Old Boys' Clubs and Upward Mobility Among the Educational Elite

Valerie Michelman, *University of Chicago*
Joseph Price, *Brigham Young University & NBER*
Seth D. Zimmerman, *Yale University & NBER*

August 10, 2021

Abstract

This paper studies how exclusive social groups shape upward mobility and whether interactions between low- and high-status peers can integrate the top rungs of the economic and social ladders. Our setting is Harvard in the 1920s and 1930s, where new groups of students arriving on campus encountered a social system centered on exclusive old boys' clubs. We combine archival and Census records of students' college lives and careers with a room-randomization design based on a scaled residential integration policy. We first show that high-status students from prestigious private high schools perform worse academically than other students but are more likely to join exclusive campus clubs. Club members go on to earn 32% more than other students and are more likely to work in finance and join country clubs, both characteristic of the era's elite. The membership premium persists after conditioning on high school, legacy status, and even family. Random assignment to high-status peers increases participation in exclusive college clubs, but overall effects are driven entirely by large gains for private school students. In the long run, a 50 percentile increase in residential peer group status raises the rate at which private school students work in finance by 40% and their membership in adult social clubs by 26%. We conclude that social interactions among the educational elite mediated access to top positions in the post-war United States but did not provide a path to these positions for underrepresented groups. Turning to recent cohorts, we show that while Harvard students differ from the past in their racial, ethnic, and gender composition, the path for high-status students from clubs to finance careers still exists, and elite university students from the highest-income families continue to earn more than others.

1 Introduction

Economic elites disproportionately come from a small number of social, educational, and business institutions, and there is growing evidence that peer interactions at these institutions affect access to and performance in top jobs.[1] Together, these two facts suggest that upward mobility to top rungs of the economic and social ladder may depend on whether people from lower-status backgrounds can access elite social networks or "old boys' clubs." Though old boys' clubs are central to qualitative accounts of elite "closure"—the idea that social groups restrict access to opportunity on the basis of shared traits and experiences—there is little quantitative evidence on who can join these groups and how they shape outcomes over the long run.[2] A key challenge is the lack of data identifying actual and potential group members and tracing their life trajectories.

This paper provides new evidence on how exclusive social groups shape upward mobility and whether interactions between low- and high-status peers can integrate top positions in the economy and society. We focus on elite universities, where talented lower-status students interact with disproportionately high-status peers and encounter the old boys' clubs upon which career opportunities may depend. We construct new data on Harvard students in the 1920s and 1930s, who go on to lead business and civic institutions at the dawn of the post-war "American Century" (Luce, 1941; Mills, 1956; Galbraith, 1958). We link records of students' social and academic lives at college to the 1940 Census and to biographical reports compiled 25 years after graduation, and pair these data with two research designs. Our first design takes a selection-on-observables approach using a rich control set, including legacy status, family fixed effects, and comparison groups of near-miss club applicants. Our second draws on a room randomization policy in which administrators assigned students to widely varying residential peer groups.

We have three main findings. First, students from high-status backgrounds are more likely to join exclusive campus clubs than low-status students, but are less likely to succeed in the classroom. Second, the labor market premium for club membership is much larger than the premium for academic success. Third, exposure to high-status college peers pushes high-status students towards high-status paths in their social and professional lives, but does not affect students from less privileged backgrounds, thus reinforcing rather than reducing inequality. We conclude that social interactions with high-status peers help pave the road to the top, but that even prolonged close contact may fail to spark these connections across group boundaries.

Harvard in the early 20th century has three features that make it ideal for studying exclusive social institutions. The first is the prominence on campus of some of the best-documented and oldest old boys' clubs in the US. Social life at Harvard centered on exclusive organizations known as final clubs, so-called because they are the last clubs one joins as a Harvard student. These

---

[1]Backgrounds of elites: Miller (1950), Useem and Karabel (1986), Temin (1999), Cappelli and Hamori (2004), Reeves et al. (2017). Top jobs: Cohen et al. (2008), Fracassi and Tate (2012), Shue (2013), Xu (2018), and Zimmerman (2019).

[2]See e.g. Weber (1922), Bourdieu (1998), Tilly et al. (1998), and Bol and Weeden (2015) or Khan (2012) for a review.

clubs, which Amory (1947) describes as the "be-alls and end-alls of Harvard social existence," are hundreds of years old and count among their members multiple US Presidents. Final clubs still exist today, and are often described in similar terms.[3]

The second feature helping our analysis is that we can observe how students from different backgrounds interact on campus, and how these interactions depend on policy choices. By the 1920s, Harvard had adopted a variety of policies aimed at enrolling students from outside its traditional constituency at high-status, high-tuition private feeder schools. Administrators supplemented these policies with residential life measures seeking to integrate social life across class lines, including randomized room assignment.[4] Though Harvard excluded women and limited enrollment by non-white and Jewish students, there was substantial *economic* diversity, and residential policies pushed students from different backgrounds into contact with one another.

The third helpful feature is that the lives of Harvard students in the 1920s and 1930s are richly and publicly documented. Records held in the Harvard archives describe students' social, academic, and career outcomes. We digitize these records for students entering Harvard between 1919 and 1935. Yearbooks and club rosters describe students' family and high school background, their residential lives at Harvard, and their engagement in social activities. Records of class rank divide students into groups based on academic performance. For long run outcomes, 25th Reunion Class Reports assemble self-reports and reports from class officers into biographical accounts of family life, occupation, and social activities. We also link the Harvard data to Census records from 1910 through 1940, using computer-assisted family history techniques to obtain high match rates.

We begin by describing how academic and social success in college vary with baseline social status. Following historical and contemporary accounts that emphasize the importance of high school background as a measure of pre-college social standing, we divide students by the kind of high school they attended. We focus on eight private feeder high schools that send many students to Harvard and are identified with high social status in qualitative accounts. Students from these schools make up 24% of our sample.

Students from private feeder schools are more socially engaged than other students, but do worse in school. On average, private feeder students participate in 1.68 extracurricular activities in their first year of college, compared to 0.85 for other students. 13.5% of private feeder students hold activity leadership positions and 19.8% participate in social committees, compared to 4.7% and 2.1%, respectively, for other students. In contrast with their social success, private feeder students are 56% more likely than other students to have grades in the lowest rank group (the

---

[3]Nir (2016) writes that final clubs are the "apex of social life at Harvard," while a 2017 faculty committee described final clubs as "impossible to escape— even for those who wish to have nothing to do with them" (CUSGGO 2017).

[4]Harvard President Charles Eliot stated in 1903 that "it is to the last degree undesirable that colleges should be accessible only to the well-to-do." In 1902, future President Lawrence Lowell cited integration of rich and poor students in dormitories as "the chief value of the College . . . for the training of character" (Karabel, 2006, p45-47).

bottom 15% of the class) and 52% less likely to have grades in the top two groups (the top 8%).

Pre-college socioeconomic status (SES) combines with early-college social and academic outcomes to drive selection into the exclusive upper year final clubs. We document a social ladder that starts with the full first-year class, continues into a prestigious sophomore society known as Hasty Pudding, and finishes with a selective final club. At each rung, the share of private feeder students rises, and the share of students with high grades falls. 73% of final club members attended private feeder high schools, compared to 1% from public feeders. Ethnic diversity also falls. The share of students with distinctively Jewish names falls from 6.6% at Harvard to 0.2% in selective final clubs, while the share with "Colonial" surnames— those common in 18th century Massachusetts relative to the contemporary population— rises sharply.

Our second contribution is to describe the labor market premium associated with final club membership and compare it to the academic success premium. The final club premium is much larger than the academic success premium. Final club members in the lowest academic rank group earn 27% more than non-members in the top two groups and are 3.3 times as likely to have topcoded earnings, corresponding to the top 0.7 percent of the population distribution.

The final club membership premium is not driven by selection into clubs on the basis of high school type, high school identity, Harvard legacy status, family, or general engagement with social life at Harvard. We identify families that sent multiple sons to Harvard during our sample period and estimate specifications with family fixed effects. Brothers who are members of selective final clubs earn 42% more than brothers who are not, and are 2.6 times as likely to report topcoded incomes. Specifications comparing final club members to "near-missers" who join Hasty Pudding but are not selected for final club membership yield similar conclusions.

In addition to earnings levels, final club membership is associated with different career and social outcomes 25 years after graduation. Members of selective final clubs are 2.9 times more likely to work in finance and 49% less likely to work in medicine. Final club members are also more likely to join social organizations like country clubs— important features of mid-century life (Mills, 1956; Putnam, 2000). Overall, our descriptive evidence shows that while students from rich families tend to be socially successful, social success is distinct from wealth in its long-run effects, and may expand gaps in long-run outcomes by baseline social status.

Our third contribution is to assess whether policies adopted in pursuit of social cohesion on campus affect the distribution of club membership and long-run career outcomes. Our approach relies on a room randomization scheme for freshman dormitories that Harvard administrators used to integrate campus social life. Students submitted housing applications indicating acceptable prices and room types. Students could apply with roommates, so roommate assignments were not random, but conditional on room price and occupancy, room assignments were made by lot. Our analysis of this randomized design provides a test of both the specific policy that Harvard used to promote cross-group interaction and of the general proposition that social in-

teractions shape high-stakes outcomes.

We identify high- and low-priced peer neighborhoods based on dorm maps, and evaluate the effect of assignment to a high-priced neighborhood on short-, medium-, and long-run outcomes. One way to think of this research design, which leverages systematic differences in peer neighborhoods driven by dorm layouts, is as a Moving to Opportunity experiment on the Harvard campus (Katz et al., 2001; Ludwig et al., 2013; Chetty et al., 2016). Consistent with archival accounts of the assignment process, controls for randomization blocks (defined by room type and price) eliminate the cross-sectional relationship between the average neighborhood room price and *own* baseline characteristics. Neighborhood price is strongly correlated with peer attributes and varies widely within randomization blocks. A 50 percentile increase in neighborhood price— well within the support of our random variation— raises students' peer private feeder share by 10.1 percentage points, 32.1% of the sample mean.

We find that exposure to high-status peers helps students achieve social success in college, but that overall effects are driven entirely by large gains for private feeder students. A 50 percentile shift in the room price distribution raises membership in selective final clubs by 3.2 percentage points in the full sample (34.1% of the mean). For private feeder students, the same shift raises membership by 8.4 percentage points (37.7%), while effects for other students are a precise zero. These effects build on similar patterns we observe starting in students' first year at college. A 50-percentile increase in neighborhood price raises the count of first-year activities by 11.2% overall, with larger gains for private feeder students and small, statistically insignificant effects for others.

The effects of college peers persist over the long run. 25 years after graduation, a 50 percentile change in peer neighborhood price raises the chance that students participate in adult social organizations by 8.7%. As with on-campus clubs, the overall long-run effects are driven entirely by large gains (26.0%) for private feeder students, with near-zero effects for others. Turning to occupations, a 50 percentile change in neighborhood price rank raises the share of private feeder students in finance by 7.2 percentage points, 39.6% of the group mean. This change is offset by small declines in academic, medical, and legal careers. For other students, these effects are reversed: exposure to high-status peers pushes them away from finance and towards medicine.

A thread running through our findings is that exposure to high-status peers pushes private school students, but not other students, to act in private school-typical ways. The campus activities, adult social organizations, and occupations towards which private school students shift are those where private feeder students are disproportionately represented. We summarize this behavior using standardized linear indices of the extent to which different outcomes predict that a student attended a private feeder school. A 50 percentile change in peer neighborhood price raises the private school index for first year activities by 0.25σ for private feeder students. The same shift raises the upper year club index by 0.18σ, the long-run social index by 0.16σ, and

the occupation index by 0.22σ. For other students, index effects are near zero for each outcome except for the occupation index, which shifts *down* by 0.08σ. Because of its asymmetric effect on private feeder students, exposure to high-status peers reinforces social and career segregation.

The final part of the paper examines how outcomes for Harvard students have changed over time. We extend our demographic, academic, and career data through the 2010s, and augment it with membership lists for one of the six selective final clubs through 2015. Harvard changes profoundly, enrolling women and many more non-white students. These changes do not carry over to the final club, which integrates Jewish students but remains all male and disproportionately white. Cross-group differences in outcomes persist as well. Club members perform worse academically and are more likely to work in finance than other students across the 20th century.

Overall, students from high-income families at elite universities continue to earn more than other students at those universities. We illustrate this using public data from Chetty et al. (2020) on earnings for students born in the 1980s who enroll at Ivy+ colleges, split by parent income percentile. While child income is flat in parent income for most of the income distribution, it rises steeply above the 95th percentile. Ivy+ students with parent incomes between the 99th and 99.9th percentile earn 33% more on average than students with parents between the 90th and 95th percentiles. Students with top 0.1% parents earn 76% more. These findings do not prove that the causal mechanisms we document in Greatest Generation cohorts persist through the present. However, they are consistent with recent descriptions of social interactions among high-status students at elite schools affecting career outcomes (Rivera, 2016; Jack, 2019).

Our findings contribute to several strands of work. First, a large body of research studies how the colleges students attend affect how much they earn.[5] We contribute by unpacking the production function within elite colleges and providing evidence that the determinants of career success are different for children from the richest families. Our within-institution results are not in tension with the common finding that the returns to between-college differences in selectivity are large for low-SES students. What our findings do suggest is that gains for lower-SES students from attending elite colleges may be mediated more by professional careers (like medicine) than by elite business jobs. This is consistent with Zimmerman (2019), who shows that the gains from admission to elite business programs accrue only to high-SES students, while admission to medical programs raises earnings across SES groups but does not increase top income attainment. A focus on top outcomes is important because top income shares are large (Alvaredo et al., 2013, 2017) and because people in top positions set firm policies and serve as role models (Matsa and Miller, 2013; Bertrand et al., 2019). Whether policies that diversify selective universities ultimately diversify top jobs may depend on social integration as well as academic match.[6]

Second, our findings suggest limits on what policies expanding intergroup contact can achieve

---

[5]See Dale and Krueger (2002, 2014); Black and Smith (2006); Hoekstra (2009); Zimmerman (2014); Hoxby (2018); Dillon and Smith (2017); Chetty et al. (2020); Sekhri (2020); Mountjoy and Hickman (2020); Bleemer (2020).

[6]See Rothstein and Yoon (2008a,b); Arcidiacono et al. (2016); Arcidiacono and Lovenheim (2016).

when the stakes are high. A growing literature provides proof of concept that cooperative intergroup interactions can increase cross-group socializing and survey measures of intergroup cohesion in low- and medium-stakes settings (Carrell et al., 2019; Rao, 2019; Lowe, 2020; Mousa, 2020), consistent with Allport (1954)'s "contact hypothesis."[7] We find no evidence that increased cross-group exposure helps lower-status students access exclusive social groups. Further, high status students assigned low status peers are no more likely to join social groups with low status students; they just participate less in high status-typical social groups. Our results have more in common with Zimmerman (2019)'s finding that college peers from high-SES backgrounds serve on firm leadership teams together, but not with lower-SES peers.

Third, this paper presents the first evidence (to our knowledge) on how residential assignment at school shapes students' long-run outcomes. Many studies use college room randomization designs to explore how peers affect outcomes measured while students are in school (Sacerdote, 2001; Zimmerman, 2003; Stinebrickner and Stinebrickner, 2006; Lyle, 2007; Carrell et al., 2009; Mehta et al., 2019; Zárate, 2019; Jones and Kofoed, 2020). We elevate this work by showing that the short-run shifts matter in the long run. Our finding that exposure to high-status peers augments baseline group differences is consistent with previous studies of homophily in college peer groups (Marmaros and Sacerdote, 2006; Mayer and Puller, 2008; Carrell et al., 2013).[8]

Fourth, we provide quantitative evidence linking the economic literature on assimilation in the pre-1920s "age of mass migration" with studies of a persistent Protestant elite in the post-war period (Abramitzky et al., 2012, 2014, 2020a,b; Baltzell, 1964; Zweigenhaft and Domhoff, 1982; Davidson et al., 1995). We highlight how social institutions limited the assimilation of Jews and Catholics into the post-war elite. Our findings help explain the sustained media (New York Times, 1907; McWilliams, 1948; Auchincloss, 1958; Rimer, 1993; Flanagan, 2016), academic (Marmaros and Sacerdote, 2002; Popov and Bernhardt, 2012), and literary (Fitzgerald, 1920) footprint of final clubs and similar groups, such as secret societies at Yale and eating clubs at Princeton.

2 Institutions

We study Harvard students in the 1920s and 1930s. This setting provides a richly-documented vantage point on the role of social sorting in the creation of post-war elites. We start by describing the institutions that shaped Harvard during this period, many of which persist today. We focus our discussion on two institutional features. The first is Harvard administrators' effort to increase the economic diversity of on-campus interactions through admissions and housing policies while continuing to attract private school applicants. The second is the role of exclusive

---

[7]See also Boisjoly et al. (2006), Corno et al. (2019), Finseraas and Kotsadam (2017), and Paluck et al. (2019).

[8]Shue (2013) studies classroom peers at Harvard Business School, but focuses on management practices for students in top jobs, not on how peers effect career paths. Carrell et al. (2018) and Einiö (2019) consider the long-run effects of primary school and military peers, respectively.

private clubs in campus social life. We draw on primary sources, histories of admissions policy (Karabel, 2006; Synott, 1979), and Amory's (1947) account of the Harvard club system.[9]

Harvard administrators in the 1920s and 1930s considered both academic and social factors when designing admissions policy. In practice, this meant trying to bring in more students from public schools while limiting the number of Jewish students, who they viewed as socially undesirable and likely to make recruiting private school students more difficult.[10] One policy of this type was the "Top Seventh" plan, adopted in 1923, which guaranteed admission to any high school student in the top seventh of his class (RPTHC 1922–23, p. 290). In 1926, unable to achieve its goals with academic admissions requirements alone, Harvard adopted non-academic criteria for the first time (RPTHC 1925–1926, p. 298). This approach persists: the 2020 Harvard admissions office considered questions like "would other students want to room with you, share a meal. . . or collaborate in a closely-knit extracurricular group?" (HCAFA, 2020).

Once students were on campus, residential policy was the main lever for promoting cross-group student interactions. Concerned that wealthy students were isolating themselves in off campus houses, Harvard opened new dormitories for first-year students in 1914.[11] Rooms at different price levels were often in close proximity, a design feature that administrators tied to goals of social integration (RPHC 1929–1930 p. 101). As an additional step toward this goal, rooms were assigned at random from at least 1922 through 1941. We describe the randomization process in Section 5.1 and present documentation in Online Appendix B.

Administrative efforts to integrate social life through admissions and residential policy contrast with accounts of student life at Harvard emphasizing the importance of exclusive organizations as measuring sticks for social success in college, determinants of post-college outcomes, and drivers of inequality within the university. Key organizations from this perspective are the Hasty Pudding Institute of 1770 and a set of upper-year student societies known as final clubs.

To understand the importance of final clubs, it is helpful to see how students, reporters, and historians describe them. Franklin Delano Roosevelt (Harvard class of 1904) remarked that one of "the greatest disappointments of his life" was not being elected to his preferred club.[12] Amory described final clubs as the "be-alls and end-alls of Harvard social existence," while Nir (2016) described final clubs as "the apex of social life at Harvard" in the *New York Times*. Amory calls Hasty Pudding, a theater-focused sophomore society, a "proving ground" for final clubs that

---

[9]Amory's work received positive reviews in popular and academic outlets upon release (Jones, 1947; Low, 1948).

[10]For example, Harvard President Lawrence Lowell remarked that increasing numbers of Jewish students would "not intermingle with the rest," and that "[Jews] drive away the Gentiles" (Karabel 2006 pp. 88–89, 107).

[11]This goal is clearly stated in contemporary sources. For example, College Dean Alfred Hanford reflected on the first fifteen years of the Freshman Halls: "The man of limited means and the rich, the high school and private school graduates, the son of the banker, and the son of the farmer were thrown together. Freshmen coming from different schools and of varying origins were to be given an opportunity for making new contacts, social distinctions were to be broken down, and a democratic class spirit developed." (RPTHC 1929–1930, p. 100–101).

[12]Delano Roosevelt ultimately joined a different final club. Roosevelts in our sample include James, Franklin, Jr. and John, the three sons of FDR, as well as Kermit Jr., Cornelius, and Theodore III, grandsons of Theodore Roosevelt.

provides "an index of social seniority almost as authoritative as the old colonial ranking."

Qualitative accounts emphasize that these clubs expanded pre-existing inequality of opportunity by student background, in particular high school background. Amory describes how

> the question of being club material at Harvard boils down to a boy's having graduated from one of a small number of socially correct Eastern private schools. Of the five hundred or so public-school graduates [...] entering Harvard each year, rare indeed is the boy who manages to break into the purple pale of its club Society.... [Even students from the most exclusive private schools] find themselves in the position of waiting anxiously for the call to Harvard clubdom[.]

