NBER WORKING PAPER SERIES

OLD BOYS' CLUBS AND UPWARD MOBILITY AMONG THE EDUCATIONAL ELITE

Valerie Michelman
Joseph Price
Seth D. Zimmerman

Working Paper 28583
http://www.nber.org/papers/w28583

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2021

Previously circulated under the title "The Distribution of and Returns to Social Success at Elite Universities." We thank Ran Abramitzky, Anjali Adukia, Joseph Altonji, Marianne Bertrand, Jeff Denning, Susan Dynarski, Peter Ganong, Paul Goldsmith-Pinkham, Larry Katz, Christopher Neilson, Santiago Perez, Bruce Sacerdote, and numerous seminar participants for feedback. We thank Jan Knuf, Nathan Kyn, Jordan Rosenthal-Kay, Mckay Jensen, Katherine Stevens, Merrill Warnick, and Jacob van Leeuwen for research assistance. Michelman gratefully acknowledges support from the Institute of Education Sciences, U.S. Department of Education, through Grant Number: R305B140048 at the University of Chicago. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. Zimmerman gratefully acknowledges support from the Richard N. Rosett Faculty Fellowship at the University of Chicago Booth School of Business. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Old Boys' Clubs and Upward Mobility Among the Educational Elite
Valerie Michelman, Joseph Price, and Seth D. Zimmerman
NBER Working Paper No. 28583
March 2021
JEL No. I24,I26,J24

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 ladder. 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 long-run 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 much more likely to join exclusive campus clubs. The club membership premium is large: members 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 raises the rates at which students join exclusive social groups on campus, but overall effects are driven entirely by large gains for private school students. In the long run, a shift from the 25th percentile of residential peer group status to the 75th percentile raises the rate at which private school students work in finance by 41% and their membership in adult social clubs by 26%. We conclude that social interactions among the educational elite mediate access to top positions in the economy and society, but may not provide a path to these positions for underrepresented groups. Differences in academic and career outcomes by high school type persist through at least the class of 1990, suggesting that this causal channel remains relevant at contemporary elite universities.

Valerie Michelman
Harris School of Public Policy
The University of Chicago
1307 East 60th Street
Chicago, IL 60637
vmichelman@uchicago.edu

Seth D. Zimmerman
Yale School of Management
Yale University
165 Whitney Ave
New Haven, CT 06511
and NBER
seth.zimmerman@yale.edu

Joseph Price
Department of Economics
Brigham Young University
162 FOB
Provo, UT 84602
and NBER
joseph_price@byu.edu

A data appendix is available at http://www.nber.org/data-appendix/w28583

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 their career opportunities may depend. We construct new data on Harvard students in the 1920s and 1930s, linking records of students' social and academic experiences at college to the 1940 Census and to biographical reports compiled 25 years later. We pair these data with two research designs. Our first research 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 Harvard 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 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, and thus reinforces rather than reduces inequality. We conclude that social interactions with high-status peers push people forward on the road to the top, but that even prolonged close contact may fail to spark these interactions 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 clubs, which Amory (1947) describes as the "be-alls and end-alls of Harvard social existence,"

[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.

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

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 attempted a variety of measures aimed at bringing students from outside its traditional constituency at high-status, high-tuition private feeder schools to campus, including automatic admission based on high school class rank. They supplemented these policies with residential life measures aimed at integrating social life across class lines. Randomized room assignment was one such measure.[3] 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 close contact with one another.

The third helpful feature is that, unlike more recent cohorts of students at elite private universities, 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 for each member of the entering class. These reports cover topics like family life, occupation, and social activities. We also link the Harvard data to census records from 1910 through 1940, using crowdsourced family trees from the FamilySearch platform to obtain high match rates at reduced cost.

Our first contribution is to describe how academic and social success at college vary with baseline socioeconomic 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 high-status private feeder schools are more socially engaged than other students, but do worse in school. On average, private feeder students participate in 1.67 extracur-

[3] At Harvard Commencement in 1903, Harvard President Charles Eliot stated that "it is to the last degree undesirable that colleges should be accessible only to the well-to-do," while in 1902 future Harvard President Lawrence Lowell described integration of rich and poor students in dormitories as "the chief value of the College as a place for the training of character." Eliot: Karabel (2006): p. 45. Lowell: Karabel (2006), p. 47. More than 100 years later, Harvard President Drew Faust described how "a very important agenda item for me ... has been expanding access ... and really emphasizing the importance of diversity to excellence" (Walsh, 2017).

ricular activities in their first year of college, compared to 0.85 for other students. 13% of private feeder students hold activity leadership positions and 20% participate in social committees, compared to 5% and 2%, respectively, for other students. Differences in membership in selective upper-year final clubs are even more pronounced: 21% of private feeder students join selective final clubs, compared to 2% of other students. In contrast with their social success, private feeder students are 55% more likely than other students to have grades in the lowest rank group (the bottom 15% of the class), and 51% less likely to have grades in the top two groups (the top 8%).

Our second contribution is to describe the labor market premium associated with club membership, and compare it to the academic success premium. The membership premium is much larger than the academic success premium. Members of selective final clubs in the lowest academic rank group earn 26% more than non-members in the top two academic ranks, and are 3.2 times more likely to have topcoded earnings. This corresponds to the top 0.7 percent of the population distribution. Conditioning on final club membership drives the private school earnings premium towards zero: private feeder students who are not in final clubs have similar earnings to other non-members.

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 44% 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 the social organizations that feed into final clubs 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 have careers in finance and 48% less likely to have careers in medicine. Final club members are more likely to participate in social organizations like country clubs and gentleman's clubs— important features of mid-century social 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 the way it affects long-run outcomes, 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 interactions shape high-stakes outcomes in the medium- and long-run.

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- and long-run outcomes. This design leverages systematic differences in peer neighborhoods driven by dorm layouts, rather than random fluctuations in peer attributes across ex ante identical groups (Angrist, 2014). One way to think of our setting 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 strong 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— equivalent to a move from the 25th to the 75th percentile of the cross-sectional distribution, and well within the support of our random variation— raises students' peer private feeder share by 10.0 percentage points, 31.8% 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 (34.2% 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 11.1% overall, with a 19.1% gain for private-feeder students. For other students, effects are again small and not statistically different from zero. Looking across activities, neighborhood effects are largest for leadership roles, where baseline gaps in participation by high school type are also largest.

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 8.2%. As with on-campus clubs, the overall long-run effects are driven entirely by large (25.6%) gains for private feeder students, with near-zero effects for others. Looking across organization types, most of the gain comes from increases in country club membership. Turning to occupations, a 50 percentile change in neighborhood price rank raises the share of private feeder students in finance by 7.5 percentage points, 41.1% of the group mean. This change is offset by small declines in higher education, medical, and legal careers. For other students, these effects are reversed: exposure to higher-status college peers pushes them away from finance careers and towards careers in 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 0.17σ, the long-run social index by 0.15σ, and the occupation index by 0.21σ. 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.

We study Harvard in the 1920s and 1930s to shed light on the general phenomena of old boys' clubs and elite closure. We conclude the paper with a comparison to recent Harvard cohorts. We extend our time series of academic and occupational outcomes through the graduating class of 1990, for whom we observe career outcomes in 2015. Students from public feeder schools academically outperform students from private feeder schools over the entire period. Differences in occupation choice persist as well. Together with a large body of qualitative evidence (see section 2), the persistence of the academic and occupation gaps over time suggests that social sorting remains an important determinant of outcomes for elite university students. The distributional stakes also persist: using data from Chetty et al. (2020), we show that Harvard students in recent cohorts end up at similar points in the income distribution to those in the 1920s.

Our findings contribute to several strands of work. First, we provide evidence that the college determinants of career success are different at the top. A growing body of research shows that the colleges students go to affect how much they earn, and that the return to college selectivity is often larger for lower-income students.[4] While elite institutions make important contributions to upward mobility overall, we show that unequal access to exclusive social groups limits their contribution to upward mobility to top corporate and social positions. Our results are consistent with Zimmerman (2019), who shows that elite business programs raise earnings for high-status students but not for others; we innovate by unpacking the social and academic channels that give rise to this kind of inequality. Understanding how people access top positions is important because top income shares are large and growing (Alvaredo et al., 2013, 2017). Further, people in top positions make decisions about firm policy and may serve as role models for future students (Matsa and Miller, 2013; Bertrand et al., 2019). The success of policies that diversify selective universities in diversifying top jobs may depend on social as well as academic match.[5]

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

[4]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).

[5]For social match, see Rivera (2016); Jack (2019). For academic match, see Rothstein and Yoon (2008a,b); Arcidiacono et al. (2016); Arcidiacono and Lovenheim (2016)

when the stakes are high. A growing literature uses lab-in-the-field techniques to provide proof of concept that cooperative intergroup interactions can increase low-stakes cross-group socializing and survey measures of intergroup cohesion (Rao, 2019; Lowe, 2020; Mousa, 2020), consistent with Allport (1954)'s "contact hypothesis."[6] Similarly, Carrell et al. (2019) show that assignment to residential groups with high-ability Black peers raises the probability that white students at the US Air Force Academy will choose Black roommates. 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 evidence from Zimmerman (2019) that peers from high-SES backgrounds at top business programs often serve on firm leadership teams together, but not with lower-SES peers.

Third, this paper presents the first evidence (to our knowledge) on how room assignment at school shapes students' long-run outcomes. Many studies use college room randomization designs to explore how school 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).[7]

Fourth, we provide the first evidence on the return to membership in a small but economically and culturally important group: the top social strata at elite universities. The large earnings premium associated with membership, the prominence of members in high-prestige career paths, and the impermeability of group boundaries to efforts at integration provide a rationale for the pronounced media (New York Times, 1907; McWilliams, 1948; Auchincloss, 1958; Rimer, 1993; Flanagan, 2016; Nir, 2016), academic (Amory, 1947; Mills, 1956; Karabel, 2006), and literary (Johnson, 1912; Fitzgerald, 1920) footprint of final clubs and other groups like them, such as secret societies at Yale and eating clubs at Princeton. Our results may also help explain the social importance of other selective social organizations in higher education, such as fraternities and sororities (Marmaros and Sacerdote, 2002; DeSantis, 2007; Popov and Bernhardt, 2012).

2 Institutions

We study Harvard students starting school in the 1920s and 1930s. We choose this setting because it combines a useful vantage point on social sorting by socioeconomic status with rich

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

[7] 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 effects of peers on long-run outcomes in primary school classrooms and military settings, respectively.

documentation of student outcomes. While we are interested in the general phenomenon of social sorting, rather than details of the setting, many of the institutions that shaped Harvard in the 1920s and 1930s persist today in strikingly similar form. 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 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.[8]

Harvard administrators in the 1920s and 1930s considered both academic and social factors when designing admissions policy. In practice, this meant trying to bring 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.[9] 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 evaluates applicants on the basis of questions like "would other students want to room with you, share a meal…or collaborate in a closely-knit extracurricular group?"[10]

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.

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

[9] 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).

[10] Harvard College Admissions and Financial Aid. `https://college.harvard.edu/admissions/apply/what-we-look`. Accessed 4/28/2020. Arcidiacono et al. (2019a,b, 2020) describe Harvard admissions institutions.

[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." (RPHC 1929–1930, p. 100–101).

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

Qualitative accounts also emphasize that these clubs expanded pre-existing inequality of opportunity by student backround, 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 these 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.

[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.

3 Data sources

3.1 Harvard archival records

Our analysis uses microdata on the Harvard entering classes from 1919 through 1935. To build these data 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 merge them together. For a more detailed description see Online Appendices B.1 and B.2.

The first record type we draw from are contemporaneous reports of who is enrolled in Harvard and what they are doing. 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. Our analysis focuses on first-year grades because these are available for more cohorts. We further link to records of upper-year club participation published between 1920 and 1938. These tell us whether students were members of Hasty Pudding and which final club they participated in, if any.

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 a combination of 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 based on dorm layouts. 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 publicly-available Census records from 1910 through 1940 using the linking methods and FamilySearch genealogical database described in Price et al. (2019). The features of our Harvard data that allow for this merge are the parent name and birth location fields in the pre-1934 Class Reports, as well as knowledge of approximate birth year from college cohort. Because we rely on the pre-1934 Class Reports for the Census match, we only attempt

a match for pre-1934 students with Class Report data. FamilySearch is a wiki-style platform in which 12 million users contribute to a shared Family Tree. We supplement the FamilySearch data with hand linking techniques and ensure precision by hand-checking existing links.

3.3 Merge statistics

Figure 1 and Table 1 report statistics on sample size and data availability. We observe 14,382 individuals in our sample universe. As shown in Panel A of Figure 1, sample size increases from 563 in 1919 to 989 in 1935. Panel B shows that 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 89% of Red Book records for the 1920 through 1935 cohorts to class reports. 87% of students matched to class reports (78% of students in the sample universe) have non-missing occupation fields. As shown in Panel C of Figure 1, match rates to Class Report work outcomes are steady between 1920 and roughly 1933 before a decline in 1934 and 1935.[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]

The FamilySearch data help us obtain the high match rates typically associated with hand-linking at a cost similar to automated methods. Our match rates are higher than match rates of 10-30% for automated census and census-to-patent linking (Abramitzky et al., 2020; 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 acknowledged as a gold standard of record linking (Bailey et al., 2020).

