The Distribution of and Returns to Social Success at Elite Universities

Valerie Michelman
University of Chicago

Joseph Price
Brigham Young University & NBER

Seth D. Zimmerman
Yale University & NBER

September 28, 2020

Abstract

This paper studies social success at elite universities: who achieves it, how much it matters for students' careers, and whether policies that increase interaction between rich and poor students can integrate the social groups that define it. Our setting is Harvard University in the 1920s and 1930s, where students compete for membership in exclusive social organizations known as final clubs. We combine within-family and room-randomization research designs with new archival and Census records documenting students' college lives and career outcomes. We find that students from prestigious private high schools perform better socially but worse academically than others. This is important because academic success does not predict earnings, but social success does: members of selective final clubs earn 32% more than other students, and are more likely to work in finance and to join country clubs as adults, both characteristic of the era's elite. The social success premium persists after conditioning on high school, legacy status, and even family. Leveraging a scaled residential integration policy, we show that random assignment to high-status peers raises rates of final club membership, but that overall effects are driven entirely by large gains for private school students. Residential assignment matters for long-run outcomes: more than 25 years later, a 50-percentile shift in residential peer group status raises the rate at which private school students work in finance by 37.1% and their membership in adult social clubs by 23.0%. We conclude that the social success premium in the elite labor market is large, and that its distribution depends on social interactions, but that the inequitable distribution of access to high-status social groups resists even vigorous attempts to promote cross-group cohesion.

We thank Nathan Kyn, Jordan Rosenthal-Kay, Mckay Jensen, Katherine Stevens, Merrill Warnick, and Jacob van Leeuwen for research assistance. We thank Joseph Altonji, Christopher Neilson and seminar participants at the University of Chicago for valuable comments. 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. All errors are our own. Michelman: vmichelman@uchicago.edu. Price: joseph_price@byu.edu. Zimmerman: seth.zimmerman@yale.edu. Online Appendix available at https://faculty.som.yale.edu/sethzimmerman/mpz_appendix/.

1 Introduction

Economic elites are disproportionately drawn from a small number of social, educational, and business institutions, and there is growing evidence that ties forged at these institutions affect performance in and access to top jobs.[1] This raises concerns about elite "closure" – the idea that outsiders may have difficulty reaching top rungs of the social ladder if they do not share the traits or experiences of incumbents– and suggests that access to elite social groups may be an important ingredient for upward mobility to top positions.[2] Relative to other elite institutions like country clubs or corporate boards, elite universities have a potentially important role to play in broadening access to these social groups because identifying and expanding career opportunities for talented but disadvantaged students has long been central to their mission.[3] Attending an elite university is a chance for less-advantaged students to take demanding classes and to join the "old boys'" clubs upon which career advancement may depend. The extent to which elite universities integrate downstream institutions hinges on which students succeed socially and academically once they arrive, and on how different kinds of college success affect long-run outcomes. There is little quantitative evidence on these questions, likely due to difficulty accessing data on the college lives and long-run careers of students at top private universities.

This paper studies the distribution and long-run consequences of social success at elite universities using newly assembled data on students entering Harvard between 1919 and 1935. We link archival records of students' social and academic lives while at college to the 1940 Census and to extensive biographical reports compiled more than 25 years after college entry to paint a detailed picture of who excels in different aspects of college life and how this maps to long-run outcomes. We use two complementary research designs: a within-family analysis that effectively compares the smart brother to the popular brother in families that sent multiple sons to Harvard, and a room randomization design in which Harvard administrators seeking to integrate campus social life assigned students to dormitories with widely varying residential peer groups. Our central findings are a) that high-status students are more socially successful than low-status students but do worse in class, b) that the labor market premium for social success is larger than the premium for academic success, which is near zero, and c) that exposure to high-status peers affects long-run social and career outcomes for students from high-status families, but does not integrate exclusive social groups or benefit students from less privileged backgrounds.

Though Harvard in the 1900s was an all male and almost all white institution, it shared at least one important similarity with Harvard in the 2000s: the tension between an institutional

[1] Backgrounds of elites: Miller (1949, 1950), Useem and Karabel (1986), Temin (1999), Cappelli and Hamori (2004), Reeves et al. (2017). Job performance/access: 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.

[3] See Hoxby and Avery (2013), Hoxby et al. (2013),and Chetty et al. (2020) for recent evidence and Karabel (2006) for historical context with respect to both universities and private high schools.

commitment to integrating campus and a campus social life characterized by exclusive student-run organizations. 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."[4] 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). Despite the opinions of Harvard presidents, the same set of exclusive social organizations known as final clubs that comprised the "be-alls and end-alls of Harvard social existence" in the early 20th century (Amory, 1947) are described as the "apex of social life at Harvard" in the New York Times in 2016 (Nir, 2016) and as "impossible to escape–even for those who wish to have nothing to do with them" in a 2017 Harvard faculty report that unsuccessfully called for their abolition (CUSGGO 2017). The final clubs and the policies adopted to offset their influence on campus life form the center of our empirical analysis.

Archival data allow us to overcome challenges inherent in studying elite private institutions, which do not typically make administrative records available to researchers, and to incorporate informal outcomes like club membership that are not present in administrative datasets. We define our sample universe using the freshman yearbooks for classes entering between 1919 and 1935. Published each spring, entries for each student give their name, high school, home address, campus address, and college social activities. We measure academic performance using class rank records that divide students into six groups based on grades. We collect membership in final clubs from rosters published by the student council and reports in Class Albums. For long run outcomes, we use 25th Reunion Class Reports, in which students and class officers compile biographical records of each member of the entering class. These include standardized fields for topics like education, family life, occupation, and social club memberships, as well as narrative accounts. Using data on parents' names and places of birth from the class reports, we link our archival data to decennial Census records from 1910 through 1940, leveraging wiki-style family tree records from the FamilySearch platform to obtain precision and high match rates at reduced researcher cost (Price et al., 2019).

The first part of the paper describes how academic and social success at college vary with family background. 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 the sample.

Students from high-status private feeder schools are more socially engaged than other stu-

[4]Eliot: Karabel (2006): p. 45. Lowell: Karabel (2006), p. 47.

dents, but do worse in school. On average, private feeder students participate in 1.67 extracurricular 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. We take membership in selective upper-year final clubs as our main measure of social success at college, and here differences 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%).

These differences matter because the social success premium in the labor market is large, while the academic success premium is near zero. Members of selective final clubs in the lowest academic rank group earn 26% more than non-members in the top two academic rank groups, and are 3.2 times more likely to have topcoded earnings, corresponding to the top 0.8 percent of the population distribution. Mean earnings and topcode rates do not change with academic performance. Once one conditions on final club membership, the private school earnings premium goes away: private feeder students who are not final club members have similar earnings to other non-members.

The social success premium is not driven by selection into social success on the basis of high school type, high school identity, Harvard legacy status, or even family. 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. In contrast, academic rank is not related to either of these outcomes. In addition to affecting 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, this evidence indicates that while students from rich families tend to be socially successful at college, social success is distinct from family wealth in the way it affects long-run outcomes, and in fact is a key mediator of long-run gaps in outcomes between students from rich and less-rich backgrounds.

We next ask whether the policies colleges adopt in pursuit of social cohesion on campus affect the distribution of social sucess and long-run career outcomes. This is important for two reasons. First, it provides a test of whether social success and the long-run outcomes that flow from it depend on college policy at all. An alternative hypothesis not ruled out by our within-family analysis is that both social success and the social success premium reflect a return to pre-college social skills. Second, it helps understand whether policies adopted to promote equitable

outcomes and cross-group cohesion succeed in doing so when the stakes are high.

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 price ranges and room types. Students could apply together with roommates, so roommate assignments were not random, but conditional on room price and occupancy, room assignments were. 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.[5] We validate the room randomization design by showing that controls for randomization blocks defined by room type and price eliminate the strong cross-sectional relationship between neighborhood price rank and own characteristics. Further, wide variation in peer neighborhood price rank and peer attributes persists within randomization blocks and strongly predict observable peer attributes. A 50 percentile increase in neighborhood price rank– well within the observed support of within-block variation– raises the peer private feeder share by 9.6 percentage points (30.6% of the mean).

We find that exposure to high-status peers *does* affect which high-status students achieve social success, but *does not* integrate exclusive organizations. A 50 percentile shift in the room price distribution raises membership in selective final clubs by 2.9 percentage points in the full sample (30.3% of the sample mean), with the entirety of the gain driven by a 7.3 percentage point (33.0%) increase for private feeder students. This builds on differential effects we observe starting in students' first year at college. A 50-percentile increase in neighborhood mean room price raises the count of activities by 0.15 overall and 0.34 for private feeder students, equal to 11.6% and 19.9% of the population and private-feeder sample means, respectively. However, effects for students not from private feeders are economically small and not statistically different from zero. Effects are the biggest and cross-group differences the largest for first-year leadership roles, where baseline gaps in participation by high school type are the largest. In contrast to social outcomes, exposure to high-status peers has little effect on academic performance.

The effects of college peers persist over the long run. 25 years after graduation, assignment to a higher-priced peer neighborhood causes students from private feeder schools to pursue finance careers and obtain membership in exclusive adult social groups. A fifty percentile change in neighborhood price rank raises the share of private feeder students in finance by 6.8 percentage points, 37.1% of the group mean and 74.5% of the gap between private feeder and other students. This change is offset by small declines in higher education, medical, and legal careers. The same change in peer neighborhood price raises the chances private feeder students participate in adult

[5]This design does *not* rely on random fluctuations in the attributes of students assigned to residential or classroom units. Instead, we study assignment to neighborhoods with systematic differences driven by dorm layouts. Our approach is not subject to the Angrist (2014) weak instruments critique of peer effect estimation. It has more in common with the Moving to Opportunity experiment (Katz et al., 2001; Ludwig et al., 2013; Chetty et al., 2016), which moved low-income families to high-income neighborhoods. In our setting, all movement takes place on the Harvard campus, and both high- and low-income students are randomly assigned to peer groups.

social organizations by 10.7 percentage points (23.0% of the sample mean and 87.9% percent of the gap in participation rate between private feeder and other students. This effect is driven largely by increases in country club membership. For students not from private feeder schools, whose college social outcomes were not affected by residential assignment, we observe no effects on adult social participation or occupation. *Among* private feeder students, it is the academic high achievers whose college and long-run outcomes are most affected.