In 1988, faculty member Alan Dershowitz described final clubs as "where Harvard students learn to discriminate." A 2017 faculty report stated that "final clubs reinforce existing campus inequities" and "are at odds with the [...] view that student body diversity is essential to Harvard College's pedagogical objectives[.]" Many clubs still do not accept female members.

What is the appeal of final clubs? Qualitative accounts emphasize opportunities for career advancement. Amory describes a student whose "well-connected Porcellian friends saw to it their adopted brother had an opportunity to marry well and take a good position in an old-line Boston firm." A student writing in the Harvard Crimson in 2020 notes that "[j]ust as being a Harvard student grants us access to an unparalleled alumni network, so too does being a member of one of these final clubs" (Premaratne, 2020). Mills (1956) sums up the perceived connection between high schools, social success at Harvard, and social success later in life:

> It is the prestige of a properly certified secondary education followed by a proper club in a proper Ivy League college that is the standard admission ticket to the world of urban clubs and parties in any major city of the nation.

We take three insights from the institutional setting. First, final clubs are arguably the primary form taken by old boys' clubs at Harvard. Second, high school background mediates entry into these clubs. Third, policymakers used room randomization to encourage social integration.

3 Data sources

3.1 Harvard archival records

We construct microdata on the Harvard entering classes from 1919 through 1935. We digitize and merge data from several sources, all of which are publicly available in the Harvard Archives. This section gives an overview of the data sources we use and how we combine them. For a more detailed description see Online Appendices B.1 and B.2.

The first record type we draw from are contemporaneous reports of enrollment and social activities. We define our sample universe using Freshman Registers known as "Red Books." These are yearbooks for the freshman class, published in the spring of each academic year. The Red Books contain information on home addresses, college addresses, high school background, and first-year activities. We link the Red Book data to records of academic class rank. These lists aggregate grades across all courses in an academic year and coarsen them to numerical groups between one and six. Students who have incomplete coursework at the time of publication or grades too low to advance are not included on the rank lists; not showing up on a list is an outcome of interest. Harvard published lists for all non-graduating students between 1920 and 1932 and for freshmen only from 1932. We focus on first-year grades because they are available for more cohorts. We further link to records of upper-year club participation published between 1920 and 1938. These report Hasty Pudding and final club membership.

The second record type we draw from are 25th Reunion Class Reports. Class officers compile class reports 25 years after scheduled graduation (roughly age 47) using student self-reports and administrative records. Reports contain standardized fields for family (noting those who also attended Harvard), occupation, adult club memberships, and other honors.

We augment the microdata with records of room attributes. We use floor maps of freshman dorms to define residential peer groups. The maps allow us to determine whether peer groups are organized "vertically" by stairwell or "horizontally" by floor. We describe rooms and peer groups using price and occupancy data from pamphlets distributed to incoming students.

3.2 Census records

We merge our Harvard data to Census records from 1910 through 1940. For the initial Census link, we augment family history techniques with ML hinting algorithms on the FamilySearch and Ancestry.com genealogy platforms (Costa et al., 2020; Price et al., 2021; Feigenbaum and Gross, 2020). We link across Censuses using techniques closest to Price et al. (2021), which incorporates methods from Abramitzky et al. (2019). Key features of the Harvard data for the Census link are parent name, birth location, and birth year. These are available in pre-1934 Class Report records. We therefore attempt matches only for pre-1934 students with Class Report data.

3.3 Merge statistics

Table 1 reports statistics on sample size and data availability. There are 14383 individuals in our sample universe. Almost all students report what high school they attended (98%) and their college address (95%), except in 1926, when address was not listed. We merge 91% of Red Book records for the 1920 through 1935 cohorts to class reports. 87% of students matched to class

reports (79% of students in the sample universe) have non-missing occupation fields.[13]

We use the 1940 Census to measure post-college outcomes. To ensure students have time to complete schooling and launch their careers, we exclude students entering college after 1930 from the Census outcome sample.[14] We match 69% of students in entering cohorts from 1920 through 1930 with Class Report data to 1940 Census records. We use pre-1940 Censuses to describe students' pre-college backgrounds. We match 66% of students in 1920–1933 cohorts to pre-1940 Census records.[15] For students who match to multiple pre-college records, we use the most recent. Before 1940, Census records do not include measures of income.[16]

These match rates are high in the context of the literature. For example, our match rates are higher than match rates of 10-30% for automated Census-to-Census and Census-to-patent linking (Abramitzky et al., 2020c; Sarada et al., 2019). Compared to this baseline, our procedure falls about two-thirds of the way to the 88% match rate achieved when hand-linking children of Union army soldiers to the 1910 census (Costa et al., 2020), often cited as a gold standard of record linking (Bailey et al., 2020).

Though our match rates to individual datasets are high, sample sizes decline as we impose sample restrictions and require matches to multiple data sources. For example, 9,343 students live in on-campus rooms subject to random assignment. Of these, 5,218 (55.8%) are in the 1930 and earlier cohorts we match to Census outcome records, 3,428 (65.7% of 5,218) are matched to the 1940 Census, and 2,485 (72.5% of 3,428) report their wages in the Census. This issue surfaces mainly in room-randomization analyses of Census outcomes and motivates our choice to focus the long-run component of that analysis on outcomes observed in Class Reports.

3.4 Categorization schemes

3.4.1 Harvard clubs and private high schools

The social outcome of primary interest to us is membership in a selective upper-year final club. To operationalize this concept, we break final clubs into groups based on prestige following Amory's ordered list, labeling Porcellian, A.D., Fly, Spee, Delphic, and Owl 'selective final clubs.'[17]

Contemporary accounts emphasize the importance of a small number of high-status, high-priced private high schools in driving social outcomes. Combining these accounts with our

---

[13]Online Appendix Figure A.1 reports how sample sizes and match rates vary across cohorts.

[14]Online Appendix Figure A.2 reports how school enrollment and labor market outcomes change with potential experience. Six years after expected graduation (i.e., the 1930 entering cohort), less than 10% of students are in school and more than 90% are in the labor force. Average earnings rise steadily through 16 years post-graduation.

[15]For pre-college measures, students in cohorts 1920–1930 are matched to 1910 and 1920 Census records, while students in the 1931–33 cohorts are matched to 1910 through 1930 Census records.

[16]Some analyses of historical Census data use occupation-specific wage scores. These are essentially predictions of income based on occupation and demographics. They are not useful in our setting because they require extrapolation from the broader population to our highly selected sample of Harvard students. See Online Appendix B.3.

[17]The remaining clubs on Amory's list and active during our period are Fox, D.U., and, starting in 1930, Phoenix.

microdata, we identify the eight private boarding schools that sent the most students to Harvard over our period, and label these institutions private feeder schools. The schools in this group are Exeter, Andover, Milton, Middlesex, Groton, St. Paul's, St. Mark's, and St. George's. The first seven schools sent more students to Harvard than any other private schools. The eighth, St. George's, sent fewer students than three Boston-area day schools but is included because it is part of the "St. Grottlesex" group emphasized in historical accounts (Amory 1947; Karabel 2006).

Our private feeder designation has a strong basis in the historical record. However, we also identify a broader group of private schools that includes other boarding schools and day schools, and discuss findings for this group as well. In addition, we identify public feeder schools that sent multiple students to Harvard over the sample period.[18] Any high school that sent at least twenty students to Harvard across our cohorts is classified as either public or private. Online Appendix Figure A.3 displays student counts for each classified school.

3.4.2   Residential peer groups

We describe residential peer neighborhoods using the average per-occupant room price. For each peer neighborhood in each entering cohort, we use floor plan and room price data to compute the occupant-weighted mean room price. Then, because we are interested in relative rank within each cohort rather than dollar values, we use the mean prices to rank neighborhoods on a zero-to-one scale, with zero being the lowest-ranked neighborhood in a cohort and one the highest. We assign neighborhoods with the same mean price the rank at the midpoint of the interval.

3.4.3   First-year activities, occupations, and adult organizations

To describe students' social lives during their first year in college, we enumerate common activities and then place activities into coarse groups. Activity groups include sports, music, and schoolwide social committees.[19] We also construct summary measures of activity participation: an indicator for participation in any activity, the count of total activities, and an indicator for holding a leadership role in an activity (e.g. president of a club, team captain).

We follow a similar procedure to describe 25-year occupation and social club outcomes using data from the Class Reports. For occupations, we create indicators for coarse job types and identify a set of text strings associated with each. In the main text we report results for 11 relatively common categories: Finance, Accounting, Medicine, Law, Higher Education, Primary/secondary teaching, Government, Art or Publishing, Retail, Senior management, and Middle/lower management.[20] Because reported work outcomes reflect both industry (such as investment banking or medicine) and occupation or office type ("partner" or "vice-president"),

---

[18]The public school sending the most students to Harvard is Boston Latin. Most public feeders are near Boston.

[19]Examples of social committees include "Regatta Committee" and "finance committee.'

[20]We include all categories with at least 100 students in index construction.

these outcomes are not mutually exclusive. One can be a senior manager and work in finance. We treat data for individuals in cohorts 1920–1935 who are not linked to Class Report records or do not report work outcomes as missing.

We divide adult clubs into three categories: social clubs, professional organizations, and honor societies. Social clubs are gentleman's clubs (e.g. the Knickerbocker Club), country clubs or sports clubs (e.g. the Brookline Country Club), and fraternal organizations. Professional associations include the American Medical Association and the American Economic Association. We treat data for those not linked to Class Reports as missing. See Online Appendix B.1 for details of club and occupation classification.

While qualitative reports can guide our interpretation of categorical outcomes (e.g., that working in finance was characteristic of the era's elite), we also present a data-driven alternative. We construct private feeder indices for college activities, occupations, and adult associations. These indices capture the extent to which a bundle of outcomes is indicative of high status and are standardized (mean zero, standard deviation one) to facilitate cross-outcome comparisons. We compute the indices as predicted values from regressions of an indicator for private high school status on indicators for participation in different activities (occupations, associations) and cohort fixed effects, using a Lasso for variable selection and excluding one's own cohort from the sample. See Online Appendix B.4 for details.

3.4.4   Cultural and religious background

We follow Abramitzky et al. (2020a) and Abramitzky et al. (2020b) and use names as measures of cultural background. We construct a Jewish name index for each first and last name based on the frequency of that name among Yiddish speakers relative to non-Yiddish speakers. For a given (first or last) name n, the name index is computed as

$$\text{Jewish Index}_n = \frac{\frac{\#\,\text{Yiddish speakers}_n}{\#\,\text{Total Yiddish speakers}}}{\frac{\#\,\text{Yiddish speakers}_n}{\#\,\text{Total Yiddish speakers}} + \frac{\#\,\text{non-Yiddish speakers}_n}{\#\,\text{Total non-Yiddish speakers}}},$$

with counts computed in the 1920 and 1930 Censuses. We use an analogous approach to construct indices for Catholic names and "Colonial" names. The Catholic name index is based on the relative frequency of names among immigrants from Italy, Ireland, and Spain. The Colonial name index is based on the relative frequency of last names (not first names) among people born in Massachusetts before 1800 in the 1850 Census, compared to the 1920 and 1930 population. Online Appendix Table A.1 reports the names with the highest scores for each index.

Continuing to follow Abramitzky et al. (2020b), we classify students as having Jewish names if the average value of the index over their first and last name in the 1920 and 1930 Census data is greater than 0.7. We maintain a cutoff of 0.7 for describing an individual as having a Catholic or

Colonial name. Name indices are an imperfect proxy for cultural background. However, as we show below, they predict sorting into final clubs, matching in the marriage market, and changes in the student body following admissions policies designed to reduce the share of Jews. Online Appendix B.5 provides additional detail on name classification.

4 College behavior and long-run outcomes

4.1 Student background by high school type

We now turn to a descriptive analysis of the relationship between student background, social and academic success in college, and long-run outcomes. We first describe students' pre-college backgrounds and how they differ by high school type, our main measure of baseline social status.

Panel A of Table 2 reports student background characteristics from Harvard sources. Overall, 46% of students come from private high schools, 24% from private feeder schools, and 23% from public feeder high schools. 7.0% of all students have a father who attended Harvard, and 20.4% report having a brother who attended Harvard. 6.6% of all Harvard students have names we identify as Jewish, 5.5% have Catholic names, and 29.3% have Colonial names. The share of students with Jewish names in our data is lower than the share of Jewish students reported in contemporary Harvard sources, which reached 20-25% in the early 1920s before changes in admissions policies.[21] Our interpretation is that the 0.7 cutoff in the name index used in Abramitzky et al. (2020b) is conservative in the Harvard context.

Students from private feeder schools come from higher-status families with stronger Harvard ties. 14.3% of private feeder students have a Harvard father and 29.0% have a Harvard brother, compared to 4.6% and 17.6%, respectively, for other students. Private feeder students are more likely to have Colonial names and less likely to have Jewish names than other students. Turning to Census data (Panel B of Table 2), we see that private feeder students are less likely to come from immigrant backgrounds and more likely to have fathers who are doctors or lawyers.

Once at Harvard, most students share a common residential setting. Panel C of Table 2 describe students' first year living environments. Overall, 80.0% of students live on campus. Almost all private feeder students live on campus (96.4%), and a large majority (74.9%) of other students do as well. Many of the students living off campus are from Boston and live at their home address. On average, private feeder students live in rooms that are ranked 14 percentiles higher in the own-room price distribution than other students and in peer groups that are ranked 10 percentiles higher. However, the IQRs for peer group ranks in the two groups mostly overlap, consistent with qualitative reports that residential life at Harvard pushed together students from different backgrounds. Section 5 discusses residential peer groups in more detail.

---

[21] Karabel reports internal Harvard estimates of 21.6% in 1922 and 27.6% for the entering class of 1925 (pp. 96, 105). Shares of Jewish students fall starting in 1926 with the adoption of non-academic admissions criteria. See section 6.5.

4.2   Social vs. academic success in college

Private feeder students do worse in the classroom than other students. This is illustrated in Panel D of Table 2. As first-year class rank rises from the lowest rank group (rank 6) to the highest (rank 1), the share of students from private feeder schools falls sharply. Students from private feeder schools are twice as likely as other students to be in the bottom rank group, and less than half as likely to be in the top rank group.

Private feeder students are much more successful outside the classroom. Panels E and F of Table 2 report the shares of students engaging in different first-year and upper-year activities. Private feeder students are 1.5 times more likely to participate in at least one activity than other students in their first year at college (70.9% vs. 46.8%), and participate in twice as many activities on average (1.68 vs. 0.85). Cross-group differences are even more pronounced for leadership activities. Private feeder students are three times as likely to have leadership roles and 4.3 times more likely to be members of social committees.

4.3   The social "funnel" and selection into final clubs

Initial gaps in activity participation and social leadership persist through to exclusive upper-year clubs. 39.4% of private feeder students join the Hasty Pudding sophomore society compared to 7.7% for other students. 13.6% of students join a final club of any kind, with a 37.0% rate for private feeder students and a 6.2% rate for other students. 7.0% of students join a selective final club. The 21.5% rate for private feeder students is nearly nine times the rate for other students.

Columns 4 and 5 of Table 2 describe the "funnel" into selective final clubs, with Hasty Pudding membership as an intermediate step. 91% of Hasty Pudding members attended private schools, and 62% attended private feeder schools. Less than 1% of Hasty Pudding members have Jewish names, and 47% have Colonial names. Hasty Pudding members are 63% more likely to have grades in the lowest rank group than students in the full sample and are 35% as likely to have grades in the highest two groups. However, they are more than four times as likely to lead first-year social committees. 44% of Hasty Pudding members go on to join selective final clubs, more than six times the population rate.

Sorting into selective final clubs is even more extreme. 95% of final club members attended private high schools, and 73% attended private feeders. Out of 941 students in final clubs, we classify two as having Jewish names, and both appear to be false positives.[22] 51% have Colonial

---

[22]One is Harry E.D. Pollock, who was born in Utah and attended an east coast private school before becoming an archaeologist (Society for American Archaeology, 1983). The other is Samuel Cabot Jr., whose father attended Harvard, and who appears to have been a member of the well-known Cabot family in Boston (Jensen, 1972). Online Appendix Figure A.4 shows how final club membership varies across the distribution of the name indices. Final club membership rates are low for Jewish index values above 0.4, well below the 0.7 cutoff from Abramitzky et al. (2020b). Extreme sorting is consistent with internal documents; a 1922 Harvard report studying the classes of 1912-18 found that no Jews were chosen for selective final clubs over the period (Karabel, p. 98).

surnames. In terms of grades, we observe zero final club members in the top rank group, and they are 84% more likely to have grades in the bottom rank group than the population as a whole. Final club members are less likely to be unranked than other students. It may be difficult for students who do not advance with their class to join upper-year clubs. Finally, consistent with qualitative reports, membership in the Hasty Pudding is close to a prerequisite for joining a selective final club: 95% of final club members were also Hasty Pudding members.

Supplemental analyses describe how pre-college characteristics interact with academic and social performance early in college to shape upper-year club membership. Our findings reflect the sorting on high school type, grades, and cultural background documented in Table 2. The additional insight is that final club membership hinges on the *combination* of private feeder background with first year social leadership. See Online Appendix B.6.1 for details.

4.4 The boys' club premium

We now turn to long-run outcomes. Table 3 describes long-run outcomes by high school type. We observe large differences in adult associations, occupations, and earnings outcomes by high school type. Looking at occupations, what stands out is that private feeder students are more likely to work in finance than other students (17.5% vs. 7.7%) and are less likely to work in medicine, law, or higher education. Private feeder students are more likely to join country clubs and gentleman's clubs, but less likely to join fraternal orders or professional associations. Private feeder students and other students are similarly likely to report non-zero wage income, but private feeder students report an average earnings of $2961, 10% more than other students. They are 41% more likely to report at least $50 of non-wage income, and 62% more likely to report the maximum value of earned income, $5,000. The topcode value corresponds to the top 0.7 percent of the earnings distribution for men aged 27–37 in the 1940 Census.

Social success in college predicts long-run outcomes. Columns 4 and 5 of Table 3 display adult outcomes for Hasty Pudding and Final club members. Selective final club members are 72.1% more likely to belong to social clubs as adults than the full population, 2.6 times more likely to work in finance, and 46.5% less likely to work in medicine. Final club members have wage income that is 29.2% higher than the population average, and are 2.4 times more likely to have topcoded wage income. Outcomes for Hasty Pudding members fall between those for the full population and those for final club members, but are generally closer to final club members.

The strong relationship between final club membership and labor market success contrasts with what we see for academic performance. Panels A, B, and C of Figure 1 display income measures from the 1940 Census split by first-year academic rank group and membership in a selective final club. Because the top rank group is small and few students in selective final clubs have high grades, we pool rank groups 1 and 2 into one group, and omit cells with fewer than twenty observations. Panel A shows that students in selective final clubs earn about $851 more

than other students in the same academic rank group, 32% of the no-club mean. In contrast, conditional on club membership, the relationship between first-year grades and earned income is flat. Panels B and C show that students in selective final clubs are also more likely to have topcoded wage income and to report non-wage income across all rank groups.

4.5 Assessing selection effects

Final club members might have high earnings for reasons other than club membership. This section assesses the role of selection on different margins in driving the final club premium.

The first explanation we consider is selection on the basis of high school type. This story seems plausible: we have already seen that most club members attended private feeder schools, and most private feeder students come from wealthy families. However, it is hard to reconcile with the data. Panels D through F of Figure 1 repeat Panels A through C, but split by both club membership and high school type. Panels D and E show that the gap in earnings by club membership is not diminished by controls for school type. This contrasts with findings for *unearned* income. Panel F shows that private feeder students are more likely to report having unearned income than non-feeder students with the same club membership status.

What about selection on other attributes? We conduct a series of analyses that test stories about selection on the basis of high school identity, family background, and social engagement. These specifications have the form

$$Y_i = S_i \beta_s + R_i \beta_r + X_i \beta_x + \theta_{c(i)} + e_i. \tag{1}$$

Y_i is an outcome of interest for individual i, S_i is final club membership, R_i is academic rank group, and X_i are additional controls that vary across specifications. $\theta_{c(i)}$ are cohort fixed effects. Class rank is in rank-group units with the sign reversed so that positive coefficients indicate that earnings outcomes rise with academic standing. Table 4 presents our results. Each panel corresponds to a different outcome and each column to a different specification.

We first test for balance on selection into the earnings sample. We place an indicator for successful match to a wage record in the 1940 Census on the left side of equation 1. These results are in Panel A of Table 4. The sample is all enrolling students, so non-matches may arise from failure to match to Class Reports, failure to match from Class Reports to the Census, or non-reporting of Census wage income. The relationship between wage match and covariates of interest is in general economically small and differs statistically from zero at the five percent level in only one out of 16 cases. This limits concerns about differential censoring.