Though our match rates to individual datasets are high, our sample sizes decline as we impose sample restrictions and require matches to multiple data sources. For example, 9,342 students live in on-campus rooms subject to random assignment. Of these, 5,218 (55.8%) are in

[13]The decline in quality of Class Report data continues past our analysis window. Reports for more recent cohorts contain much less information than those we use here.

[14]Online Appendix Figure A.2 reports how rates of school enrollment, labor force participation, and other variables change with years since (predicted) year of graduation. By six years after expected graduation (corresponding to the 1930 entering cohort), less than 10% of students are in school and more than 90% of students are in the labor force. Earnings levels continue to rise 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.

the 1930 and earlier cohorts we match to Census outcome records, 3,376 (64.7% of 5,218) are matched to the 1940 Census, and 2,445 (72.4% of 3,376) report wage records. 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 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.1 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.

[17]The remaining clubs are Fox, D.U., and, starting in 1930, Phoenix. We do not include the tenth club listed by Amory, the Iroquois club, because it did not become a final club until 1941.

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

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") 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 and belonging to a country club were 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 represent 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 procedure for variable selection and excluding one's own cohort from the sample. See Online Appendix B.4 for details.

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

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

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. The first step is to describe how students' pre-college backgrounds differ by high school type, our main measure of baseline social status.

Students from private feeder schools come from richer families with stronger ties to Harvard. Panel A of Table 2 reports student background characteristics from Harvard sources. Overall, 42% of students come from private high schools, 24% from private feeder schools, and 27% from public feeder high schools. 7% of all students have a father who attended Harvard, and 20% report having a brother who attended Harvard. 14% of private feeder students have a Harvard father and 29% have a Harvard brother, compared to 5% and 18%, respectively, for other students. Private feeder students are also less likely to come from Massachusetts and more likely to come from New York. Panel B of Table 2 reports background characteristics from the Census. Private feeder students live in homes with an average of 1.88 servants, compared to 0.30 servants for public feeder students. 1% of private feeder students are first- or second-generation immigrants from Eastern or Southern Europe, compared to 13% for other students.

Once at Harvard, most students share a common residential setting. Panel C of Table 2 describe students' first year living environments. Overall, 80% of students live on campus. Almost all private feeder students live on campus (96%), and a large majority (75%) 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.

4.2 The distribution of social and academic success

College outcomes differ dramatically by high school type. We describe these differences in Panels D through F of Table 2 and Figure 2. Our first result is that private feeder students do badly in the classroom compared to others. Panel A of Figure 2 illustrates this. 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.[21]

[21] As reported in Panel D of Table 2, private feeder students and other students are similarly likely to not show up in the class rank data, indicating incomplete or non-passing grades.

Private feeder students are much more successful outside the classroom. In freshman year, students from private feeder schools participate in more and different campus activities than other students. As shown in Panel B of Figure 2, private feeder students are 1.5 times more likely to participate in at least one activity than other students (71% vs. 47%), and participate in twice as many activities on average (1.67 vs. 0.85). Cross-group differences are even more pronounced for leadership activities. As reported in Panel C of Figure 2, private feeder students are three times as likely to have leadership roles (like team captain) and 4.2 times more likely to be members of social committees.

Initial gaps in activity participation and social leadership persist through to exclusive upper year clubs. As reported in Panel D of Figure 2, 39% of private feeder students join the Hasty Pudding sophomore society compared to 8% for other students. 11% of students join a final club of any kind, with a 30% rate for private feeder students and a 5% rate for other students. 7% of students join one of the selective final clubs. The 21% rate for private feeder students is nearly nine times more than the rate for all other students.

Final clubs are much more segregated than campus residential assignments. One way to see this is to calculate the exposure of private feeder students to students from other backgrounds in their first-year residential peer groups, and compare it to the same statistic for final clubs. The average private feeder student lives in a residential peer group in which 60% of students are from non-private feeder backgrounds. The average private feeder student in a selective final club has 26% of club peers from non-private feeder backgrounds.

4.3 The boys' club premium

We now turn to long-run outcomes. Table 3 reports descriptive statistics for 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 (18% vs. 8%), and 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 $2958, 10% more than other students. They are 42% 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 labor market success in the long run. Panels A, B, and C of Figure 3 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 relatively 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 $835 more than other students in the same academic rank group, 31% of the no-club mean. In contrast, conditional on club membership, the relationship between first-year grades and earned income is flat. These differences in Census records of earned income likely understate true differences in income. Panels B and C show that students in selective final clubs are more likely to have topcoded wage income and to report non-wage income.

4.4 Assessing selection effects

Many final club members might have had high earnings regardless of club membership. We use data on students' pre-college backgrounds to assess the importance of selection on different margins in driving the observed 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 3 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. In fact, once one conditions on club membership, differences in earnings by school type decline. 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.

Though high school type captures important information about student background, students may sort into final club membership on the basis of other attributes as well. We conduct a series of analyses that test stories about selection on the basis of high school identity, family background, family identity, 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 specification, and each column to a different outcome.

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 the first column. 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 13 cases. This limits concerns about differential censoring in this analysis.

We next consider earned income, reported in the second column. Panel A of Table 4 shows a baseline specification with controls for private feeder status and class rank, but not final club membership. Before controlling for final club membership, we observe a sizeable private school premium of $293, or 11% of the non-private feeder mean.

Panel B of Table 4 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 3, there is a large final club membership premium even after conditioning on high school type. Students in selective final clubs earn $762 more than other students, a 28% premium above the non-member mean. This is 16 times the size of the academic rank group premium. Also consistent with what we observed in Figure 3, the private feeder earnings premium falls by half once we control for membership in a selective final club.

The next story we consider is selection on the basis of family background. Private feeder students selected for final club membership may come from different family backgrounds than other private feeder students, and this, not club membership, could drive earnings differences. Panel C of Table 4 tests this hypothesis by limiting the sample to students from private feeder schools and adding fixed effects for each high school, as well as a student-specific indicator for Harvard legacy status. The selective final club earnings premium remains large after adding these controls, and the academic success premium remains small.

Panel D of Table 4 continues to explore the role of selection on family background, this time by controlling directly for family fixed effects. We limit the sample to families who send more than one son to Harvard during our sample period.[22] 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 44% of the brothers' sample mean for earnings and 161% of the brothers' sample mean for topcodes. The class rank premium remains close to zero. It is hard for differential selection across families to explain the final club premium.

Panel E of Table 4 considers a third 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 test this hypothesis by restricting the sample to an approximation of the applicant pool for selective final clubs: students selected for Hasty Pudding membership in their sophomore year. Students in this sample are engaged with Harvard high society, but differ in whether they make it to the very top tier. The final club premium remains large within this group, despite the fact that the comparison group of non-members 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.

[22]Students in these families are disproportionately from private feeder schools. See Online Appendix B.5.

The third and fourth columns 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 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, and is not diminished as much as for earned income by the inclusion of additional controls. The second is that while the final club premium for unearned income is large in panels B, C, and E, it is negative and imprecisely estimated in specifications that include family fixed effects. Our interpretation is that unearned income is more closely tied to family background than earned income, and that 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 both overall and relative to the academic success premium. 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 stories about selection on the basis of family attributes or engagement with social institutions at Harvard. Results in Online Appendix B.5 show that similar results hold 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. An important question this analysis leaves open is whether the final club premium reflects a return to pre-existing social skills, or 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.

4.5 Occupations and adult social outcomes

In addition to income, social success predicts career paths and social engagement in the long run. Panels A through D of Figure 4 show how occupation 25 years after graduation varies with class rank and final club membership. We display results for four illustrative career types: finance, medicine, higher education, and law. 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. Higher education and law are both more common at better academic ranks, but are not as strongly related to club membership within rank.

Panels E and F of Figure 4 show how adult participation in social and professional organiza-

tions varies with academic and social outcomes in school. Selective final club members are nearly twice as likely to participate in adult social organizations like country clubs, gentleman's clubs, and fraternal organizations. Participation in such organizations is weakly related to grades. In contrast, selective final club members are less likely to participate in professional organizations (like the American Medical Association or the American Economic Association), and participation in these groups rises as academic rank improves.

As was the case with income, differences in long run career and social outcomes by social success are not simply the product of selection into final clubs on the basis of high school type. The lower row of graphs in Figure 4 displays the same outcomes as the upper row, but split by both club membership and high school type. Within high school type, final club membership predicts a higher likelihood of pursuing a finance career at all levels of academic achievement, and a lower likelihood of pursuing a career in medicine at all but the lowest achievement levels. Both private feeder and other students who are club members are more likely to be members of social organizations as adults. In contrast, membership in professional associations seems more closely related to high school type than social success in college.

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

5.1 Peer groups and room prices

In this section, 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.[23]

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

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

[23] See Online Appendices B.1.5 and B.1.6 for documentation.

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 5 provides an example of one experiment. Panel A maps the main freshman dorms prior to 1931 and their 1920 room prices. There are rooms at a variety of price levels on each floor, and rooms at different price points are often adjacent. Panel B displays a low-priced randomization block: doubles priced at \$175 per student. Peer neighborhoods are outlined and colored by neighborhood mean price. Students assigned to a \$175 per student double may be assigned to any of the rooms outlined in blue, which span a wide range of neighborhood price levels.

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 otherwise indistinguishable peer group units. Our approach is therefore not subject to the Angrist (2014) weak instruments critique; it has more in common with Moving to Opportunity designs in which treatment is the opportunity to relocate 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.

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% of first-year students live on campus, of whom 89% live in freshman halls that are part of the randomization scheme.[24] The rightmost column of Table 2 reports 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 6 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 6 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 6 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 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. This heterogeneity is consistent with administrators' stated goal to integrate residential life by student background.

Turning to the randomized design, we first show that there is wide variation in peer group assignments within randomization blocks. Panel A of Figure 7 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 maximium 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 neighbor-

[24] 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.

hood 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 68% 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 Figure 8 and 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 predicted values of 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 29.4 percentage points more likely to have attended a private feeder school. They are 7.1 percentage points (80.8% of the sample mean) more likely to have a Harvard father and 11.0 percentage points (48.0% 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=0.855). This null relationship persists when we add the full set of randomization blocks and high school dummies (p=0.385). 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 B of Figure 8 and Panel C of Table 5 report 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 26.7 percentage points higher (on an in-sample mean of 54.0) than students assigned to the lowest-price neighborhoods, and private feeder high school peer shares that are 20.0 percentage points higher (on a mean of 31.4), conditional on randomization block. These effects are precisely estimated, with standard errors of 3.4 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 31.8% 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 5.0 percentage points lower. Exposure to rooms in the top 50% and top 10% of the room price distribution 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 0.15, 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. Panel A of Figure 9 displays coefficient estimates for outcomes that are central to our discussion. 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 0.282 (SE=0.110; mean of 1.266), 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 0.205 standard deviations (SE=0.081), equal to 33.8% of the gap in activity index between private feeder students and other students in the full sample. Effects on activities by type suggest increases in participation in social committees (4.9 percentage points) as a key channel for increased participation.

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 0.656 (SE=0.222), and the private school activity index by 0.480 (SE=0.170). 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 18.4 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 higher-status residential peer groups shift away from activities with lower-status students. Effects are positive or near zero across the board, even for activities where the share of private feeder students is relatively low, such as music or those in the "other activities" category. For these students, 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. Results for final clubs are very similar to those for first-year activities. We report our findings in Panel B of Table 6 and Panel C of Figure 9. A one-unit change in room rank raises the rate at which students in the pooled sample join selective final clubs by 6.5 percentage points (SE=2.1; on a full sample mean of 9.5). This effect is again driven entirely by a 16.8 percentage point effect for private feeder students (SE=5.6), with precise zero effects 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, a preliminary step to membership in final clubs, also rises for private feeder students but not for other students, as does the upper-year club index measure. 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 Hasty Pudding participation and the feeder index.

The social effects we see are economically large. A fifty percentile increase in peer neighborhood price rank raises the count of activities private feeder students participate in by 19.1% of the sample mean in that group and 53% of the gap between private feeder and other students. The equivalent figures for social committees are 50% and 73%; for selective final clubs they are 37.7% and 45.1%.

Panel A of Figure 10 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 of selective final club membership 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.

