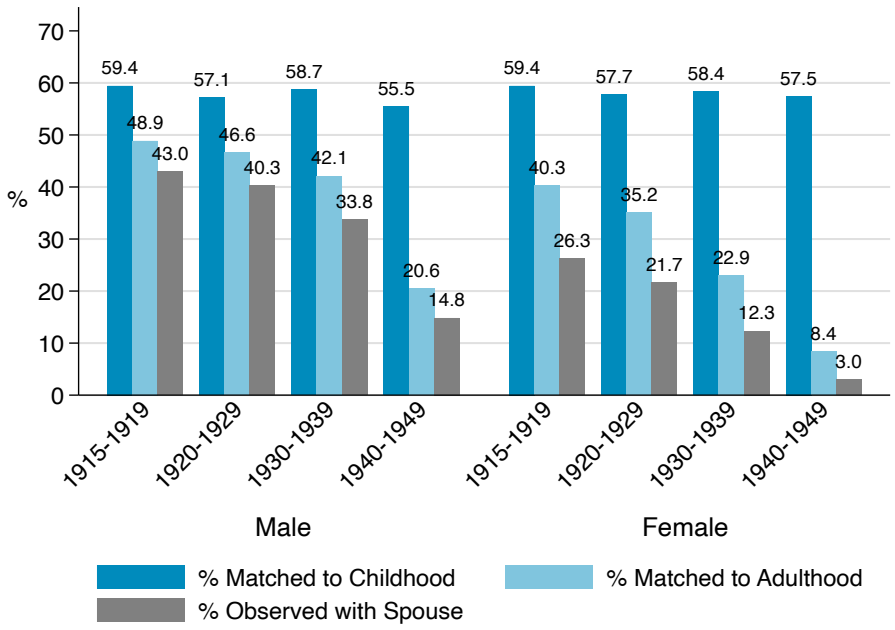
Notes: This figure illustrates the procedure that we use to identify the college attended by parent-child pairs. Our procedure first links college registers *backwards* to the two closest population censuses, where we observe students in their childhood households. We then link students' census records *forward* where we observe them as adults potentially with children of their own.