Our central findings are that there is a large long-run premium to social success at elite universities accruing mostly to students from rich families, and that university policies can affect who achieves this success but not necessarily in ways that make access to it more equitable. These results contribute to several strands of literature. First, our finding that social success is more important than academic success for career outcomes provides a look inside the production function at top universities and in doing so addresses a puzzle in the literature on returns to college selectivity. Several studies find that gaining admission to highly selective programs raises incomes, but does not seem to be correlated with learning or completion outcomes (Zimmerman, 2019; Sekhri, 2020). Other studies show that cross-institution income differences partially reflect causal effects, but are not correlated with academic selectivity, especially for richer students (Dale and Krueger, 2002, 2014; Chetty et al., 2020).[6] A simple explanation is that social success is a key element of the earnings production function for students at elite universities, while academic success is not. If this is true, the success of policies that diversify selective universities may depend as much on social integration (Rivera, 2016; Jack, 2019) as on academic match (Rothstein and Yoon, 2008b,a; Arcidiacono et al., 2016; Arcidiacono and Lovenheim, 2016).

Second, our finding that cross-group exposure does not integrate exclusive social groups or careers suggests that policymakers should be cautious in extrapolating results from lower-stakes studies of intergroup exposure and social cohesion to high-stakes settings. A growing literature in economics and psychology uses lab-in-the-field techniques to provide proof on concept that cooperative intergroup interactions can increase low-stakes cross-group socializing and survey measures of intergroup cohesion (Rao, 2019; Lowe, 2020; Mousa, 2020).[7] Carrell et al. (2019) show that assignment to residential groups with high-ability black peers raises the probability white students at the US Air Force Academy will choose black roommates the following year. This is an important paper because roommate choice has a concrete effect on students' lives in at least the medium run. However, it does not have the status connotation or long-run career implications that club membership does in our setting. Our contrasting findings may be because residential integration at Harvard does not foster collaborative contact, or because collaborative contact does not matter as much when the stakes are high. Our results are consistent with evidence from Zimmerman (2019) that peers from high-SES backgrounds at top business programs often serve

[6] This literature is still evolving; see also Hoxby (2018) and Mountjoy and Hickman (2020). Abdulkadiroğlu et al. (2014) report small effects of academic peer quality at the high school level.

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

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 room randomization designs to explore how school peers affect outcomes measured while students are in school (Zimmerman, 2003; Stinebrickner and Stinebrickner, 2006; Lyle, 2007; Mehta et al., 2019; Jones and Kofoed, 2020).[8] A limitation of this research is that it does not tell us whether the observed academic, social, and preference shifts matter in the long run. Our results show that peers have long-lasting impacts on life trajectories. Turning to a direct comparison of short-run results, our findings of null effects on academic outcomes are consistent with Zárate (2019), but in some tension with Sacerdote (2001) and Carrell et al. (2009). That exposure to high-status peers augments pre-existing differences is consistent with findings on homophily in college peer groups (Marmaros and Sacerdote, 2006; Mayer and Puller, 2008; Carrell et al., 2013).[9]

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 such as finance, and impermeability of group boundaries to efforts at integration that we observe may provide some 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. They may also help explain the social importance of other selective social organizations in higher education, such as fraternities and sororities, which claim large shares of business and political leaders as members (Marmaros and Sacerdote, 2002; DeSantis, 2007; Popov and Bernhardt, 2012). Many papers describe the relationship between membership in social groups and business success (Taussig and Joslyn, 1932; Miller, 1949, 1950; Warner and Abegglen, 1955; Useem and Karabel, 1986), or study how connections between group members influence business or policy outcomes (Cohen et al., 2008; Fracassi and Tate, 2012; Shue, 2013; Fracassi, 2017; Xu, 2018). Our contribution is to compare outcomes for group members to an informative comparison group of non-members to see how membership helps people obtain positions of influence in the first place.

[8] See Sacerdote (2011) for a review.

[9] 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.

2 Institutions

2.1 Admissions criteria and campus diversity

We study the determinants and long-run effects of social success for Harvard students starting school in the 1920s and 1930s. We choose this setting because it provides a useful vantage point on social sorting by baseline socioeconomic status. Our claims do not hinge upon the specifics of the Harvard environment remaining unchanged over the past 100 years. However, many of the social institutions and policy choices that shaped Harvard in the 1920s and 1930s persist through the present in strikingly similar form, and resemble those at other highly selective US colleges. We focus our discussion on two features of the institutional environment. The first is Harvard administrators' effort to increase the economic diversity of on-campus interactions through admissions and housing policies while maintaining the university's appeal to private school applicants. The second is the important role of exclusive private clubs in campus social life, and the extent to which these clubs tended to push back against administrators' goals for social integration. We draw on primary sources, histories of Harvard admissions policy (Karabel, 2006; Synott, 1979), and Amory's (1947) account of the Harvard club system.

Harvard administrators in the 1920s and 1930s took both academic and social considerations into account when designing admissions policy. In practice, this meant trying to bring more students from public schools and other schools beyond a small set of traditional private feeder schools without admitting too many Jewish students, who administrators viewed as socially undesirable and likely to make recruiting private school students more difficult.[10] One policy of this type was the "Top Seventh" plan, adopted in 1923, which guaranteed admission to any high school student graduating in the top seventh of his class (RPTHC 1922-23, p. 290). In 1926, unable to achieve their goals for the mix of students on campus with purely academic admissions requirements, Harvard administrators adopted non-academic admissions criteria for the first time (RPTHC 1925-1926, p. 298). This approach persists through present; the 2020 Harvard admissions office evaluates applicants on the basis of questions like "would other students want to room with you, share a meal, be in a seminar together, be teammates, or collaborate in a closely-knit extracurricular group?"[11] Arcidiacono et al. (2019a,b, 2020) describe contemporary Harvard admissions institutions in detail.[12]

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

[11] Harvard College Admissions and Financial Aid. `https://college.harvard.edu/admissions/apply/what-we-look`. Accessed 4/28/2020.

[12] Admissions rates for high-status vs. low-status students provide an informative point of cross-time comparison. Even as admission rates have fallen dramatically over time, students from privileged backgrounds retain an admissions advantage. Karabel reports that in 1940 only one out of 77 applicants from the elite "St. Grottlesex" private schools was rejected, compared to 14 out of 59 applicants from Boston Latin, an academically rigorous public school with historical ties to Harvard (p. 174). Arcidiacono et al. (2019b) reports that applicants who are Harvard legacies, "development" cases whose admission may lead to substantial donations, or children of Harvard faculty are a) much

2.2 Residential life as a social lever

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 specifically for first-year students in 1914. In 1930, Dean of Harvard College Alfred Hanford reflected on the first fifteen years of 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. These aims have been realized with remarkable success" (RPHC 1929-1930, p. 100-101).

Reflecting the goal of social integration through housing policy, first-year students were assigned dormitory rooms at random. Randomization worked as follows. Rooms available to freshman were assigned different prices, depending on size, occupancy, and quality. First-year students were asked to fill out a housing application indicating their acceptable prices. 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." 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). This procedure was in place from at least 1922 through 1941. Section 5.1 describes how we use room randomization in our empirical work. Online Appendix B presents detailed documentation of the process.

2.3 Social clubs at Harvard

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

To understand the importance of final clubs to campus life, it is helpful to lay out some of the ways students, reporters, and historians have described them over the past century. Franklin Delano Roosevelt (Harvard class of 1904) remarked that one of "the greatest disappointments of his life" was not being elected to Porcellian, which Amory identifies as the most prestigious of the final clubs and of which his cousin Theodore Roosevelt was a member. Delano Roosevelt instead

less likely to come from disadvantaged backgrounds, b) a minimum of four to five times more likely to be admitted than other students within each each decile of academic achievement, and c) account for 21% of all admitted students and 33% of all admitted white students .

joined the Fly club, which Amory identifies as the third-most prestigious.[13] Amory described final clubs as the "be-alls and end-alls of Harvard social existence," while Nir described them in the *New York Times* in 2016 as "the apex of social life at Harvard." Amory describes Hasty Pudding, a nominally theater-focused sophomore society, as 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 expand 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 unsuccessfully recommending the dissolution of final clubs 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 and institutional mission." Many final 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 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 away three insights from the institutional setting. The first is that selective club membership is a good measure of social success at Harvard. The second is that private school background is an important mediator of opportunity to enter these groups. The third is that room randomization was a lever policymakers used to push back against students' tendency to socialize in homogenous groups. In our empirical work we test claims from the qualitative literature about differences in access to exclusive clubs by student background, limits on the ability of

[13] 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.

residential policies to integrate exclusive clubs, and the relationship between club membership and long run career success. This analysis requires fine classification of clubs and high schools. We rely on Amory to divide final clubs into groups by perceived status, and a combination of Amory and empirical data to identify private feeder high schools.

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 different data sources we use, and how we merge them together. For a more detailed description, as well as additional example images, see Online Appendix B.

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. They consist of individual entries recording students' home addresses, college addresses, high school backgrounds, and activities during the first year of college. They also include activity pages with further details on group membership. We link activity pages by name to student records from the same year to recover full records of activity participation for all students in the Red Books.

Figure 1 and Table 1 report statistics on sample size and the availability of different data types. We observe 14382 individuals in the Red Book universe. As shown in Panel A of Figure 1, entering class size increases over time from 563 in 1919 to 989 in 1935. Panel B shows that almost all students report what high school they attended (98.4%). Except for 1926, when the Red Books did not include the field, almost all students report their address while at college (95.0%).

We measure academic performance using class rankings that Harvard published for all non-graduating students between 1920 and 1930, and then for freshmen only from 1930 and forward. These lists aggregate grades across all courses within the academic year and coarsen them to numerical groups between one and six, with one indicating "Highest Distinction" and six indicating "Low Pass." Students who have incomplete coursework at the time of publication or grades too low to advance are not included on the rank lists, meaning that not showing up on a list is an outcome of interest. We focus our analysis on first-year grades, which we match by name and class year to Red Book records.

We observe student participation in Hasty Pudding and final clubs through two sources. For academic years 1927 and earlier, we rely on lists of club members in the Harvard University Register, published by the student council of Harvard College. For academic years 1928 through 1938, we use student reports of club memberships published in Class Albums (senior

yearbooks).[14] In both cases we merge to Red Book records using name and class year. We code students who do not appear in these records as not being members of these organizations.

We measure long run outcomes using 25th Reunion Class Reports. Class officers compile class reports 25 years following graduation (roughly age 47) using a combination of student self-reports and administrative records. They contain standardized fields for family (including parent's names, spouse, and children, and which family members are also Harvard graduates) birthplace, high school, place of work and occupation, offices held, club memberships, and other honors. They also contain extended narrative accounts. We digitize report data and merge it to Red Book records by name within class year. We merge 89.3% of Red Book records for the 1920 through 1935 cohorts to class reports. We use class reports to describe work outcomes and adult social organizations. 77.8% of students in our universe (87.1% of students matched to class reports) have non-missing occupation fields. Panel C of Figure 1 displays rates of data availability for long-run outcomes. Match rates to Class Report work outcomes are steady between 1920 and roughly 1933 before a decline in 1934 and 1935, when the Class Report records become noticably less detailed.