We next consider earned income, reported in Panel B of Table 4. Column 1 shows a baseline specification with controls for private feeder status and class rank, but not final club membership. We observe a private school premium of $297, 11% of the non-private feeder mean.

Column 2 adds an indicator for membership in a selective final club. This specification corresponds to our first selection story— selection on high school type. As we saw in Figure 1, there is a large final club membership premium even after conditioning on high school type. Students in selective final clubs earn $777 more than other students, a 29% premium above the non-member mean. This is 15 times the size of the academic rank group premium. The private feeder earnings premium falls by half once we control for membership in a selective final club.

We next consider selection on family attributes beyond high school type. Column 3 adds Census controls for family attributes: fixed effects for father's occupation, fixed effects for father's and mother's state or non-US country of birth, family size, parental presence, home ownership, the presence of domestic employees, and whether the home was on a farm. The restriction to observations where these data are available reduces the sample size by about half, but the relationship between final club membership and earnings does not change. The academic success premium also remains small. In contrast, the coefficient on private feeder school falls to zero. In other words, controlling for both social success at college and pre-college attributes fully accounts for private school students' earnings advantage.

We conduct additional tests of selection on pre-college attributes using covariates from the Harvard records. Column 4 limits the sample to students from private feeder schools and adds fixed effects for each high school, as well as for Harvard legacy status. The final club earnings premium remains large, and the academic success premium remains small.

Column 5 controls directly for family fixed effects. We limit the sample to families who send more than one son to Harvard during our sample period.[23] These regressions essentially compare the smart brother in each family to the popular brother. The final club premium is larger in these specifications than in the others we estimate, equal to 42% of the brothers' sample mean for earnings and 156% of the mean for topcodes. The class rank premium remains close to zero. It is hard to explain the final club premium with differential selection across families.

Column 6 considers a different story: that the observed final club premium reflects a return to social engagement in general and accrues to individuals who would like to join a selective final club, whether they succeed or not. We restrict the sample to Hasty Pudding members, who are engaged with Harvard high society, but differ in whether they make the jump to the top tier. The final club premium remains large, despite the fact that the comparison group is socially successful in its own right. The earnings gap between final club members and near-missers suggests that social engagement does not explain the success premium.

Panels C and D of Table 4 repeat the same set of specifications for two additional outcomes: an indicator that a student's earned income is topcoded at the Census maximum of $5,000, and an indicator that the student has at least $50 in unearned income. The patterns for topcoded earnings are very similar to those we observe for wage earnings. The final club premium is large

---

[23]Students in these families are disproportionately from private feeder schools. See Online Appendix B.6.

and persists across all specifications we consider, while the academic success premium is near zero. High topcode rates for final club members suggest that our earnings specifications may tend to *understate* the true final club premium.

Results for non-wage income differ in two ways from those for earned income. The first is that the private feeder premium is large across all specifications. It is not diminished as much by the inclusion of additional controls. The second is that while the final club premium for unearned income is large in columns 2, 3, 4, and 6, it is negative and imprecisely estimated in specifications in column 5 that include family fixed effects. Our interpretation is that unearned income is more closely tied to family background than is earned income. While there may be a final club premium for unearned income, the evidence is not as strong as for earned income.

This descriptive analysis establishes three facts. First, the social success premium is large. Second, differences in social success account for much of the difference in earned income by private school background. Third, the social success premium is hard to explain with selection on the basis of family attributes or engagement with social institutions at Harvard. Online Appendix B.6 shows similar patterns when we impute income on the basis of home value, when we impute income values for occupations where business income may be particularly important, and when we use third-year rather than first-year class rank, and when we include students with missing class rank. A question this analysis leaves open is the extent to which the final club premium reflects a return to pre-existing social skills versus a return to social skills or networks acquired in college. Understanding the causal role of peer inputs in college is the focus of section 5.

4.6   Occupations and adult social outcomes

As was the case with income, club membership predicts long-run career and social outcomes within academic rank and high school type. Panels A and B of Figure 2 show how the shares of students working in finance and medicine 25 years after graduation vary with class rank and final club membership. Students in final clubs are more likely to go into finance than others in the same rank group. In the lowest rank group, 31% of club members pursue finance careers compared to 14% of non-members. Rates of finance careers decline with class rank regardless of membership status. The reverse is true for medicine: rates are higher for non-members, and rise with academic rank. Panel C of Figure 2 shows that final club members are nearly twice as likely to participate in adult social organizations across the grade distribution.

The lower row of graphs in Figure 2 displays the same outcomes as the upper row, but split by both club membership and high school type. Within high school type, final club members are more likely to work in finance at all levels of academic achievement, and less likely to work in medicine at all but the lowest achievement levels. Both private feeder and other students who are final club members are more likely to join social organizations as adults. Online Appendix Figure A.5 shows parallel results for other career types and for professional organizations.

5 Random room assignment in the short- and long-run

5.1 Peer groups and room prices

We use room randomization to assess whether increased residential contact with higher-status neighbors can spark social success at college and alter students' long-run career and social trajectories. This exercise tests both the general proposition that social interactions shape high-stakes medium- and long-run outcomes and the specific policy that Harvard (and many other universities) uses to promote cross-group interaction on campus.

Room randomization worked as follows. Freshman rooms were assigned different prices, depending on size, occupancy, and quality. First-year students were asked to fill out a housing application indicating their acceptable price and number of roommates. Room assignments were then made "by lot, from rooms of the price indicated in the application blank," except that "students coming in considerable numbers from any one school are distributed among the various halls, and the cheapest rooms are reserved for men of limited means." Randomization took place at the level of the *room*, not the individual, because students wanting to live in the same room could apply together. See Online Appendices B.1.5 and B.1.6 for documentation.

With this design in mind, our experimental specifications take the form

$$Y_i = \beta_0 + \beta_1 RP_{p(i)} + \theta_{r(i)} + \tau_{h(i)} + e_i. \tag{2}$$

Y_i is an outcome for student i, $RP_{p(i)}$ is the price rank of i's residential dormitory neighborhood $p(i)$, $\theta_{r(i)}$ are randomization block fixed effects, and $\tau_{h(i)}$ are indicator variables for each feeder high school (public or private). Following the randomization design, we define the $\theta_{r(i)}$ as fully-saturated interactions between entering cohort, room price, and room occupancy. When computing standard errors, we allow for clustering at the level of realized peer group $p(i)$. We consider alternate approaches to inference in Section 5.7.

This specification gives rise to experiments of the following form: two pairs of students apply for rooms of occupancy size two and price $175 per student. One pair is assigned to a peer neighborhood where the other rooms are more expensive and the neighborhood average price is $240 per student, and the other to a room where the other rooms are less expensive and the neighborhood average price is $125 per student. We then compare outcomes across the pairs. Figure 3 shows a map of peer neighborhoods and gives an example of one such experiment.

The identities of "high-priced" and "low-priced" neighborhoods are predetermined by observable and systematic differences in floor plans. They do not depend on the outcomes of assignment. This contrasts with many roommate and classmate designs, which rely on fluctuations in group-level means across ex ante indistinguishable units. Our approach has more in common with "Moving to Opportunity" designs, in which treatment is the opportunity to relo-

cate to a wealthier neighborhood (Katz et al., 2001; Ludwig et al., 2013; Chetty et al., 2016). As in studies of MTO, we do not rely on peer attributes for our econometric analysis other than to describe how assignment to a high-priced neighborhood alters peer composition.[24]

A corollary to this point is that our experimental specifications identify the effect of assignment to a higher-priced neighborhood, as mediated by a variety of peer attributes and individual behaviors. These specifications do not identify the specific peer attributes that alter student outcomes, nor do they reveal how the short- and medium-run effects of assignment (such as final club membership) individually contribute to long-run outcomes like career paths. As we discuss below, neighborhood price is strongly correlated with a variety of measures of the socioeconomic status of neighborhood peers. We use the phrase "exposure to high-status peers" to mean being placed in residential proximity to fellow students who are higher status along a variety of measures, observable and potentially unobservable.

5.2 Peer groups and randomization blocks

The first step in our analysis of the room-randomization quasi-experiment is to describe the distribution of first-year students on campus. 80.0% of first-year students live on campus, of whom 89% live in freshman halls that are part of the randomization scheme.[25] The rightmost column of Tables 2 and 3 report descriptive statistics for this sample. Campus residents are more likely to come from private feeder schools and less likely to come from Massachusetts, but overall there is broad coverage across high school type and other background characteristics. Students in the freshman halls live in residential peer groups that are generally integrated by room price and high school background. As an example, Panel A of Figure 4 plots the distribution of own room prices by high school type for the 1928 entering cohort. Private school students tend to live in higher-priced rooms, but the distributions overlap except at the very bottom and very top.

Residential peer groups are modest in size and highly heterogeneous even for students living in rooms with similar prices. Panel B of Figure 4 displays a histogram of residential peer group sizes. Mean group size is 9.7, with the middle 50% of the distribution falling between 7 and 12. Panels C and D of Figure 4 show the variation in peer attributes within groups defined by own room price. Panel C displays the mean and 90–10 spread of peer neighborhood price rank by ventile of the own-price distribution within each year. Peer mean room rank rises steadily through roughly the fiftieth percentile of the own-price distribution, at which point it plateaus. Except for the bottom-most ventile of the own-price distribution, students at each level of the own-price distribution experience a wide range of residential peer price levels. As shown in the

---

[24]Our strategy is therefore not subject to the Angrist (2014) weak instruments critique, which applies to settings where differences in mean attributes across peer groups disappear as group size grows large. Our focus on predetermined, systematic differences in neighborhood attributes is what Angrist (2014) recommends.

[25]These numbers exclude students in the 1926 cohort for whom we do not have address records. Students were sometimes assigned to other housing when there was excess demand for spots in freshman halls.

right panel of the graph, the same is true for peer private high school shares. There are students in the top ventile of the own-price distribution whose peer groups consist entirely of students not from private feeder schools, and students in the bottom-most ventile whose peers are almost half private feeder school students.

Turning to the randomized design, we first show that there is wide variation in peer group assignments within randomization blocks. Panel A of Figure 5 shows the distribution of block sizes. The average block consists of 34.1 individuals; the 25th percentile of the block size distribution is 11 and the 75th percentile is 52. Panels B and C describe how peer neighborhood attributes vary within block. In both panels, each vertical unit represents an individual. Individuals are sorted vertically by randomization blocks, with blocks in Panel B (C) sorted by the maximum value of neighborhood mean price (neighborhood private high school share) in the block. The shaded area for each individual corresponds to the range between the minimum and maximum value of neighborhood price (neighborhood private high school share) in the block.

Panel B shows that within-block variation spans nearly the entire range of peer neighborhood prices. The vertical lines show the 10th and 90th percentiles of the neighborhood price distribution. There are blocks in which prices range from the maximum price to the median, and many that span the middle 50% of the distribution. The only region of the neighborhood price space over which there is little within-block variation is the very bottom of the distribution. The cheapest rooms at Harvard are physically separated from the rest; students asking for the cheapest rooms cannot be placed in high- or even mid-priced peer neighborhoods. The right graph shows that most blocks span a wide range of neighborhood private high school shares, with many blocks spanning the full range of possible shares, from zero to one. Overall, 24% of within-year variance in peer neighborhood price and 67% of within-year variance in peer private feeder share occurs within randomization blocks.

5.3 Balance tests and first stage effects

For the residential randomization design to provide evidence on the effects of exposure to high-status peers, assignment to more expensive rooms within a randomization block must affect the peer environment but be uncorrelated with students' baseline characteristics. Results reported in Panel A of Table 5 show that controlling for randomization blocks eliminates the strong cross-sectional relationship between neighborhood price rank RP_p and predetermined individual characteristics. We estimate versions of Equation 2 with the individual covariates X_i listed in the rows as the dependent variable and expanding sets of fixed effects as we move from left to right. Each cell reports estimates of the effect of RP_p. Recall that RP_p is a rank variable that ranges from zero to one, so coefficients can be interpeted as the change in the dependent variable associated with moving from the lowest-ranked peer neighborhood in a cohort to the highest.

The first column of Table 5 reports estimates from specifications which control only for entry

year dummies, i.e., not for randomization blocks. Students at the bottom of the peer neighborhood price distribution differ dramatically from those at the top. Students at the top of the distribution pay $207 more on average for their rooms (97.0% of the sample mean), are 39.4 percentage points more likely to have attended a private school, and 29.6 percentage points more likely to have attended a private feeder school. They are 7.1 percentage points (81.6% of the sample mean) more likely to have a Harvard father and 11.0 percentage points (47.8% of the sample mean) more likely to have a Harvard brother. They are less likely to report southern or eastern European heritage and more likely to have fathers who are doctors or lawyers. We reject the joint null of no relationship between neighborhood price and own attributes at the $p < 0.001$ level.

The second column of Table 5 adds fixed effects for each interaction of per-occupant price and entering cohort, while the third column adds the full set of randomization block fixed effects and main feeder high school dummies as in Equation 2. With the addition of block controls the relationship between neighborhood mean price and own attributes becomes economically and statistically insignificant. Note that price per student effects in columns two and three and the private feeder effect in column three are mechanically zero due to the control set. After controlling for per-occupant price by cohort blocks, we cannot reject the null that all effects are zero (p=0.929). This null relationship persists when we add the full set of randomization blocks and high school dummies (p=0.601). Our findings are consistent with primary source descriptions of the assignment process as conditionally random.

Panel B of Table 5 reports the relationship between neighborhood price rank and links to other data sources. These outcomes are downstream of room assignment and in principle could be affected by it. However, after conditioning on randomization block, we do not observe a relationship between neighborhood price rank and links to Class Report or census data sources. This result mitigates concerns related to differential censoring.

Students assigned to higher-priced neighborhoods have higher-status peers. Panel C of Table 5 reports estimates of Equation 2 with peer mean attributes on the left hand side. When measuring neighborhood mean attributes for student i, we leave out individual i and other students in i's room. Conditional on randomization block and high school fixed effects, students assigned to the highest-price neighborhoods have peer private high school shares that are 28.2 percentage points higher (on an in-sample mean of 59.3) than students assigned to the lowest-price neighborhoods, and private feeder high school peer shares that are 20.2 percentage points higher (on a mean of 31.4), conditional on randomization block. These effects are precisely estimated, with standard errors of 3.3 and 3.2, respectively. A 50 percentile increase in peer neighborhood rank— easily within the support of observed random variation— increases peer private feeder share by 32.1% of the sample mean. We see similar effects across other student attributes. Peer legacy shares are 6.3 percentage points higher, and peer eastern European immigrant shares are 4.9 percentage points lower. Exposure to rooms in the top 50% and top 10% of the room price distri-

bution also rises dramatically.

The in-college and post-college outcomes of residential peers also change with RP_p. We report these findings in Panel D of Table 5. These estimates should be interpreted cautiously because they capture both differences in peer background across groups and endogenous social effects (Manski, 1993). Conditional on randomization blocks and high school dummies, a zero-to-one increase in neighborhood rank raises the mean count of first-year activities residential peers participate in by 0.34, the share of peers participating in social activities by 0.05, and the peer activity index by 0.30. It raises the mean peer academic rank group by 0.16, corresponding to reduced academic achievement. It raises the share of peers joining selective final clubs by 0.11—our main measure of social success. Residential peers are more likely to report finance careers, have a higher peer occupation index, and higher wage income. Our overall interpretation of Table 5 is that this is a promising setting for learning about the effects of exposure to high-status peers. Assignment to higher-priced residential neighborhoods is conditionally uncorrelated with predetermined student attributes but produces large changes in peer environment in terms of pre-college background, in-college activities, and post-college careers.

5.4 Social and academic success at college

Table 6 reports the effects of assignment to higher priced peer neighborhoods on social and academic outcomes while at college. We present separate estimates for the full sample, for private feeder students, and for other students. The 'Test' column reports p-values from statistical tests of the null that effects for private feeder students and other students are equal.

Placement in a higher-priced peer neighborhood raises participation in social activities while in college. Panel A of Table 6 reports effects of peer neighborhood price on freshman extracurricular activities. A change in peer neighborhood rank from zero to one raises the share of students who report any activity by 7.0 percentage points (SE=3.4; relative to a residential sample mean of 60.5%), the activity count by 0.283 (SE=0.110; mean of 1.267), and the percent of students reporting leadership roles in any activity by 4.3 percentage points (SE=2.0; mean of 8.4%). The aggregate effect of changes in quantity and type of activities is to raise the activity private high school index by 0.208 standard deviations (SE=0.082), equal to 34.1% of the gap in activity index between private feeder students and other students in the full sample. Looking across activities, we observe the largest point estimates for social committees (4.9 percentage points) and music.

The full sample effects are entirely driven by students from private feeder schools. A zero-to-one change in neighborhood rank raises the count of activities that private feeder school students participate in by 0.648 (SE=0.222), and the private school activity index by 0.494 (SE=0.171). The activity types that drive the index effect for private feeder students are leadership positions and schoolwide social committees. In particular, social committee participation rises by 18.3 percentage points (SE=5.7) with a one-unit change in room rank. For students not from private

feeder schools, the effects of placement in higher-priced peer neighborhood on any activity, the count of activities, activity leadership, and all specific activity types are economically small and not statistically different from zero at conventional levels. The p-values from tests of the null of equal effects for private feeder and other students are roughly 0.01 for the summary activity count and activity index measures.

We do not find evidence that private feeder students assigned to high-status residential peer groups shift away from activities with lower-status students. Effects are positive or near zero across the board, even for activities with relatively low shares of private feeder students, such as music or those in the "other activities" category. Placement in high-priced neighborhoods leads to increased participation in high-status activities, not substitution between activities.

We now turn to our main measure of social success in college: selective upper-year final clubs. Our findings, reported in Panel B of Table 6, are similar to those for first-year activities. In the full sample, a one-unit change in neighborhood rank raises the rate at which students join selective final clubs by 6.5 percentage points (SE=2.1; sample mean of 9.5). This effect is driven entirely by a 16.7 percentage point effect for private feeder students (SE=5.6), with a precise zero effect for other students. We reject the null of equal effects across the two groups (p=0.004).

We see small negative effects for private feeder students on participation in less selective final clubs, suggesting that some of the increase in selective final club participation is an intensive margin effect (i.e., participating in a more exclusive club rather than a less exclusive one). Hasty Pudding participation and the upper-year club index also rise for private feeder students but not for other students. Statistical tests of differences between the effects for private feeder students and other students reject the null at p-values of roughly 0.05 for both outcomes.

The social effects we see are economically large. A fifty percentile increase in peer neighborhood rank raises activity counts for private feeder students by 18.8% of the sample mean in that group and 48.7% of the gap between private feeder and other students. The equivalent figures for social committees are 45.2% and 66.2%; for selective final clubs they are 37.7% and 45.0%.

Panel A of Figure 6 uses a binscatter graph to show how the gap in selective final club membership by high school type grows as randomized exposure to high-status peers increases. We regress final club membership and neighborhood price rank on randomization block and high school fixed effects as in equation 2, splitting the sample by high school type. We then plot the mean selective final club residual against the neighborhood price rank mean residual at each decile of its distribution. To capture level differences across high school types, we add the sample mean back to the residuals for each group. For private feeder students, final club membership rates grow steadily with neighborhood price rank. For other students, club membership rates stay flat at a lower level. The gap between the two groups grows by about two thirds as we move from the bottom decile to the top decile of the within-block neighborhood rank distribution. This corresponds to a roughly 50 percentile increase in neighborhood rank.

24

In contrast to our findings for social activities, assignment to a higher-priced peer neighborhood has no effect on academic rank. Results reported in Panel C of Table 6 show that the effects of assignment to higher-priced peer neighborhoods on academic rank group membership are close to zero in most cases and not statistically significant at conventional levels. The effect of assignment to higher-priced peer neighborhoods on a continuous class rank measure is also near zero. Panel B of Figure 6 shows that the gap between private feeder students and their better-performing peers from other school types is stable across the distribution of neighborhood rank.

We *do* see evidence that assignment to a higher-priced neighborhood raises the rates at which students show up in the rank lists, indicating that they do not have incomplete or non-passing grades. In the full sample, a fifty percentile change in neighborhood rank raises the rate of list appearance by 3.7 percentage points, 4.6 percent of the sample mean. We observe effects of similar size for private feeder and other students. Overall, we interpret these results as evidence that higher-status peers do not generally affect grades but may modestly reduce the risk of very bad academic outcomes or failure to complete a term on time.