Figure A3: Sample Size and Matching Rates, by Entry Decade

(a) Sample size

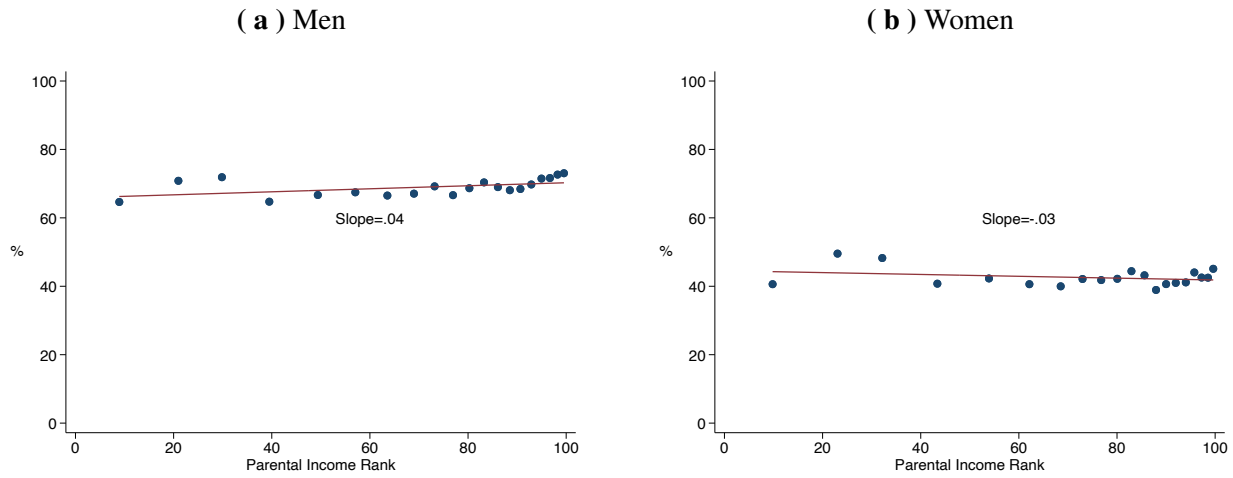


(b) Matching rate



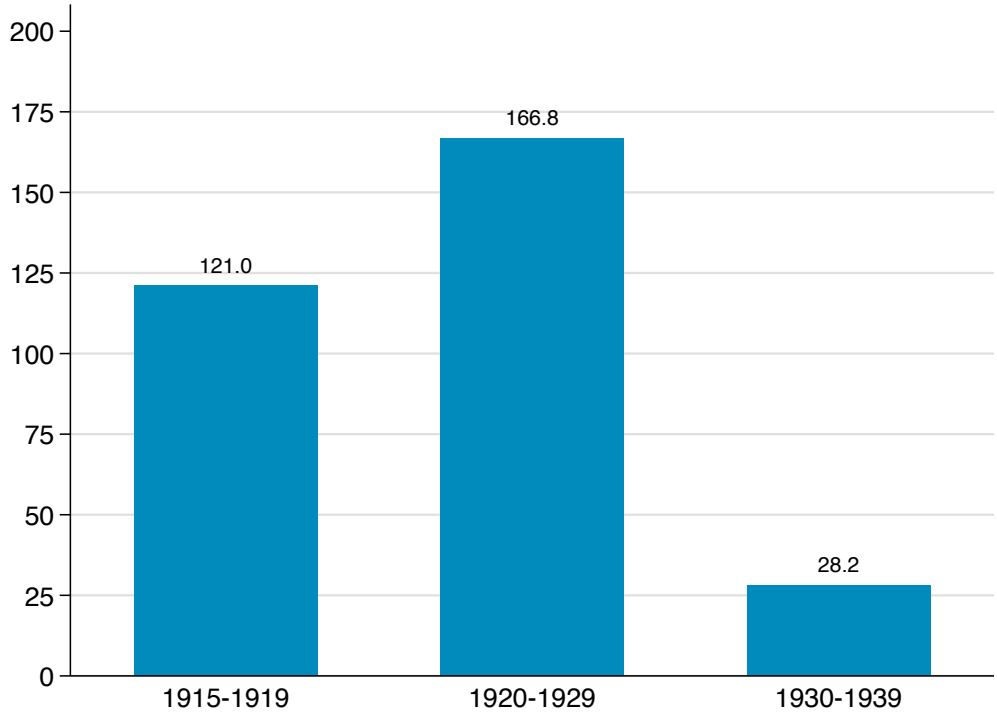
Notes: Panel (a) shows the total number of students in our sample, the number that are matched to a childhood census, the number that are matched to an adult census, and the number that are observed with a spouse, by decade of first attendance. Panel (b) shows the proportion of students that are matched to a childhood census, the proportion of students that are matched to *both* a childhood and an adult census, and the proportion of students who are matched to both a childhood census and an adult census in which they are observed with a spouse, by decade of first attendance.

Figure A4: Parental Income Rank and Probability of Matching to Adult Census



Notes: This figure shows a binned scatter plot relating parental income ranks (x axis) to the likelihood of matching a person from a childhood to at least one adult census. Panel (a) focuses on men whereas Panel (b) focuses on women.

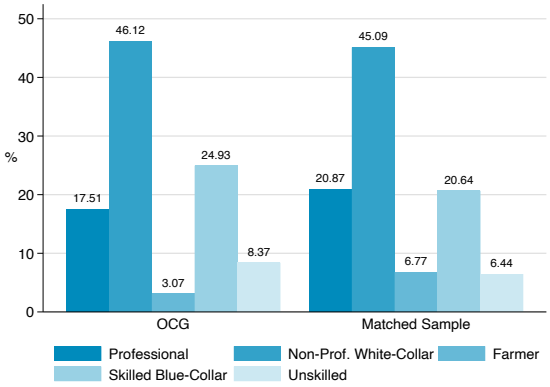
Figure A8: Number of Matched Children, by Parental Entry Cohort (in Thousands)



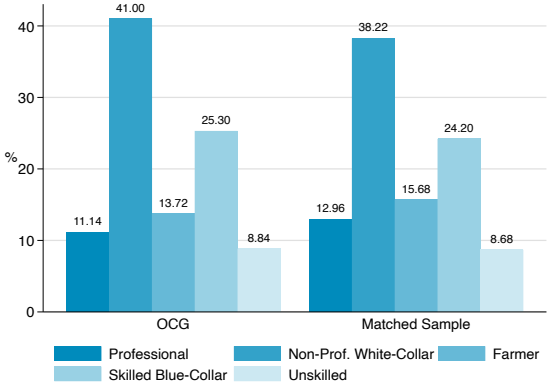
Notes: This figure shows the number of children of the original students that we observe in a childhood census, by entry cohort of the original student.