3.2 Census records

We merge our Harvard data to publicly-available Census records from 1910 through 1940 using the record linking methods and FamilySearch genealogical database described in Price et al. (2019). The features of our Harvard data that allow for this merge are records of parents' names and birth location in the Class Reports for cohorts entering 1933 and earlier, as well as knowledge of approximate birth timing inferred from each college cohort. Because we rely on the pre-1933 Class Report for the Census match, we only attempt a match for individuals merged from Red Books to Class Reports in these cohorts. FamilySearch is a wiki-style platform in which 12 million registered users contribute to a shared Family Tree that includes the profiles of over 1.2 billion deceased individuals. We use the crowed-sourced FamilySearch data to obtain the high match rates typically associated with hand-linking (Costa et al., 2020) at a cost more similar to automated methods (Abramitzky et al., 2020). 59.9% of the students in our sample already had a profile on FamilySearch; 69.5% of these profiles were already connected to a 1940 census, and 70.5% were connected to a pre-Harvard census record. We raise our match rates by supplementing the existing FamilySearch data with hand linking techniques and ensure precision by hand-checking existing links. See Online Appendix B.2.5 for details.

We use the 1940 Census to measure post-college outcomes. Because our goal is to capture outcomes for students once they have completed schooling and launched their careers, we exclude

[14] We were in the process of collecting data for academic year 1939 when the Harvard Archives were shut down due to Covid-19. This is the year in which students in the 1935 entering cohort would have reported their final club memberships in their Class Album. We exclude the 1935 cohort from our analysis of final club outcomes.

students entering college after 1930 (and thus graduating after roughly 1934) from the Census outcome sample.[15] We match 63.5% of students in entering cohorts of 1920 through 1930 to 1940 Census records. We use pre-1940 censuses to describe students' pre-college backgrounds. We match 60.9% of students in 1920-1933 cohorts to pre-1940 Census records.[16] For students who match to multiple pre-college census records, we take data from the most recent census. Before 1940, census records do not include measures of income, but do include occupation.

The match rates we achieve are higher than previously-documented match rates of 10-30% for automated census linking (Abramitzky et al., 2020) and 10% for linking patents to censuses (Sarada et al., 2019). However, they are lower than 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). Abstracting from differences in context that may also affect match rates, our approach falls about two thirds of the way towards this gold standard from a 10% base for automated linking.

3.3 Multiple matches and match balance

Our match rates to individual datasets are high. Further, as we discuss in detail below, match rates are not correlated with individual attributes of interest to us descriptively or with attributes of the residential peer neighborhoods to which on-campus students are randomly assigned. However, it is important to note that observation counts decline as we impose sample restrictions and require matches to multiple data sources. This is mainly an issue for analyses that require matches to 1940 Census wage reports *and* other data types. For example, 9342 students live in on-campus rooms subject to random assignment. Of these, 5218 (55.9%) are in cohorts we match to Census outcome records, 3376 (64.7% of 5218) are matched to the Census, and 2445 (72.4% of 3376) report wage records. In short, though we are able to provide rich descriptive analyses of archival and Census outcomes, and quasi-experimental analysis of archival outcomes, our statistical power when conducting quasi-experimental analyses of Census outcomes is limited.

[15] Online Appendix Figure A.3 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. Even the oldest students in our data have not reached the peak of their age-earnings profiles by 1940.

[16] Students in cohorts 1920-1930 are matched to 1910, 1920, and 1930 Census records, while we only match students in cohorts 1931-1933 to 1910 and 1920 Census records to ensure we are capturing features of the students' childhood homes prior to entering Harvard.

3.4 Categorization schemes

3.4.1 Private clubs

Our main measure of social success is membership in a selective upper-year final club. We focus on Amory's list of ten primary final clubs. We observe membership at all of them, and break them down into groups based on prestige following Amory's ordered list. We label Porcellian, A.D., Fly, and Spee as 'selective final clubs.' The remaining clubs are Delphic, Owl, Fox, D.U., Phoenix, and Iroquois. We observe Hasty Pudding membership as well.

3.4.2 High schools

Contemporary accounts emphasize the importance of a small number of private high schools in driving social outcomes. We divide high schools into three categories: private feeder schools, public feeder schools, and other schools. The group of private feeder schools consists of the eight private boarding schools that send the largest number of students to Harvard over the period: Exeter, Andover, Milton, Middlesex, Groton, St. Paul's, St. Mark's, and St. George's. The first seven schools in this list sent more students to Harvard than any other private schools. The eighth, St. George's, sent fewer students than three Boston-area private day schools, but is included because it is part of the "St. Grottlesex" group emphasized in writings from the time (Amory 1947; Karabel 2006).

While our private feeder designation has a strong basis in student count data and the historical record, one could reasonably argue for the inclusion of some private day schools or other boarding schools. To this end, we also identify a broader group of private schools that includes other boarding schools and day schools, and discuss key findings for this group as well. We identify public feeder schools as public schools that send multiple students to Harvard over this period. The public school sending the most students to Harvard by far is Boston Latin, and most public feeders are from the Boston area. Any high school which send 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 public and private school.

3.4.3 Residential peer groups

To describe residential peers, we augment the Red Book records of college address with data on room attributes. First, we use archival floor maps of each freshman dorm to define peer groups based on dorm layouts. The maps allow us to determine whether peer groups are organized "vertically" by stairwell or "horizontally" by floor. For more details on peer group classification, see Online Appendix Section B.3. Second, we use room-specific price and occupancy data printed in the Official Register of Harvard University (1920, 1932-41) to describe the the room

in which each student is housed and attributes for rooms in the peer neighborhood. These data come from pamphlets distributed to incoming students making housing choices.

We describe peer neighborhoods using the average per-occupant price of rooms. For neighborhood p in entering cohort c, we define mean neighborhood room price MP_{cp} as

$$MP_{cp} = \frac{1}{N_{cp}} \sum_{s \in p} price_{scp}$$

where N_{cp} is the number of spaces available in neighborhood p in cohort c, s indexes each available space, and $price_{scp}$ is the price of that space. We then compute RP_{cp} as the rank of MP_{cp} out of all neighborhoods in cohort c, with zero being the lowest ranked room and one being the highest ranked room. In cases where multiple neighborhoods have the same value of MP_{cp} we assign all neighborhoods the rank at the midpoint of the interval.

3.4.4 First-year activities

We use detailed records of students' social activities in their first year to describe of their social lives early in college. We enumerate common activities, and then place activities into coarse groups. Online Appendix Table B.3 lists the activities in each group.

We focus on several summary measures of first-year activity participation: an indicator for participation in any activity, the count of total activities, an indicator for holding a leadership role in an activity (e.g. president of a club, captain of a team) and indicators for participation common activity types including sports (the most common activity type), music, and schoolwide social committees. Examples of social committees include "Regatta Committee" and "finance committee.' The goal of the leadership and social committee categories is to identify student leaders early in their time at Harvard.

To summarize how changes in activity descriptors relate to integration across social groups, we compute the predicted probabilities that an individual comes from a private school background, given their activity participation. We estimate specifications of the form

$$P_i = X_i\beta + \gamma_{c(i)} + e_i \tag{1}$$

where P_i is an indicator variable equal to one if individual i in entry cohort c attended a private feeder high school, X_i is a vector describing the coarse activity indicators, and $\gamma_c(i)$ are cohort fixed effects. To avoid using one's own schooling background as an input to predicted values, we estimate this specification for a given cohort c^* using all cohorts $c \neq c^*$. The X_i consist of indicators for each activity category. To improve prediction accuracy we select coefficients with a Lasso procedure, using the EBIC for model selection. We compute predicted values from these specifications, and then standardize (mean zero, standard deviation one) to facilitate cross-

outcome comparisons. Online Appendix Table A.1 reports estimation results from the prediction procedure estimated using data from all cohorts, and Online Appendix Figure A.2 displays the distribution of the $\hat{P}_i$.[17]

3.4.5 Occupations and adult social organizations

We describe 25-year career outcomes using text descriptions of jobs in the Class Reports. We create indicators for 11 coarse job types, and identify a set of text strings associated with each. The coarse categories are Finance, Accounting, Medicine, Law, Higher Education, Primary/secondary teaching, Government, Art or Publishing, Retail, Senior management, and Middle/lower management. See Online Appendix Section B.1.7 for details and list of strings associated with each group. 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 who are linked to Class Reports but do not report work outcomes as missing. As with activities, we create a standardized index relating occupations to private feeder school background using equation 1. See Online Appendix Table A.1 and Figure A.2.

We follow a similar appraoch for adult social organizations, dividing them into two groups. The first, social clubs, consists of of gentleman's clubs (e.g. the Knickerbocker Club), country clubs or sports clubs (the Brookline Country Club, the Union Boat Club), and fraternal organizations (Masons, Elks, Kiwanis). The second consists of professional, honor, religious, and political associations. Examples include the American Medical Association, the American Academy of Arts and Sciences, B'nai B'rith (a Jewish Service organization). See Online Appendix Section B.1.7 for details and a list of club codes.

3.5 Student background and residential environment by high school type

Students from private feeder schools come from richer families with stronger ties to Harvard. Panel A of Table 2 reports student background characteristics from archival records, split by high school type. 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 did the same. 14% of private feeder students have a Harvard father and 29% have a Harvard brother, compared to rates of 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 than others. Panel B of Table 2 reports

[17] For consistent reporting across outcomes we also construct this score for upper-year social clubs.

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 non-private feeder students.

Most Harvard students live on campus. Panel C of Table 2 and Figure 2 describe students' first year living environments. 80% of students live on campus overall. Almost all (96%) of private feeder students live on campus, 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 Boston-area home address. We observe room attributes for 89% of campus residents. The students for whom we do not observe room attributes are those assigned to non-freshman housing, which the college sometimes used when there was excess demand for spots in freshman halls.

The rightmost column of Table 2 reports descriptive statistics for the sample of students who live on campus in freshman halls. This sample is the focus of the room randomization analysis in Section 5. These students are somewhat more likely to come from private feeder schools and less likely to come from Massachusetts, but overall we have broad coverage across high school type and other background characteristics.

Once on campus, first-year students live in dorms and within-dorm peer groups that are integrated by room price and high school background. The average per-occupant room price for on-campus students is $209, and the average within-year own-price rank is 0.49, on a zero to one scale. Private feeder students live in more expensive rooms, with an average price of $234 and rank of 0.58 compared to $198 and 0.45 for other students, but there is substantial overlap across all but the very bottom and very top of the distribution. This is visible in Panel A of Figure 2, which plots the histogram of own room prices by high school type.

Turning to peer neighborhoods, Panel B of Figure 2 displays a histogram of residential peer group sizes. Mean group size is 9.7, with the middle 50% of the distribution falling between 7.0 and 12.0. The mean peer neighborhood price is $214, and the mean peer neighborhood within-year rank is 0.49. Gaps in peer neighborhood attributes by high school type are somewhat smaller than those for own-room attributes. The mean peer group price (within-year rank) for private feeder students is $228 (0.55) compared to $208 (0.46) for other students. Differences in peer group room prices by high school type are 56% the size of differences in own-room prices.