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.[25] The effect of assignment to higher-priced peer neighborhoods on a continuous class rank measure is also near

[25]The rank group indicators we take as outcome variables here include unranked students as zero values.

zero. Panel B of Figure 10 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, 4.7 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, with Panel B of Figure 9 displaying the coefficients of interest. In the full sample, a zero-to-one change in peer neighborhood rank raises the rate of membership in any adult social organization by 6.3 percentage points (SE=3.6, sample mean of 38.1%). The full-sample increase reflects the combination of a 23.8 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=0.008). A 50 percentile change in peer neighborhood price raises rates of participation in adult social clubs for private feeder students by 25.6% of their sample mean and 98.1% of the gap by high school type. Panel C of Figure 10 shows the dramatic 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 rank.

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 21.6 percentage points (SE=6.6). We see no effects for membership in professional or honor societies. These findings again parallel descriptive results from Figure 4 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 and Panel C of Figure 9. 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 neighborhood price raises the occupation index by 0.470 (SE=0.183), while for other students it lowers the occupation index by 0.173 (SE=0.095). The p-value from a test of the null 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 15.0 percentage points (SE=5.9), on a base of 18.3% 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, and a test of the null that the finance effect is the same for private feeder students and other students returns a p-value of 0.005. There are smaller shifts across other occupation types, with private feeder students shifting slightly away from medicine, higher education, and teaching, and other students shifting towards medicine and lower management.

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 41.1% of their sample mean and 82.8% of the gap by high school type. For the occupation index, this effect is equal to 72.4% of the gap by high school type. Panel D of Figure 10 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.

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 of 24.9%, 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 outcomes from the 1940 Census to construct an earnings index that captures how much we might expect census 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 Report occupation, we regress earnings from the 1940 Census on indicators for each broad occupation and adult social activity category. As with our private school indices, we select covariates using a Lasso.[26] Our index is predicted values from this regression, which we can compute for all students in the Class Report occupation sample, not just those matched to

[26] We again use EBIC for model selection.

the census. Units are dollars. We then place this Class Report earnings index on the left hand 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, 2018a; Saavedra and Twinam, 2020; Abramitzky et al., Forthcoming). 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 \$65 (SE=25). This is equal to 8.5% of the descriptive selective final club premium reported in Panel B of Table 4, which is similar to the 8.4 percentage point increase in selective final club membership this shift induces, as reported in Table 6. 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.

5.7 Extensions and robustness

Additional analyses explore effects for different subpopulations and the sensitivity of our quasi-experimental analysis to alternate modeling decisions. Online Appendix Tables A.1 through A.4 show that our findings are similar 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.5 and A.6 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.7 and A.8 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.9 and A.10 report results using randomization inference as opposed to clustered standard errors to conduct statistical tests; our results are unaffected.

Finally, Online Appendix Figure A.3 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. As in linear specifications, we observe a pattern of 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 Heterogeneous effects and the opportunity costs of finance careers

One of our main findings is that exposure to high-status peers shifts private feeder students towards finance careers. An important question for interpreting this shift is whether it reflects a socially productive reallocation of talent (Baumol, 1990; Murphy et al., 1991; Philippon, 2010). A full answer to this question is beyond the scope of this paper, but we can describe two key inputs to the debate: the nature of the jobs student shift between, and how academically successful the students are who shift.

Compared to finance jobs in other periods, finance jobs in mid-century were characterized by relatively low job complexity and skill intensity (Philippon and Reshef, 2012).[27] Stylized depictions of the finance industry from the 1950s through the 1970s tell a story in which regulations limited competition, and incumbent banks benefited from monopoly rents. One manifestation of limited competition was the so-called "3-6-3" rule describing bankers' business strategies and hours during the period: borrow at 3 percent, lend at 6 percent, on the golf course by 3pm (Walter, 2006).

Is it academic high achievers who are routed into finance jobs? Rather than answering this question by running an experimental analysis that conditions on grades, which are causally downstream of room assignment, we study how assignment to higher-priced peer neighborhoods affects the joint distribution of first-year academic performance and career outcomes. We define indicator variables for *both* having good (or bad) grades *and* engaging in a given career or social outcome. We then take these indicators as outcomes of interest in estimates of equation 2. We define good grades as being in the fourth rank group (roughly the median of the grade distribution) or above.

Results from this exercise show that the career and social shifts we observe are driven by academic high-achievers. We present these findings in Table 8. A zero-to-one change in neighborhood price rank raises the rate at which students both have high grades and are members of selective final clubs by 12.9 percentage points. It raises the rate at which students have high grades and enter finance careers by 11.9 percentage points. And it reduces the rates at which students enter careers in higher education and medicine by 4.9 and 4.5 percentage points. Rates

[27] Specifically, finance jobs in mid-century had similar skill intensity to the rest of the economy; finance jobs in earlier and later periods were higher-intensity.

at which students hit these career and social milestones and have low grades are not affected by residential assignment. Though we do not know whether the "3-6-3" depiction of finance careers reflects the jobs for Harvard students pursuing finance in our sample, this evidence suggests that whatever the finance jobs were, their social opportunity cost in terms of the allocation of top academic talent was high.

6 Harvard students in historical context

6.1 Academic achievement and occupational sorting over the 20th century

Our main results provide direct evidence on the role of old boys' clubs and interactions between high-status peers in the formation of American elites in the middle of the 20th century. As described in sections 1 and 2, qualitative reports suggest this phenomenon is important in many contexts. This section focuses on the relevance of our results for elite universities in the contemporary US. The goal is to provide guidance about which elements of the context we study have changed over time, which have stayed stable, and what this means for the portability of our findings across settings.

Our first step is to extend the time series for key outcomes through the late 20th century. To do this, we link Freshman Red Books for each graduating class through 1990 to 25th Reunion Class Reports at five year intervals; i.e., 1945, 1950, and so on, through 1990. For students in the graduating class of 1990, we observe labor market outcomes in 2015. Census outcomes are not available over this longer period. However, several useful data fields *are* recorded consistently. These include the high schools students attend, the honors designations with which they graduate, subsequent degrees they earn, and the occupations in which they are employed. We use these data to describe how academic and career outcomes vary with high school type for the Harvard graduating classes between 1924 and 1990. Online Appendix B.7 describes the construction of the long-run series, and provides evidence that the data we use allows for consistent cross-time comparisons and matches aggregate statistics on attributes of Harvard students in the years these statistics are available.

Our long-run analysis compares students from private feeder schools to the full population of students, and to students from public feeder schools. The goal is to contrast outcomes for groups of high-status and low-status students that we can identify across the full period. On the high status side, numerous sources confirm that the schools we identify as private feeders in the 1920s and 1930s retain their status.[28] We keep our list of private feeder schools fixed for the long-run analysis. As we discuss further in section 6.2, the share of high-status students on the Harvard campus remains high through the present; we view students from private feeder

[28]Cookson Jr. and Persell (1985, p. 43-44) discuss multiple prestige rankings of private high schools; these lists of 15-20 schools nationally include all eight of the private feeder schools we use here.

high schools as an identifiable subgroup within this broader category. On the low-status side, we extend our classification of public feeder schools to include schools that send more students to Harvard after the conclusion of our main sample period in the 1930s. See Online Appendix B.7 for details.

As Harvard's national footprint expands over the 20th century (Hoxby, 2009), the share of students from feeder high schools declines. Panel A of Figure 11 shows how the share of students from different high school types evolves over the period. Each point on the graph is an average within a five-year bin. The share of students from the "other HS" category rises from about one third in the 1920s to 66% by 1990. The biggest jump takes place during and after World War II.

Panel B of Figure 11 shows that the share of students graduating with any kind of honors designation rose from 27% in the early 1920s to 76% in in the late 1980s, with the fastest expansion coming in the 1960s. To smooth over smaller samples in the Class Report data, each point on this graph and those that follow displays means over graduating classes within 2.5 years of the centered value. For example, the 1982.5 data point is an average of the 1980 and 1985 class years. Our findings closely match aggregate data from Harvard administrative sources reported in Healy (2001); we display the Healy data on the graph as well. Because most students graduate with honors by the late 20th century, we use a definition of academic high achievement that includes only students awarded a *magna* or *summa cum laude* degree.[29] By this definition, the share of academic high achievers rises from 8% in 1925 to 31% in 1990.

Panel C of Figure 11 presents the first main result of our long-run analysis: private feeder students perform worse in the classroom than other students over the entire 1924-1990 period. In the early 1920s, public feeder students were two times as likely to graduate with high honors than private feeder students (8% vs. 4%). In the 1980s and 1990s, they were 1.5 times more likely (34% vs. 23%). To facilitate consistent comparisons across time and with our main analysis, Panel C displays honors attainment for male students only. However, as reported in Online Appendix Figure A.4, we see similar results when we include female students.[30]

We next describe the evolution of career choices by high school type, continuing to focus on male students. For medical and law careers, we focus on the receipt of MD and LLB/JD degrees, respectively, as our main measures of interest. We do this because the use of degree type rather than text descriptions of jobs makes cross-time comparisons more straightforward for careers that correspond closely to specific degrees. We discuss career classification in the extended time series in Online Appendix B.7, and provide evidence that our findings are consistent across different classification approaches.

Gaps in career choice by high school type persist over time. We report our findings in Figure 12. Looking first at finance, we see that private feeder students are more likely to pursue finance

[29] Healy (2001) attributes the increase in honors degrees to grade inflation during the Vietnam War.

[30] Harvard began to admit women as undergraduates in the 1970s. See Online Appendix B.7 for gender shares by year.

careers than public feeder students over the full period, but that the size of the gap declines over time. The most notable decline occurs in the late 1970s, when public feeder students begin to pursue finance careers at higher rates. This corresponds to the increase in skill intensity in finance reported in Philippon and Reshef (2012). Full-sample finance shares fall in between the private feeder and public feeder shares, a pattern that repeats across other career types.

Turning to higher education careers, we see growing gaps by high school type over time. Public and private feeder students pursue careers in higher education at similar rates prior to the 1940s. Starting in the 1940s, rates grow for both groups, but they grow more rapidly for public feeder students. By the end of the period, public feeder students are 60% more likely than private feeder students (24% vs. 15%) to report higher education as a career path. Public feeder students are also more likely to pursue medical careers. Both the level and the cross high school gap in MD receipt are relatively stable over the full period. In contrast, public and private feeder students receive law degrees at similar rates, with the level for both groups growing over time. Rates of MBA receipt are low for both private and public feeder students and grow slowly until the late 1970s, when rates for private feeder students jump sharply.

The overall picture is one where career paths of Harvard students shift over time, but large differences by baseline social status persist. We see some evidence that the role of the finance sector as a differentiator of career outcomes by baseline status diminished for cohorts graduating in the 1970s and 1980s as the skill content of finance jobs shifted. However, given the persistent gap we observe in medical careers and the widening gap in higher education careers and MBA receipt, we see little to suggest that this reflects overall convergence in career outcomes.

Our quantitative finding that differences in career outcomes persist across the 20th century coheres with the extensive qualitative documentation of social divides by baseline status through the present day described in section 2. The combination of qualitative and quantitative evidence extending from the 1920s through the 2010s suggests that interactions between high-status peers at elite universities are an important determinant of who holds top jobs over much of the 20th century and into the present day.

In supplementary analyses we consider sorting into careers on the basis of academic achievement, and career sorting in the full sample of Harvard students, including women. Patterns for the full sample are similar to those we observe for men only. Splitting by academic achievement shows that high-performing students tend to go into higher education, while lower-performing students go into finance and get MBA degrees. See Online Appendix Figures A.5 and A.6.

6.2 Academic performance and admissions standards

Our finding that students from high-status backgrounds tend to perform less well academically than their low-status peers is consistent with the data the available data on admissions standards. Admission rates have fallen dramatically over time, but students from the most privileged back-

grounds appear to have retained substantial admissions advantages. 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. (2019b) reports that applicants who are legacies, likely donors, or children of faculty (LDC) are a) less likely to come from disadvantaged backgrounds, and b) much more likely to be admitted to Harvard than other students whose academic performance is better. For example, a student in the LDC category in the third decile of Harvard's academic performance index is more likely to be admitted than a student outside of these groups in the 10th decile. Though the recruiting and admissions processes have changed over the past 80 years, the ingredients that lead privileged students to be disproportionately represented at the lower end of the academic performance distribution remain.

In addition to reinforcing our findings on the persistence of cross-group gaps, admissions records shed light on how the share of highly-privileged students has changed over time. We report in Figure 11 that the share of students from private feeder students drops over time. However, given the expansion of Harvard's geographic reach over the period and the potential emergence of other high-status schools, the share of high-status students from outside the private feeders may be rising. Consistent with this hypothesis, Arcidiacono et al. (2019b) report that LDC students make up 19% of all Harvard admits in the classes of 2014-2019. The 19% share of recent cohorts is fairly similar to the share of private feeder students we observe enrolling in Harvard in our main sample period (24%). The LDC designation may capture elements of family background similar to those represented by the private feeder designation.