5.5   Adult social lives and career paths

Twenty-five years after graduation, assignment to higher-priced peer neighborhoods raises adult participation in social organizations, but again only for private feeder students. Panel A of Table 7 reports the effects of residential peer group rank on participation in adult social organizations. In the full sample, a zero-to-one change in peer neighborhood rank raises the rate of membership in any adult social organization by 6.6 percentage points (SE=3.6, sample mean of 38.1%). The full-sample increase reflects the combination of a 24.1 percentage point increase for private feeder students (SE=6.7) and a zero effect for others. We reject the null that the two effects are equal (p=0.003). A 50 percentile change in peer neighborhood price raises rates of participation in adult social clubs for private feeder students by 26.0% of their sample mean. Panel C of Figure 6 shows the increase in the gap in adult social club participation by high school type as one moves from the bottom to the top of the distribution of random variation in neighborhood price.

Within the social organization category, increased participation rates for private feeder students are driven mostly by country clubs, for which participation rates rise by by 21.9 percentage points (SE=6.6). We see no effects for membership in professional or honor societies. These findings again parallel descriptive results from Table 3 showing that members of selective final clubs were more likely to participate in adult social clubs but not professional or honor societies.

Exposure to higher-status peers causes students from private feeder schools to pursue careers with higher private school shares but pushes other students in the opposite direction. We report these findings in Panel B of Table 7. Focusing first on the occupation private high school index, we see an effect near zero in the full sample. This reflects offsetting effects for private feeder students and other students. For private feeder students, a zero-to-one change in peer neigh-

borhood price raises the occupation index by 0.450 (SE=0.181). For other students it lowers that index by 0.164 (SE=0.093). The p-value from a test that these two effects are equal is 0.002.

Finance is the key driver of shifts in the private school index. Assignment to a higher-priced neighborhood raises the share of private feeder students going into finance by 14.4 percentage points (SE=5.8), on a base of 18.1% in the residential sample. Recall that finance is the occupation with the largest share of private feeder students. Exposure to higher status peers decreases the rates at which other students go into finance. A test of the null that the finance effect is the same for private feeder students and other students returns a p-value of 0.002. There are smaller shifts across other occupations, with private feeder students shifting away from medicine, higher education, and teaching, and other students shifting towards medicine. Both groups shift modestly towards an aggregate "business" category combining management, accounting, and retail.[26]

As with social outcomes, effects on occupational outcomes are economically large. A 50 percentile increase in peer mean room price raises the share of private feeder students pursuing finance careers by 39.6% of their sample mean and 80.1% of the gap by high school type. For the occupation index, this effect is equal to 69.5% of the gap by high school type. Panel D of Figure 6 shows how the gap in the occupation index increases for private feeder students and decreases for others as neighborhood price rank increases. As was the case for short- and long-run social outcomes, exposure to higher-status peers tends to expand gaps by baseline social status.[27]

5.6 Census outcomes

We next estimate the effects of residential peer neighborhood on outcomes from the 1940 Census. Sample restrictions required by the Census match procedure reduce sample size by roughly two thirds relative to specifications with Class Report outcomes. This smaller sample produces confidence intervals that are too large to be informative. We report our findings in Panel C of Table 7. To take one example, the standard error of our estimate of the effects of a one-unit change in peer price rank on wage income topcoding for private feeder students is 9.9 percentage points. Given our point estimate of approximately zero, the 95% CI spans increases and decreases in topcode rates of close to 20 percentage points in each direction, on a private feeder sample mean

---

[26]This shift is driven by senior management and accounting for private feeder students, and by lower management for others. See Online Appendix Table A.2.

[27]The decrease for low-status students differs somewhat from zero effects for this group we observe in most other domains. One possible rationale for the shift away from private school typical careers among non-private feeder students is social identity theory (Tajfel et al., 1971; Goette et al., 2006, 2012), which argues that making group distinctions salient can produce group-favoring behavior and expand cross-group gaps. Carrell et al. (2013) similarly provide an example where placing low-achieving students with high-achieving peers leads to endogenous segregation and harms low achievers. Looking across outcomes, we see some support for this idea, in the form of modest (though typically statistically insignificant) shifts away from private feeder typical activities such as social committees and Hasty Pudding. Exposure to high-status peers also modestly reduces marriage rates for lower-status students (see section 5.8) . However, the presence of null effects on most other outcomes for lower-status students and the difficulty of ruling out other causal channels argues for a cautious interpretation.

of 25.0%, so we cannot rule out a doubling in topcode rate or a fall to near zero.

A more informative approach is to combine outcomes from the Class Reports with Census earnings data to construct an earnings index that captures how much we might expect earnings to change based on the shifts in occupation and adult social outcomes that we observe in the Class Reports. Using the set of students matched to both Census earnings records and Class Reports, we regress earnings on indicators for each broad occupation and adult social activity category. As with our private school indices, we select covariates using a Lasso, using EBIC for model selection. Our index is the predicted value from this regression, which we can compute for all students in the Class Report occupation sample, not just those matched to the Census. Units are dollars. We place this Class Report earnings index on the left side of equation 2.

This approach can be thought of as a Harvard-specific version of the occupation indices often used to analyze historical census records (Olivetti and Paserman, 2015; Feigenbaum, 2018; Saavedra and Twinam, 2020; Abramitzky et al., Forthcoming).[28] At minimum, it provides insight into whether the kinds of shifts in Class Report outcomes we see in response to exposure to high-status peers are associated with higher earnings in the cross section. Under additional (strong) assumptions, it may provide a guide to what we would see if we could observe earnings for more students in our Harvard sample. These assumptions are a) that the relationship between earnings and Class Report outcomes for students not matched to the census are the same as those who are matched, and b) that earnings responses to shifts in Class Report outcomes induced by peer environment affect earnings as in the cross section

We report our findings in the bottom row of Table 7. A 50-percentile shift in peer neighborhood rank raises the Class Report earnings index for private feeder students by $64 (SE=24). The effect for other students is a precise zero. While our sample of census-matched students in the room randomization analysis is small and yields imprecise estimates, the shifts we see in Class Report outcomes are generally associated with higher earnings.[29]

5.7 Alternate modeling choices

Additional analyses explore the sensitivity of our quasi-experimental findings to alternate modeling decisions. Online Appendix Tables A.3 through A.6 show that our findings are similar

---

[28]Recall that standard occupation scores are ill-suited to our context because the business and finance careers that Harvard students pursue are poorly categorized in Census data. See section 3.3 and Online Appendix B.3.

[29]The size of the implied changes in earnings is similar to what we observe in cross-sectional data. Consider an instrumental variables regression in the sample of private feeder students in which neighborhood price rank is the instrument, selective final club membership is the endogenous regressor, and the Class Report wage index is the outcome. The IV estimate of the final club premium is $767 (=$129/0.167), where the $129 is from Table 7 and the 0.167 first-stage effect is from Table 6. This is similar to our $777 OLS estimate of the final club premium in our base descriptive specification (Panel B of Table 4). The IV exclusion restriction requires that the changes in college social lives we document in Table 6 do not affect earnings independently of final club membership. This is a strong assumption, so we interpret the IV estimate cautiously.

when we exclude high school fixed effects or when we both exclude high school fixed effects and limit the randomization blocks to interactions between per-occupant price and year, dropping interactions with occupant count. As suggested by the balance analysis in Table 5, year-specific room price dummies are the key controls. Online Appendix Tables A.7 and A.8 report findings for alternate definitions of the private feeder group and for private non-feeder schools. Our findings for private feeders are similar when we consider more or less expansive group definitions. Effects for private non-feeder students are similar to those for other non-feeder students. Online Appendix Tables A.9 and A.10 report results that drop the 1919, 1920, and 1921 entering cohorts. These students arrived on campus before the first reference to the room-randomization scheme we find in primary source documents, so it is possible they were subject to a different procedure. Dropping them does not affect our findings. Online Appendix Tables A.11 and A.12 report results using randomization inference as opposed to clustered standard errors to conduct statistical tests; our results are unaffected. Finally, Online Appendix Figure A.6 displays results from specifications in which we replace our linear measure of peer neighborhood rank with dummies for terciles of rank, measuring the effects of placement in the second or third tercile of neighborhood price rank relative to the first. We observe rising effects across terciles on selective social and career outcomes for private feeder students, while the effects for students not from private feeder schools are near zero across all outcomes.

5.8 Mechanisms and heterogeneous effects

Several supplemental analyses shed light on the pathways through which residential peers shape long-run outcomes. We summarize our findings here. See Online Appendix B.8 for details.

Academic achievement. The shifts in social and career outcomes that we observe for private feeder students are driven by students who are also academic high achievers. The shift towards finance careers comes from relatively high achieving high status students who otherwise would have become doctors or professors, not from lower-achieving students.

The Great Depression. The students in classes of 1930 and later graduated into the teeth of the Great Depression, which could affect the importance of social ties for early-career job finding (Kramarz and Skans, 2014). However, we see little evidence that the effects of residential peers vary with the timing of Depression exposure.

Major choice and career intent. School peers may shape long run outcomes by changing academic specialization in college, by changing career goals at the time of college completion, or by changing how careers play out. Consistent with Sacerdote (2001), we find no evidence that peers change students' major choices. We observe some evidence of modest effects on stated career intent at the time of graduation. Changes in career paths due to residential peer assignment appear to come mostly from changes in outcomes conditional on specialization and goals.

The marriage market. Exposure to high-status residential peers may also shape outcomes in the

marriage market. Historical accounts of Harvard social clubs such as Amory (1947) describe their importance for marriage outcomes, and studies in other settings emphasize the importance of marriage for the intergenerational transmission of social status (Ager et al., 2019). We document strong assortative matching on religious/cultural background. We then show that assignment to high-status residential peers raises the chances that high-status students marry wives with Colonial birth names, increasing marital assortativity.

6 Harvard students across the 20th century

6.1 The long-run Harvard time series

Our main results speak to the role of social interactions among high-status peers in the formation of the mid-century American elite. This section extends the descriptive analysis through Generation X and Millennial cohorts in the 2010s labor market. The goal is to understand how patterns of enrollment, academic performance, and career choice have changed, and what this says about upward mobility among the educational elite across the 20th century.

We extend the time series for key outcomes through the present. To describe student attributes, we use Freshman Red Books for each graduating class through 2015. For academic and career outcomes, we link the Red Books to 25th Reunion Class Reports digitized at five year intervals through 1990; i.e., 1945, 1950, and so on. Census data are not available after 1940, but Latin honors at graduation and career descriptors are. We use first names to classify students on the basis of gender and extend our race/ethnicity classification scheme to include Black, Latino, and Asian students based on names and hand-codes of headshots.[30] We retain the same set of private feeder schools for the long-run analysis, while extending our classification of public feeders to include schools that send many students to Harvard after the 1930s. To facilitate cross-time comparisons, we classify medical and law careers based on degree receipt.

While systematic reporting of social activities is not available for recent cohorts, we were able to access a "club catalog" for the A.D. club— one of the selective final clubs we identify in section 3 (A.D. Club, 2015). This document lists all A.D. members through the class of 2015.

Online Appendix B.9 describes data construction and presents evidence that the data allow for consistent cross-time comparisons of group outcomes. Several pieces of evidence indicate the data we construct is of high quality. We match publicly reported aggregate statistics on academic honors, gender mix, and the share of Asian students almost exactly. Shares of under-represented minority (URM) students closely follow published aggregates through the 1980s,

---

[30]Because hand-coding is costly, we code three out of every five years after 1955, focusing on blocks surrounding each five-year interval. e.g. 1989, 1990, and 1991 for the 1990 block. We drop uncoded cohorts from race/ethnicity analyses. Because fashions in first names may change, race/ethnicity codes in the long-run series use only surnames. Note that while name and visual codes are common in practice (Bertrand and Mullainathan, 2004; Fiscella and Fremont, 2006; Adukia et al., 2021), they are imperfect measures of self-reported racial identity (Feliciano, 2016).

though they fall below published values after 1990 (after the termination of our Class Report series). Shares of students pursuing different graduate degrees over time are similar to survey reports in Goldin and Katz (2008), and we see little evidence of differential reporting across groups. In sum, these data capture big-picture trends in other data sources, extend the time horizon on available aggregates, and allow for new comparisons of outcomes across groups.

6.2 Harvard diversifies

Harvard became dramatically more diverse in terms of gender and race/ethnicity between the 1920s and 2010s. Table 8 describes the gender and race/ethnicity of Harvard students in four blocks, each 20 to 25 years long. Panel A of Table 8 reports statistics for Harvard students as a whole. We start observing female students in our data in the class of 1975. By the early 2000s, the class is split evenly between men and women. The share of students with Jewish surnames rises after World War II, then falls again in the modern era. Shares of Asian and URM students are near zero through the late 1960s, then rise sharply. As reported in Panels B and C, students from private and public feeder schools diversify in parallel with the broader population. The share of students from private feeders declines as Harvard's national footprint expands (Hoxby, 2009).

The picture is different for A.D., the selective final club. Increases in female and non-white enrollment at Harvard do not carry over to club members. The club remains all male through 2015, and only 9% of post-1990 members are non-white, compared to 29.5% of students overall. The club *does* appear to assimilate Jewish students by the 1990s, with shares similar to Harvard as a whole. The share of members from private feeder schools falls sharply, from 81% before 1945 to 9.0% after 1990 (though the 9% value is more than double the overall share). For club membership, distinctions between white ethnic groups and the dominance of the private feeders decline in importance over time but divides by race and gender emerge as Harvard integrates.

6.3 Grades and career paths over time

We next consider academic performance, taking Latin honors at graduation as our measure of interest. Because nearly 80% of students graduate with some kind of honors degree by the late 20th century, we focus on *magna* or *summa cum laude* graduates. By this definition, the share of academic high achievers rises from 8% in the mid 1920s to 33% in the late 1980s. Healy (2001) attributes the increase in honors degrees to grade inflation during the Vietnam War.

There are persistent differences in academic performance by high school type, race/ethnicity, and club membership. Panel A of Figure 7 shows that students from public feeder high schools outperform students from private feeder high schools over the full 1924-1990 period. Panel B shows that students with Jewish surnames outperform students with Colonial surnames over the full period as well. In more recent cohorts, students we categorize as Asian are more likely

to graduate with high honors, and students we categorize as URMs are less likely. Finally, as shown in Panel C, members of the A.D. final club are much less likely to graduate with high honors than other students across the 20th century.

We next describe the evolution of career choices. Our first result is that the shares of students pursuing finance careers by high school type and race/ethnicity converge over time. Panels D and E of Figure 7 show a steady decline in the finance gap between private and public feeder schools, and between students with Jewish and Colonial surnames. For the classes between 1970 and 1990, the shares of students pursuing finance by race/ethnicity cluster between 6.3% (for Jewish students) and 11.1% (for Colonial students), with rates for other white students, URM students, and Asian students all between 8.7% and 10.8%.

Our second result is that members of the A.D. final club are much more likely than other students go into finance over the full 1924-1990 period. As shown in Panel F of Figure 7, this gap expands over time. 33% of club members in 1970s and later cohorts work in finance, compared to 9.2% of other students. Table 9 reports how other career outcomes vary over time. Club members are less likely than other students to work in medicine and higher education over the full sample period, and more likely to get MBA degrees in the post-war period.

The patterns of career choice, academic performance, and demographic composition that we observe for the A.D. club between 1923 and 2015 recall those we observed for all selective final clubs in our main sample. We note that our long-run findings may not apply to other final clubs, and that the continuity of descriptive patterns does not imply the continuity of causal mechanisms. Our findings do, however, yield an existence result: the "social path" for high-status, academically low-performing students to final clubs and then to finance careers remains operative across the 20th century, with some of the same institutions involved.

Other cross-group comparisons reported in Table 9 are also of interest. For example, we observe large gaps in finance careers by gender, emerging gaps in academic careers and MBA receipt by high school type, and a persistent Jewish-Colonial gap in medical careers.

6.4 Income mobility for recent cohorts

To what extent does family background continue to predict income for students at elite universities? We address this question using data from Chetty et al. (2020). These data describe the joint distribution of parent and child income for children born between 1980 and 1982, split by university type. We focus on the Ivy+ tier, which consists of the eight Ivy League universities (including Harvard), Duke, Stanford, MIT, and UChicago.

Panel A of Figure 8 displays the histogram of parent income among students at Ivy+ universities. As reported in Chetty et al. (2020), Ivy+ students are drawn disproportionately from the top of the income distribution. 2.2% come from the top 0.1% of the parent income distribution, 14.5% from the top 1%, and 68% from the top quintile. Understanding the child-parent SES

gradient within Ivy+ universities requires zooming in on high parent incomes.

Panels B and C of Figure 8 show that, within Ivy+ universities, the children of the richest parents have much higher incomes than other children. The figures display mean income and (age-adjusted) top 1% share for children within bins defined by parent income percentile. We focus on the top quintile of the parent income distribution, and show separate points for each percentile within the top 5% of that distribution, as well for as the top 0.1%.

Mean earnings for children are fairly flat between the 80th and 95th percentiles of the parent income distribution, which accounts for 26.9% of Ivy+ student enrollment. Ivy+ students with parents between the 80th and 90th percentile of the income distribution earn $135,000 on average, similar to the $132,000 average for students with parents between the 90th and 95th percentiles.[31] Starting around the 95th percentile, the slope ticks up. Students with parents between the 99th and 99.9th percentiles earn $176,000 on average, 33% more than students with parents in the 90-95th percentiles. Students with parents in the top 0.1% earn an average of $233,000, 76% more than students from 90th-95th percentile parents. For top 1% income shares, the story is similar. Roughly 25% of students from top 1% parents are in the top 1% of the income distribution themselves, 40% higher than for students from 90-95th percentile parents.[32]

The gaps in top income shares between Ivy+ students from top 1% parents and other Ivy+ students among 1980s birth cohorts are similar to the differences by high school type that we observe in the historical Harvard data. In our main analysis sample, 24.7% of Harvard students from private feeder schools have incomes in the top 1% of the earnings distribution for men in their birth cohorts ($3945 or higher), 62% higher than the 15.2% for other Harvard students. Private feeder attendance and parent income are not perfect proxies for one another. However, these findings do indicate that gaps in child outcomes by parent income for older Millennials are large in historical terms— i.e., they are of fairly similar size those generated by the extensive social, academic, and career differences we document among Greatest Generation cohorts.

6.5 University institutions and policies

This section summarizes how various university institutions and policies may mediate the long-run patterns that we observe. See Online Appendix B.10 for details.

---

[31]This relationship remains flat for lower parent incomes. See Online Appendix Table A.13.

[32]These findings may appear to differ from results reported in Chetty et al. (2020), who describe a flat, linear relationship between average child income *rank* and parent income rank across the income distribution. In fact, our findings are consistent with their results— the linear relationship between average child rank and parent income rank coexists with the sharp rise in average income and top 1% shares above the 95th percentile. This occurs because the share of children not working or with low incomes also rises at the top of the parent income distribution, and approximately balances out the increasing shares of very high-income children when averages are taken over ranks (as opposed to dollars). Online Appendix Figure A.7 reproduces the flat, linear rank-rank relationship reported in Figure III.C of Chetty et al., following Chetty et al. by aggregating across ventiles of parent income. Looking within the top ventile of the parent income distribution, mean child rank has a hump shape, first rising and then falling steeply for children of the highest-income parents. See Online Appendix Table A.13 for descriptive statistics.

Admissions policy. That students from high-status backgrounds tend to perform less well academically than others is consistent with available data on admissions standards. Students from privileged backgrounds appear to have retained substantial admissions advantages over the 20th century. Karabel reports that in 1940 only one out of 77 applicants from the elite private schools was rejected, compared to 14 out of 59 applicants from Boston Latin (p. 174). Arcidiacono et al. (2019) reports that applicants who are legacies, likely donors, or children of faculty (LDC) are more likely to be admitted to Harvard than other students whose academic performance is better. For example, LDC students in the third decile of Harvard's academic performance index are more likely to be admitted than non-LDC students in the 10th decile. The 19% LDC share in recent Harvard cohorts is similar to the 24% private feeder share in our main sample period.

Harvard introduced non-academic admissions criteria in 1926. Online Appendix Figure A.8 shows that the share of students with Jewish names fell substantially following this change.

Academic success. We observe a final club membership premium that is much larger than the academic success premium. A comparison to Goldin and Katz (2008) suggests that the returns to academic performance at Harvard may have risen over time, but that the social success premium we observe is large even compared to returns to academic success for recent cohorts.

Field of study. Consistent with recent evidence (Goldin and Katz, 2008; Bleemer and Mehta, Forthcoming), descriptive analyses show that majoring in economics offers a sizeable earnings return and a path to finance careers for pre-war Harvard cohorts. However, major choice does not mediate the career effects of final club membership. The academic path towards finance is separate from the social path in the cohorts we study.