Panels C and D of Figure 2 present a visualization of the variation in peer attributes within groups defind by own room price. Panel C displays the mean and 90-10 spread of peer neighborhod room prices by ventile of the own-price distribution within each year. Peer mean room price 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 variety of range of residential peer price levels. For example, the tenth percentile of peer mean price distribution for the top-most

ventile is below the mean peer price level at the all but the lowest ventile. 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 is consistent with archival accounts of differences in pre-college background by school type and of administrator's attempts to bridge these gaps through on-campus housing.

4 College behavior and long-run outcomes

4.1 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 3. The first main point here is that private feeder students do badly in the classroom compared to others. Panel A of Figure 3 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.[18]

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 3, private feeder students are 1.5 times more likely to participate in any activity (71% vs. 47%) than other students, participate in twice as many activities in total (1.67 vs. 0.85), and are three times as likely to have activity leadership roles. Looking across activity types, we see that rates of participation in music groups are fairly similar, and that private feeder students are about 1.8 times as likely to participate in sports, the most common activity type. Differences in participation in schoolwide social committees are especially stark, with private feeder students 4.2 times more likely to participate than other students (20% vs. 5%).

Initial gaps in activity participation and social leadership persist through to our main measures of social success: exclusive upper year social clubs. Panel C of Figure 3 reports these findings. 39% of private feeder students join the Hasty Pudding sophomore society compared to 8% for other students. 14% of students join a final club of any kind, with a 37% rate for private feeder students and a 6% rate for other students. And 7% of students join one of the selective final clubs, with a 21% rate for private feeder students that is nearly nine times more than the rate for all other students.

[18] 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.

Upper-year social 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 that to the exposure of private feeder students to non-feeder students in Hasty Pudding and 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 participating in Hasty Pudding has 38% of group peers from non-private feeder backgrounds, and the average private feeder student in a selective final club has 27% of peers from non-private feeder backgrounds.

Breaking out rates of freshman social leadership and membership in upper-year selective clubs underscores the extent to which a small number of high schools stand out from the rest. Panels D through F of Figure 3 show rates of participation in selective activities by school, with the count of students by high school on the horizontal axis and the share of students participating in social committees on the horizontal axis. The private feeder schools, and in particular Groton, St. Mark's, and St. Paul's, stand out. More than a third of students at these three schools participate in first-year social committees, more than 60% join Hasty Pudding, and more than 40% join a selective final club. Andover and Exeter send a large number of students to Harvard, and these students are reasonably likely to join freshman social committees, but less likely to gain entrance to exclusive upper-year groups.

4.2 The returns to social vs. academic successs

Social success in college predicts labor market success in the long run, but academic success does not. Panels A, B, and C of Figure 4 display income measures from the 1940 Census split by class rank group (on the horizontal axis) 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 from the figures. Panel A shows that students in selective final clubs earn about $835 more than other students in the same honor group, 31% of the no-club mean. In contrast, conditional on club membership, the relationship between grades and earned income is flat. Final club members in the lowest rank group earn $770 (26%) more than non-members in the top rank group. 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 (corresponding to roughly the 99.2nd percentile of the income distribution of similarly aged men in 1940) and to have non-wage income. 52% of club members in the lowest rank group have topcoded income, compared to 16% of non-members in the top two groups.

A natural hypothesis is that the relationship between club membership and later life outcomes is due to the way clubs select members, in particular on the basis of high school type. Students from private feeder schools make up the bulk of the membership in these clubs, and

tend to come from wealthier families. Descriptive statistics reported in Table 3 show that students from private feeder high schools also have different careers and higher incomes as adults.

This plausible story turns out not to be true. A simple visual analysis indicates that selection on the basis of high school type cannot explain differences in earned income by club membership. Panels D through F of Figure 4 repeat Panels A through C, but split by both selective club membership and high school type. Panels D and E show that the gap in measures of wage income by club membership are not diminished by controls for school type. Instead, once one conditions on club membership, differences in earnings by school type mostly disappear. 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.

4.3 Assessing selection effects using a within-family design

Though high school type is an important feature of historical accounts, students may sort into final club membership on the basis of other attributes that are also correlated with earnings. To assess the importance of selection for the social success premium more generally, we conduct a series of analyses that control for high school type, high school identity, legacy status, and family identity. These specifications have the form

$$Y_i = S_i\beta_s + R_i\beta_r + X_i\beta_x + \theta_{c(i)} + e_i \tag{2}$$

Y_i is an outcome of interest for individual i, S_i are measures of i's social success, R_i is student i's academic rank group, and X_i are sets of the controls listed above. $\theta_{c(i)}$ are fixed effects for entering cohorts c, and are included in all specifications. Class rank is in rank-group units with the sign reversed so that positive coefficients indicate that earnings outcomes rise as academic standing improves.

The goal of this analysis is to evaluate stories in which social success is related to some non-social student attribute, such as family wealth, and it is that omitted variable that drives the relationship between social success and long-run outcomes. These specifications do not tell us whether the long-run social success premium reflects something students gain at college or the return to pre-existing social skills. We return to this point in Section 5.1 when we discuss room randomization.

Table 4 presents our results. Each panel corresponds to a different set of covariates. The first column is an indicator for successful match to a wage record in the 1940 Census. Looking across all specifications, the relationship between wage match and covariates of interest is are in general economically small, and is statistically different from zero at the five percent level in only one out of fourteen cases– the coefficient on participation a less selective final club in Panel

C. This suggests that concerns related to differential censoring are limited as related to our main variables of interest: class rank and membership in a selective final club.

Panel A of Table 4 shows a baseline specification with controls for private feeder status and class rank, but not social success. Private feeder students earn $293 more than other students, 11% of a non-private-feeder mean of $2680. This is about 8 times larger than the $35 premium associated with a one-group improvement in class rank. Private feeder status is associated with a 9.1 percentage point increase in rates of earnings topcoding (60% of the non-private-feeder mean), and a 17.8 percentage point increase in reports of unearned income. Academic standing is negatively correlated with earnings topcoding and unrelated to the presence of non-wage income.

Panel B adds controls for membership in a selective final club. Students in selective final clubs earn $762 more than other students (a 28% premium above the no-selective club mean), are 23.0 percentage points more likely to have topcoded earnings (total topcoded rate 2.5 times higher than amoung non-club men), and are 17.3 percentage points more likely to have unearned income (36% above the no-club mean). For mean earnings, the final club premium is 16 times the academic rank group premium; there is no academic rank premium for topcoded earnings. After conditioning on selective final club membership, the private feeder premium falls by 50.5%.

Panel C follows Panel B but adds a separate dummy for membership in less selective final clubs. The mean earnings premium for membership in a less selective final club is about half the premium for more selective clubs, and the topcode premium is about one third. After controlling for membership in both selective and less selective final clubs, the private feeder effect on mean earnings falls to 27% of its baseline level in Panel A and becomes statistically insignificant. The private feeder effect on earned income topcoding is also reduced by about two thirds. In contrast, the private feeder premium for unearned income retains two-thirds of its size in Panel A and is highly statistically significant. These regression results underscore visual intuition in Figure 5 that private feeder students who do not make into the top social tier at Harvard look similar to other students in terms of their earned income.

Panels A through C of Table 4 show that it is not the selection of private feeder students as a group that drives the observed social success premium. However, the social success premium may arise from the selection of students from specific feeder schools or students from particular backgrounds within these schools. Panels D and E explore this possibility. Panel D limits the sample to students from private feeder schools and adds fixed effects for each school, as well as a student-specific indicator for Harvard legacy status to address within-school selection. We know from Figure 3 that cross-school differences account for much of the variation in club membership. We find that the selective final club premium remains large after adding these controls. The academic success premium is near zero.

Panel E of Table 4 limits the sample to families who send more than one son to Harvard over

our sample period and adds family fixed effects. These regressions essentially compare the smart brother in each family to the popular brother. The final club premium for earned income is larger in the within-family specifications than in the other specifications we estimate (equal to 44% of the brothers sample mean for earnings and 161% of the brothers sample mean for topcodes), while the class rank premium remains close to zero.

The social success premium for non-wage income is negative and imprecisely estimated. Paired with the observation from Panel C that gaps in unearned income by high school type persist after conditioning on final club participation while gaps in earned income do not, this evidence suggests social success may contribute more to earned income while family background drives differences in unearned income.

4.4 Social success vs. academic success in the long run

In addition to income, social success predicts career paths and social engagement in the long run. Panels A through D of Figure 5 show how occupation 25 years after graduation varies with class rank and selective final club membership. We display results for four illustrative career types: finance, medicine, higher education, and law. Students in social clubs are much 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 for both members and non-members. The reverse is true for medicine: rates are higher for non-members than members, and rising with academic rank. Higher education and law both are more common for students with better academic ranks, but not strongly related to club membership within rank.

Panels E and F of Figure 5 show how adult participation in social and professional organizations 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, difference in long run career and social outcomes by social success are not the product of selection into final clubs on the basis of high school type. The lower row of graphs in Figure 5 display the same outcomes as in the upper row, but split by both club membership and high school type. Within high school type, selective 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 higher education career at all but the lowest achievement levels. Both private feeder and other students who are members of social clubs are more likely to be members of social organizations as adults across all achievement levels. In contrast, mem-

bership 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

Our within-family and within-school research designs provide strong evidence that the return to academic success at Harvard is close to zero, and that the observed earnings premium for social success is not driven by omitted non-social student attributes, such as family wealth. Students from rich families tend to be more socially successful at college, but social success is distinct from family wealth in the way it affects long-run outcomes. What these approaches do not resolve is whether the social success premium reflects a labor market return to pre-existing social skills, or the return to social attributes that are developed or acquired at college. 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.

As described in Section 2, historical records of the assignment process state that room assignment is random within groups defined by a) entering cohort, b) the price range indicated by the applicant or applicants, c) room type (single, double, etc.), and d) high schools sending large numbers of students. Recall that randomization takes place at the level of the *room*, not the individual, because students may apply in groups.

Our experimental specifications take the form

$$Y_i = \beta_0 + \beta_1 RP_{p(i)} + \theta_{r(i)} + \tau_{h(i)} + e_i \tag{3}$$

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). In accordance with qualitative reports of the randomization design, we define the $\theta_{r(i)}$ as fully-saturated interaction 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 5.9.

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 pair.

Figure 6 illustrates how this random assignment works. The Panel A shows maps of the main freshman dorms prior to 1931 and their 1920 room prices. There are rooms at a variety of price

levels on each dorm floor, and rooms at very different price points are often adjacent. Panel B gives an example of a low-priced randomization block: doubles priced at $175 per student. All 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 building 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 what our experimental specifications identify is the effect of assignment to a higher-status neighborhood, as mediated by a variety of possible 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) contribute individually to long-run outcomes like career path.