6.3 Harvard students in the income distribution

The labor market position of Harvard students in the cohorts we study resembles the position of more recent cohorts. Figure 13 plots histograms of the earned income distribution for all men 27–37 in the 1940 Census (the age range for Harvard students in our census sample), for all men in that age range with four years of college, and for Harvard students in our sample. The mean income for Harvard students of $2747 is more than double the mean income in the full population ($1167) and more than thirty percent higher than the mean for college graduates ($2043). 18.1% of Harvard students are among the top 1% of the earnings distribution (including non-earners as 0s) of men in their birth cohorts ($3945 or higher), three times the rate for men with at least four years of college (6.0%).

Three-quarters of a century later, Harvard students are overrepresented at the top of the earnings distribution at similar rates. As reported in Chetty et al. (2020), 21.1% of Harvard students born 1980–1982 (roughly corresponding with entering cohorts of 1998–2000) have earnings in the top 1 percent of the age-specific income distribution. The relative stability of this statistic over time suggests that what happens at Harvard today may shape the formation of future elites in a way similar to what we observe in our data.

6.4 The return to academic success

We observe a final club membership premium that is much larger than the academic success premium. Comparing our findings 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 relatively recent cohorts. Goldin and Katz (2008) find that a one standard deviation increase in college grades raises earnings by 15 to 20%.[31] Converting our estimates to standard deviation units, the largest effects we observe across multiple specifications correspond to a 5 to 7% premium. We report these results in columns 5 and 6 of Panel C in Online Appendix Table B.11, which focuses on third-year rather than first year class rank. In particular, column 6 adopts the Goldin and Katz approach of scaling topcoded earnings values by 1.4. See Online Appendix B.5.3 for details. The difference between our estimates of grade effects and the Goldin and Katz estimates may reflect differences in measurement (they use survey and administrative reports of income and cumulative GPA; we use census income data and year-specific class rank). The 28% final club premium that we estimate in our main specifications is equivalent to about a 1.5 to 2σ change in grades using the Goldin and Katz estimate, and a 4σ change using our largest in-sample estimate.

7 Conclusion

This paper uses data on Harvard students to show that social interactions among the educational elite shape the top rungs of the economic and social ladders over the long run, but that the career effects of exposure to high-status peers only accrue to students from high-status backgrounds. Even close and prolonged residential contact with high-status peers is not sufficient to help lower-status students access exclusive old boys' clubs and the long-run 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. The key value of our exercise is to test both the general proposition that social interactions shape high-stakes outcomes and the specific policies that Harvard and other universities use to promote cross-group interaction on campus.

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

[31] These coefficients are estimated but not reported in the paper. We thank Larry Katz for providing these results.

them access influential positions in the economy and society. However, our findings suggest that expanding formal access to elite higher education beyond its traditional constituencies does not by itself integrate the informal old boys' clubs upon which access to at least some top social and economic positions seems to depend. What policymakers at universities and beyond might do to promote broader access to these groups is a subject for future work.

References

Abramitzky, Ran, Leah Platt Boustan, Elisa Jácome, and Santiago Pérez, "Intergenerational Mobility of Immigrants in the US over Two Centuries," *American Economic Review*, Forthcoming.

_ , 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.

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.

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, "Divergent: The Time Path of Legacy and Athlete Admissions at Harvard," Working Paper 26315, National Bureau of Economic Research September 2019.

_ , _ , and _ , "Legacy and Athlete Preferences at Harvard," Working Paper 26316, National Bureau of Economic Research September 2019.

_ , _ , and _ , "Asian American Discrimination in Harvard Admissions," Working Paper 27068, National Bureau of Economic Research April 2020.

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.

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

Bertrand, Marianne, 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.

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.

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 post-war 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.

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

DeSantis, Alan, *Inside Greek U.: Fraternities, sororities, and the pursuit of pleasure, power, and prestige*, University Press of Kentucky, 2007.

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

Feigenbaum, James J, "Multiple measures of historical intergenerational mobility: Iowa 1915 to 1940," *The Economic Journal*, 2018, *128* (612), F446–F481.

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.

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.

Folkman, Tyler, Rey Furner, and Drew Pearson, "GenERes: A Genealogical Entity Resolution System," in "2018 IEEE International Conference on Data Mining Workshops (ICDMW)" IEEE 2018, pp. 495–501.

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

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.

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.

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.

Johnson, Owen, *Stover at Yale*, Grossett & Dunlap, 1912.

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.

Jr., Peter W. Cookson and Caroline Hodges Persell, *Preparing for power: America's elite boarding schools*, Basic Books, 1985.

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.

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.

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 reevaluated," *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.

_ **and Ariell Reshef**, "Wages and human capital in the US finance industry: 1909–2006," *The Quarterly Journal of Economics*, 2012, *127* (4), 1551–1609.

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," Working Paper 26227, National Bureau of Economic Research September 2019.

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.

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.

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

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.

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.

Walsh, Colleen, "For Faust, the Road Ahead," *The Harvard Gazette*, September 2017.

Walter, John R, "The 3-6-3 rule: an urban myth?," *FRB Richmond Economic Quarterly*, 2006, *92* (1), 51–78.

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.

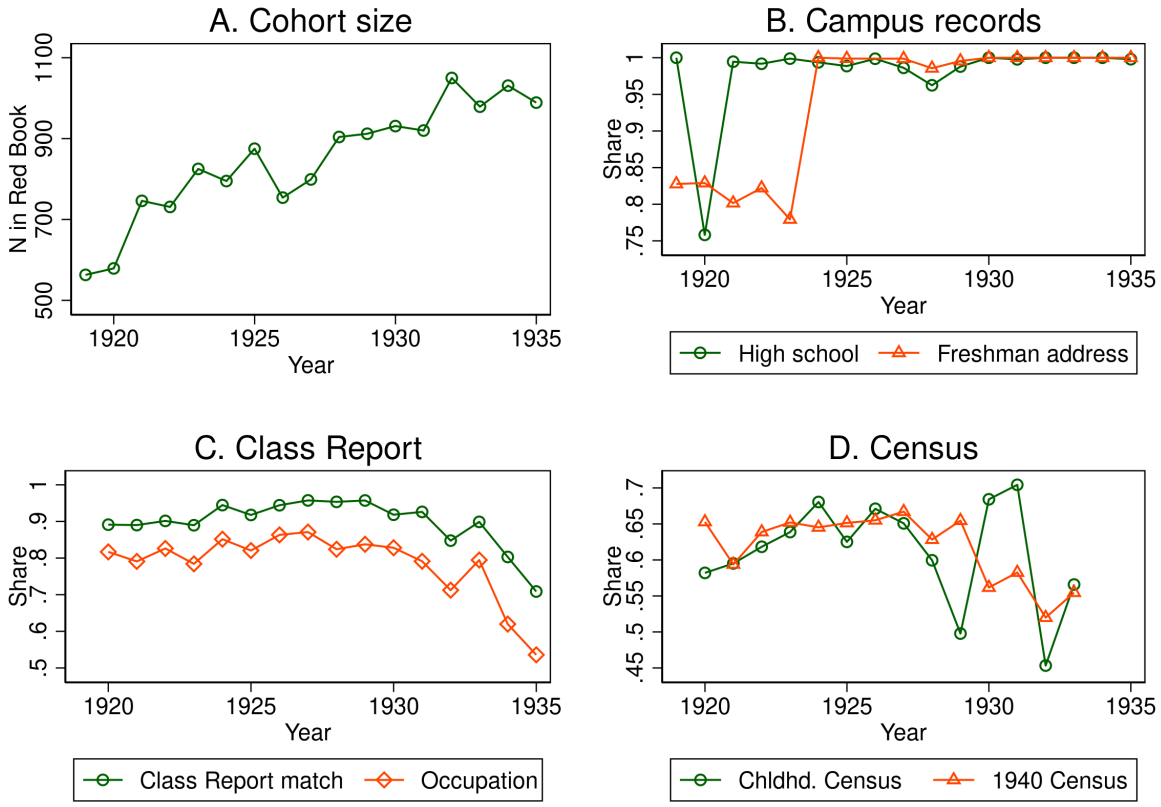
Zimmerman, David J, "Peer effects in academic outcomes: Evidence from a natural experiment," *Review of Economics and statistics*, 2003, *85* (1), 9–23.

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.

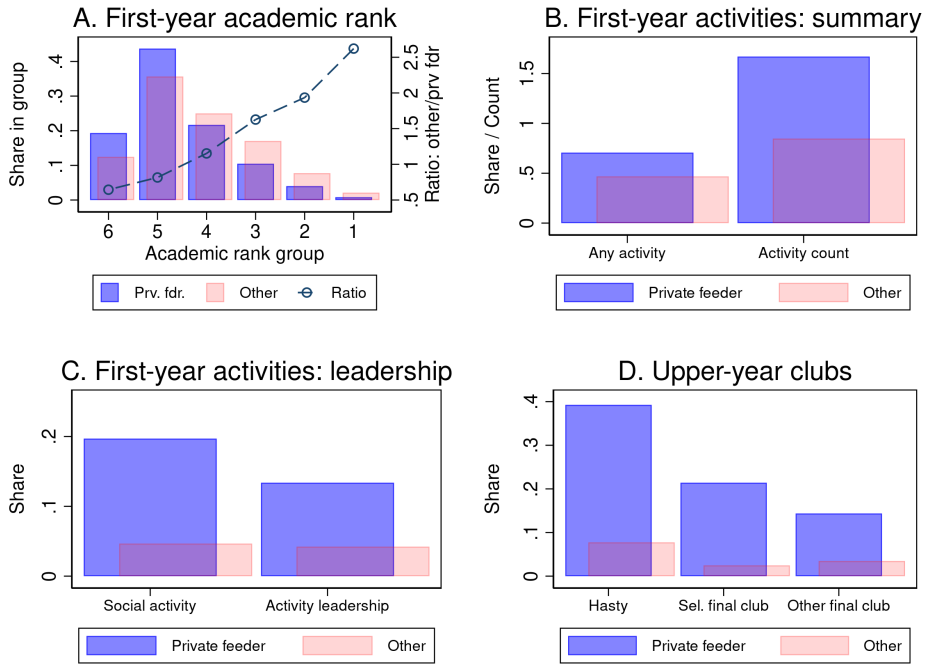
Figures

Figure 1: Data availability by cohort



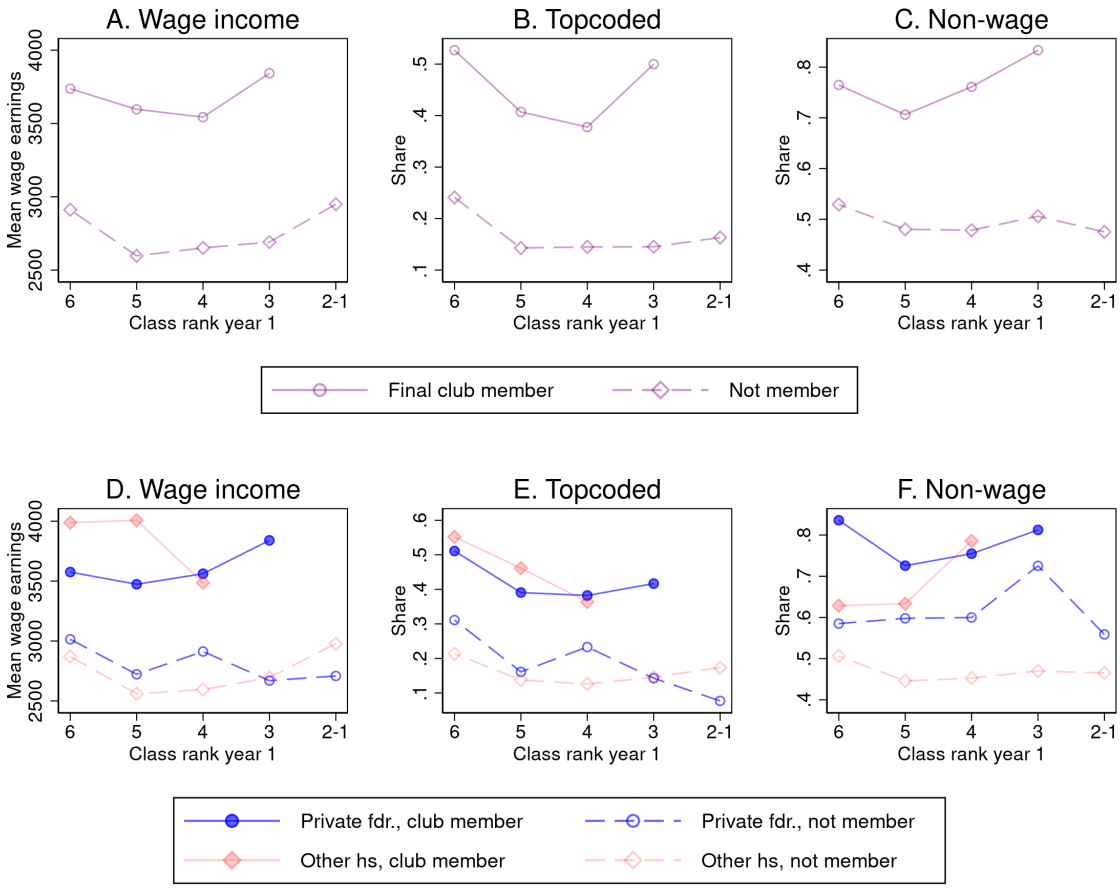
Sample counts and data availability by year entering Harvard. Panel A: Count of students by year. Panel B: Share with data on high school attended and campus address. Panel C: Share matched to 25th Reunion Class Reports and share who are both matched to Class Reports and have non-missing occupation data. 1919 cohort excluded because Class Reports are not available. Panel D: Share matched to childhood and adult Census records (overall, not conditional on Class Report match). We exclude cohorts 1919, 1934, and 1935 from our Census match because Class Reports for those cohorts do not include variables needed for the match.