Figure A5: Comparison Between Matched Sample and OCG

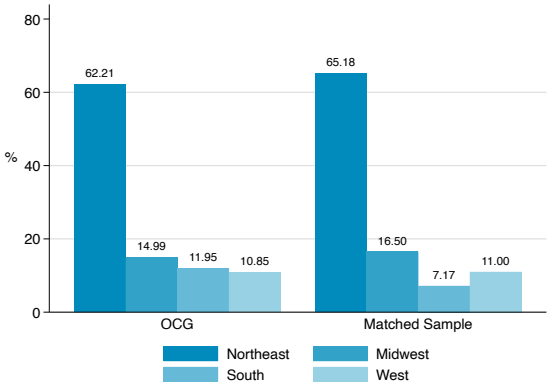
(**a**) Parental Occupations, Private



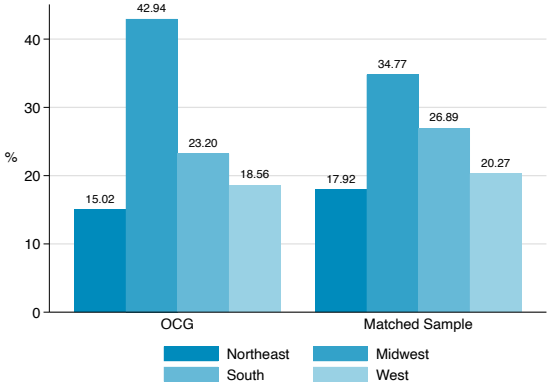
(**b**) Parental Occupations, Public



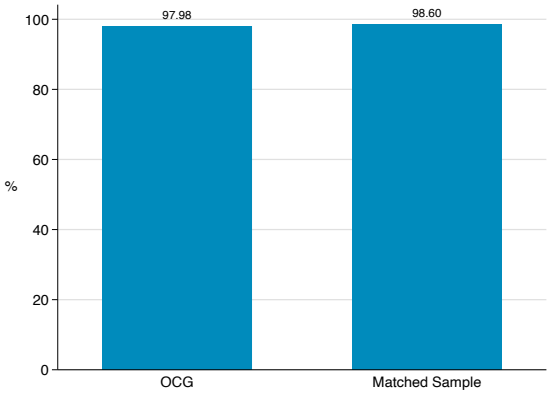
(**c**) Region of Birth, Private



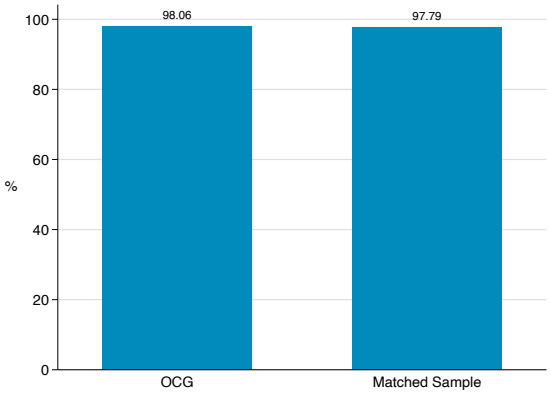
(**d**) Region of Birth, Public



(**e**) Proportion of Whites, Private

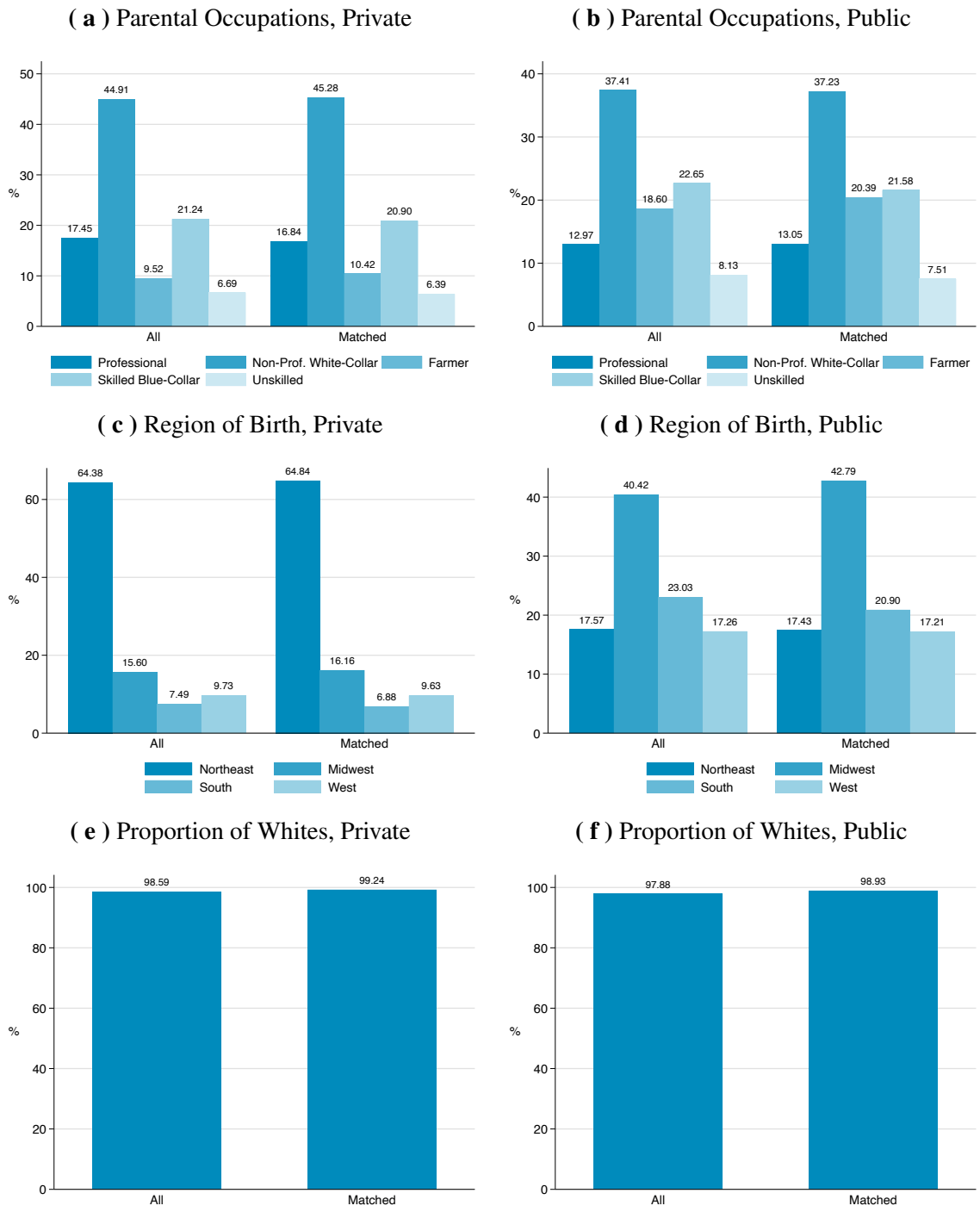


(**f**) Proportion of Whites, Public



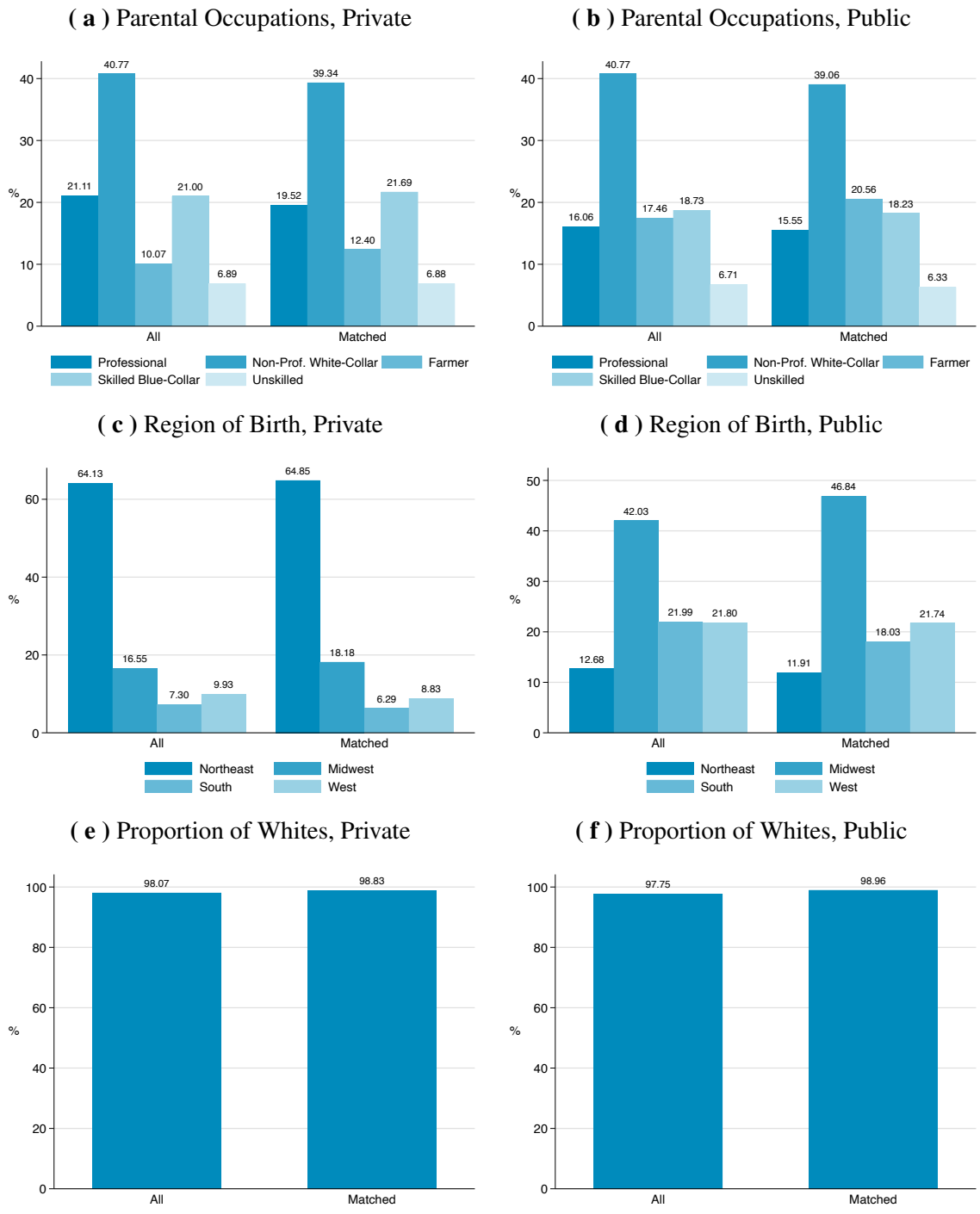
This figure compares the background characteristics of students in our sample matching student registers to the population census to the background characteristics of individuals attending the same set of colleges from the 1973 module of the Survey of Occupational Change in a Generation (OCG). Panels (a) and (b) focus on parental occupations, Panels (c) and (d) on region of birth, and Panels (e) and (f) on the proportion of White students. Panels (a), (c), and (e) focus on elite private colleges whereas panels (b), (d), and (f) focuses on public colleges.

Figure A6: Comparison Between Childhood Sample and Adult Sample, Men



This figure compares the background characteristics of students who match to an adult census to the background characteristics of all students. Panels (a) and (b) focus on parental occupations, Panels (c) and (d) on region of birth, and Panels (e) and (f) on the proportion of White students. Panels (a), (c), and (e) focus on elite private colleges whereas panels (b), (d), and (f) focuses on public colleges.