5.2 Peer groups and randomization blocks

There is substantial variation in potential peer group assignments within randomization block. We describe the randomization blocks in Figure 7. Panel A shows the distribution of block sizes. The average block consists of 9.7 individuals; the 10th percentile of the block size distribution is 5.0 and the 90th percentile is 66.0. 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 MP_p (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 MP_p (neighborhood private high school share) in the block. The graphs thus help visualize the within-block variation in peer environment.

Panel B shows that within-block variation spans nearly the entire range of peer neighborhood prices. The vertical lines show the 10th and 90th percentiles of the neighborhood price distribution. There are blocks in which prices range from the maximum price to the median, and where prices range from the 90th percentile down to below the tenth. 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 full the 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. We conduct tests of covariate balance first. Results in Panel A of Table 5 show that controlling for randomization blocks eliminates the strong cross-sectional relationship between neighborhood price rank RP_p and predetermined observable individual attributes. We estimate versions of Equation 3 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 of the table 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 dependent variable associated with moving from the lowest-ranked peer neighborhood 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. We observe that 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 $193 more on average for their rooms (90.7% of the sample mean), are 38.9 percentage points more likely to have attended a private school, and 28.9 percentage points more likely to have attended a private feeder school. They are 6.7 percentage points (76.6% of the sample mean) more likely to have a Harvard father and 11.4 percentage points (49.8% of the sample mean) more likely to have a Harvard brother. They are less likely to be or descend from southern or eastern European immigrants and more likely to have fathers who are doctors or lawyers. We reject the joint null hypothesis 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 3. With the addition of block controls the relationships between neighborhood mean price and own attributes become economically and statistically insignificant across the board. 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=0.900). This null relationship persists when we add the full set of randomization

blocks and high school dummies (p=0.505). Our findings here 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 principal 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 mitigates concerns related to differential censoring in specifications with Class Report or Census outcomes.

Assignment to neighborhoods with higher mean prices predicts peer composition. Panel C of Table 5 reports estimates of Equation 3 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 26.1 percentage points higher (on a in-sample mean of 54.0) than students assigned to the lowest-price neighborhoods and private feeder high school peer shares that are 19.2 percentage points higher (on a in-sample mean of 31.4), conditional on randomization block. These effects are precisely estimated, with standard errors of 3.2 and 3.1, respectively. Peer legacy shares are 6.3 percentage points higher, and peer eastern European immigrant shares shares are 4.7 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. Conditional on randomization blocks and high school dummies, a one-unit increase in room rank raises the mean count of first-year activities residential peers participate in by 0.31, the share of peers participating in social activities by 0.05, and the peer activity index by 0.28. It raises the mean peer academic rank group by 0.14, 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. The overall picture here is one where assignment to higher-priced residential neighborhoods is conditionallly 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 outcoms 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 and other students are equal.

Placement in a higher-priced peer neighborhood raises participation in social activities while in college. Panel A of Table 6 reports effects of peer neighborhood price on freshman extracurricular activities. A change in peer neighborhood rank from zero to one raises the share of students who report any activity by 6.9 percentage points (SE=3.3; relative to a residential sample mean of 60.5%), the activity count by 0.293 (SE=0.106; relative to a mean of 1.265), and the percent of students report leadership roles in any activity by 4.4 percentage points (SE=1.9; sample mean of 8.3%). The aggregate effect of changes in quantity and type of activities is to raise the activity private high school score by 0.207 standard deviations (SE=0.079), equal to 34.1% of the the gap in activity score between private feeder students and other students in the full sample. Effects on activities by type suggest increases in participation in social committees (5.1 percentage points) as a key channel for increased participation.

The full sample effects are entirely driven by students from private feeder schools. 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 aggregate effect of these small changes on the activity score is also zero. In contrast, placement in higher-priced peer neighborhood raises the count of activities that private feeder school students participate in by 0.684 (SE=0.216), and the private school activity index by 0.490 (SE=0.165). We reject the null that the index effect for private feeder students is equal to those for other students at the 5% level. Effects on participation in any activity are similar to those in the full sample and do not differ statistically from zero for private feeder students, suggesting that the intensive margin of activity participation is important here.

The activity types that drive the index effect are leadership positions and schoolwide social committees. For private feeder students, social committee participation rises by 18.2 percentage points (SE=5.6) with a one-unit change in room rank, compared to a precise zero for non-private students. Assignment to higher-income, more private-school heavy peer groups makes private feeder students more likely to hold leadership positions in their first year at school, but does not have the same effect on other students.

Increases in first-year leadership activities carry over to our main measure of social success in college, selective upper year social clubs. We report these results in Panel B of Table 6. A one-unit change in room rank raises the rate at which students in the pooled sample join highly selective final clubs by 5.7 percentage points (SE=2.1; on a full sample mean of 9.5). This effect is again driven entirely by large effects for private feeder students, with precise zero effects for other students. We reject the null of equal effects across the two groups (p=0.009). 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, generally a preliminary step to membership in final clubs of any kind, also rises for private feeder students but not others. The upper-year club index measure rises only for private feeder students. Statistical tests of differences between effects for private feeder and other students reject the null at p-values near 0.05 for both Hasty Pudding participation and the feeder index. Overall, these results indicate that first-year residential assignment can alter the composition of the most selective social groups at Harvard but does not succeed in integrating these groups by student backround.

In contrast to 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.[19] The effect of assignment to higher-priced peer neighborhoods on a continuous class rank measure is near zero in the full sample and in samples split by high school type. 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 one-unit change in neighborhood rank raises the rate of list appearance by 7.9 percentage points, with 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 reduce the risk of very bad academic outcomes.

5.5 Career paths and adult social lives

25 years after graduation, assignment to a higher-priced peer neighborhood causes students from private feeder schools to pursue careers with higher private school shares but, if anything, pushes other students in the opposite direction. We report these findings in Panel A of Table 7. Paralleling results from Figure 5 showing large gaps in finance career paths by final club membership, we find that assignment to a higher-priced neighborhood raises the share of private feeder students going into finance by 13.6 percentage points (SE=5.8), on a base of 18.3% in the residential sample. Effects for other students, whose social outcomes were not affected by residential assignment, are slightly negative; a test of the null that the two effects are equal returns a p-value of 0.005. Shifts for other career tracks are smaller in magnitude, with positive but statistically insignificant shifts for private feeder students towards relatively high private school share upper management positions and away from low private feeder fields like medicine, higher education, and teaching.

Aggregating across occupation categories using the occupation private feeder index, we see that assignment to a higher-priced peer neighborhood has a large positive effect for private

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

feeder students and a negative effect for other students. For private feeder students, assignment to a high-priced peer neighborhood raises the occupation index by 0.423 (SE=0.180), while for other students it lowers the occupation index by 0.156 (SE=0.091). These effects are large relative to baseline gap in the occupation score between private feeder and other students of 32.5. The p-value from a test of the null that these two effects are equal is 0.004. Residential exposure to high-status peers pushes high-status students towards career paths where high-status students are more common.

Assignment to higher-priced peer neighborhoods raises adult participation in social organizations, but again only for private feeder students. Panel B of Table 7 reports the effects of residential peer group rank on participation in adult social organizations. In the full sample, a one-unit change in peer neighborhood rank raises the rate of membership in any adult social organization by 6.6 percentage points (standard error of 3.5, sample mean of 38.1%). This reflects the combination of a 21.4 percentage point increase for private feeder students (SE=6.6) and a zero effect for other students. We reject the null the two effects are equal (p=0.008). 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 19.3 percentage points (standard error of 6.5). In contrast, we see no effects for membership in professional or honor societies. These findings again parallel results from Figure 5 showing that members of selective final clubs were more likely to participate in adult social clubs but not professional or honor societies.

5.6 Census outcomes

We next estimate the effects of residential peer neighborhood on 1940 Census outcomes, with results reported in Panel C of Table 7. The need to restrict to earlier entering cohorts and the requirement that observations be matched to Census reports of wage records reduce sample sizes relative to our analysis of the effects of room assignment on long-run outcomes from Class Reports by roughly two thirds. This 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 9.7. Given our point estimate of approximately zero, the 95% CI spans increases and decreases in topcode rates of close to 20 percentage points 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.

5.7 Interpreting effect sizes

Estimated effects suggest that residential assignment is a quantitatively important determinant of social outcomes in college and career outcomes in the long run. One way to see this is to scale

reported effects by 0.5 so that they reflect the value of a 50-percentile shift in peer neighborhood price rank; for example, from the 25th to 75th percentile of the neighborhood distribution. Recall from Figure 7 that within-block differences in assigned neighborhood price of this size are common. The mean gap in rates of selective final club membership between private feeder students and other students is 18.6 percentage points (Table 2) . A 50-percentile shift for a private feeder student in peer room price distribution raises that person's membership chances by 7.3 percentage points, or about 39.4% of the mean gap in the full sample. The same change can account for 74.5% of the 9.1 percentage point gap in finance careers, and 87.9% of the 12.1 percentage point gap in adult social club participation.

5.8 Nonlinear effect estimates

Figure 8 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. These results are qualitatively similar to what we see in linear specifications. For private feeder students, placement in the third tercile of the peer neighborhood price distribution rather than the first raises the chances of participating in a first-year social committee by 4.8 percentage points, a selective final club by 4.1 percentage points, a finance career by 3.8 percentage points, and an adult social organization by 7.0 percentage points. A general pattern of rising effects across terciles on selective social and career outcomes for private feeder students stands out in the graph. Nonlinear specifications also suggest possible positive effects of peer neighborhood rank on our Census income measures. These effects are not statistically significant at conventional levels in most cases and in some tension with the very imprecise estimates in our main linear specifications, so we interpret them with caution. As in the linear analysis, effects for students not from private feeder schools are near zero across all outcomes.

5.9 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.2 through A.5 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 occupancy price and year. As suggested by the balance analysis in Table 5, year-specific room price dummies are the key controls here. Online Appendix Tables A.6 and A.7 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.8 and A.9 report results that drop the 1919, 1920, and 1921 entering cohorts, who arrived on campus before the first reference we observe to the room assignment scheme used in 1922 and later. Dropping these cohorts does not affect our findings. Finally, Online Appendix Tables A.10 and A.11 report results using randomization inference as opposed to clustered standard errors to conduct statistical tests; our results are unaffected by this change.

6 Discussion

6.1 Academic performance and the opportunity cost of finance careers

Finance in the mid-20th Century was characterized by relatively low job complexity and skill intensity compared to other periods (Philippon and Reshef, 2012). Stylized (though debated) depictions of the finance industry from the 1950s through the 1970s tell a story in which Depression-era 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).