Sports participation. Intercollegiate sports were central to social life at Harvard in the period we study. For example, Harvard and Porcellian alumnus Teddy Roosevelt equated the value of college sports to academic work (Roosevelt, 1900). Shulman and Bowen (2011) show that recent cohorts of college athletes tend to have lower grades than other students, but earn more. We observe the same patterns among pre-war Harvard students. Descriptively, roughly 50% of the earnings premium associated with sports participation is explained by athletes' social success.

7   Conclusion

This paper shows how social interactions at Harvard shaped the business and civic elite in the post-war United States. Our key finding is that the benefits of exposure to high-status peers accrue only to students from high-status backgrounds. Even prolonged residential contact with high-status peers is not sufficient to help lower-status students access exclusive old boys' clubs and the career and social outcomes with which club membership is correlated.

Our results are consistent with the broad point that social factors are an important determinant of labor market success (Granovetter, 1973; Deming, 2017; Lleras-Muney et al., 2020) and

highlight the extent to which these factors may expand baseline differences across social groups. Our exercise contributes in three ways: 1) by testing the general proposition that social interactions shape high-stakes outcomes, 2) by showing that social interactions among educational elites *did* shape the composition of the US elite at an important historical juncture, and 3) by evaluating the specific policies that Harvard used at the time to promote cross-group interaction on campus, and that many institutions continue to use today.

A large body of social science research emphasizes the importance of inclusive institutions and competition between different groups of elites for economic growth (Acemoglu and Robinson, 2000, 2006, 2008, 2012), as well as the consequences of ingroup bias for efficiency (Alesina and Ferrara, 2005; Hjort, 2014; Fisman et al., 2017). Elite universities have a potentially important role to play in identifying talented students from underrepresented groups and helping them access influential positions in the economy and society. However, our findings suggest that expanding formal access to elite higher education may not by itself integrate the old boys' clubs upon which access to some top positions seems to depend. What policymakers at universities and beyond can do to promote broader access to these groups is a subject for future work.

References

Abramitzky, Ran, Leah Boustan, and Katherine Eriksson, "Do immigrants assimilate more slowly today than in the past?," *American Economic Review: Insights*, 2020, *2* (1), 125–41.

__ , **Leah Platt Boustan, and Dylan Connor**, "Leaving the enclave: Historical evidence on immigrant mobility from the industrial removal office," Technical Report, National Bureau of Economic Research 2020.

__ , __ , **and Katherine Eriksson**, "Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration," *American Economic Review*, 2012, *102* (5), 1832–56.

__ , __ , **and** __ , "A nation of immigrants: Assimilation and economic outcomes in the age of mass migration," *Journal of Political Economy*, 2014, *122* (3), 467–506.

__ , __ , **Elisa Jácome, and Santiago Pérez**, "Intergenerational Mobility of Immigrants in the US over Two Centuries," *American Economic Review*, Forthcoming.

__ , __ , **Katherine Eriksson, James J Feigenbaum, and Santiago Pérez**, "Automated Linking of Historical Data," Working Paper 25825, National Bureau of Economic Research May 2019.

_ , **Roy Mill, and Santiago Pérez**, "Linking individuals across historical sources: A fully automated approach," *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 2020, *53* (2), 94–111.

Acemoglu, Daron and James A. Robinson, "Political losers as a barrier to economic development," *The American Economic Review*, 2000, *90* (2), 126–130.

_ **and** _ , "Economic backwardness in political perspective," *American Political Science Review*, 2006, *100* (1), 115–131.

_ **and** _ , "Persistence of Power, Elites, and Institutions," *The American Economic Review*, 2008, *98* (1), 267–93.

_ **and** _ , *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*, Crown Business, 2012.

A.D. Club, "The AD Club of Harvard University," 2015.

Adukia, Anjali, Alex Eble, Emileigh Harrison, Hakizumwami Birali Runesha, and Teodora Szasz, "What We Teach About Race and Gender: Representation in Images and Text of Children's Books," *University of Chicago, Becker Friedman Institute for Economics Working Paper*, 2021, (2021-44).

Ager, Philipp, Leah Platt Boustan, and Katherine Eriksson, "The intergenerational effects of a large wealth shock: White southerners after the Civil War," Technical Report, National Bureau of Economic Research 2019.

Alesina, Alberto and Eliana La Ferrara, "Ethnic diversity and economic performance," *Journal of economic literature*, 2005, *43* (3), 762–800.

Allport, Gordon Willard, "The nature of prejudice," 1954.

Altonji, Joseph G, Peter Arcidiacono, and Arnaud Maurel, "The analysis of field choice in college and graduate school: Determinants and wage effects," in "Handbook of the Economics of Education," Vol. 5, Elsevier, 2016, pp. 305–396.

Alvaredo, Facundo, Anthony B. Atkinson, Thomas Piketty, and Emmanuel Saez, "The Top 1 Percent in International and Historical Perspective," *Journal of Economic Perspectives*, September 2013, *27* (3), 3–20.

_ , **Lucas Chancel, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman**, "Global Inequality Dynamics: New Findings from WID.world," *American Economic Review*, May 2017, *107* (5), 404–09.

Amory, Cleveland, *The Proper Bostonians*, Vol. 1, EP Dutton, 1947.

Angrist, Joshua D, "The perils of peer effects," *Labour Economics*, 2014, *30*, 98–108.

Arcidiacono, Peter and Michael Lovenheim, "Affirmative action and the quality-fit trade-off," *Journal of Economic Literature*, 2016, *54* (1), 3–51.

___ , **Esteban M Aucejo, and V Joseph Hotz**, "University differences in the graduation of minorities in STEM fields: Evidence from California," *American Economic Review*, 2016, *106* (3), 525–62.

___ , **Josh Kinsler, and Tyler Ransom**, "Legacy and Athlete Preferences at Harvard," Working Paper 26316, National Bureau of Economic Research September 2019.

Auchincloss, Kenneth, "The Final Clubs: Little Bastions of Society In a University World that No Longer Cares," *Harvard Crimson*, November 1958.

Bailey, Martha, Connor Cole, Morgan Henderson, and Catherine Massey, "How Well Do Automated Linking Methods Perform? Lessons from US Historical Data," *Journal of Economic Literature*, December 2020, *58* (4), 997–1044.

Baltzell, Edward Digby, *The protestant establishment: Aristocracy & caste in America*, Yale University Press, 1964.

Baumol, William J, "Entrepreneurship: Productive, Unproductive, and Destructive," *Journal of Political Economy*, 1990, *98* (5 Part 1), 893–921.

Bertrand, Marianne and Sendhil Mullainathan, "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination," *American economic review*, 2004, *94* (4), 991–1013.

___ , **Sandra E Black, Sissel Jensen, and Adriana Lleras-Muney**, "Breaking the glass ceiling? The effect of board quotas on female labour market outcomes in Norway," *The Review of Economic Studies*, 2019, *86* (1), 191–239.

Black, Dan A and Jeffrey A Smith, "Estimating the returns to college quality with multiple proxies for quality," *Journal of labor Economics*, 2006, *24* (3), 701–728.

Bleemer, Zachary, "Top percent policies and the return to postsecondary selectivity," *Unpublished Manuscript*, 2020.

___ **and Aashish Mehta**, "Will Studying Economics Make You Rich? A Regression Discontinuity Analysis of the Returns to College Major," *American Economic Journal: Applied*, Forthcoming.

Boisjoly, Johanne, Greg J Duncan, Michael Kremer, Dan M Levy, and Jacque Eccles, "Empathy or antipathy? The impact of diversity," *American Economic Review*, 2006, *96* (5), 1890–1905.

Bol, Thijs and Kim A Weeden, "Occupational closure and wage inequality in Germany and the United Kingdom," *European Sociological Review*, 2015, *31* (3), 354–369.

Bourdieu, Pierre, *The state nobility: Elite schools in the field of power*, Stanford University Press, 1998.

Bowden, Witt, "War and Postwar Wages, Prices, and Hours, 1914-23 and 1939-44," Bulletin 852, United States Bureau of Labor Statistics 1946.

Cappelli, Peter and Monika Hamori, "The path to the top: Changes in the attributes and careers of corporate executives, 1980-2001," Technical Report, National Bureau of Economic Research 2004.

Carrell, Scott E, Bruce I Sacerdote, and James E West, "From natural variation to optimal policy? The importance of endogenous peer group formation," *Econometrica*, 2013, *81* (3), 855–882.

_ , **Mark Hoekstra, and Elira Kuka**, "The long-run effects of disruptive peers," *American Economic Review*, 2018, *108* (11), 3377–3415.

_ , _ , **and James E West**, "The impact of college diversity on behavior toward minorities," *American Economic Journal: Economic Policy*, 2019, *11* (4), 159–82.

_ , **Richard L Fullerton, and James E West**, "Does your cohort matter? Measuring peer effects in college achievement," *Journal of Labor Economics*, 2009, *27* (3), 439–464.

Cattaneo, Matias D, Richard K Crump, Max H Farrell, and Yingjie Feng, "On binscatter," *arXiv preprint arXiv:1902.09608*, 2019.

Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan, "Income segregation and intergenerational mobility across colleges in the united states," *The Quarterly Journal of Economics*, 2020.

_ , **Nathaniel Hendren, and Lawrence F Katz**, "The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment," *American Economic Review*, 2016, *106* (4), 855–902.

Cohen, Lauren, Andrea Frazzini, and Christopher Malloy, "The small world of investing: Board connections and mutual fund returns," *Journal of Political Economy*, 2008, *116* (5), 951–979.

Committee on Unrecognized Single-Gender Social Organizations (CUSGSO), "Final Report of the Committee on Unrecognized Single-Gender Social Organizations," Technical Report, Harvard University September 2017.

Corno, Lucia, Eliana La Ferrara, and Justine Burns, "Interaction, stereotypes and performance: Evidence from South Africa," Technical Report, IFS Working Papers 2019.

Costa, Dora L, Noelle Yetter, and Heather DeSomer, "Wartime health shocks and the postwar socioeconomic status and mortality of union army veterans and their children," *Journal of Health Economics*, 2020, *70*, 102281.

Dale, Stacy Berg and Alan B Krueger, "Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables," *The Quarterly Journal of Economics*, 2002, *117* (4), 1491–1527.

_ **and** _ , "Estimating the effects of college characteristics over the career using administrative earnings data," *Journal of human resources*, 2014, *49* (2), 323–358.

Davidson, James D, Ralph E Pyle, and David V Reyes, "Persistence and change in the Protestant establishment, 1930–1992," *Social Forces*, 1995, *74* (1), 157–175.

Deming, David J, "The growing importance of social skills in the labor market," *The Quarterly Journal of Economics*, 2017, *132* (4), 1593–1640.

Dillon, Eleanor Wiske and Jeffrey Andrew Smith, "Determinants of the match between student ability and college quality," *Journal of Labor Economics*, 2017, *35* (1), 45–66.

Einiö, Elias, "Mixing the Rich and Poor: The Impact of Peers on Education and Earnings," Technical Report 2019.

Feigenbaum, James and Daniel P Gross, "Automation and the Fate of Young Workers: Evidence from Telephone Operation in the Early 20th Century," Working Paper 28061, National Bureau of Economic Research November 2020.

Feigenbaum, James J., "Multiple Measures of Historical Intergenerational Mobility: Iowa 1915 to 1940," *The Economic Journal*, 06 2018, *128* (612), F446–F481.

Feliciano, Cynthia, "Shades of race: How phenotype and observer characteristics shape racial classification," *American Behavioral Scientist*, 2016, *60* (4), 390–419.

Finseraas, Henning and Andreas Kotsadam, "Does personal contact with ethnic minorities affect anti-immigrant sentiments? Evidence from a field experiment," *European Journal of Political Research*, 2017, *56* (3), 703–722.

Fiscella, Kevin and Allen M Fremont, "Use of geocoding and surname analysis to estimate race and ethnicity," *Health services research*, 2006, *41* (4p1), 1482–1500.

Fisman, Raymond, Daniel Paravisini, and Vikrant Vig, "Cultural Proximity and Loan Outcomes," *American Economic Review*, February 2017, *107* (2), 457–92.

Fitzgerald, F Scott, *This side of paradise*, Scribner & Sons, 1920.

Flanagan, Caitlin, "Why Harvard shouldn't push its all-male final clubs to go co-ed," *Washington Post*, April 2016.

Fracassi, Cesare and Geoffrey Tate, "External networking and internal firm governance," *The Journal of finance*, 2012, *67* (1), 153–194.

Galbraith, John Kenneth, *The affluent society* 1958.

Glaeser, Edward L, Bruce I Sacerdote, and Jose A Scheinkman, "The social multiplier," *Journal of the European Economic Association*, 2003, *1* (2-3), 345–353.

Goette, Lorenz, David Huffman, and Stephan Meier, "The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups," *American Economic Review*, 2006, *96* (2), 212–216.

_ , _ , and _ , "The impact of social ties on group interactions: Evidence from minimal groups and randomly assigned real groups," *American Economic Journal: Microeconomics*, 2012, *4* (1), 101–15.

Goldin, Claudia and Lawrence F Katz, "Transitions: Career and family life cycles of the educational elite," *American Economic Review*, 2008, *98* (2), 363–69.

Granovetter, Mark S, "The Strength of Weak Ties," *American Journal of Sociology*, 1973, *78* (6), 1360–1380.

Harvard College Admissions and Financial Aid (HCAFA), "What we look for," 2020.

Harvard University, "Adopt Lot System for Senior Rooms," *The Harvard Crimson*, December 1924.

Healy, Patrick, "Matters of Honnor: Harvard's honors fall to the merely average," *Boston Globe*, October 2001.

Hjort, Jonas, "Ethnic divisions and production in firms," *The Quarterly Journal of Economics*, 2014, *129* (4), 1899–1946.

Hoekstra, Mark, "The effect of attending the flagship state university on earnings: A discontinuity-based approach," *The review of economics and statistics*, 2009, *91* (4), 717–724.

Hoxby, Caroline M, "The changing selectivity of American colleges," *Journal of Economic perspectives*, 2009, *23* (4), 95–118.

— , "The productivity of US postsecondary institutions," *Productivity in Higher Education*, 2018, pp. 1–60.

Jack, Anthony Abraham, *The privileged poor: How elite colleges are failing disadvantaged students*, Harvard University Press, 2019.

Jensen, Michael C., "The Cabots of Boston," *The New York Times*, March 1972.

Jones, Howard Mumford, "Boston, Seen Through a "State of Mind"," Oct 1947.

Jones, Todd R and Michael S Kofoed, "Do peers influence occupational preferences? Evidence from randomly-assigned peer groups at West Point," *Journal of Public Economics*, 2020, *184*, 104154.

Karabel, Jerome, *The chosen: The hidden history of admission and exclusion at Harvard, Yale, and Princeton*, Houghton Mifflin Harcourt, 2006.

Katz, Lawrence F, Jeffrey R Kling, and Jeffrey B Liebman, "Moving to opportunity in Boston: Early results of a randomized mobility experiment," *The Quarterly Journal of Economics*, 2001, *116* (2), 607–654.

Khan, Shamus, "The sociology of elites," *Annual Review of Sociology*, 2012, *38*, 361–377.

Kramarz, Francis and Oskar Nordström Skans, "When strong ties are strong: Networks and youth labour market entry," *Review of Economic Studies*, 2014, *81* (3), 1164–1200.

Levine, Steven B., "The Rise of American Boarding Schools and the Development of a National Upper Class," *Social Problems*, 1980, *28* (1), 63–94.

Lleras-Muney, Adriana, Matthew Miller, Shuyang Sheng, and Veronica T Sovero, "Party On: The Labor Market Returns to Social Networks and Socializing," Working Paper 27337, National Bureau of Economic Research June 2020.

Low, Josiah O., "The Proper Bostonians. Cleveland Amory," *American Journal of Sociology*, 1948, *53* (5), 395–396.

Lowe, Matt, "Types of contact: A field experiment on collaborative and adversarial caste integration," 2020.

Luce, Henry R, "The american century," *Life Magazine*, February 1941.

Ludwig, Jens, Greg J Duncan, Lisa A Gennetian, Lawrence F Katz, Ronald C Kessler, Jeffrey R Kling, and Lisa Sanbonmatsu, "Long-term neighborhood effects on low-income families: Evidence from Moving to Opportunity," *American economic review*, 2013, *103* (3), 226–31.

Lyle, David S, "Estimating and interpreting peer and role model effects from randomly assigned social groups at West Point," *The Review of Economics and Statistics*, 2007, *89* (2), 289–299.

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

Marmaros, David and Bruce Sacerdote, "Peer and social networks in job search," *European economic review*, 2002, *46* (4-5), 870–879.

_ **and** _ , "How do friendships form?," *The Quarterly Journal of Economics*, 2006, *121* (1), 79–119.

Matsa, David A and Amalia R Miller, "A female style in corporate leadership? Evidence from quotas," *American Economic Journal: Applied Economics*, 2013, *5* (3), 136–69.

Mayer, Adalbert and Steven L Puller, "The old boy (and girl) network: Social network formation on university campuses," *Journal of public economics*, 2008, *92* (1-2), 329–347.

McWilliams, Carey, "Equality or Fraternities? The Role of Secret Societies in Democratic Education," *Commentary*, May 1948.

Mehta, Nirav, Ralph Stinebrickner, and Todd Stinebrickner, "Time-Use and Academic Peer Effects in College," *Economic Inquiry*, 2019, *57* (1), 162–171.

Miller, William, "The recruitment of the American business elite," *The Quarterly Journal of Economics*, 1950, pp. 242–253.

Mills, C. Wright, *The power elite*, Oxford University Press, 1956.

Mountjoy, Jack and Brent Hickman, "The Returns to College (s): Estimating Value-Added and Match Effects in Higher Education," *University of Chicago, Becker Friedman Institute for Economics Working Paper*, 2020, (2020-08).

Mousa, Salma, "Building social cohesion between Christians and Muslims through soccer in post-ISIS Iraq," *Science*, 2020, *369* (6505), 866–870.

Murphy, Kevin M, Andrei Shleifer, and Robert W Vishny, "The allocation of talent: Implications for growth," *The quarterly journal of economics*, 1991, *106* (2), 503–530.

New York Times, "Initiate Young Roosevelt," *New York Times*, February 1907.

Nir, Sarah Maslin, "Are Final Clubs Too Exclusive for Harvard?," *The New York Times*, August 2nd 2016.

Olivetti, Claudia and M. Daniele Paserman, "In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850-1940," *American Economic Review*, August 2015, *105* (8), 2695–2724.

Paluck, Elizabeth Levy, Seth A Green, and Donald P Green, "The contact hypothesis re-evaluated," *Behavioural Public Policy*, 2019, *3* (2), 129–158.

Philippon, Thomas, "Financiers versus engineers: Should the financial sector be taxed or subsidized?," *American Economic Journal: Macroeconomics*, 2010, *2* (3), 158–82.

Popov, Sergey V and Dan Bernhardt, "Fraternities and labor-market outcomes," *American Economic Journal: Microeconomics*, 2012, *4* (1), 116–41.

Premaratne, Reshini, "What the Social Scene at Harvard Could Be," *Harvard Crimson*, April 2020.

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

Putnam, Robert D, "Bowling alone: America's declining social capital," in "Culture and politics," Springer, 2000, pp. 223–234.

Rao, Gautam, "Familiarity does not breed contempt: Generosity, discrimination, and diversity in Delhi schools," *American Economic Review*, 2019, *109* (3), 774–809.

Reeves, Aaron, Sam Friedman, Charles Rahal, and Magne Flemmen, "The decline and persistence of the old boy: Private schools and elite recruitment 1897 to 2016," *American Sociological Review*, 2017, *82* (6), 1139–1166.

Rimer, Sara, "Harvard Journal; All-Male Club Opens Its Door Warily," *New York Times*, October 1993.

Rivera, Lauren A, *Pedigree: How elite students get elite jobs*, Princeton University Press, 2016.

Roosevelt, Theodore, "Character and success," *Foundations magazine*, 1900.

Rothstein, Jesse and Albert H Yoon, "Affirmative action in law school admissions: What do racial preferences do?," Technical Report, National Bureau of Economic Research 2008.

—— **and Albert Yoon**, "Mismatch in law school," Technical Report, National Bureau of Economic Research 2008.

Saavedra, Martin and Tate Twinam, "A machine learning approach to improving occupational income scores," *Explorations in Economic History*, 2020, *75*, 101304.

Sacerdote, Bruce, "Peer effects with random assignment: Results for Dartmouth roommates," *The Quarterly Journal of Economics*, 2001, *116* (2), 681–704.

Sarada, Sarada, Michael J Andrews, and Nicolas L Ziebarth, "Changes in the demographics of American inventors, 1870–1940," *Explorations in Economic History*, 2019, *74*, 101275.

Sekhri, Sheetal, "Prestige Matters: Wage Premium and Value Addition in Elite Colleges," *American Economic Journal: Applied Economics*, July 2020, *12* (3), 207–25.