Figure 2: Academic and social achievement by high school type



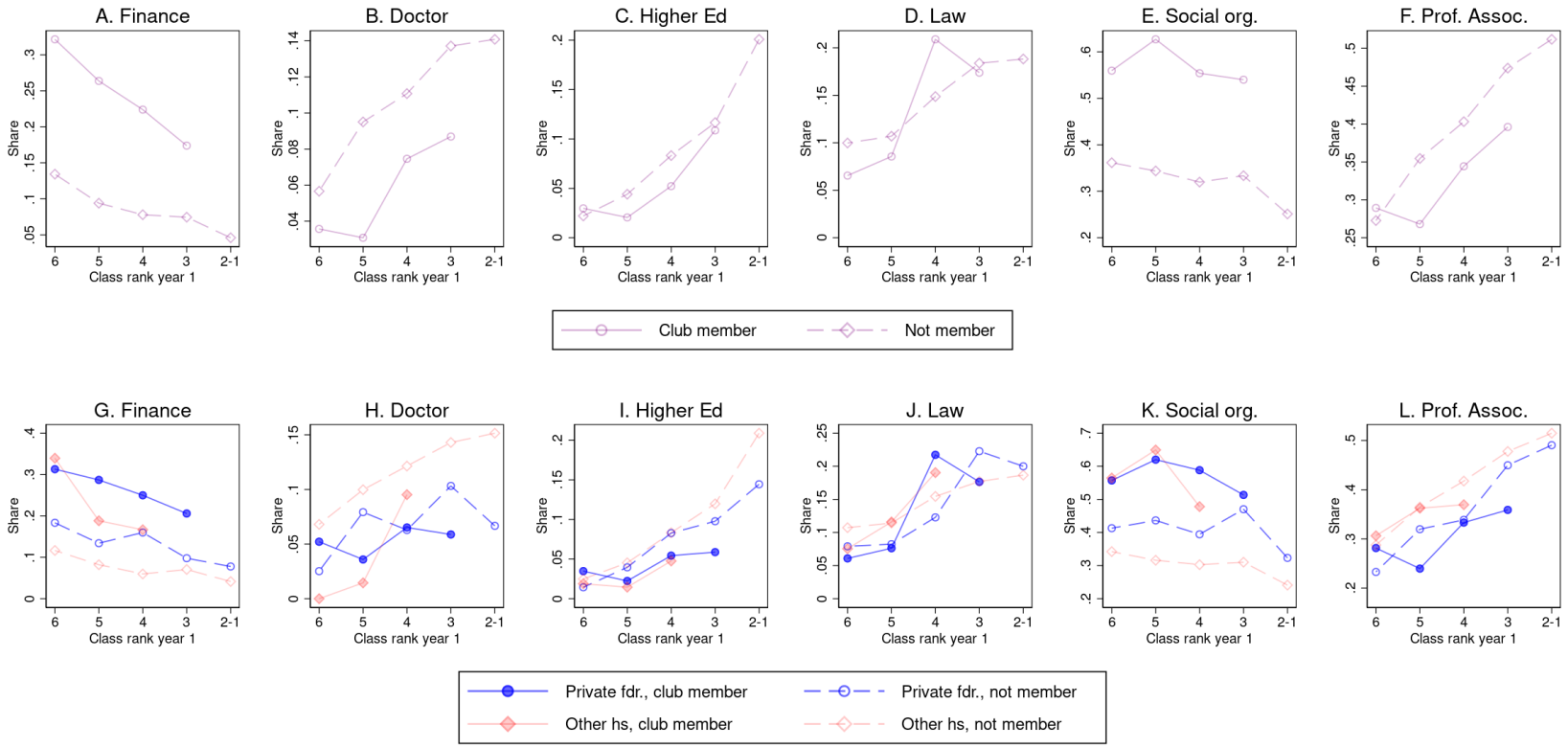
Panel A: Shares of private feeder and other students in each academic rank group. Group 1 contains the highest academic performers; group 6 contains the worst performers. The line is the ratio of other students to private feeder students. Panels B–D: Shares of private feeder and other students who participate in activities of the listed type. Panel B presents summary measures of first-year activities. Panel C focuses on leadership activities. Panel D reports results for upper-year clubs: Hasty Pudding, selective final clubs, and non-selective final clubs.

Figure 3: 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. We collapse groups 1 and 2 and 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 4: Adult career and social outcomes by academic performance and final club membership

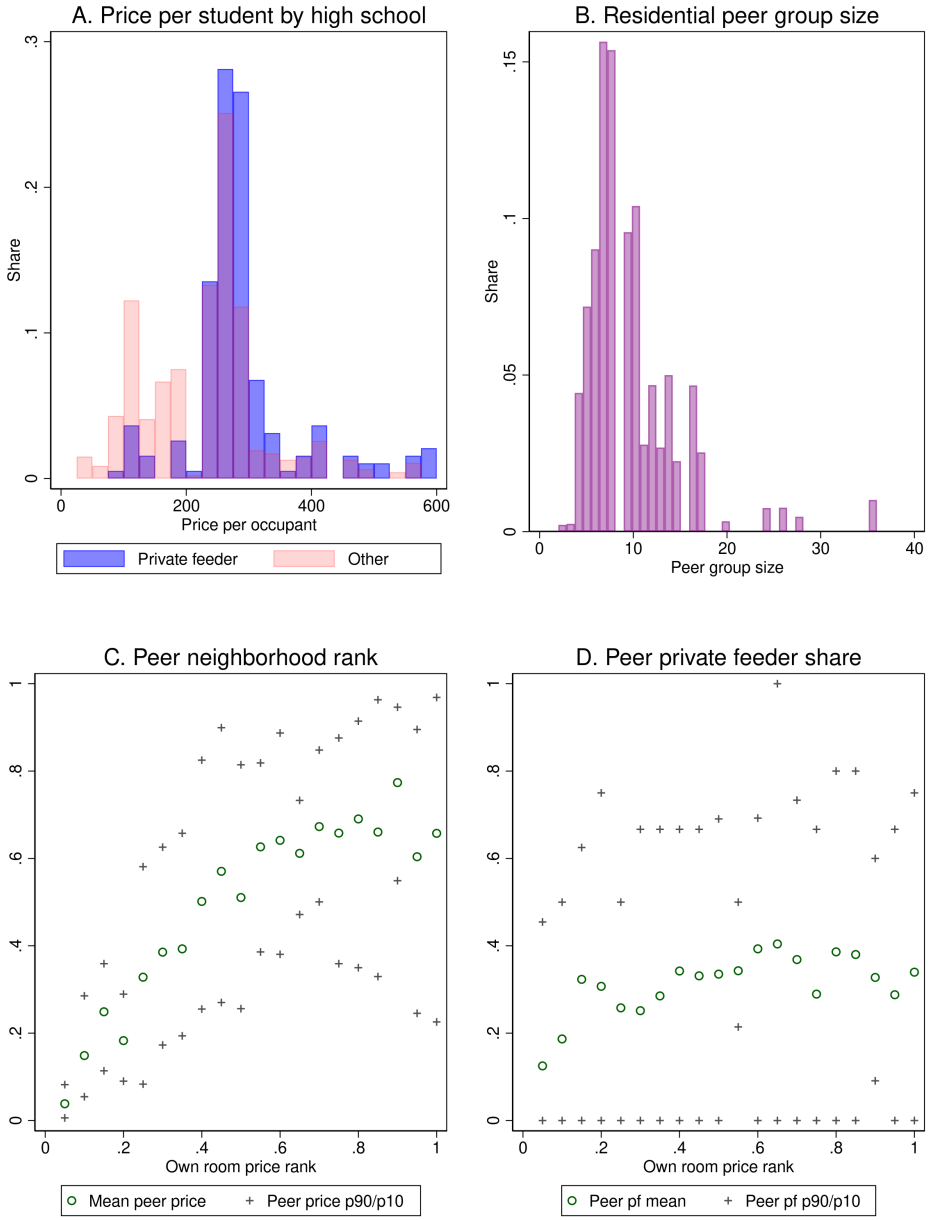


Career outcomes and adult social outcomes by academic performance and membership in a selective final club. Outcome types come from Class Report data and are listed in panel titles. Panels A–F depict adult outcomes for students by freshman academic rank group and selective final club membership. Panels G–L 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 5: Room price and randomization examples

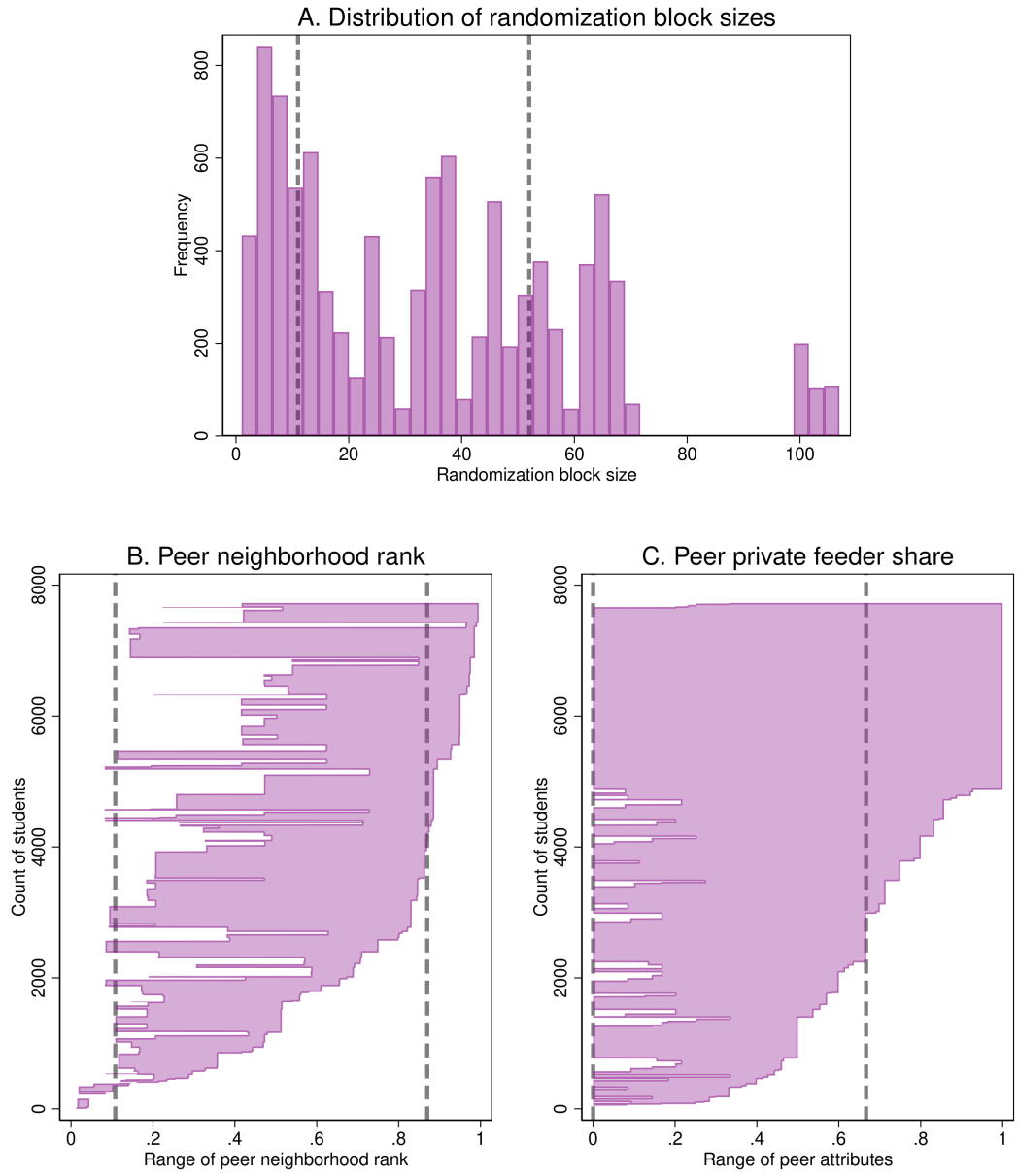
Left panel: Room prices of Harvard first-year dorms in 1920. Harvard used these dorms for first-year students until 1931. Right panel: Mean peer neighborhood prices and example of a randomization block. Rooms from an example low-price randomization block are outlined in blue. These outlined rooms all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

Figure 6: Dorm room prices and peer attributes



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.