How one interprets the shift towards finance careers for high-status students randomized into more interactions with high-status peers depends on the academic achievement of the career-switchers. Is it academically 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. To do this, 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 3. 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 one-unit change in neighborhood price rank raises the rate at which students both have high grades and are members of selective final clubs by 11.8 percentage points. It raises the rate at which students have high grades and enter finance careers by 11.2 percentage points. And it reduces the rates at which students enter careers in higher education and medicine by 5.0 and 4.1 percentage points. Rates 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, the opportunity cost in terms of human capital for academic high

achievers was high.[20]

6.2 Harvard students in national and historical context

Comparisons between Harvard and the full Census population confirm that our analysis is most relevant to understanding who makes it to the upper tail of the income distribution, and that the share of Harvard students in this upper tail has remained fairly stable over a very long time horizon. Figure 9 illustrates this point by plotting 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 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 amoung 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. 21.1% of Harvard students born 1980-1982 (roughly corresponding with entering cohorts of 1998-2000) earn in the top 1 percent of the income distribution (Chetty et al., 2020).

6.3 Implications for interdisciplinary literature

Our observational and quasi-experimental findings tell a story in which the return to social success in college is high and the return to academic success is small. Interactions with high-status college peers are an important determinant of social success in college and of long-run career and social outcomes, but only for students from high-status backgrounds. Interactions across pre-existing boundaries of social status rarely spark, even with prolonged residential contact.

Our results are consistent with Mills' (1956) "Two Harvards" story, in which college social success depends on private school background, with Amory's (1947) description of the way college social success affects later-life career and social outcomes, and with recent claims that final clubs tend to expand pre-existing social divisions between students from different backgrounds.They are harder, though not impossible, to reconcile with results from a literature in economics and psychology suggesting that collaborative intergroup contact can yield sustained changes in attitudes and activities over the short to medium run (Rao, 2019; Lowe, 2020; Mousa, 2020). We see no evidence that high-status students react to intergroup contact by joining groups favored by lower-status individuals, or that low-status students benefit from intergroup interactions through increased access to high-status groups. This may be because residential integration

[20] With respect to the 3-6-3 rule, we note the prominent role of country clubs in our analysis of the causal effects of room assignment.

at Harvard does not foster collaborative contact, or because collaborative contact does not matter as much when the stakes are high. Under any of these candidate explanations, universities face major challenges in fostering durable cross-group relationships through housing policy.

7 Conclusion

This paper shows that Harvard students from high status private school backgrounds follow different paths than other Harvard students. They do poorly in school, but excel in the pursuit of a form of social success that appears to offer large economic returns. Though Harvard administrators implemented admissions and residential life policies with the goal of integrating the social life of Harvard by income (if not other important characteristics, such as gender, religion, and race), these policies do not succeed in integrating student social groups or later life outcomes. They *do* have large effects on which students from high-status backgrounds achieve social and career milestones. Our results are consistent with the broad point that social factors are an important determinant of labor market success (Granovetter, 1973; Deming, 2017), even or perhaps especially at the top of the economic ladder.

A large body of research emphasizes the importance of inclusive institutions and competition between different groups of elites for economic growth in the long run (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). Our findings suggest that expanding access to elite higher education beyond its traditional constituencies does not by itself integrate informal, student-run organizations and social groups in college or beyond, and that cross-group contact alone may not suffice to spur cross-group cohesion or help new constituencies compete for spots at the top of social institutions. Whether and how it is possible for policymakers at universities and beyond to facilitate cross-group ties sufficiently strong to expand access to exclusive social groups and career paths is a subject for future work.

References

Abdulkadiroğlu, Atila, Joshua Angrist, and Parag Pathak, "The elite illusion: Achievement effects at Boston and New York exam schools," *Econometrica*, 2014, *82* (1), 137–196.

Abramitzky, Ran, 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.

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.

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.

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 postwar socioeconomic status and mortality of union army veterans and their children," *Journal of Health Economics*, 2020, *70*, 102281.

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

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

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.

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*, 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, "Corporate finance policies and social networks," *Management Science*, 2017, *63* (8), 2420–2438.

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

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

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

Hoxby, Caroline and Christopher Avery, "The Missing" One-Offs": The Hidden Supply of High-Achieving, Low-Income Students.," *Brookings Papers on Economic Activity*, 2013.

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

Hoxby, Caroline, Sarah Turner et al., "Expanding college opportunities for high-achieving, low income students," *Stanford Institute for Economic Policy Research Discussion Paper*, 2013, *12*, 014.

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

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

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

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

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.

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.

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, "American historians and the business elite," *The Journal of Economic History*, 1949, *9* (2), 184–208.

_ , "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.

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.

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

Philippon, Thomas 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.

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

_ , "Peer effects in education: How might they work, how big are they and how much do we know thus far?," *Handbook of the Economics of Education*, 2011, *3*, 249–277.

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.

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.

Taussig, Frank William and Carl Smith Joslyn, "American business leaders.," 1932.

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.

Warner, W Lloyd and James C Abegglen, *Occupational mobility in American business and industry, 1928-1952*, University of Minnesota Press, 1955.

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.

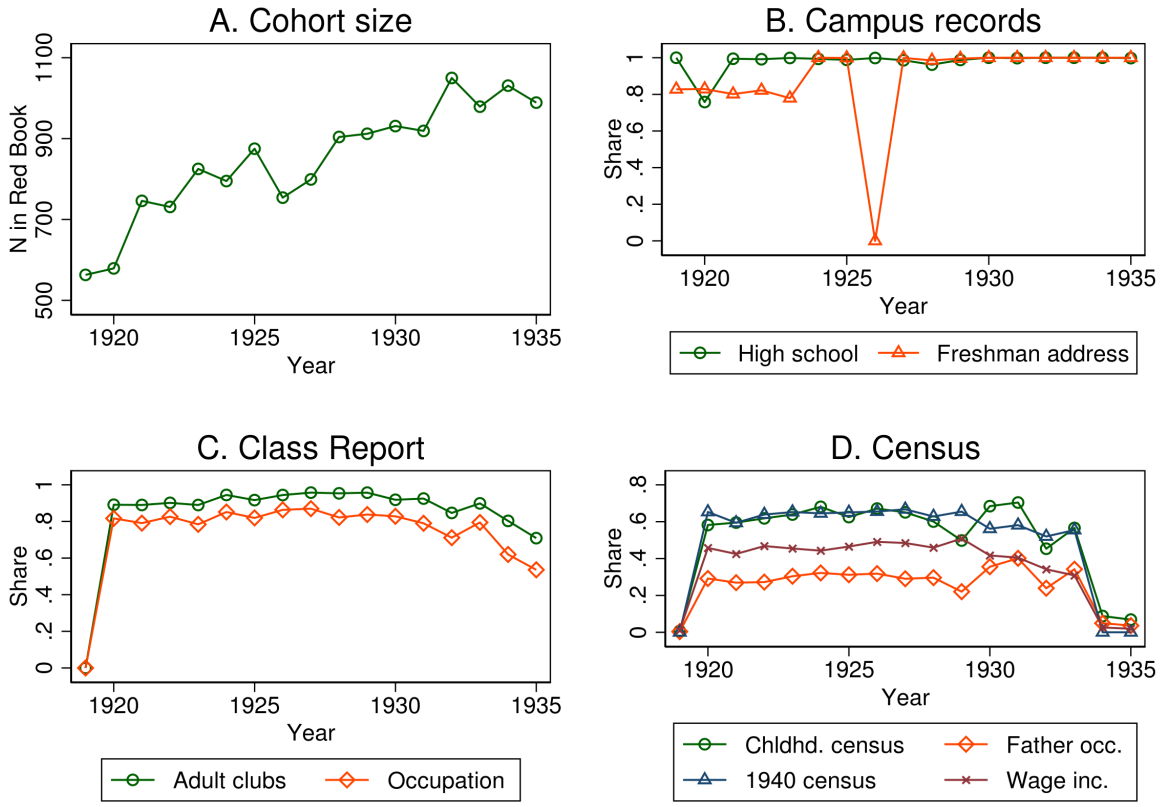
Zárate, Román Andrés, "Social and Cognitive Peer Effects: Experimental Evidence from Selective High Schools in Peru," 2019.