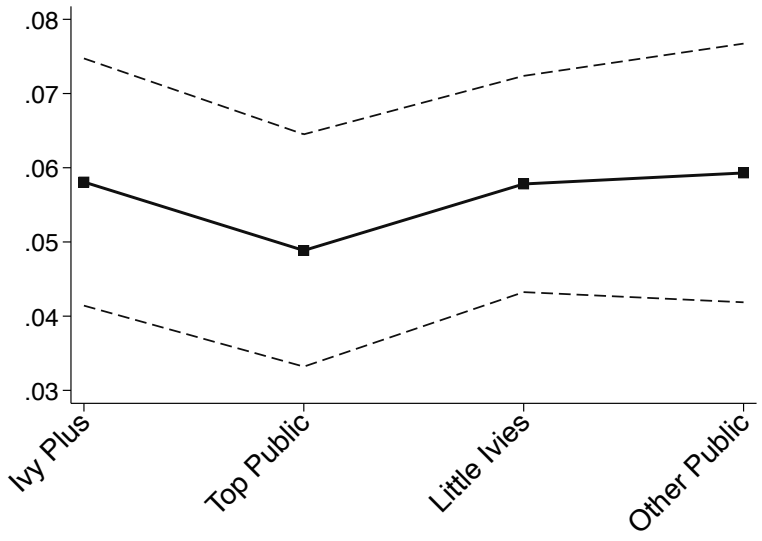
Figure A7: Comparison Between Childhood Sample and Adult Sample, Women



This figure compares the background characteristics of students who match to an adult census to the background characteristics of all students. Panels (a) and (b) focus on parental occupations, Panels (c) and (d) on region of birth, and Panels (e) and (f) on the proportion of White students. Panels (a), (c), and (e) focus on elite private colleges whereas panels (b), (d), and (f) focuses on public colleges.

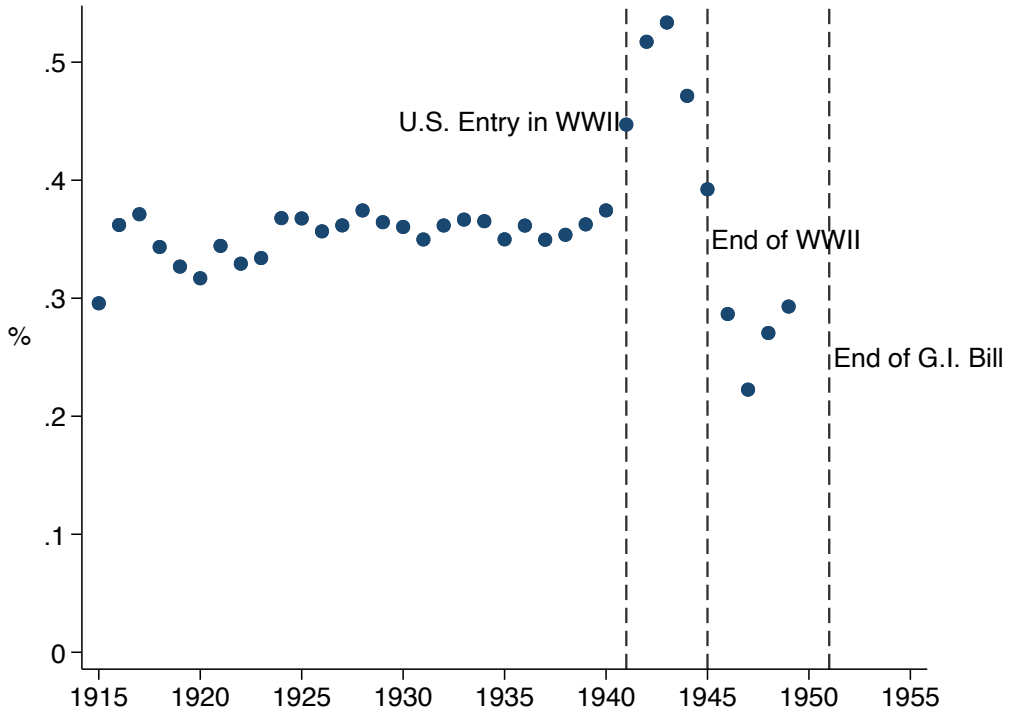
B Additional Tables and Figures

Figure B2: Share of Opposite-Sex Students in Cohort and Likelihood of Same-College Spouse, Excluding One School Group at a Time