Shue, Kelly, "Executive networks and firm policies: Evidence from the random assignment of MBA peers," *Review of Financial Studies*, 2013, *26* (6), 1401–1442.

Shulman, James L and William G Bowen, *The game of life: College sports and educational values*, Vol. 62, Princeton University Press, 2011.

Social Security Administration, "Beyond the Top 1000 Names," 2021.

Society for American Archaeology, "Harry Evelyn Dorr Pollock, 1901–1982," *American Antiquity*, 1983, *48* (4), 782.

Stinebrickner, Ralph and Todd R Stinebrickner, "What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds," *Journal of public Economics*, 2006, *90* (8-9), 1435–1454.

Synott, Martha Graham, *The half-opened door: Discrimination and Admissions at Harvard, Yale, and Princeton, 1900-1970* number 80, Contributions in American History, 1979.

Tajfel, Henri, Michael G Billig, Robert P Bundy, and Claude Flament, "Social categorization and intergroup behaviour," *European journal of social psychology*, 1971, *1* (2), 149–178.

Temin, Peter, "The stability of the American business elite," *Industrial and Corporate Change*, 1999, *8* (2), 189–209.

Tilly, Charles et al., *Durable inequality*, Univ of California Press, 1998.

Useem, Michael and Jerome Karabel, "Pathways to top corporate management," *American Sociological Review*, 1986, pp. 184–200.

Weber, Max, *Economy and society: An outline of interpretive sociology*, Vol. 1, Univ of California Press, 1922.

Xu, Guo, "The costs of patronage: Evidence from the british empire," *American Economic Review*, 2018, *108* (11), 3170–98.

Zárate, Román Andrés, "Social and Cognitive Peer Effects: Experimental Evidence from Selective High Schools in Peru," 2019.

Zimmerman, David J, "Peer effects in academic outcomes: Evidence from a natural experiment," *Review of Economics and statistics*, 2003, *85* (1), 9–23.

Zimmerman, Seth D., "The returns to college admission for academically marginal students," *Journal of Labor Economics*, 2014, *32* (4), 711–754.

_ , "Elite colleges and upward mobility to top jobs and top incomes," *American Economic Review*, 2019, *109* (1), 1–47.

Zweigenhaft, Richard L and G William Domhoff, *Jews in the Protestant establishment*, VNR AG, 1982.

Figures

Figure 1: Labor market outcomes by academic performance and final club membership
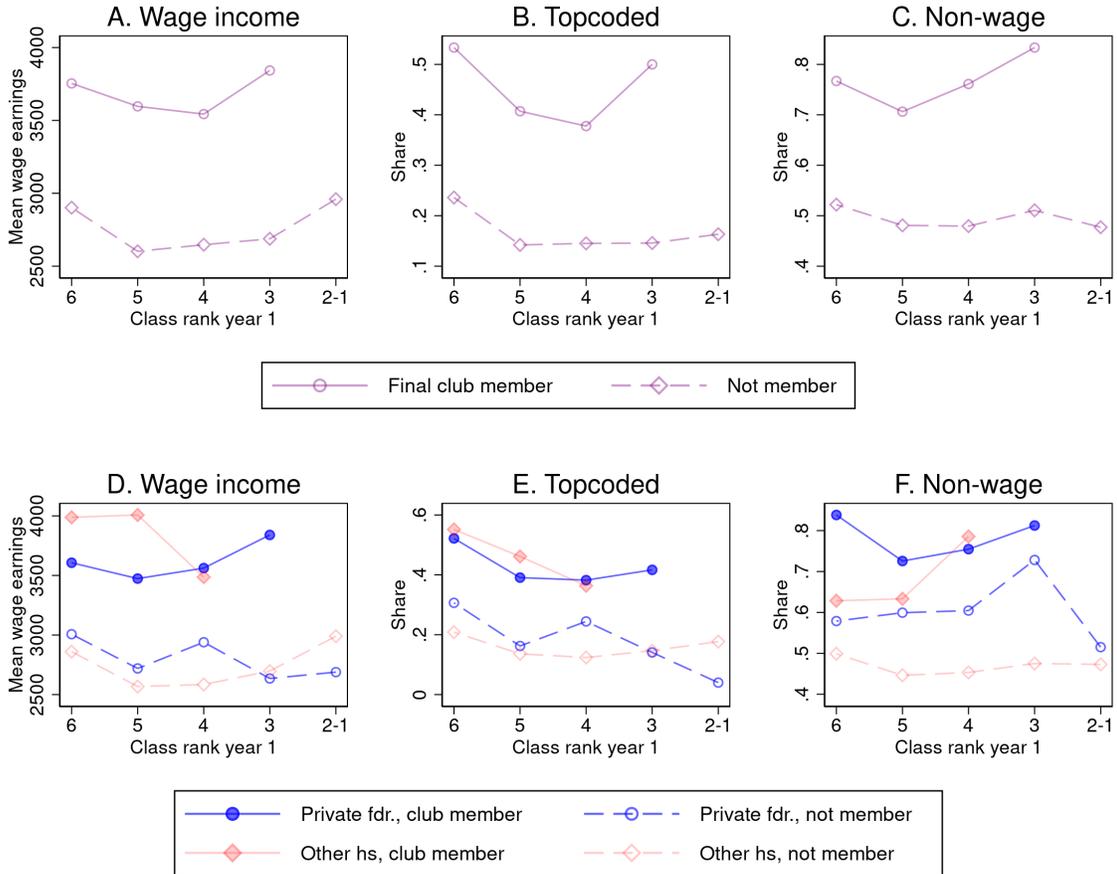
Labor market outcomes by academic performance and membership in a selective final club. Outcome types are listed in panel titles. Panels A–C depict earnings by freshman academic rank group and selective final club membership. Panels D–F present the same outcomes but also divide students by high school type. Because of small sample sizes in rank group 1, we collapse groups 1 and 2. We do not display groups with fewer than 20 students. Sample: students from cohorts 1920–1930 who matched to the 1940 census (and, for Panels A, B, D, and E, who reported wage income). Wage income is earnings in dollars. "Topcoded" is an indicator equal to one if a student reports the maximum wage income value of $5,000. "Non-wage" is an indicator equal to one if a student reports having at least $50 of non-wage income.

Figure 2: Adult career and social outcomes by academic performance and final club membership
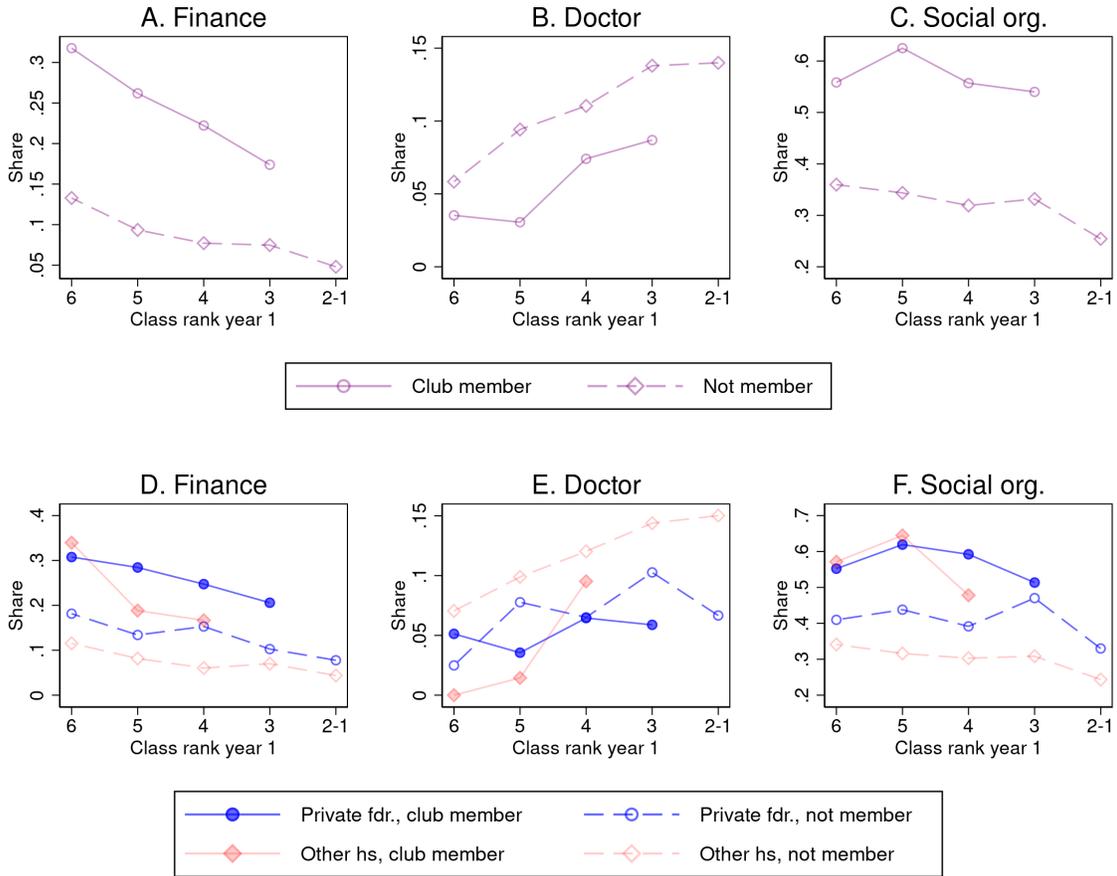
Career outcomes and adult social outcomes by academic performance and membership in a selective final club. Outcome types come from Class Report data and are listed in panel titles. Panels A–C depict adult outcomes for students by freshman academic rank group and selective final club membership. Panels D–F present the same outcomes but also divide students by high school type. We collapse groups 1 and 2 and do not display groups with fewer than 20 students. Sample: students from cohorts 1920–1934 who matched to a Class Report; for occupations we further restrict to students with non-missing occupation data.

Figure 3: Room price and randomization examples

(a) Rooms by price per student

(b) Neighborhoods by mean price per student with rooms in example randomization blocks outlined

Price Per Student
- 0 to 100
- 100 to 200
- 200 to 300
- 300 to 400
- 400 to 500

Mean Neighborhood Price
- 0 to 150
- 150 to 175
- 175 to 200
- 200 to 225
- 225 to 250
- 250 to 275

Randomization block
- $175 per student; double

Left panel: Room prices of Harvard first-year dorms in 1920. Harvard used these dorms for first-year students until 1931. Right panel: Mean peer neighborhood prices and example of a randomization block. Rooms from an example low-price randomization block are outlined in blue. These outlined rooms all have capacity for two students at the price of $175 per student and are found in both low- and high-priced neighborhoods. See section 5.1 for details.

47

Figure 4: Dorm room prices and peer attributes

A. Price per student by high school

B. Residential peer group size

C. Peer neighborhood rank

D. Peer private feeder share
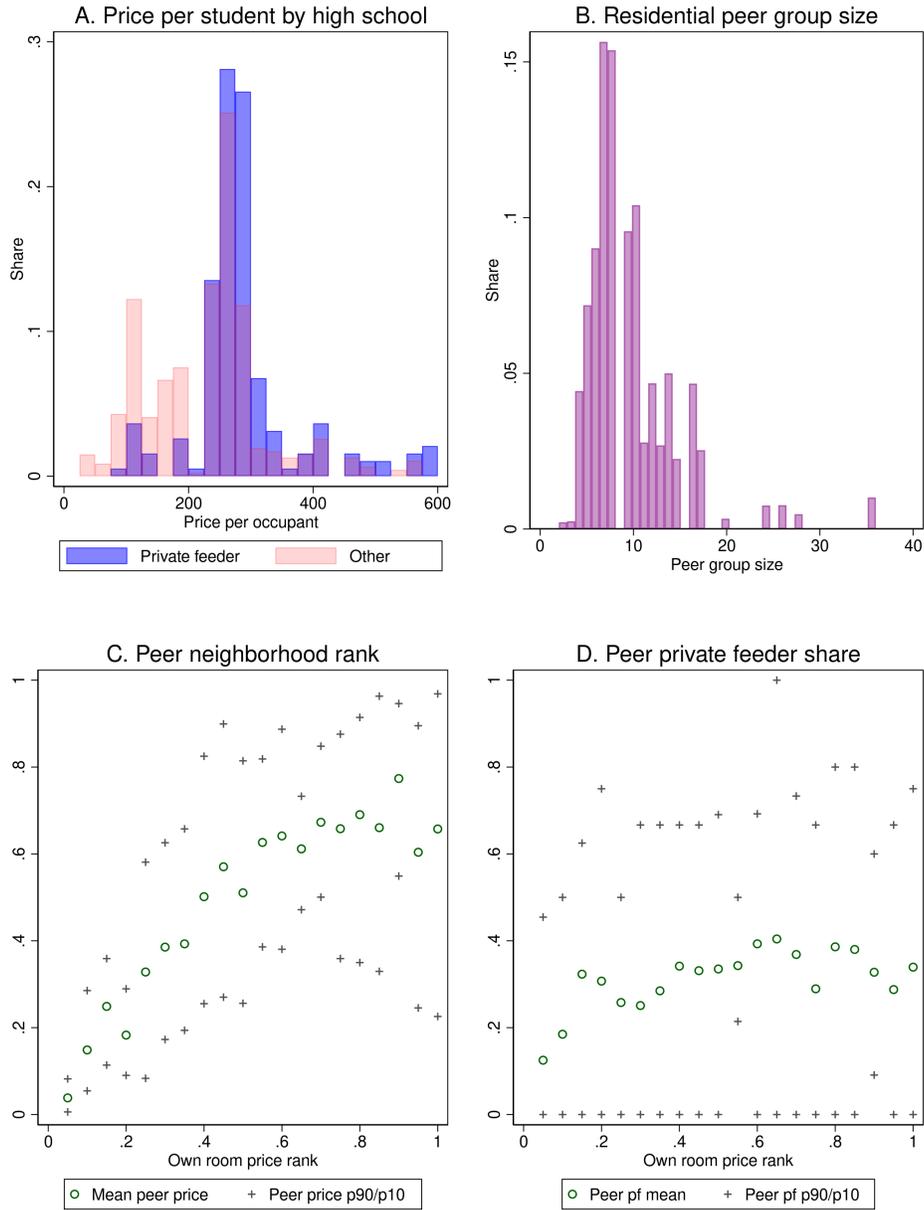
Panel A: Histogram of dorm price per student by student high school type for 1928 entering cohort. Panel B: histogram of the peer neighborhood size (summing over all rooms). Panel C: Mean, 10th, and 90th percentiles of peer neighborhood mean price rank by own room price rank. Panel D: Mean, 10th, and 90th percentiles of peer neighborhood private high school share by own room price rank, excluding own room. Each dot in panels C and D corresponds to the mean value of the listed statistic within a ventile of the own-room price distribution. The plus signs are the 10th and 90th percentiles within each ventile.

Figure 5: Randomization block size and within-block variation in peer attributes
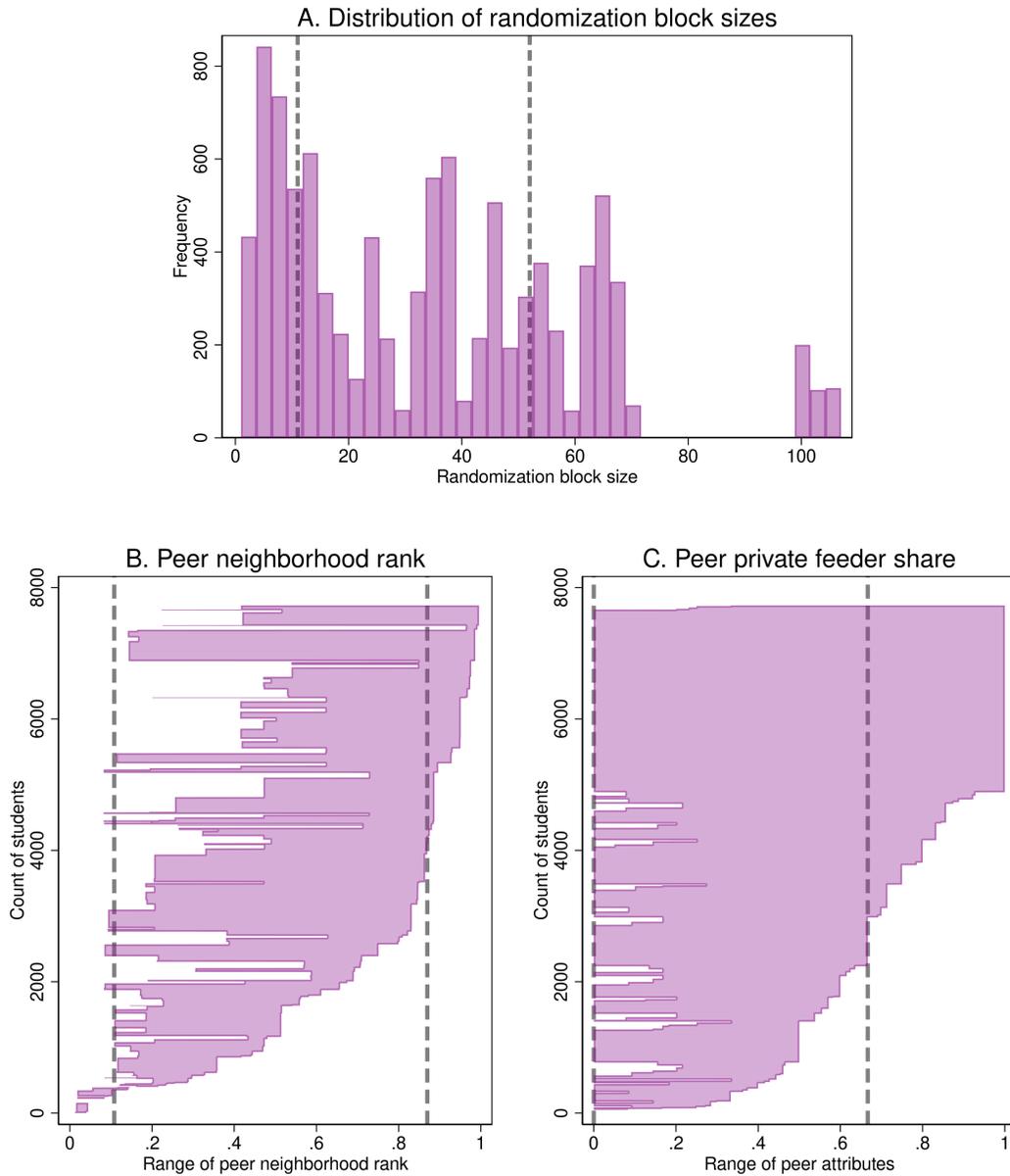
Panel A depicts a student-weighted histogram of randomization block size. Panels B and C depict the range of mean neighborhood room rank (B) and private school share (C) across neighborhoods within each randomization block. Blocks are sorted vertically by maximum room rank (B) or private school share (C) with vertical height equal to the cumulative number of students. Private feeder shares exclude individuals in the reference room. Blocks are defined by interactions between year, room size, and per-occupant room price. In panel A dashed lines denote 25th and 75th percentiles. In panels B and C dashed lines denote 10th and 90th percentiles. We exclude blocks with fewer than nine students from panels B and C.