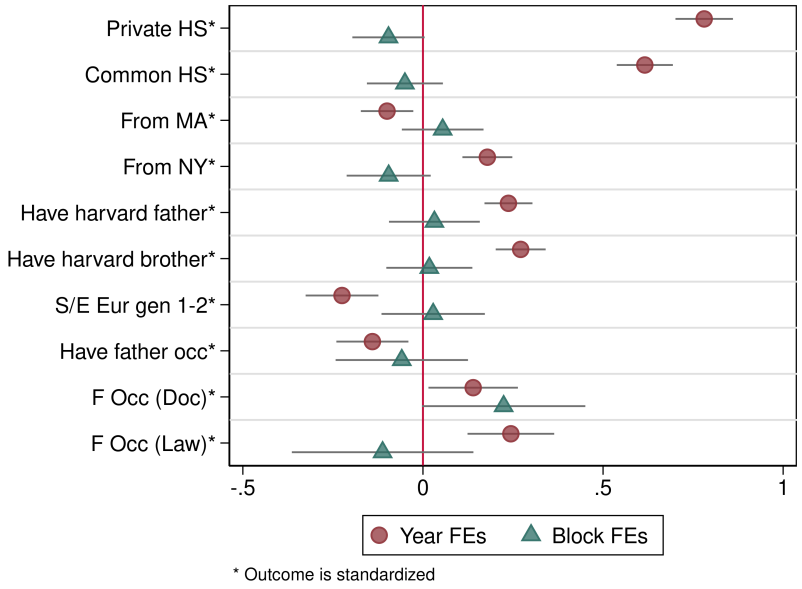
Figure 7: Randomization block size and within-block variation in peer attributes



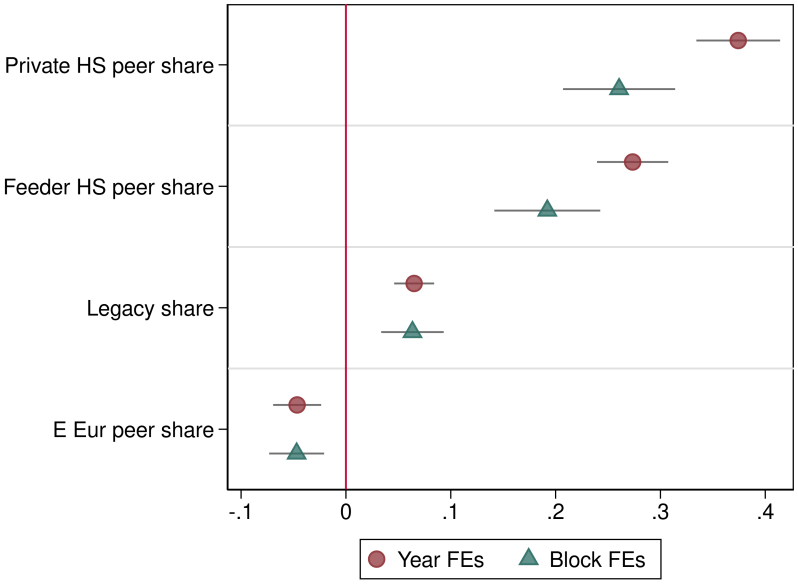
Panel A depicts a student-weighted histogram of randomization block size. Panels B and C depict the range of mean neighborhood room rank (B) and private school share (C) across neighborhoods within each randomization block. Blocks are sorted vertically by maximum room rank (B) or private school share (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.

Figure 8: Balance and First Stage Effects of peer neighborhood price

(a) Balance tests



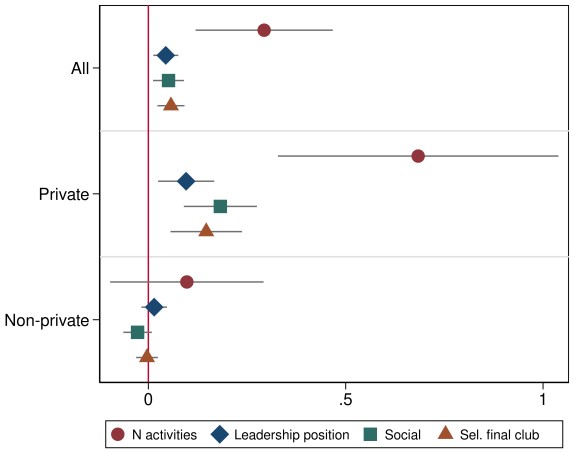
(b) Peer attributes



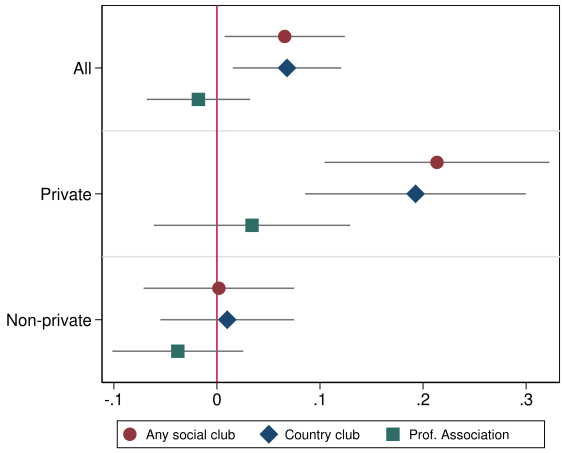
Coefficient estimates and 90% CIs for the peer neighborhood room rank variable in equation 2. Outcomes are listed on vertical axis of each panel. "Year FEs" specifications include only year effects, not randomization blocks. "Block FEs" specifications add controls for randomization blocks. Panel A: Outcomes are predetermined student attributes. Outcomes are standarized (mean zero, standard deviation one) to facilitate display. Panel B: Outcomes are attributes of the peer neighborhood to which a student is assigned. These results correspond to specifications reported in columns 1 and 2 of Table 5.

Figure 9: Effects of peer neighborhood price

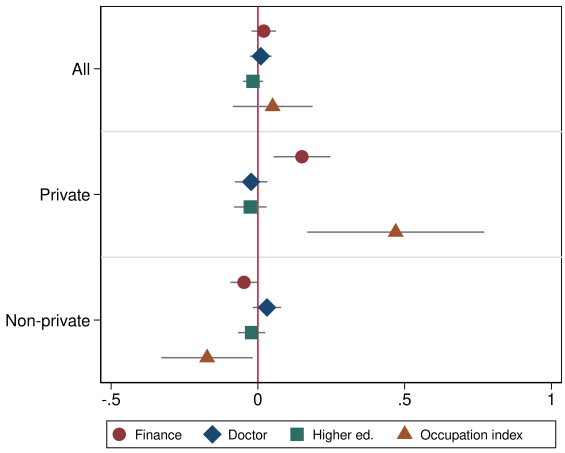
(a) College social activities



(b) Adult social activities

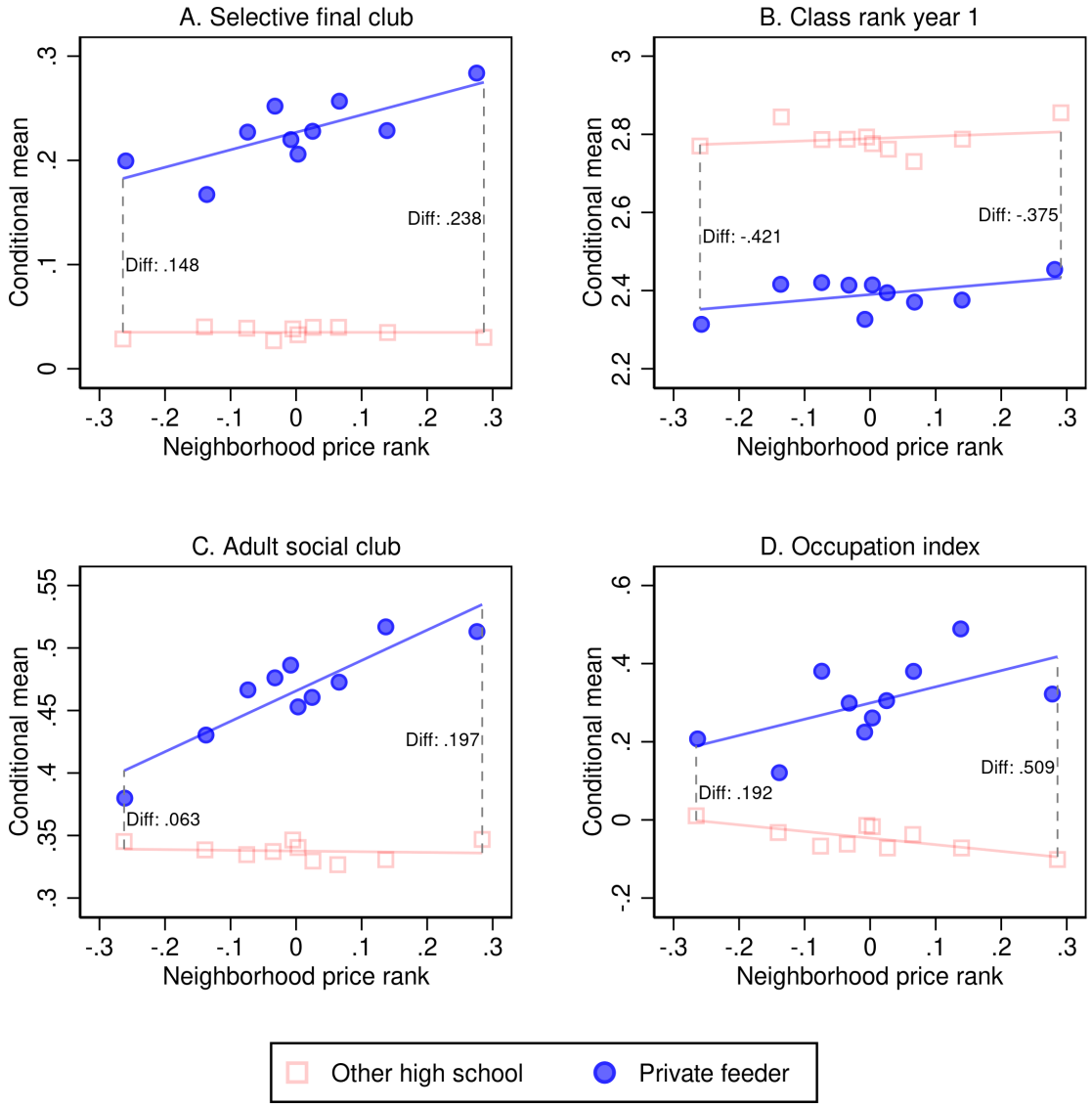


(c) Adult career outcomes



Coefficient estimates and 90% CIs for the peer neighborhood room rank variable in equation 2, in the full sample and split by high school type. Samples are listed on the vertical axis of each panel. Each point is a different outcome, noted in the legend of each panel. Each panel shows results first for the full randomization sample ('All'), then for private feeder school students only ('private'), then for other students ('non-private'). Panel A displays results for college social activities. Panel B displays adult social activities. Panel C displays adult career outcomes. See main text for detailed variable definitions. These results correspond to specifications reported in Tables 6 and 7.