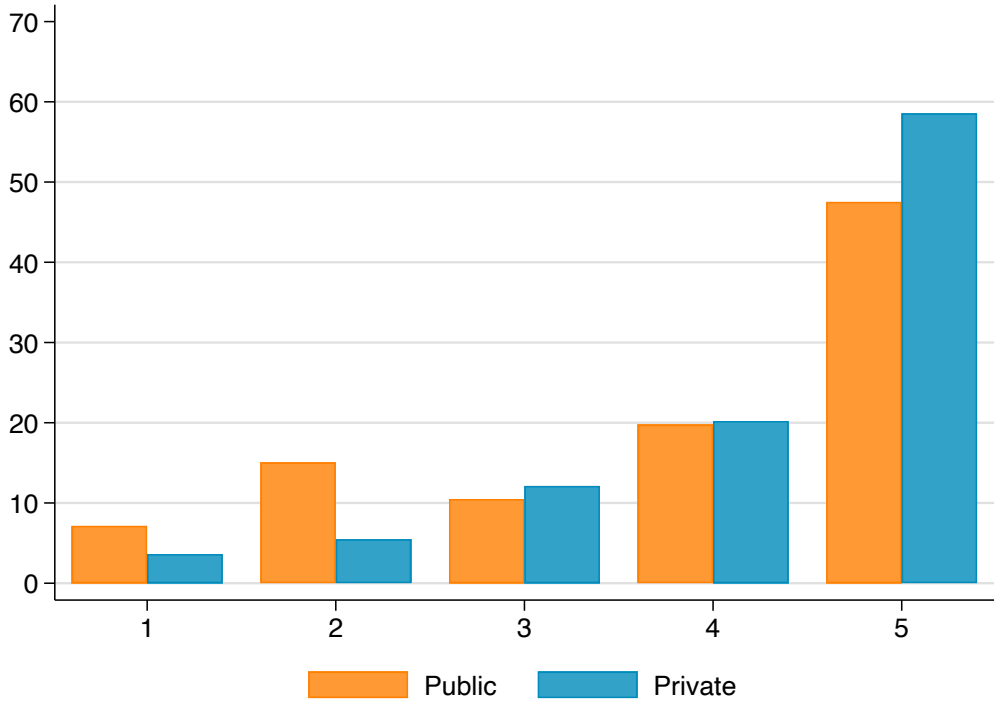
Notes: This figure shows the results of estimating equation 3 Standard errors are clustered at the college level.

Figure B1: Share of Women in Student Body, Co-Educational Colleges



Notes: This figure shows the proportion of women among students in a sample of coeducational colleges.

Figure B3: Proportion of Students by Parental Income Quintile, Private and Public Colleges



Notes: This figure shows the proportion of students by parental income quintile in the public and private colleges in our sample.

Table B1: Correlation Between Pre-Determined Characteristics and Share of Opposite-Sex Students in Cohort

	Parental Income Rank	Professional	White-Collar Non Prof	Farmer	Skilled Blue-Collar	Unskilled	White
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
% Opposite-Sex	-0.374	0.0297***	-0.0155	-0.00215	-0.0118	-0.000675	-0.00333**
	(1.426)	(0.0101)	(0.0273)	(0.0181)	(0.0117)	(0.00614)	(0.00141)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	479284	479284	479284	479284	479284	479284	479284
Mean of Dep. Var.	70.71	0.142	0.385	0.191	0.208	0.0719	0.991

Notes: Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. *% Opposite-Sex* ranges from 0 to 1 and measures the proportion of individuals in a student college cohort (defined as those starting in the same college and in the same year) who are of the opposite sex. All columns control for gender, college, and entry cohort fixed effects. The outcome in column 1 is the income rank of the student father. The outcomes in columns 2-6 are indicators that take a value of one if the student father was employed in a certain occupational category (as indicated by the column headings). The outcome in column 7 is an indicator that is one if the student is White. Standard errors are clustered at the college level.

Table B2: Share of Opposite-Sex Students in Cohort and Likelihood of Same-College Spouse, Re-Scaled Dependent Variable

	(1)	(2)	(3)
% Opposite-Sex	0.127***	0.101***	0.100***
	(0.0105)	(0.00940)	(0.00912)
Gender FEs	Yes	Yes	Yes
College FEs	No	Yes	Yes
Entry Cohort FEs	No	Yes	Yes
Individual Controls	No	No	Yes
Observations	454716	454716	454716
Mean of Dep. Var.	0.0549	0.0549	0.0549

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. This table repeats the analysis in Table 2 but re-scaling the dependent variable to account for the fact that we only observe a subset of same-college marriages. The outcome variable is an indicator that is one if the student is ever observed in an adult census with a spouse who attended the same college. *% Opposite-Sex* ranges from 0 to 1 and measures the proportion of individuals in a student college cohort (defined as those starting in the same college and in the same year) who are of the opposite sex. Column 1 shows the unconditional relationship between the two variables. Column 2 adds college, entry cohort, and gender fixed effects. Column 3 further control for father's income rank, indicators for father's occupation (professional, white-collar non-professional, farmer, skilled blue collar, with unskilled as the omitted category), and an indicator for being White. Standard errors are clustered at the college level.

Table B3: Share of Opposite-Sex Students in Cohort and Likelihood of Same-College Spouse, Robustness

	Baseline	Linear Trend	Pre-1941	College× 5-Year FEs
	(1)	(2)	(3)	(4)
% Opposite-Sex	0.0569***	0.0521***	0.0709***	0.0546***
	(0.00729)	(0.00904)	(0.00951)	(0.00929)
Gender FEs	No	No	Yes	Yes
College FEs	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes
Observations	479284	479284	392771	479284
Mean of Dep. Var.	0.0290	0.0290	0.0340	0.0290

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. The outcome variable is an indicator that is one if the student is ever observed in an adult census with a spouse who attended the same college. *% Opposite-Sex* ranges from 0 to 1 and measures the proportion of individuals in a student college cohort (defined as those starting in the same college and in the same year) who are of the opposite sex. In the *Linear Trends* columns we control for college-specific linear trends. In the *Pre-1941* columns we exclude students starting college in 1941 or after. In the *College× 5-Year FEs* columns we include college × five-year entry-cohort bins fixed effects. Standard errors are clustered at the college level.