49

Figure 6: Key outcomes by within-block decile of peer neighborhood price and high school type
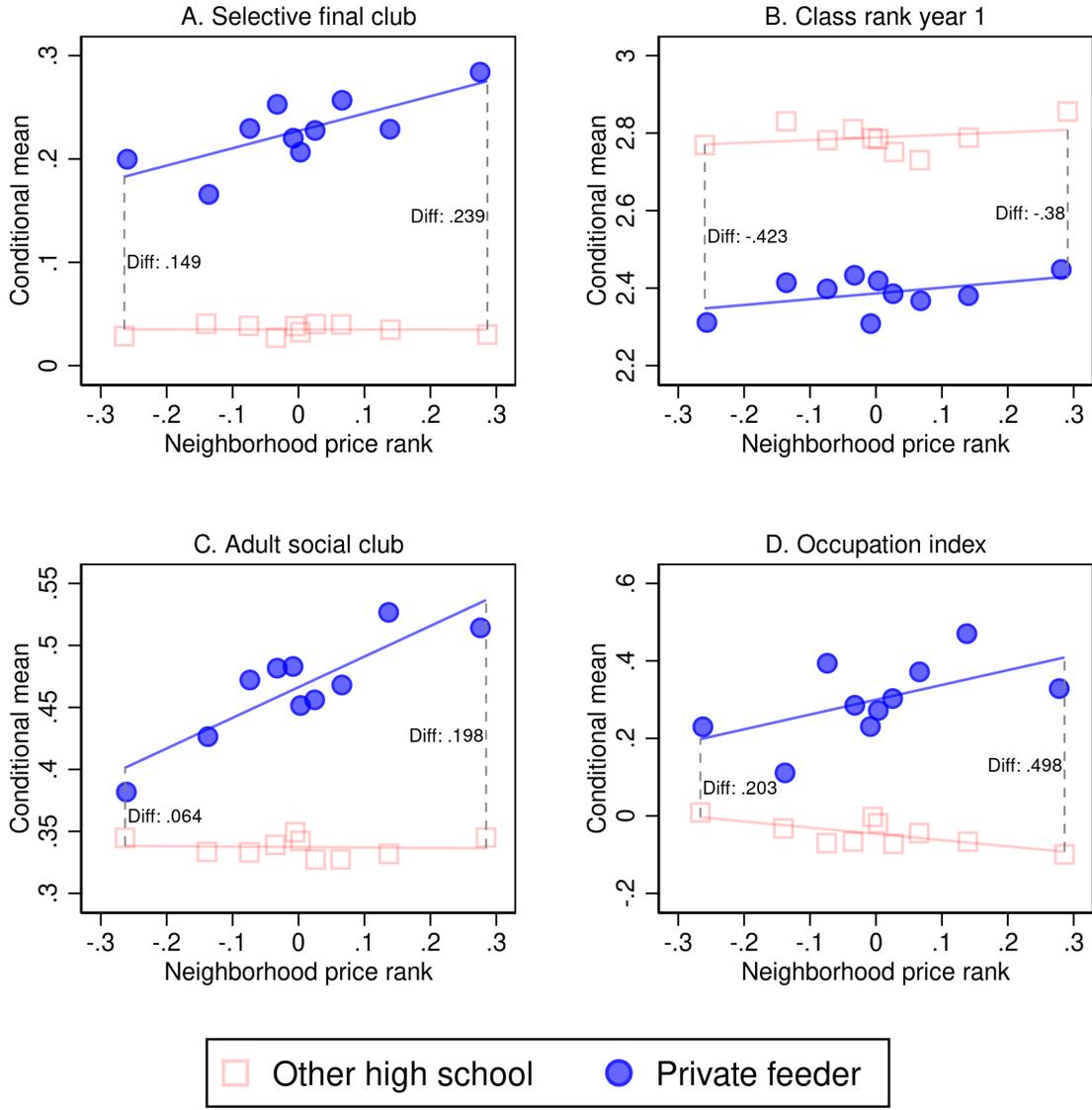
Binscatter plots showing the listed outcome by decile of peer neighborhood price rank, within randomization block. We obtain these graphs by regressing the vertical- and horizontal-axis variables on randomization block and large high school fixed effects separately by high school type. We then plot the conditional mean of the residuals for outcome variables against residual neighborhood price rank in each decile. We add sample means back to outcome variables to capture level differences by high school type. Outcome variables by panel as follows. Panel A: membership in selective final club. Panel B: first-year academic class rank. Reverse coded so that six is the highest rank and one is the lowest, with higher values corresponding to better academic performance. Panel C: membership in social clubs 25 years after graduation. Panel D: Occupation status index, 25 years after graduation. See section 3.4.3 for index definition. See section 5 for analysis details.

Figure 7: Grades and career outcomes over time
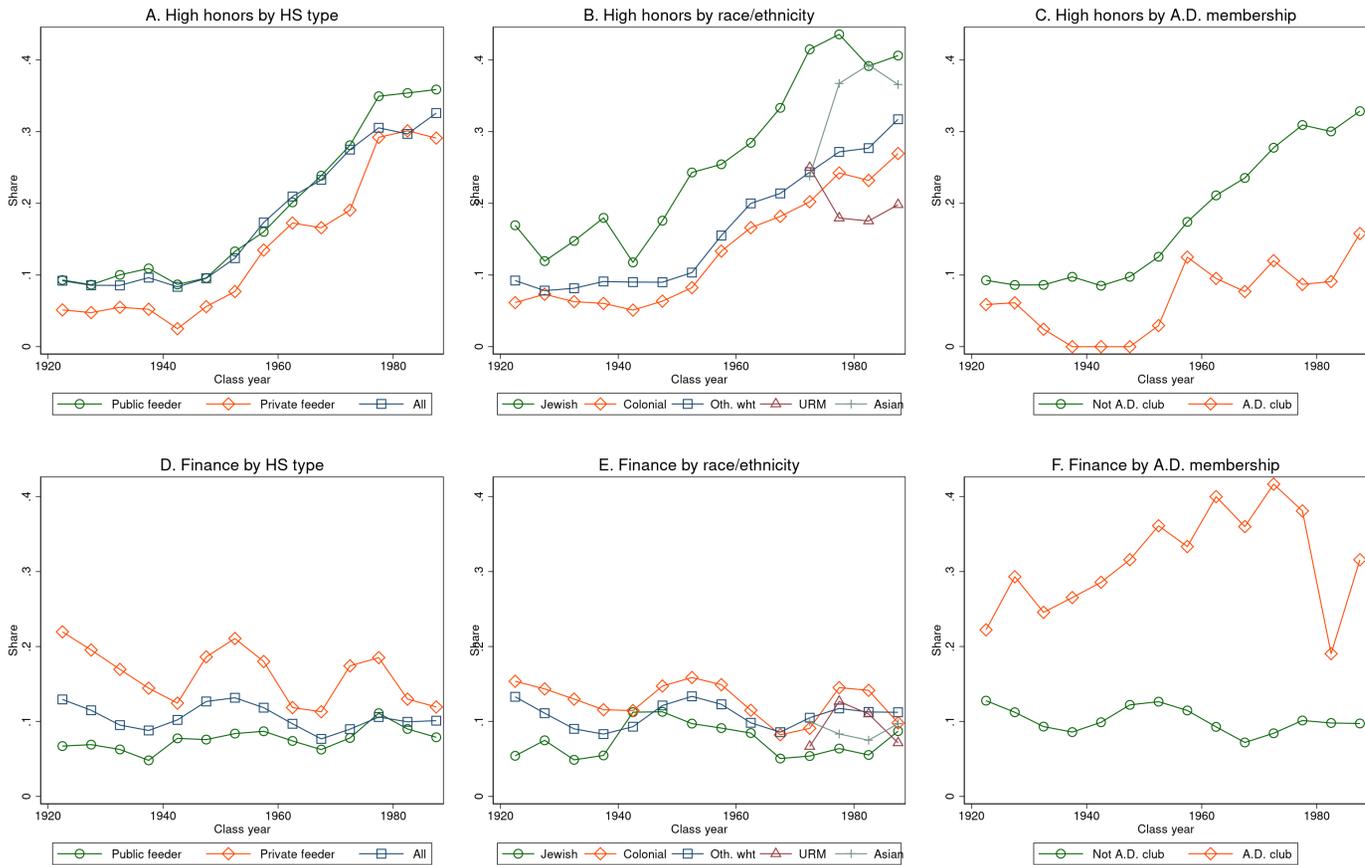
Share of students in different groups graduating with high honors (Panels A-C) or working in finance 25 years after graduation (D-F). Horizontal axis in all panels is graduating class year. "High honors" is defined as magna or summa cum laude. Academic and career outcomes are available for each class from 1923 through 1939 and then at five years intervals from 1940 through 1990. Points display means over all years within 2.5 years on either side of the centered value. For example, the 1982.5 datapoint is an average of 1980 and 1985 class years, and the 1987.5 datapoint is an average of 1985 and 1990 class years. Panels A and D split by high school type. Panels B and E split by by race/ethnicity. "White" is defined as non-URM, non-Asian, non-Jewish, non-Colonial students. "URM" category is Black and Hispanic students. Panels C and F split by membership in the A.D. club, one of the selective undergraduate final clubs at Harvard. See section 6.1 for details.

Figure 8: Child's outcomes by parent income for Ivy+ students in 1980s birth cohorts

A. Histogram of Ivy+ parent income

B. Kid mean income by parent rank

C. Kid top 1% by parent rank
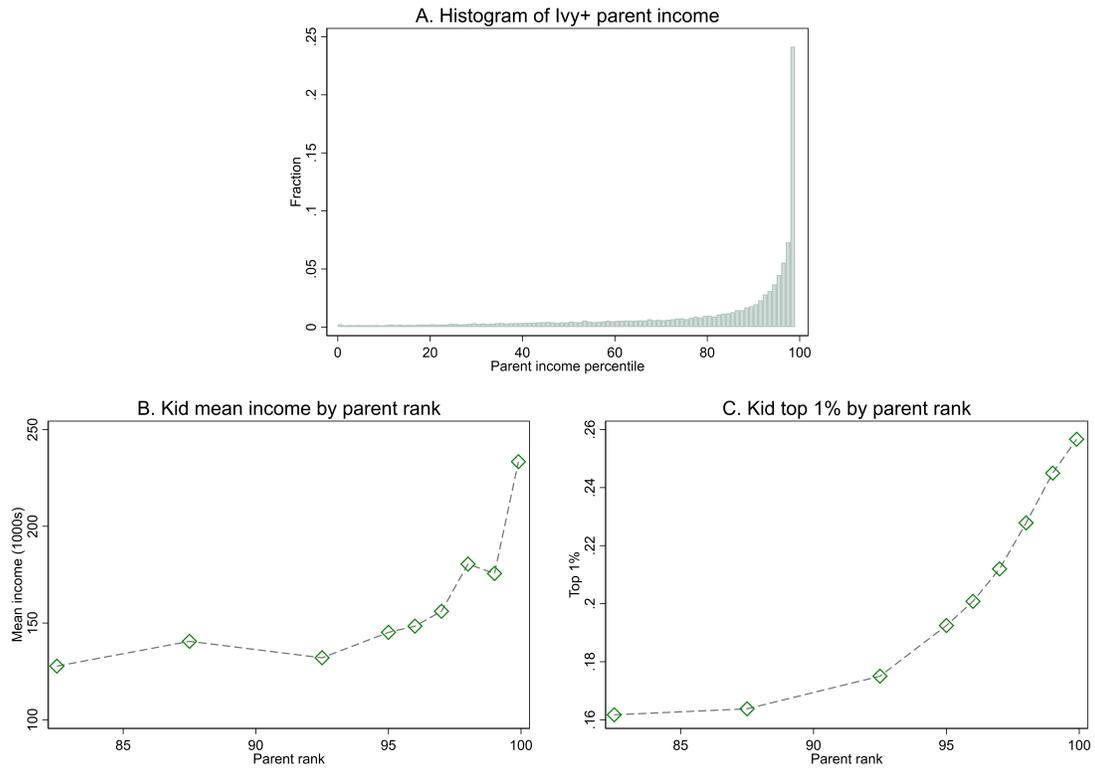
Childrens' outcomes by parent income percentile for students in 1980-82 birth cohorts who enroll in Ivy+ colleges. Data are from Chetty et al. (2020). Horizontal axis in all panels is parent income percentile. Ivy+ colleges are the eight Ivy League schools plus Duke, Stanford, MIT, and UChicago. Outcomes by panel are as follows. Panel A: histogram of parent income rank. Panel B: mean child income (1000s of USD). Panel C: share of children with top 1% incomes in age cohort. In panels B and C, points below the 95th percentile are means within centered five percentile bins. Points at the 95th percentile and above are one-percentile bins, with the top percentile split into a 99-99.9 percentile point and a top 0.1 percentile point. See section 6.4 for details and Online Appendix Figure A.7 and Table A.13 for additional college types.

Tables

Table 1: Data availability by source

Data type	Share non-missing	Universe	N
A. Match rates within available cohorts			
Freshman Red Book	1.000	Cohorts 1919-35	14383
High school	0.984	Cohorts 1919-35	14383
Campus address	0.950	Cohorts 1919-25; 1927-35	13629
25 year Class Report	0.911	Cohorts 1920-35	13820
Class Report occupation	0.793	Cohorts 1920-35	13820
1940 Census	0.647	Cohorts 1920-30	8851
Pre-Harvard census	0.619	Cohorts 1920-33	11800
B. Match rates conditional on Class Report availability			
Class Report occupation	0.871	Cohorts 1920-35 in Class Report	12592
1940 Census	0.685	Cohorts 1920-30 in Class Report	8350
Pre-Harvard Census	0.662	Cohorts 1920-33 in Class Report	11039

Match rates by data source and/or data type. The second column presents the share of students in the listed sample universe who have the data described in the first column. The third column describes the relevant universe. Cohorts for whom a particular data source was unavailable are excluded from the universe. The fourth column presents the number of students in the relevant universe. The universe in Panel A is students in Freshman Red Books, for listed cohorts. The universe in Panel B is students in Freshman Red Books who are matched to Class Report records, again for listed cohorts.

Table 2: Family background and college outcomes for Harvard students

	All	Private feeder	Non-private	Hasty Pudding	Sel. Fin. Club	Randomized
A. Demographics						
Have high school data	0.984	1.000	0.979	0.979	0.981	0.983
Any private high school	0.463	1.000	0.289	0.909	0.949	0.589
Private feeder	0.240	1.000	0.000	0.616	0.733	0.313
Any public feeder	0.226	0.000	0.297	0.019	0.012	0.080
From MA	0.509	0.453	0.527	0.473	0.439	0.387
From NY	0.172	0.255	0.145	0.249	0.293	0.218
Have Harvard father	0.070	0.143	0.046	0.179	0.216	0.087
Have Harvard brother	0.204	0.290	0.176	0.362	0.402	0.229
Jewish name	0.066	0.013	0.083	0.006	0.002	0.038
Catholic name	0.055	0.028	0.063	0.030	0.014	0.037
Colonial name	0.293	0.430	0.250	0.468	0.510	0.333
B. Census childhood household demographics						
Have Census pre-Harvard	0.619	0.708	0.592	0.728	0.743	0.651
S or E Eur. immg. gen. 1-2	0.098	0.011	0.126	0.006	0.009	0.051
Have father's occupation	0.690	0.650	0.702	0.663	0.668	0.680
Father's occupation: Doctor	0.103	0.128	0.095	0.084	0.063	0.121
Father's Occupation: Lawyer	0.115	0.167	0.099	0.189	0.176	0.141
C. First-year campus location						
Have address data	0.950	0.995	0.936	0.998	0.999	1.000
Live on campus	0.800	0.964	0.749	0.983	0.987	1.000
Have room attributes	0.712	0.910	0.650	0.933	0.919	1.000
Room price per occupant	209.2	234.5	198.1	231.3	232.7	213.0
Peer neighborhood price	213.9	228.0	207.5	227.6	233.9	213.9
25th pctile neighborhood rank	0.256	0.359	0.215	0.386	0.444	0.256
75th pctile neighborhood rank	0.711	0.768	0.691	0.768	0.805	0.711
D. Academic class rank groups						
Rank group 1	0.018	0.008	0.021	0.004	0.000	0.018
Rank group 2	0.069	0.039	0.078	0.027	0.023	0.064
Rank group 3	0.155	0.104	0.170	0.088	0.069	0.151
Rank group 4	0.242	0.218	0.250	0.204	0.197	0.241
Rank group 5	0.375	0.436	0.357	0.448	0.451	0.376
Rank group 6	0.141	0.194	0.124	0.229	0.259	0.149
Not ranked year 1	0.208	0.216	0.205	0.138	0.130	0.207
E. First-year activities						
Have any activity	0.526	0.709	0.468	0.868	0.918	0.605
N activities	1.046	1.679	0.847	2.473	2.791	1.267
Activity leadership position	0.064	0.135	0.042	0.224	0.269	0.084
Sports	0.367	0.552	0.309	0.728	0.791	0.430
Social	0.083	0.198	0.047	0.358	0.481	0.108
Music	0.133	0.163	0.123	0.208	0.187	0.156
First-year activity index	0.000	0.539	-0.170	1.122	1.501	0.159
F. Upper-year social clubs						
Hasty Pudding	0.153	0.394	0.077	1.000	0.950	0.209
Sel. Fin. Club	0.070	0.215	0.025	0.435	1.000	0.095
Any final club	0.136	0.370	0.062	0.744	1.000	0.185
Upper-year club index	0.000	0.727	-0.229	2.138	2.778	0.152
N	14383	3449	10934	2053	941	9343

Descriptive statistics across sample definitions. Columns are samples, rows are variables, cells display variable means unless otherwise specified. Columns as follows. All: full Red Book sample universe. Private feeder: students who attended private feeder high schools. All non-private: all students not in the private feeder category. Hasty Pudding and Sel. Fin. Club: Hasty Pudding club and selective final club members, respectively. Randomized: students living in on-campus housing with price and occupancy records in a randomization block with more than one room. See Section 5 for more on room randomization. Panel A: Demographic variables from Red Books. Panel B: Demographics from pre-college Census records. We exclude the 1919, 1934, and 1935 cohorts from this sample due to data availability. Panel C: Campus address data from Red Books and merged room characteristics; 1926 entering class omitted. D: Academic first-year class rank groups; one is highest achievement, six is lowest. Panel E: College activities from Red Books. Panel F: Membership in upper year social clubs. 1935 cohort excluded due to data availability. See text for details.

Table 3: Adult outcomes for Harvard students

	All	Private feeder	Non-private	Hasty Pudding	Sel. Fin. Club	Randomized
A. Adult associations						
Any social club	0.343	0.450	0.308	0.549	0.590	0.381
Country club	0.242	0.363	0.203	0.459	0.490	0.287
Gentleman's club	0.112	0.210	0.081	0.293	0.367	0.138
Fraternal order	0.103	0.078	0.111	0.062	0.054	0.091
Any honor/prof group	0.369	0.311	0.387	0.323	0.320	0.362
Adult association index	0.000	0.413	-0.133	0.694	0.894	0.115
B. Occupations						
Have occupation	0.871	0.858	0.875	0.896	0.883	0.864
Finance	0.101	0.175	0.077	0.212	0.259	0.120
Accounting	0.108	0.119	0.105	0.109	0.098	0.106
Medicine	0.091	0.060	0.101	0.053	0.049	0.081
Law	0.123	0.099	0.130	0.113	0.102	0.124
Higher ed.	0.070	0.048	0.077	0.035	0.036	0.072
Teach	0.080	0.071	0.083	0.067	0.065	0.074
Government	0.034	0.029	0.035	0.035	0.038	0.034
Art/pub	0.074	0.075	0.073	0.070	0.067	0.076
Senior management	0.213	0.236	0.205	0.257	0.281	0.230
Low management	0.119	0.105	0.123	0.115	0.107	0.115
Retail	0.137	0.126	0.140	0.126	0.124	0.131
Occupation index	0.000	0.276	-0.087	0.388	0.543	0.065
C. Adult census						
In school	0.030	0.025	0.031	0.021	0.021	0.030
In labor force	0.954	0.951	0.954	0.953	0.952	0.948
Wage income	2747	2961	2681	3346	3549	2857
Has wage income	0.726	0.712	0.731	0.720	0.715	0.725
Non-wage inc. 50+	0.503	0.646	0.457	0.669	0.732	0.557
Wage inc. 5000+	0.174	0.247	0.152	0.357	0.418	0.210
N	14383	3449	10934	2053	941	9343

Descriptive statistics across sample definitions. Columns are samples, rows are variables, cells display variable means. Columns as follows. All: Full Red Book sample universe. Private feeder: students who attended private feeder high schools. All non-private: All students not in the private feeder category. Hasty Pudding and Sel. Fin. Club: Hasty Pudding club and selective final club members, respectively. Randomized: Students living in on-campus housing with price and occupancy records in a randomization block with more than one room. See Section 5 for more on room randomization. Panel A: Adult social club and professional association data from 25th Anniversary Class Reports. Cohort 1919 excluded. Panel B: Occupation data from 25th Anniversary Class Reports. Cohort 1919 excluded. Panel C: Adult outcomes from the 1940 census. Cohorts 1919 and 1931–35 excluded due to data availability. See main text for details.