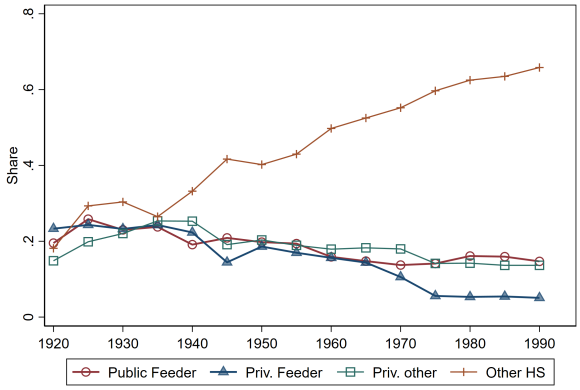
Figure 10: Key outcomes by within-block decile of peer neighborhood price and high school type



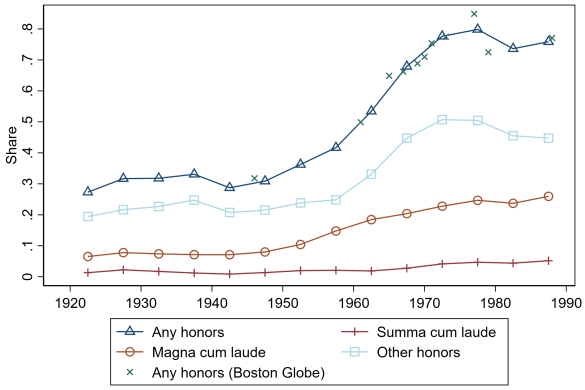
Binscatter plots showing the listed outcome by decile of peer neighborhood price rank, within randomization block. We obtain these graphs by regressing the vertical- and horizontal-axis variables on randomization block and large high school fixed effects separately by high school type. We then plot the 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 11: Long-run trends in high school background and academic achievement

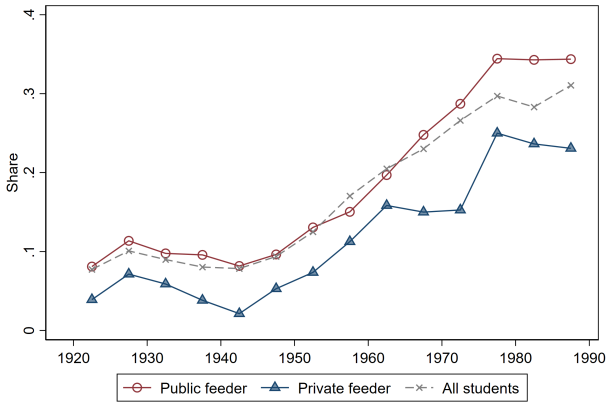
(a) High school background



(b) Honors pooled

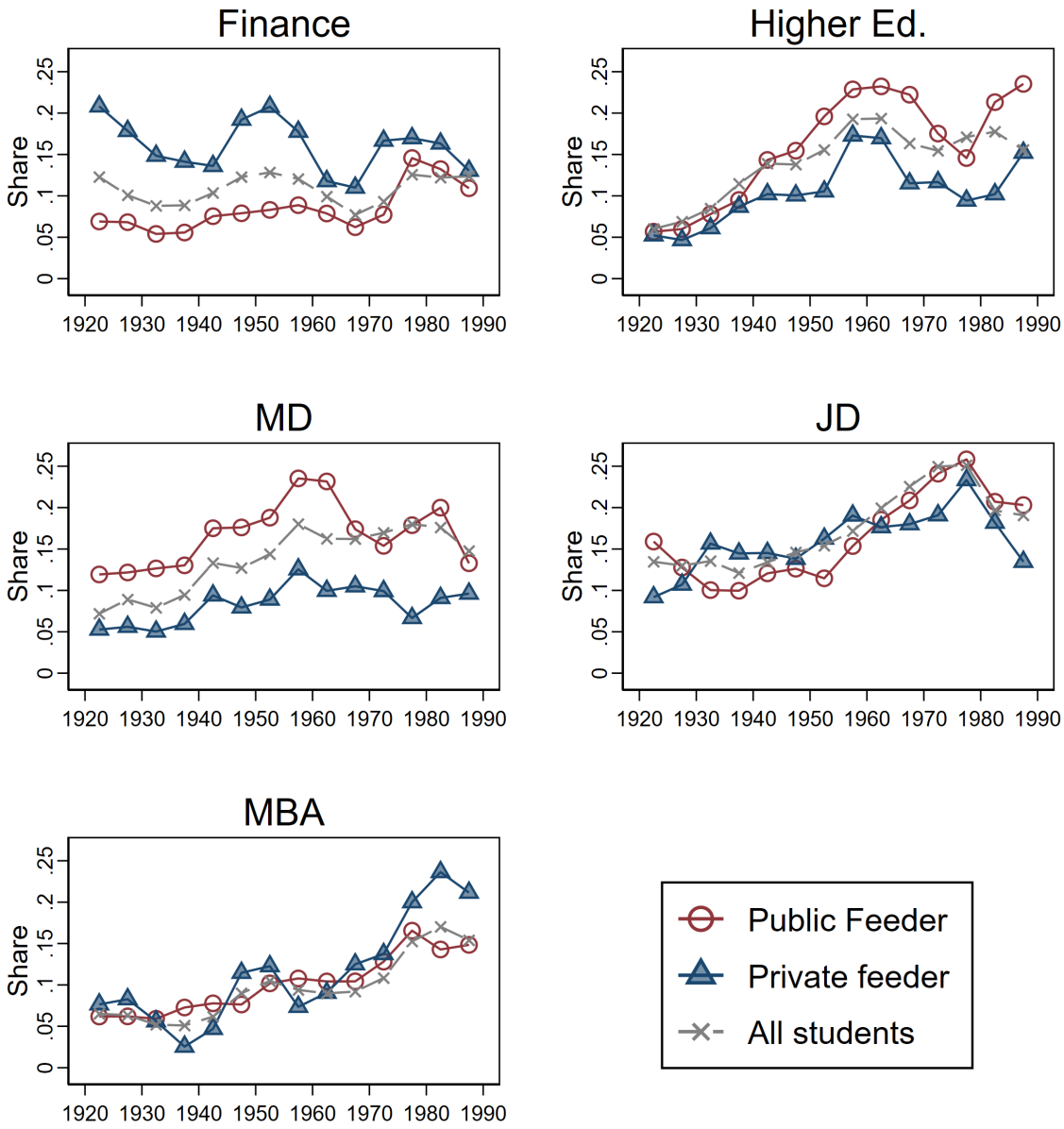


(c) High honors (*magna* or *summa*) by high school type



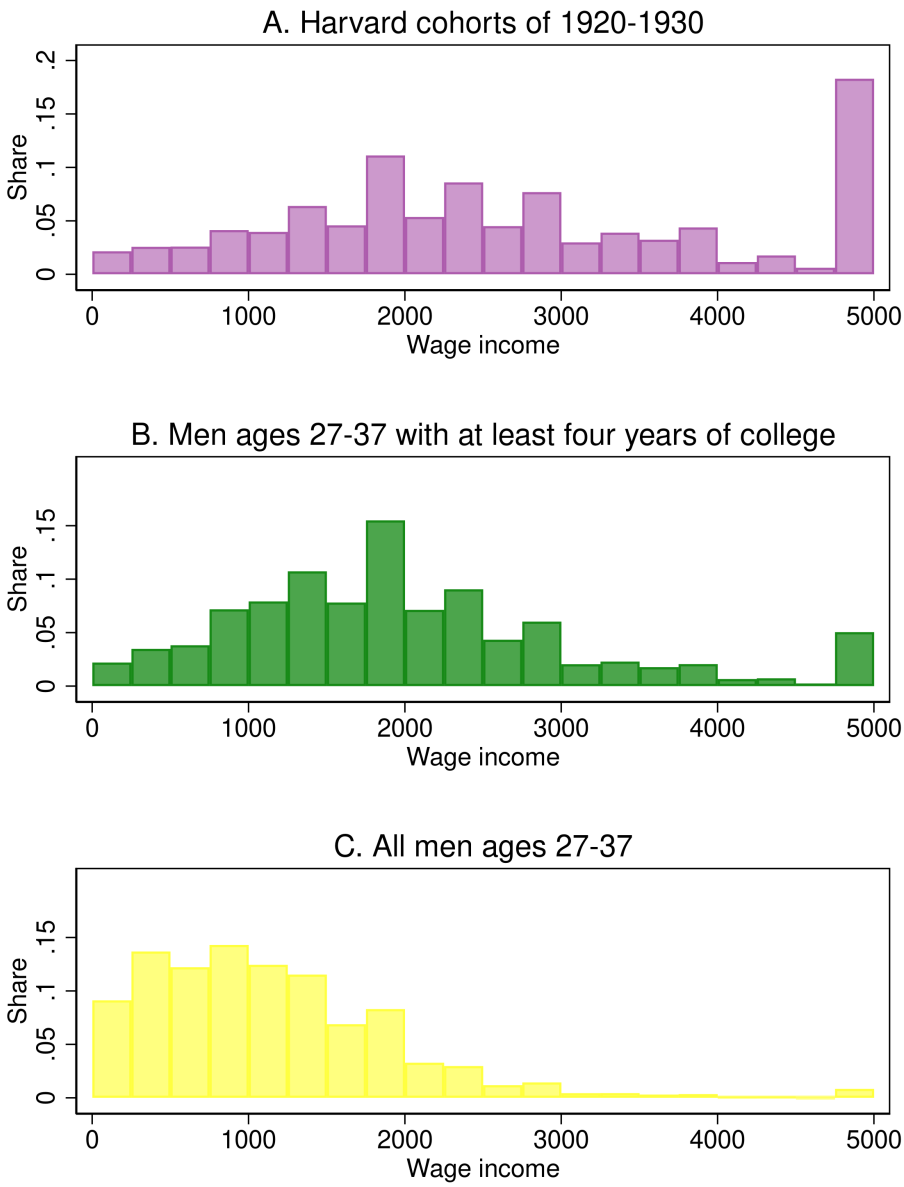
Horizontal axis in all panels is graduating class year. Panel A: share of students with different high school classifications. High school classifications are available for all class years. Points are averages within left-justified five-year bins. For example, the 1925 point includes data from 1925-1929. 1920 point includes 1924 only. Panel B: shares of students graduating with different honors designations. The Boston Globe series is from Healy (2001). Panel C: share of students with "high honors"– either magna or summa cum laude– for all students as well as separately for private feeder and public feeder students. Sample is male students. Panels B and C use data from Class Reports that is available for each graduating class from 1924 through 1939 and then at five year intervals starting in 1940. Points are 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. See section 6.1 for details.

Figure 12: Long-run trends in career outcomes by high school background



Share of students in graduating class in listed occupation or with listed degree type. Horizontal axis in all panels is graduating class year. All panels present rates for the full sample as well as split by high school type. Sample is male students. "JD" panel includes both LLB and JD degrees. All use data from Class Reports that is available for each graduating class from 1924 through 1939 and then at five year intervals starting in 1940. Points are 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. See section 6.1 for details.

Figure 13: Wage income of Harvard students and similarly aged men



Distributions of wage income among men reporting non-zero wages in the 1940 census. Panel A: Harvard students in 1920–1930 entering cohorts, or approximately ages 27–37 in 1940. Panel B: 27–37 year-old men with at least four years of college. Panel C: All 27–37 year-old men. Wage income is topcoded at $5000.

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	0.893	Cohorts 1920-35	13820
Class Report occupation	0.779	Cohorts 1920-35	13820
1940 Census	0.635	Cohorts 1920-30	8851
Pre-Harvard census	0.609	Cohorts 1920-33	11800
B. Match rates conditional on Class Report availability			
Class Report occupation	0.872	Cohorts 1920-35 in Class Report	12344
1940 Census	0.686	Cohorts 1920-30 in Class Report	8193
Pre-Harvard Census	0.664	Cohorts 1920-33 in Class Report	10815

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	All non-private	Randomized
A. Demographics				
Have high school data	0.984	1.000	0.979	0.983
Any private high school	0.423	1.000	0.235	0.536
Private feeder	0.241	1.000	0.000	0.313
Any public feeder	0.266	0.000	0.351	0.132
From MA	0.509	0.455	0.526	0.387
From NY	0.172	0.254	0.146	0.218
Have Harvard father	0.071	0.143	0.047	0.088
Have Harvard brother	0.204	0.287	0.177	0.230
B. Census childhood household demographics				
Have Census pre-Harvard	0.609	0.699	0.581	0.640
N servants	0.937	1.880	0.621	1.188
S or E Eur. immg. gen. 1-2	0.098	0.010	0.127	0.051
Have father's occupation	0.689	0.650	0.701	0.680
Father's occupation: Doctor	0.101	0.128	0.094	0.120
Father's Occupation: Lawyer	0.109	0.152	0.097	0.136
C. First-year campus location				
Have address data	0.950	0.995	0.936	1.000
Live on campus	0.800	0.961	0.749	1.000
Have room attributes	0.712	0.906	0.650	1.000
Room price per occupant	209.2	234.3	198.1	213.0
Peer neighborhood price	213.9	227.9	207.5	213.9
25th pctile neighborhood rank	0.256	0.359	0.217	0.256
75th pctile neighborhood rank	0.711	0.768	0.691	0.711
D. Academic class rank groups				
Rank group 1	0.018	0.008	0.021	0.018
Rank group 2	0.069	0.040	0.077	0.064
Rank group 3	0.155	0.105	0.170	0.151
Rank group 4	0.242	0.217	0.250	0.241
Rank group 5	0.376	0.437	0.356	0.376
Rank group 6	0.141	0.193	0.125	0.149
Class rank year 1	4.311	4.615	4.217	4.338
Not ranked year 1	0.209	0.216	0.207	0.209
E. First-year activities				
Have any activity	0.526	0.707	0.468	0.605
N activities	1.046	1.671	0.847	1.266
Activity leadership position	0.064	0.134	0.042	0.084
Sports	0.367	0.548	0.310	0.430
Social	0.083	0.197	0.047	0.108
Music	0.133	0.163	0.123	0.156
Other activities	0.174	0.249	0.150	0.211
First-year activity index	0.000	0.532	-0.169	0.158
F. Upper-year social clubs				
Hasty Pudding Inst. 1770	0.153	0.393	0.077	0.209
Selective final club	0.070	0.214	0.025	0.095
Any final club	0.136	0.369	0.062	0.185
Upper-year club index	-0.000	0.721	-0.229	0.152
N	14383	3466	10917	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. See main text for more on high school groups. 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	All non-private	Randomized
A. Adult associations				
Any social club	0.343	0.449	0.309	0.381
Country club	0.242	0.362	0.203	0.287
Gentleman's club	0.113	0.209	0.082	0.138
Fraternal order	0.103	0.078	0.111	0.091
Any honor/prof group	0.361	0.306	0.379	0.355
Prof. Association	0.337	0.278	0.356	0.330
Honor society	0.074	0.064	0.077	0.072
Adult association index	-0.000	0.407	-0.133	0.115
B. Occupations				
Have occupation	0.872	0.858	0.876	0.865
Finance	0.102	0.177	0.078	0.120
Accounting	0.108	0.118	0.105	0.105
Doctor	0.091	0.060	0.101	0.081
Law	0.122	0.099	0.129	0.123
Higher ed.	0.070	0.048	0.077	0.072
Teach	0.081	0.072	0.084	0.075
Government	0.034	0.029	0.035	0.034
Art/pub	0.074	0.074	0.073	0.076
Senior management	0.213	0.236	0.205	0.229
Low management	0.118	0.105	0.122	0.115
Retail	0.136	0.127	0.139	0.130
Occupation index	-0.000	0.277	-0.088	0.065
C. Adult census				
In school	0.030	0.026	0.032	0.030
In labor force	0.953	0.951	0.954	0.947
Wage income	2747	2960	2680	2853
Has wage income	0.726	0.710	0.731	0.724
Non-wage inc. 50+	0.503	0.646	0.457	0.556
Wage inc. 5000+	0.175	0.247	0.153	0.210
N	14383	3466	10917	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. See main text for more on high school groups. 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