Table B4: Share of Opposite-Sex Students in Cohort and Likelihood of Marriage, Robustness

	Baseline	Linear Trend	Pre-1941	College× 5-Year FEs
	(1)	(2)	(3)	(4)
% Opposite-Sex	-0.0463*	-0.0374	-0.0603**	-0.0396
	(0.0240)	(0.0240)	(0.0281)	(0.0247)
Gender FEs	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes
Observations	479284	479284	392771	479284
Mean of Dep. Var.	0.717	0.717	0.764	0.717

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. The outcome variable is an indicator that is one if the student is ever observed in an adult census with a spouse. *% Opposite-Sex* ranges from 0 to 1 and measures the proportion of individuals in a student college cohort (defined as those starting in the same college and in the same year) who are of the opposite sex. In the *Linear Trends* columns we control for college-specific linear trends. In the *Pre-1941* columns we exclude students starting college in 1941 or after. In the *College× 5-Year FEs* columns we include college × five-year entry-cohort bins fixed effects. All columns control for father's income rank, indicators for father's occupation (professional, white-collar non-professional, farmer, skilled blue collar, with unskilled as the omitted category), and an indicator for being White. Standard errors are clustered at the college level.

Table B5: Correlation Between Pre-Determined Characteristics and Share of Opposite-Sex Students in Cohort, Married Sample

	Parental Income Rank	Professional	White-Collar Non Prof	Farmer	Skilled Blue-Collar	Unskilled	White
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
% Opposite-Sex	-1.653	0.00382	-0.0274	-0.0116	0.0219	0.0136	-0.00262**
	(2.015)	(0.0162)	(0.0324)	(0.0267)	(0.0156)	(0.00828)	(0.00116)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	109834	109834	109834	109834	109834	109834	109834
Mean of Dep. Var.	69.63	0.139	0.364	0.249	0.182	0.0631	0.998

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. *% Opposite-Sex* ranges from 0 to 1 and measures the proportion of individuals in a student college cohort (defined as those starting in the same college and in the same year) who are of the opposite sex. All columns control for gender, college, and entry cohort fixed effects. The outcome in column 1 is the income rank of the student father. The outcomes in columns 2-6 are indicators that take a value of one if the student father was employed in a certain occupational category (as indicated by the column headings). The outcome in column 7 is an indicator that is one if the student is White. Standard errors are clustered at the college level.

Table B6: Share of Opposite-Sex Students in Cohort and Likelihood of Same-Quintile Spouse, Robustness

(a) Private Colleges, Top 20%

	Baseline	Pre-1941	Linear Trend	College× 5-Year FEs
	(1)	(2)	(3)	(4)
% Opposite-Sex	0.0388*	0.0395	0.0567**	0.0482*
	(0.0198)	(0.0229)	(0.0241)	(0.0245)
Gender FEs	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes
Observations	8974	8687	8974	8974
Mean of Dep. Var.	0.656	0.658	0.656	0.656

(b) Private Colleges, Bottom 80%

	Baseline	Pre-1941	Linear Trend	College× 5-Year FEs
	(1)	(2)	(3)	(4)
% Opposite-Sex	-0.0689**	-0.0435	-0.0622*	-0.0721**
	(0.0284)	(0.0265)	(0.0309)	(0.0296)
Gender FEs	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes
Observations	6350	6061	6350	6350
Mean of Dep. Var.	0.656	0.315	0.314	0.314

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. The outcome variable is an indicator equal to one if the student is married to a spouse in the same quintile of the parental income distribution. *% Opposite-Sex* ranges from 0 to 1 and measures the proportion of individuals in a student college cohort (defined as those starting in the same college and in the same year) who are of the opposite sex. All columns control for father's income rank, indicators for father's occupation (professional, white-collar non-professional, farmer, skilled blue collar, with unskilled as the omitted category), and an indicator for being White. Standard errors are clustered at the college level.

Table B7: Correlation Between Pre-Determined Characteristics and Share of Opposite-Sex Students in Cohort, Intergenerational Sample

	Parental Income Rank	Professional	White-Collar Non Prof	Farmer	Skilled Blue-Collar	Unskilled	White
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
% Opposite-Sex	-1.084	0.0267*	-0.0318	-0.00362	-0.00177	0.0110	-0.00340
	(1.877)	(0.0159)	(0.0304)	(0.0252)	(0.0172)	(0.00935)	(0.00271)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	315979	315979	315979	315979	315979	315979	315979
Mean of Dep. Var.	68.50	0.132	0.350	0.246	0.196	0.0740	0.993

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. *% Opposite-Sex* ranges from 0 to 1 and measures the proportion of individuals in a student college cohort (defined as those starting in the same college and in the same year) who are of the opposite sex. All columns control for gender, college, and entry cohort fixed effects. The outcome in column 1 is the income rank of the student father. The outcomes in columns 2-6 are indicators that take a value of one if the student father was employed in a certain occupational category (as indicated by the column headings). The outcome in column 7 is an indicator that is one if the student is White. Standard errors are clustered at the college level.