Table 4: Labor market outcomes by academic performance in first year and social success

	Baseline	Final clubs	Census controls	High school FEs	Family FEs	Applicant pool
A. Has earnings						
Private feeder	0.012	0.017	-0.006			-0.048
	(0.014)	(0.015)	(0.024)			(0.032)
Class rank	0.000	0.000	0.006	0.005	0.008	0.005
	(0.005)	(0.005)	(0.008)	(0.012)	(0.022)	(0.015)
Selective final club		-0.021	-0.012	-0.058	0.005	-0.071
		(0.024)	(0.038)	(0.032)	(0.100)	(0.031)
Have Harvard father				0.008		
				(0.031)		
Sample mean	0.480	0.480	0.524	0.501	0.504	0.504
Observations	7113	7113	3388	1650	1026	1123
B. Earnings (dollars)						
Private feeder	297	145	-34			89
	(56)	(58)	(89)			(126)
Class rank	39	50	67	-9	-20	62
	(20)	(19)	(30)	(45)	(107)	(61)
Selective final club		777	711	657	1250	449
		(94)	(150)	(129)	(632)	(120)
Have Harvard father				-15		
				(121)		
Sample mean	2771	2771	2760	2978	2970	3362
Observations	3417	3417	1737	827	291	566
C. Topcoded earnings flag						
Private feeder	0.093	0.048	0.003			0.033
	(0.016)	(0.016)	(0.025)			(0.040)
Class rank	-0.007	-0.004	0.004	-0.027	-0.006	0.012
	(0.006)	(0.005)	(0.008)	(0.013)	(0.031)	(0.019)
Selective final club		0.230	0.226	0.184	0.371	0.130
		(0.031)	(0.048)	(0.041)	(0.147)	(0.039)
Have Harvard father				0.007		
				(0.036)		
Sample mean	0.181	0.181	0.184	0.253	0.237	0.359
Observations	3417	3417	1737	827	291	566
D. Non-wage earnings flag						
Private feeder	0.178	0.142	0.067			0.118
	(0.017)	(0.018)	(0.027)			(0.037)
Class rank	0.005	0.007	0.002	0.024	0.001	0.011
	(0.006)	(0.006)	(0.010)	(0.013)	(0.032)	(0.016)
Selective final club		0.174	0.135	0.073	-0.175	0.128
		(0.027)	(0.043)	(0.035)	(0.139)	(0.035)
Have Harvard father				0.042		
				(0.034)		
Sample mean	0.511	0.511	0.500	0.645	0.616	0.675
Observations	4611	4611	2343	1114	513	754

Estimates of equation 1. Panels correspond to outcome variables, columns correspond to specifications, and rows correspond to covariates. With each panel, each column reports estimates from a single specification. All specifications restrict to students in the 1920–30 entering cohorts who are matched to records of first-year academic rank group. Panels B and C additionally restrict to students who report wage earnings in the 1940 Census, and Panel D to students matched to the 1940 Census (regardless of whether they report wage earnings). Outcome variables are as follows. Panel A: "Has Earnings" is an indicator equal to one if a student is matched to the 1940 Census and reports wage earnings. Panels B and C: "Earnings" is 1940 Census wage earnings in dollars, and "Topcoded" is an indicator for whether a student reports the maximum possible earnings value of $5,000. Panel D: "Non-wage" is an indicator for whether a student reports at least $50 of non-wage earnings. All specifications include cohort fixed effects (not reported). Moving across columns, specifications differ as follows. Baseline: private feeder and class rank controls. Final clubs: adds indicator for selective final club membership. Census controls: adds controls for father's occupation, father's and mother's state or non-US country of birth, family size, parental presence, home ownership, presence of domestic employees, and farm status; sample is restricted to students for whom these records are non-missing. High school FEs: restricts the sample to students who attended private feeder high schools and includes fixed effects for each high school. Family FEs: restricts the sample to students from families where multiple brothers attended Harvard during our sample period, and includes family fixed effects. Applicant pool: includes only students who are members of the Hasty Pudding club. See text for details. The sign on class rank groups is reversed in all specifications so that coefficients reflect the effect of a one rank group improvement in academic performance. Robust standard errors in parentheses. Standard errors for the Family FEs are clustered at the family level.

Table 5: Test of room randomization

	Year FEs		Price * Year FEs		Block FEs	
A. Balance test						
Room price per occupant	206.6	(3.3)	–	–	–	–
Any private high school	0.394	(0.024)	-0.044	(0.037)	-0.064	(0.032)
Private feeder high school	0.296	(0.021)	0.022	(0.035)	–	–
Have HS code	0.297	(0.022)	-0.025	(0.034)	-0.030	(0.031)
From MA	-0.056	(0.021)	0.007	(0.036)	0.028	(0.035)
From NY	0.083	(0.017)	-0.028	(0.029)	-0.037	(0.030)
Have Harvard father	0.071	(0.011)	0.005	(0.019)	0.016	(0.022)
Have Harvard brother	0.110	(0.018)	0.008	(0.028)	-0.005	(0.031)
Jewish name	-0.014	(0.008)	0.005	(0.013)	0.013	(0.015)
Catholic name	-0.031	(0.008)	-0.008	(0.012)	-0.006	(0.013)
Colonial name	0.104	(0.019)	0.004	(0.031)	-0.014	(0.032)
S or E Eur. immg. gen. 1-2	-0.053	(0.014)	-0.004	(0.018)	-0.000	(0.020)
Have father's occupation	-0.074	(0.027)	-0.009	(0.047)	-0.035	(0.051)
Father's occupation: Doctor	0.034	(0.024)	0.032	(0.042)	0.030	(0.045)
Father's Occupation: Lawyer	0.109	(0.024)	-0.016	(0.045)	-0.023	(0.051)
Joint balance test [p-value]	[0.000]		[0.929]		[0.601]	
B. Links to other data sources						
Class report link	0.002	(0.010)	-0.013	(0.016)	-0.012	(0.017)
Have occupation	-0.002	(0.014)	-0.020	(0.024)	-0.008	(0.027)
1940 Census	0.029	(0.014)	0.026	(0.022)	0.034	(0.027)
C. Peer and neighborhood attributes						
Private HS peer share	0.370	(0.024)	0.306	(0.031)	0.282	(0.033)
Feeder HS peer share	0.272	(0.021)	0.221	(0.028)	0.202	(0.032)
Legacy share	0.062	(0.012)	0.063	(0.015)	0.063	(0.018)
E Eur. peer share	-0.050	(0.014)	-0.050	(0.015)	-0.049	(0.016)
Share rooms > median price	1.142	(0.018)	0.800	(0.019)	0.786	(0.018)
Share rooms > 90th pctile price	0.309	(0.016)	0.436	(0.016)	0.472	(0.018)
D. Endogenous peer outcomes						
Peer mean acad. rank	0.390	(0.058)	0.264	(0.072)	0.157	(0.076)
Peer activity index	0.306	(0.049)	0.303	(0.065)	0.297	(0.072)
Peer mean activity count	0.302	(0.066)	0.293	(0.085)	0.336	(0.094)
Peer mean social act.	0.055	(0.014)	0.061	(0.018)	0.054	(0.021)
Peer more sel. final	0.136	(0.014)	0.128	(0.019)	0.110	(0.021)
Peer occ. index	0.378	(0.050)	0.499	(0.068)	0.438	(0.075)
Peer mean finance	0.107	(0.015)	0.143	(0.021)	0.128	(0.023)
Peer mean wage inc.	362.4	(95.3)	419.5	(132.9)	359.0	(143.7)
N	9343		9343		9343	

Coefficients on peer neighborhood price rank from regressions of the form given in equation 2. The sample consists of students in non-singleton randomization blocks. Rows are dependent variables. Standard errors clustered at peer neighborhood level in parentheses to the right of point estimates. Each pair of columns corresponds to a different control set. "Year FEs": controls are year dummies only. "Price X Year FEs": controls for all interactions between per-occupant price and year dummies. "Block FEs": controls for all interactions between per-occupant price, year dummies, and room occupancy, as well as indicators for each private feeder high school and other large feeder schools. Panel A: Room attributes and predetermined student characteristics. Harvard family data is from Class Reports. Immigrant and father occupation variables are from Census data; non-matched individuals are excluded from those rows. "Joint test" is test of the null hypothesis that all coefficients in Panel A are zero. Panel B: Matches to later-life data sources. Panel C: Predetermined attributes of students and rooms in the assigned neighborhood. Peer attribute statistics omit one's own room. "Share of rooms>median price" is the share of rooms in the assigned neighborhood with a price above the median; "Share of rooms >90th ptile price" is the share of rooms with a price above the 90th percentile. Panel D: Endogenous peer outcomes, realized freshman year or later. See text for detailed definitions of variables.

Table 6: Peer neighborhood effects on short-run outcomes

	All	Private	Non-private	Test
A. First-year activities				
Have any activity	0.070	0.080	0.064	0.816
	(0.034)	(0.057)	(0.043)	
N activities	0.283	0.648	0.108	0.030
	(0.110)	(0.222)	(0.124)	
Activity leadership position	0.043	0.098	0.012	0.071
	(0.020)	(0.044)	(0.021)	
Social	0.049	0.183	-0.028	0.001
	(0.024)	(0.057)	(0.023)	
Sports	0.029	0.085	-0.007	0.234
	(0.035)	(0.064)	(0.043)	
Music	0.046	0.095	0.029	0.283
	(0.027)	(0.052)	(0.033)	
Other activities	0.021	-0.027	0.059	0.179
	(0.031)	(0.058)	(0.035)	
First-year activity index	0.208	0.494	0.025	0.014
	(0.082)	(0.171)	(0.087)	
N	9343	2824	6370	
B. Upper-year social clubs				
Selective final club	0.065	0.167	0.001	0.004
	(0.021)	(0.056)	(0.017)	
Less selective final club	-0.028	-0.075	-0.007	0.206
	(0.020)	(0.050)	(0.021)	
Hasty Pudding Inst. 1770	0.019	0.105	-0.018	0.070
	(0.029)	(0.062)	(0.030)	
Upper-year club index	0.122	0.357	-0.021	0.047
	(0.078)	(0.176)	(0.078)	
N	8589	2603	5847	
C. First-year academic rank				
Rank group 1	-0.000	-0.012	0.003	0.503
	(0.011)	(0.015)	(0.015)	
Rank group 2	0.018	0.017	0.014	0.932
	(0.016)	(0.025)	(0.021)	
Rank group 3	0.022	0.006	0.029	0.634
	(0.024)	(0.039)	(0.030)	
Rank group 4	-0.008	0.090	-0.055	0.011
	(0.030)	(0.046)	(0.036)	
Rank group 5	0.022	-0.067	0.073	0.049
	(0.032)	(0.061)	(0.039)	
Rank group 6	0.021	0.013	0.006	0.887
	(0.023)	(0.046)	(0.027)	
Rank listed year 1	0.073	0.047	0.070	0.708
	(0.029)	(0.051)	(0.035)	
Class rank year 1	0.024	0.101	0.020	0.706
	(0.104)	(0.180)	(0.129)	
N	7035	2085	4802	

Coefficients on peer neighborhood price rank from regressions of the form given in equation 2. The first three columns denote samples. Rows are outcome variables. All specifications include randomization block and dummies for large feeder high schools; see section 5.1 for details. "Test" column reports the p-value from a test of the null that the coefficients reported in the private and non-private columns are equal. Panels A and B report first-year activity outcomes and upper-year club outcomes. "First-year activity index" and "upper-year club index" are the standardized indices of the association between activities and private high school background described in section 3.4.3. Other variables are indicators for a given activity type unless stated otherwise. Panel C describes academic outcomes in the first year. Class rank is a continuous variable from one through six, with one the best and six the worst. The other outcomes are dummies for having grades in the listed rank group and being listed at all. Rank group indicators treat unranked students as zero values. Sample sizes vary across panels due to data availability. Private and non-private column sample sizes add up to slightly less than the "all" column sample size because there is no variation in peer neighborhood assignment for private-feeder or non-private-feeder students within some randomization blocks; students in these blocks are omitted from split sample regressions. Standard errors clustered at peer neighborhood level.

Table 7: Peer neighborhood effects on long-run outcomes

	All	Private	Non-private	Test
A. Adult associations				
Any social club	0.066	0.241	-0.003	0.003
	(0.036)	(0.067)	(0.046)	
Country club	0.073	0.219	0.006	0.006
	(0.032)	(0.066)	(0.040)	
Gentleman's club	0.019	0.053	0.003	0.423
	(0.026)	(0.056)	(0.028)	
Fraternal order	0.002	0.020	-0.002	0.644
	(0.022)	(0.037)	(0.029)	
Any honor/prof group	-0.014	0.039	-0.024	0.394
	(0.032)	(0.060)	(0.041)	
Prof. Association	-0.021	0.032	-0.037	0.350
	(0.032)	(0.060)	(0.040)	
Honor society	-0.017	0.020	-0.019	0.361
	(0.019)	(0.035)	(0.024)	
Adult association index	0.139	0.328	0.040	0.135
	(0.078)	(0.171)	(0.087)	
N	8178	2476	5551	
B. Occupation choice				
Finance	0.016	0.144	-0.052	0.002
	(0.025)	(0.058)	(0.028)	
Medicine	0.001	-0.018	0.017	0.425
	(0.023)	(0.034)	(0.030)	
Higher ed.	-0.019	-0.026	-0.023	0.952
	(0.021)	(0.034)	(0.028)	
Law	-0.032	-0.009	-0.036	0.640
	(0.027)	(0.044)	(0.035)	
Business	0.067	0.071	0.046	0.786
	(0.040)	(0.077)	(0.049)	
Teach	-0.014	-0.017	-0.005	0.795
	(0.021)	(0.038)	(0.027)	
Government	0.005	0.012	0.009	0.940
	(0.015)	(0.031)	(0.018)	
Art/pub	-0.005	-0.017	0.008	0.565
	(0.021)	(0.035)	(0.027)	
Occupation index	0.056	0.450	-0.164	0.002
	(0.081)	(0.181)	(0.093)	
N	7039	2094	4797	
C. Adult income				
Wage income	-26.1	-440.5	191.0	0.138
	(191.4)	(367.8)	(236.3)	
Wage inc. 5000+	0.020	-0.029	0.073	0.369
	(0.050)	(0.099)	(0.061)	
Non-wage inc. 50+	0.053	0.081	0.044	0.753
	(0.052)	(0.099)	(0.064)	
Class Report wage index	33.6	128.8	-17.2	0.011
	(26.7)	(48.1)	(33.8)	
N	2396	697	1611	

Coefficients on peer neighborhood price rank from estimates of equation 2. Rows: outcome variables. All specifications include randomization block and high school FEs. The first three columns are samples."Test" column: p-value from a test that the coefficients in the private and non-private columns are equal. Panel A: adult social clubs. Sample: students matched to Class reports. Panel B: Occupation outcomes. Sample: matches to Class Reports w/ non-missing occupation. Table A.2 disaggregates "business" occupation category. "Adult association index" and "Occupation index:" standardized indices of association between adult outcomes and private high school background. See section 3.4.3. Other variables: indicators for listed outcomes. Panel C: 1940 Census income. Sample: Census-matched students in 1920–30 cohorts with non-missing outcome variable. "Class report wage index:" earnings-weighted index of Class Report variables. See section 5.6. Subsample sizes add up to less than the "All" column sample size because there is no variation in neighborhood assignment within some high school type by randomization blocks. Standard errors clustered at peer neighborhood level.

	1923-1944	1945-1966	1967-1990	1991-2015
A. All				
Male	1.000	1.000	0.720	0.539
Jewish (surname only)	0.139	0.169	0.194	0.134
Colonial	0.273	0.185	0.098	0.058
Oth. wht	0.569	0.621	0.569	0.513
URM	0.015	0.018	0.084	0.102
Asian	0.004	0.007	0.054	0.193
Private feeder	0.238	0.170	0.084	0.042
Public feeder	0.231	0.191	0.155	0.133
Share	1.000	1.000	1.000	1.000
N w/ race codes	19095	20897	17944	16627
B. Private feeder HS				
Male	1.000	1.000	0.830	0.531
Jewish (surname only)	0.046	0.077	0.112	0.118
Colonial	0.404	0.314	0.193	0.099
Oth. wht	0.541	0.591	0.604	0.534
URM	0.007	0.014	0.048	0.069
Asian	0.002	0.004	0.044	0.181
Share	0.238	0.170	0.084	0.042
N w/ race codes	4484	3494	1470	687
C. Public feeder HS				
Male	1.000	1.000	0.681	0.507
Jewish (surname only)	0.270	0.281	0.289	0.205
Colonial	0.178	0.119	0.072	0.056
Oth. wht	0.520	0.569	0.503	0.460
URM	0.027	0.023	0.063	0.063
Asian	0.004	0.008	0.072	0.217
Share	0.231	0.191	0.155	0.133
N w/ race codes	4352	3919	2701	2171
D. In A.D. club				
Male	1.000	1.000	0.985	0.970
Jewish (surname only)	0.012	0.015	0.060	0.133
Colonial	0.490	0.457	0.214	0.109
Oth. wht	0.486	0.524	0.697	0.667
URM	0.008	0.004	0.020	0.042
Asian	0.004	0.000	0.010	0.048
Private feeder	0.813	0.646	0.355	0.090
Public feeder	0.012	0.015	0.051	0.084
Share	0.013	0.013	0.011	0.010
N w/ race codes	255	273	201	169

Demographic characteristics of Harvard students over time. Columns are time periods, split by graduating class. Panels are samples. "All" is all students. "Private feeder" and "Public feeder" restrict the sample by high school type. "In A.D. club" is members of the selective A.D. final club. Within each panel the "Male" through "Public feeder" rows are shares of the listed variable with the sample defined by the panel and time period defined by the column. Data are based on name indices and visual codes; see section 6.1 for details. Gender codes are based on first names for graduating classes of 1975 and later; gender is imputed as male prior to coeducation. Jewish and Colonial categories are based on surnames. Note that this differs from the Jewish student definition in the historical analysis, which uses first and last names. URM (under-represented minority) groups Black and Hispanic students. "Other white" are students who are not Black, Hispanic, or Asian, and also not Colonial or Jewish. "Share" is the share of students listed in panel in the full Harvard sample in listed time period.

Table 9: Long run trends in grades and career paths

	All	Male	Female	Priv. Fdr.	Pub. Fdr.	Jewish	Colonial	Oth. wht	URM	Asian	A.D. club	Not A.D. club
A. 1924-1940												
High hon.	0.090	0.090		0.050	0.102	0.155	0.065	0.083	0.131		0.032	0.090
Finance	0.103	0.103		0.176	0.060	0.059	0.137	0.097	0.063		0.282	0.100
Higher ed.	0.077	0.077		0.053	0.072	0.077	0.068	0.080	0.116		0.026	0.077
MD	0.090	0.090		0.058	0.134	0.145	0.073	0.082	0.152		0.024	0.091
JD	0.153	0.153		0.143	0.146	0.224	0.129	0.147	0.159		0.137	0.153
MBA	0.066	0.066		0.075	0.070	0.065	0.074	0.063	0.055		0.040	0.066
N	11500	11500	0	2737	2776	1555	3347	6365	189	36	156	11344
B. 1945-1965 (5-year intervals)												
High hon.	0.148	0.148		0.103	0.143	0.241	0.103	0.135			0.042	0.150
Finance	0.115	0.115		0.166	0.078	0.094	0.138	0.115			0.347	0.111
Higher ed.	0.169	0.169		0.131	0.205	0.246	0.128	0.161			0.069	0.171
MD	0.154	0.154		0.105	0.203	0.239	0.091	0.148			0.056	0.156
JD	0.173	0.173		0.164	0.149	0.218	0.131	0.173			0.083	0.174
MBA	0.096	0.096		0.107	0.097	0.085	0.099	0.099			0.181	0.095
N	3819	3819	0	723	727	671	773	2323	34	14	72	3747
C. 1970-1990 (5-year intervals)												
High hon.	0.300	0.295	0.315	0.236	0.331	0.411	0.237	0.274	0.184	0.365	0.123	0.303
Finance	0.096	0.109	0.044	0.146	0.090	0.063	0.111	0.108	0.087	0.093	0.327	0.092
Higher ed.	0.171	0.161	0.208	0.131	0.204	0.220	0.159	0.158	0.103	0.193	0.055	0.173
MD	0.169	0.169	0.169	0.097	0.165	0.211	0.117	0.148	0.189	0.282	0.053	0.171
JD	0.218	0.228	0.184	0.167	0.211	0.245	0.195	0.215	0.245	0.160	0.193	0.219
MBA	0.132	0.139	0.107	0.181	0.137	0.118	0.139	0.139	0.142	0.094	0.333	0.129
N	3184	2522	653	260	524	733	334	1745	195	161	55	3129

Shares of students obtaining graduating honors and pursuing various career paths over time, by demographic background and college club membership. Panels split by graduating class. All cells are shares of students in the category listed in the row for the sample defined by the column. "N" is the count of students for whom we observe occupation data. "All" is the full sample. "Male" and "Female" columns split by gender. Gender codes are based on first names for graduating classes of 1975 and later; gender is imputed as male in the period before coeducation. Private feeder and public feeder columns split by HS type. URM students are defined as Black and Hispanic. Jewish and Colonial categorizations are based on last names. Note that this differs from the Jewish student definition in the historical analysis, which uses first and last names. "Other white" columns consists of non-Asian, non-URM, non-Jewish, and non-Colonial students. "A.D. club" sample includes members of the A.D. club, one of the selective final clubs at Harvard. "Not A.D. club" is all other students. Academic and career data are from Class Report records compiled 25 years after graduation. They are available every year before 1940 and then every five years thereafter. See section 6.1 for details