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

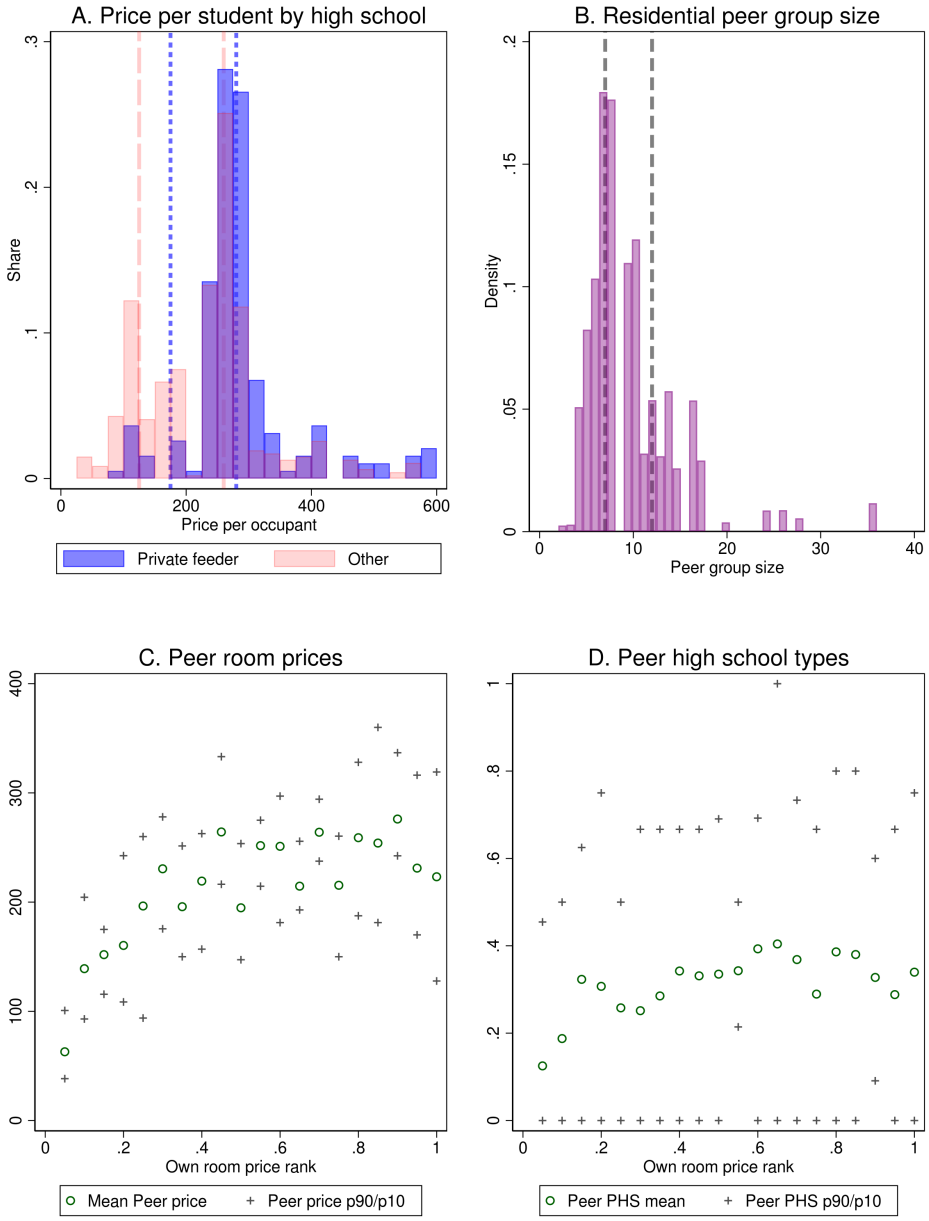
Figures

Figure 1: Data availability by cohort



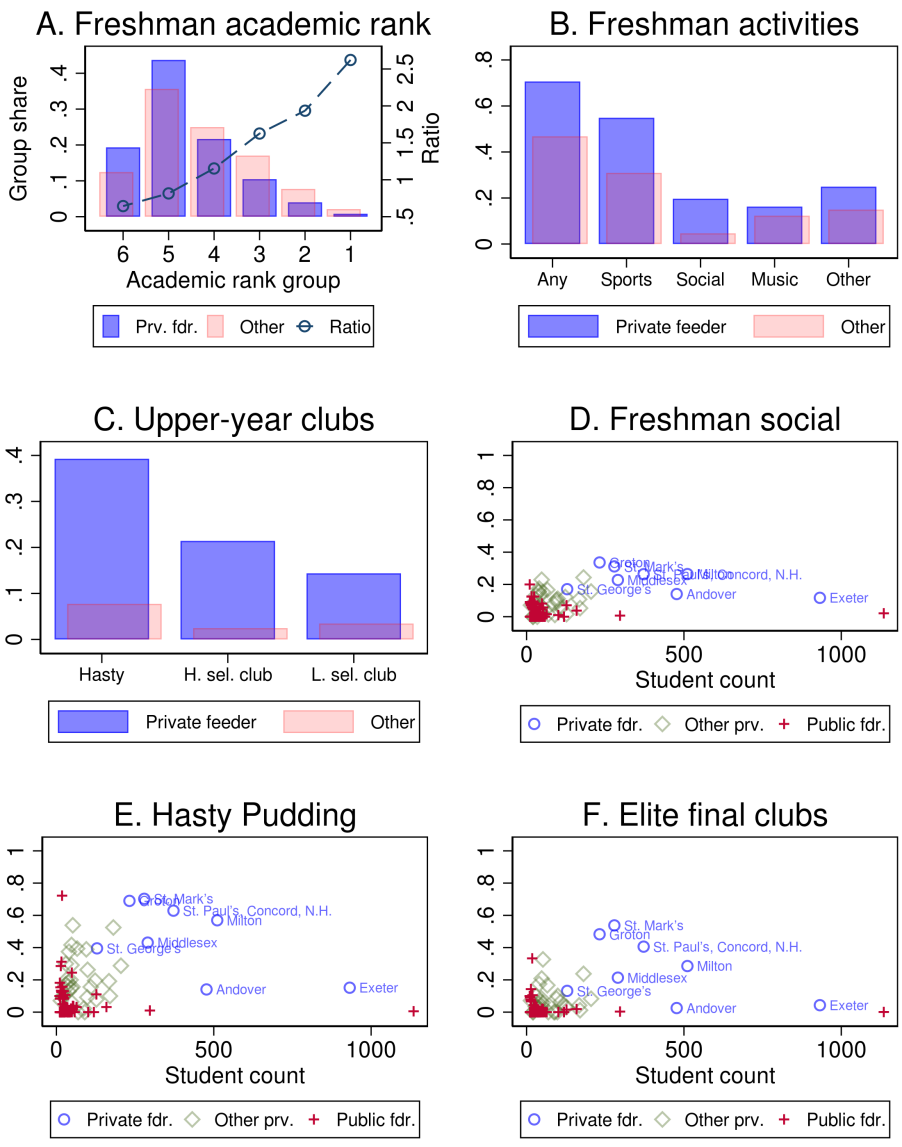
Cohorts are defined by presence in the freshman class Red Book. All data availability shares are presented unconditionally, e.g., the share with positive wage income in the 1940 census is the share of the entire cohort not only of those matched to a 1940 census record. Childhood census is the presence of a match to at least one of the 1910 or 1920 Censuses for incoming cohorts of 1920-1929 or a match to either 1910, 1920 or 1930 Census for incoming cohorts of 1930-1933. 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: Dorm room prices and peer attributes



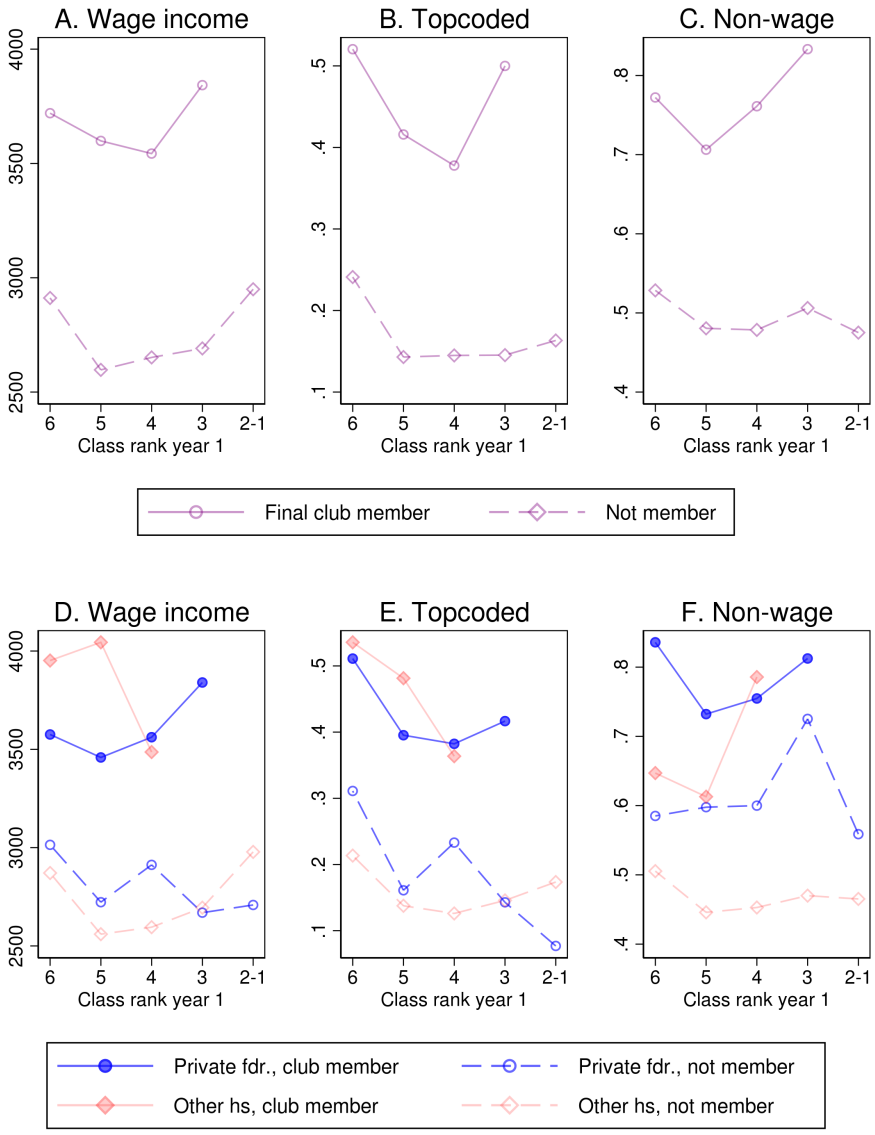
Panel A presents a histogram of dorm price per student by student high school type. The incoming cohort of 1928 is depicted as an example. Lines of long and short dashes the denote 25th and 75th percentiles for private feeder students and other students respectively. Panel B depicts a histogram of the combined room capacity of all rooms in a residential neighborhood. Gray dashed lines mark the 25th and 75th percentiles. Panels C and D present mean, 10th, and 90th percentiles of neighborhood per-student room price distribution excluding own room (C) and peer neighborhood private feeder high school share excluding own room (D) by own room price ventile defined within year (horizontal axis). Across years, all ventiles except one have minimum and maximum private feeder shares of 0 and 1 respectively (not shown).