Table B8: Share of Opposite-Sex Students in Cohort and Likelihood of College Education in the Next Generation, Heterogeneity

	Own Gender		Parent Gender		Public/Private		Grandparent Income Rank	
	(1) Women	(2) Men	(3) Women	(4) Men	(5) Public	(6) Private	(7) Bottom 80%	(8) Top 20%
% Opposite-Sex	0.0421***	0.0152	0.0158	0.0198	0.0238**	0.0228	0.0286***	0.0298**
	(0.00743)	(0.0133)	(0.0229)	(0.0147)	(0.0108)	(0.0206)	(0.00928)	(0.0121)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	153554	162425	78594	237385	262653	53326	166095	149884
Mean of Dep. Var.	0.0840	0.118	0.0856	0.106	0.101	0.102	0.0863	0.118

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. Each column in this table corresponds to a different sample, as indicated by the column headings. The outcome variable is an indicator that is one if the child attended any of the 95 colleges in our extended sample. *% Opposite-Sex* ranges from 0 to 1 and measures the proportion of individuals in a student college cohort (defined as those starting in the same college and in the same year) who are of the opposite sex. All columns control for children's gender and birth year fixed effects. All columns control for father's income rank, indicators for father's occupation (professional, white-collar non-professional, farmer, skilled blue collar, with unskilled as the omitted category), and an indicator for being White. Standard errors are clustered at the college level.

Table B9: Share of Opposite-Sex Students in Cohort and Likelihood of College Education in the Next Generation, Robustness

	Same College as Parents		Other College		Any College	
	(1) Linear Trend	(2) College× 5-Year FEs	(3) Linear Trend	(4) College× 5-Year FEs	(5) Linear Trend	(6) College× 5-Year FEs
% Opposite-Sex	0.0186***	0.0182**	0.00532	0.00458	0.0239**	0.0228**
	(0.00695)	(0.00717)	(0.00436)	(0.00439)	(0.00973)	(0.00999)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	315979	315979	315979	315979	315979	315979
Mean of Dep. Var.	0.0502	0.0499	0.0512	0.0513	0.101	0.101

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. The outcome variable is an indicator that is one if the child attended any of the 95 colleges in our extended sample. In the *Linear Trends* columns we control for college-specific linear trends. In the *College× 5-Year FEs* columns we include college × five-year entry-cohort bins fixed effects. All columns control for children's gender and birth year fixed effects. All columns control for father's income rank, indicators for father's occupation (professional, white-collar non-professional, farmer, skilled blue collar, with unskilled as the omitted category), and an indicator for being White. Standard errors are clustered at the college level.

Table B10: Same-College Couples and Fertility Outcomes, Married Sample

	Number of children			1+ Children		
	(1)	(2)	(3)	(4)	(5)	(6)
Same-College Spouse	0.103***	0.0365**	0.0540***	0.0545***	0.0416***	0.0425***
	(0.0183)	(0.0176)	(0.0170)	(0.00416)	(0.00399)	(0.00397)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	No	Yes	Yes	No	Yes	Yes
Entry Cohort FEs	No	Yes	Yes	No	Yes	Yes
Individual Level Controls	No	No	Yes	No	No	Yes
Observations	343781	343781	343781	343781	343781	343781
Mean of Dep. Var.	1.657	1.657	1.657	0.794	0.794	0.794

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. *Same-College Spouse* is an indicator that is one if the student married another student from their same college. Columns 3 and 6 control for father's income rank, indicators for father's occupation (professional, white-collar non-professional, farmer, skilled blue collar, with unskilled as the omitted category), and an indicator for being White. Standard errors are clustered at the college level.

Table B11: Share of Opposite-Sex Students in Cohort and Spousal Education, Robustness

	Years of Schooling			4+ Years of College		
	(1) Pre-1941	(2) Linear Trend	(3) College× 5-Year FEs	(4) Pre-1941	(5) Linear Trend	(6) College× 5-Year FEs
% Opposite-Sex	0.636***	0.622***	0.649***	0.165***	0.158***	0.162***
	(0.154)	(0.145)	(0.147)	(0.0224)	(0.0204)	(0.0200)
Gender FEs	Yes	Yes	Yes	Yes	Yes	Yes
College FEs	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	207753	216761	216761	207753	216761	216761
Mean of Dep. Var.	12.66	12.61	12.61	0.292	0.288	0.288

Notes: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1. The outcome variable in columns 1 to 2 is the years of schooling of the student's spouse. The outcome variable in columns 3 to 4 is an indicator that is one if the student's spouse has at least one year of college. The outcome variable in columns 5 to 6 is an indicator that is one if the student's spouse has four or more years of college. In the *Pre-1941* columns we exclude students starting college in 1941 or after. In the *Linear Trends* columns we control for college-specific linear trends. In the *College× 5-Year FEs* columns we include college × five-year entry-cohort bins fixed effects. All columns control for father's income rank, indicators for father's occupation (professional, white-collar non-professional, farmer, skilled blue collar, with unskilled as the omitted category), and an indicator for being White. Standard errors are clustered at the college level.