	Has earnings	Earnings	Topcoded	Non-wage
A. Baseline				
Private feeder	0.010	293	0.091	0.178
	(0.014)	(56)	(0.016)	(0.017)
Class rank	0.001	35	-0.009	0.002
	(0.005)	(20)	(0.006)	(0.006)
Sample mean	0.473	2770	0.182	0.511
N	7097	3360	3360	4532
B. Add most elite final clubs				
Private feeder	0.014	145	0.047	0.144
	(0.015)	(58)	(0.017)	(0.018)
Class rank	0.000	46	-0.006	0.005
	(0.005)	(20)	(0.005)	(0.006)
Selective final club	-0.017	762	0.230	0.173
	(0.024)	(94)	(0.032)	(0.027)
Sample mean	0.473	2770	0.182	0.511
N	7097	3360	3360	4532
C. Private feeders with HS FEs, legacy indicators				
Class rank	0.005	-9	-0.025	0.024
	(0.012)	(45)	(0.013)	(0.013)
Selective final club	-0.055	640	0.185	0.075
	(0.032)	(130)	(0.041)	(0.035)
Have Harvard father	0.012	-39	-0.002	0.036
	(0.031)	(119)	(0.036)	(0.034)
Sample mean	0.499	2975	0.252	0.646
N	1635	816	816	1103
D. Within family				
Class rank	0.009	-1	-0.004	-0.007
	(0.022)	(111)	(0.033)	(0.032)
Selective final club	0.022	1312	0.384	-0.182
	(0.100)	(631)	(0.148)	(0.142)
Sample mean	0.505	2968	0.239	0.618
N	996	285	285	505
E. Within Hasty Pudding (approximate applicant pool)				
Private feeder	-0.052	85	0.033	0.123
	(0.032)	(126)	(0.041)	(0.037)
Class rank	0.001	60	0.015	0.013
	(0.015)	(63)	(0.019)	(0.016)
Selective final club	-0.069	453	0.135	0.129
	(0.031)	(120)	(0.040)	(0.035)
Sample mean	0.500	3354	0.360	0.676
N	1123	561	561	749

Estimates of equation 1. Columns are outcome variables, rows are regressors. Each panel is a different specification. All specifications restrict to students in the 1920–30 entering cohorts who are matched to records of first-year academic rank group. Columns 2 and 3 additionally restrict to students who report wage earnings in the 1940 Census, and column 4 to students matched to the 1940 Census (regardless of whether they report wage earnings). "Has Earnings" is an indicator equal to one if a student is matched to the 1940 Census and reports wage earnings. "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. "Non-wage" is an indicator for whether a student reports at least $50 of non-wage earnings. The sign on class rank groups is reversed so that coefficients reflect the effect of a one rank group improvement in academic performance. All specifications include cohort fixed effects (not reported). Panel C restricts the sample to students who attended private feeder high schools and includes fixed effects for each high school. Panel D restricts the to students from families where multiple brothers attended Harvard during our sample period, and includes family fixed effects. Panel E includes only students who are members of the Hasty Pudding club. See text for details. Robust standard errors in parentheses. Standard errors for model D 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)	-0.033	(0.039)	-0.058	(0.032)
Private feeder high school	0.294	(0.021)	0.022	(0.035)	–	–
Have HS code	0.295	(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	(0.012)	0.005	(0.020)	0.016	(0.022)
Have Harvard brother	0.110	(0.018)	0.010	(0.028)	-0.005	(0.031)
S or E Eur. immg. gen. 1-2	-0.052	(0.013)	-0.003	(0.018)	0.001	(0.020)
Have father's occupation	-0.065	(0.028)	0.004	(0.050)	-0.024	(0.054)
Father's occupation: Doctor	0.044	(0.024)	0.047	(0.041)	0.077	(0.046)
Father's Occupation: Lawyer	0.086	(0.025)	-0.039	(0.046)	-0.047	(0.054)
Joint balance test [p-value]	[0.000]		[0.855]		[0.385]	
B. Links to other data sources						
Class report link	-0.000	(0.011)	-0.017	(0.018)	-0.013	(0.019)
Have occupation	-0.004	(0.015)	-0.023	(0.024)	-0.009	(0.027)
1940 Census	0.027	(0.014)	0.026	(0.022)	0.031	(0.026)
C. Peer and neighborhood attributes						
Private HS peer share	0.369	(0.024)	0.293	(0.032)	0.267	(0.034)
Feeder HS peer share	0.271	(0.021)	0.220	(0.028)	0.200	(0.032)
Legacy share	0.062	(0.012)	0.063	(0.016)	0.063	(0.018)
E Eur. peer share	-0.047	(0.014)	-0.050	(0.015)	-0.050	(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	0.385	(0.057)	0.260	(0.072)	0.153	(0.075)
Peer activity index	0.306	(0.049)	0.304	(0.065)	0.297	(0.072)
Peer mean activity count	0.302	(0.066)	0.292	(0.085)	0.335	(0.093)
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	0.388	(0.051)	0.499	(0.068)	0.439	(0.076)
Peer mean finance	0.112	(0.015)	0.147	(0.021)	0.134	(0.023)
Peer mean wage inc.	365.9	(95.9)	425.9	(132.9)	353.1	(143.9)
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	0.820
	(0.034)	(0.057)	(0.043)	
N activities	0.282	0.656	0.105	0.027
	(0.110)	(0.222)	(0.124)	
Activity leadership position	0.043	0.099	0.012	0.070
	(0.020)	(0.044)	(0.021)	
Social	0.049	0.184	-0.028	0.000
	(0.024)	(0.057)	(0.023)	
Sports	0.028	0.085	-0.007	0.233
	(0.035)	(0.064)	(0.043)	
Music	0.047	0.094	0.030	0.295
	(0.027)	(0.052)	(0.033)	
Other activities	0.021	-0.027	0.059	0.184
	(0.031)	(0.058)	(0.035)	
First-year activity index	0.205	0.480	0.027	0.017
	(0.081)	(0.170)	(0.088)	
N	9343	2828	6367	
B. Upper-year social clubs				
Selective final club	0.065	0.168	0.001	0.004
	(0.021)	(0.056)	(0.017)	
Less selective final club	-0.028	-0.074	-0.007	0.208
	(0.020)	(0.050)	(0.021)	
Hasty Pudding Inst. 1770	0.019	0.106	-0.018	0.070
	(0.029)	(0.062)	(0.030)	
Upper-year club index	0.122	0.358	-0.021	0.047
	(0.078)	(0.176)	(0.078)	
N	8589	2606	5845	
C. First-year academic rank				
Rank group 1	-0.002	-0.012	0.002	0.504
	(0.011)	(0.015)	(0.015)	
Rank group 2	0.018	0.017	0.014	0.936
	(0.016)	(0.025)	(0.021)	
Rank group 3	0.022	0.006	0.029	0.637
	(0.024)	(0.039)	(0.030)	
Rank group 4	-0.007	0.093	-0.053	0.010
	(0.029)	(0.046)	(0.036)	
Rank group 5	0.023	-0.068	0.075	0.045
	(0.032)	(0.061)	(0.039)	
Rank group 6	0.021	0.016	0.006	0.854
	(0.023)	(0.046)	(0.027)	
Rank listed year 1	0.075	0.052	0.073	0.731
	(0.029)	(0.051)	(0.035)	
Class rank year 1	0.016	0.098	0.017	0.710
	(0.104)	(0.180)	(0.129)	
N	7020	2085	4790	

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. 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	0.063	0.238	-0.006	0.003
	(0.036)	(0.067)	(0.046)	
Country club	0.069	0.216	0.004	0.006
	(0.033)	(0.066)	(0.040)	
Gentleman's club	0.015	0.060	-0.005	0.305
	(0.026)	(0.057)	(0.029)	
Fraternal order	0.002	0.024	-0.003	0.570
	(0.022)	(0.037)	(0.029)	
Any honor/prof group	-0.012	0.037	-0.023	0.412
	(0.032)	(0.059)	(0.040)	
Prof. Association	-0.022	0.030	-0.039	0.340
	(0.031)	(0.059)	(0.040)	
Honor society	-0.014	0.025	-0.019	0.301
	(0.019)	(0.035)	(0.023)	
Adult association index	0.126	0.337	0.021	0.104
	(0.078)	(0.173)	(0.088)	
N	8046	2450	5449	
B. Occupation choice				
Finance	0.020	0.150	-0.047	0.002
	(0.025)	(0.059)	(0.028)	
Accounting	0.029	0.065	0.008	0.296
	(0.024)	(0.048)	(0.029)	
Doctor	0.010	-0.023	0.031	0.223
	(0.022)	(0.034)	(0.029)	
Law	-0.026	-0.009	-0.029	0.716
	(0.027)	(0.044)	(0.035)	
Higher ed.	-0.017	-0.026	-0.021	0.919
	(0.021)	(0.034)	(0.028)	
Teach	-0.017	-0.018	-0.011	0.885
	(0.022)	(0.038)	(0.028)	
Government	0.005	0.020	0.005	0.669
	(0.015)	(0.031)	(0.017)	
Art/pub	-0.012	-0.017	-0.001	0.700
	(0.022)	(0.035)	(0.028)	
Senior management	0.031	0.064	0.011	0.473
	(0.033)	(0.061)	(0.041)	
Low management	0.065	0.041	0.070	0.583
	(0.026)	(0.042)	(0.033)	
Retail	0.005	0.030	-0.010	0.508
	(0.028)	(0.049)	(0.037)	
Occupation index	0.050	0.470	-0.173	0.002
	(0.083)	(0.183)	(0.095)	
N	6928	2075	4712	
C. Adult income				
Wage income	6.7	-368.1	217.4	0.169
	(189.8)	(370.5)	(234.5)	
Wage inc. 5000+	0.022	-0.020	0.073	0.417
	(0.051)	(0.099)	(0.062)	
Non-wage inc. 50+	0.044	0.094	0.025	0.556
	(0.053)	(0.099)	(0.065)	
Class Report wage index	31.1	130.2	-17.6	0.011
	(27.2)	(49.2)	(34.1)	
N	2361	685	1587	

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. "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. Sample for this variable is as in Panel B. 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.

Table 8: Peer neighborhood effects on outcome-by-grade interactions for private feeder students

	Rank group 1-4		Rank group 5-6 or unlisted	
A. Upper-year social clubs				
Hasty Pudding Inst. 1770	0.091	(0.050)	-0.002	(0.058)
Selective final club	0.129	(0.039)	0.054	(0.051)
N	2460		2460	
B. Occupation choice				
Finance	0.119	(0.035)	0.031	(0.049)
Doctor	-0.045	(0.021)	0.022	(0.027)
Law	-0.004	(0.032)	-0.005	(0.033)
Higher ed.	-0.049	(0.026)	0.023	(0.020)
N	2075		2075	
C. Adult associations				
Any social club	0.162	(0.047)	0.075	(0.062)
Prof. Association	0.051	(0.043)	-0.020	(0.053)
N	2682		2682	

Coefficients on peer neighborhood price rank from regressions of the form given in equation 2. Outcomes are interactions between the social or career outcome listed in the row and indicators for academic performance levels listed in the column. All specifications include randomization block and large feeder dummies; see section 5.1 for details. See section 6.3 for description of outcome variables. The sample is private feeder students only. Standard errors clustered at peer neighborhood level are reported in the columns to the right of coefficients.