Figure 3: Campus academic and social achievement by high school



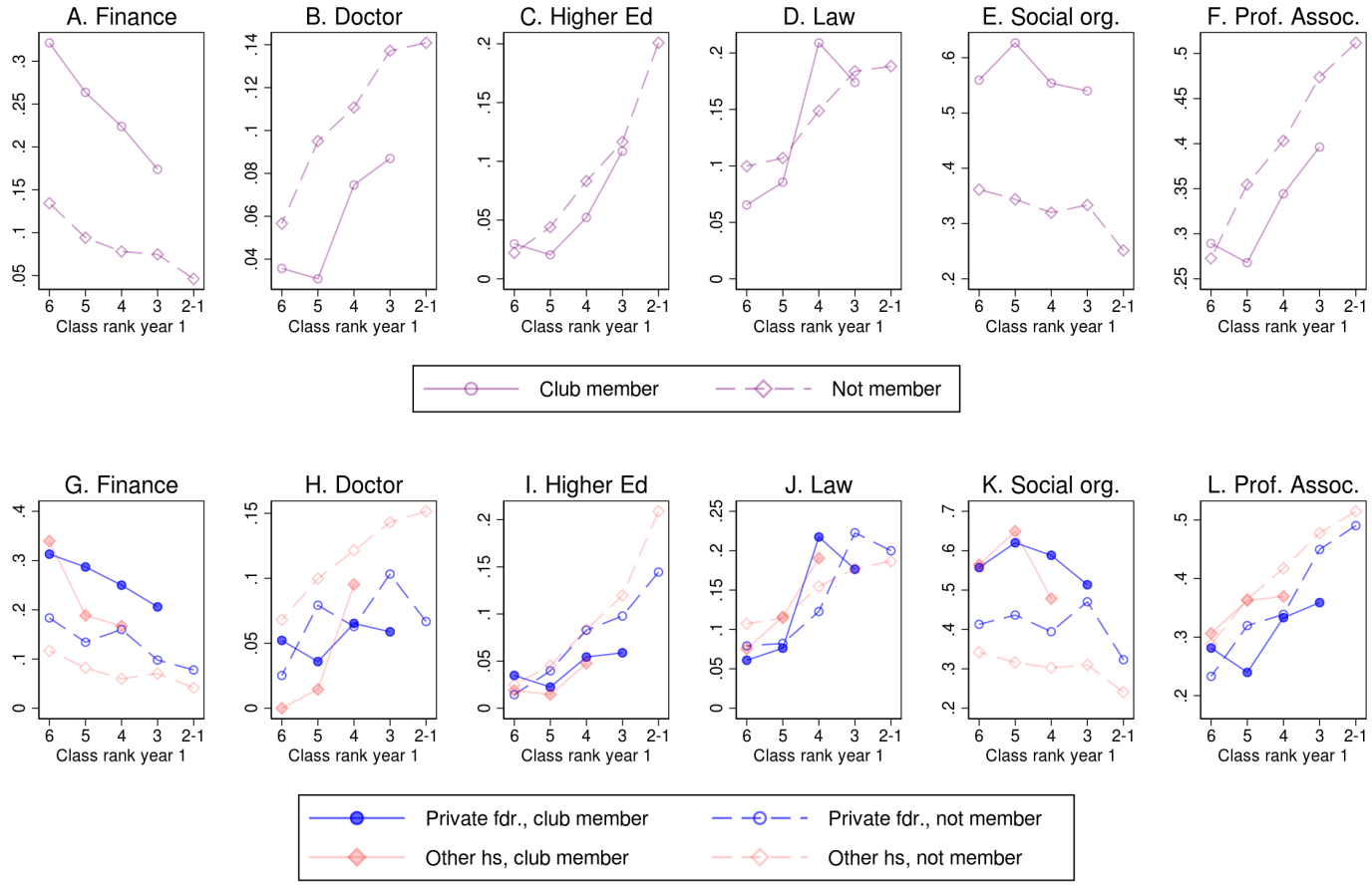
Panels A-C present shares of private feeder and non-private feeder students in each freshman academic rank group and participating in activities. Panel A also presents the ratio of non-private feeder students to private feeder students in each academic rank group. Panels D-F present share of students from an identified high school participating in freshman social activities, Hasty Pudding, and selective final clubs respectively by the number of students coming from that school. Private feeder schools are labeled by name.

Figure 4: Earnings by academic performance and final club membership



Panels A-C depict earnings by freshman academic rank group and selective final club membership. Panels D-F present the same but also divide students by high school type. Rank groups 1 and 2 are collapsed and groups with fewer than 20 students not displayed. Includes students from cohorts 1920-1930 who matched to the 1940 census (and reported wage income for Panels A and B).

Figure 5: Adult career and social outcomes by academic performance and final club membership

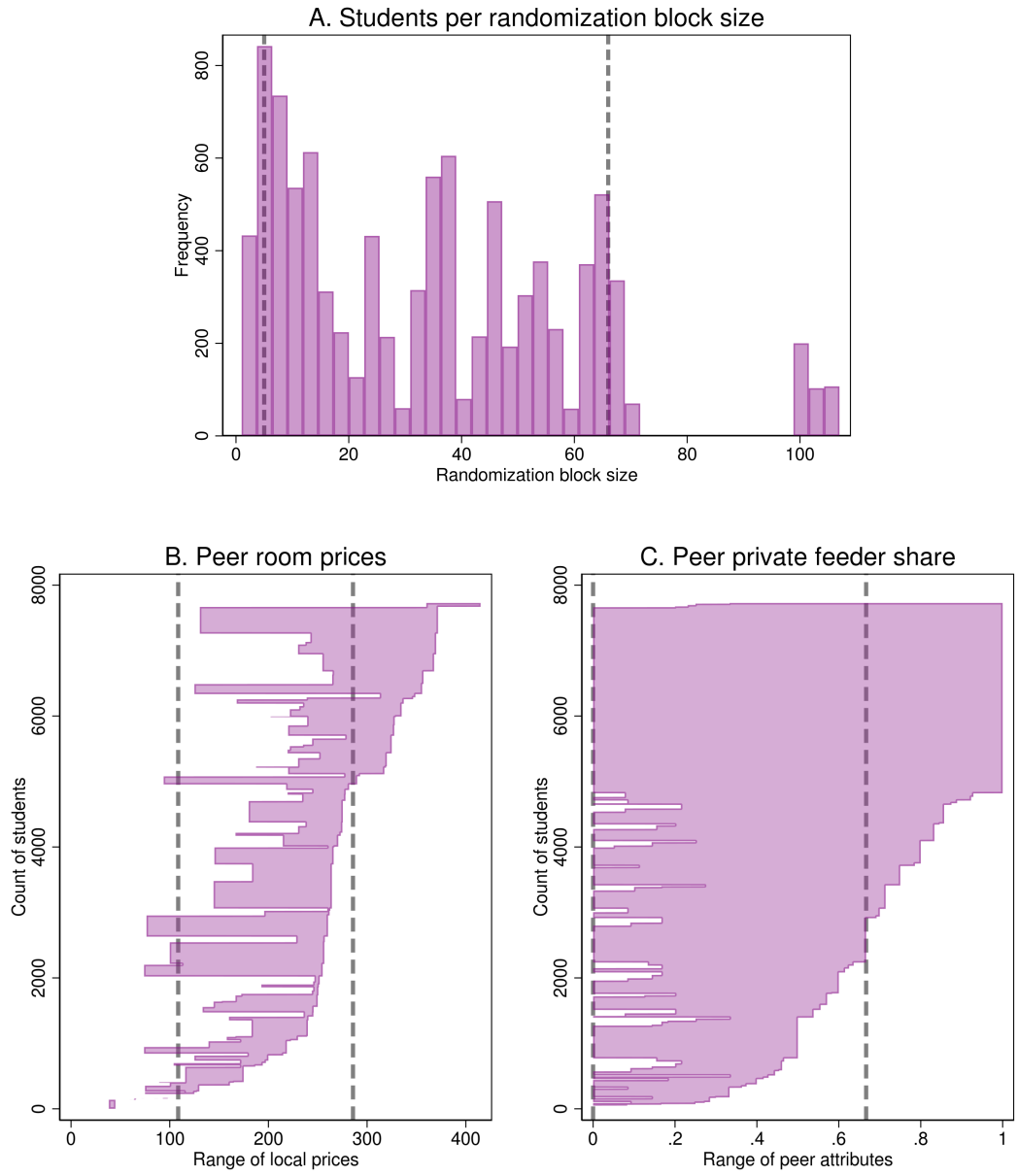


Panels A-F depict adult outcomes for students by freshman academic rank group and selective final club membership. Panels G-L present the same but also divide students by high school type. Rank groups 1 and 2 are collapsed and groups with fewer than 20 students not displayed. Includes students from cohorts 1920-1934 who matched to a Class Report.

Figure 6: Room price and randomization examples

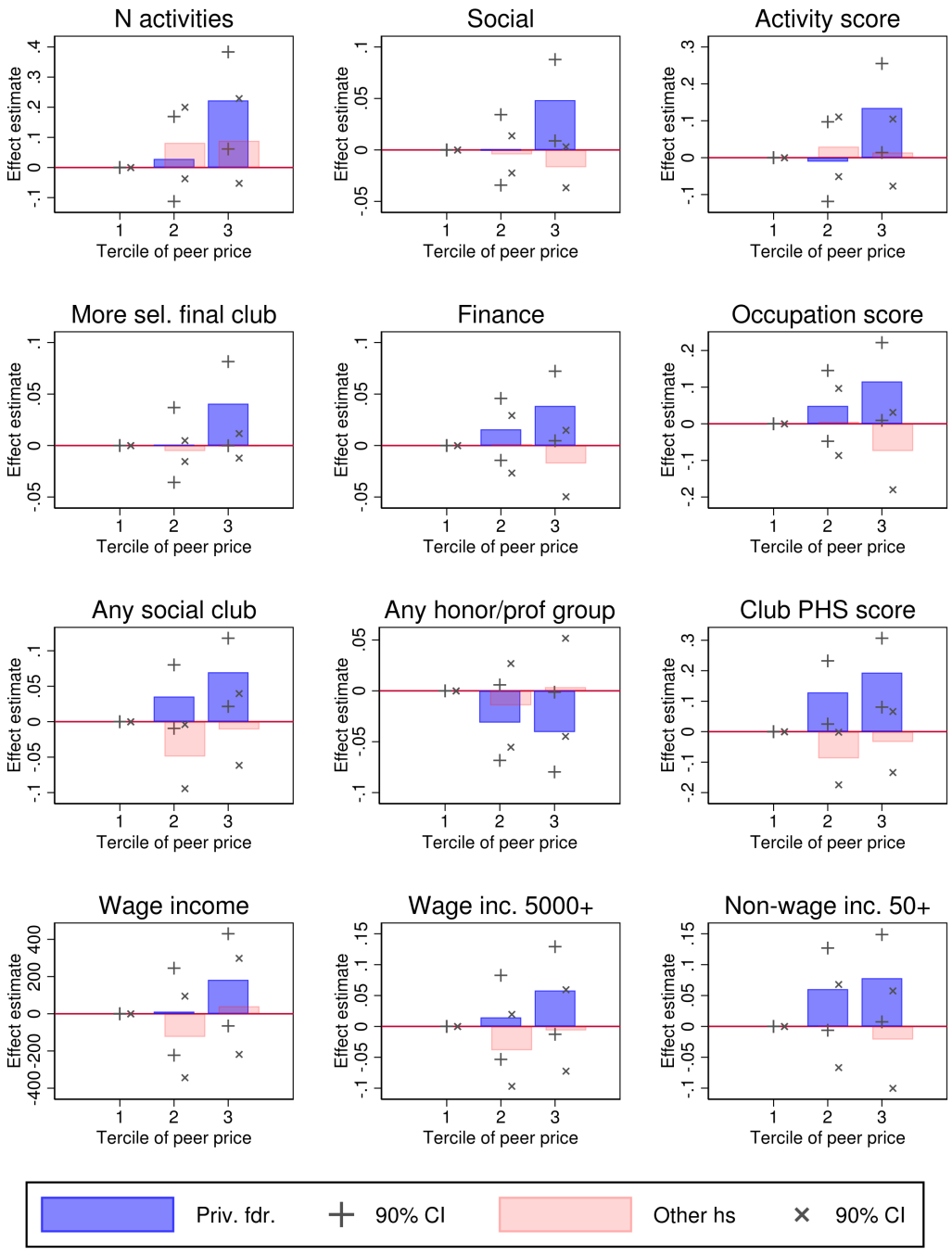
Floor plans depict floors 1-4 of the three main dorms housing freshman prior to 1931. Prices from 1920. Panel A depicts the price per student at the room level. Panel B depicts the average price per student at the peer neighborhood level. Rooms from an example low-price randomization block are outlined 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 7: Randomization support and block size



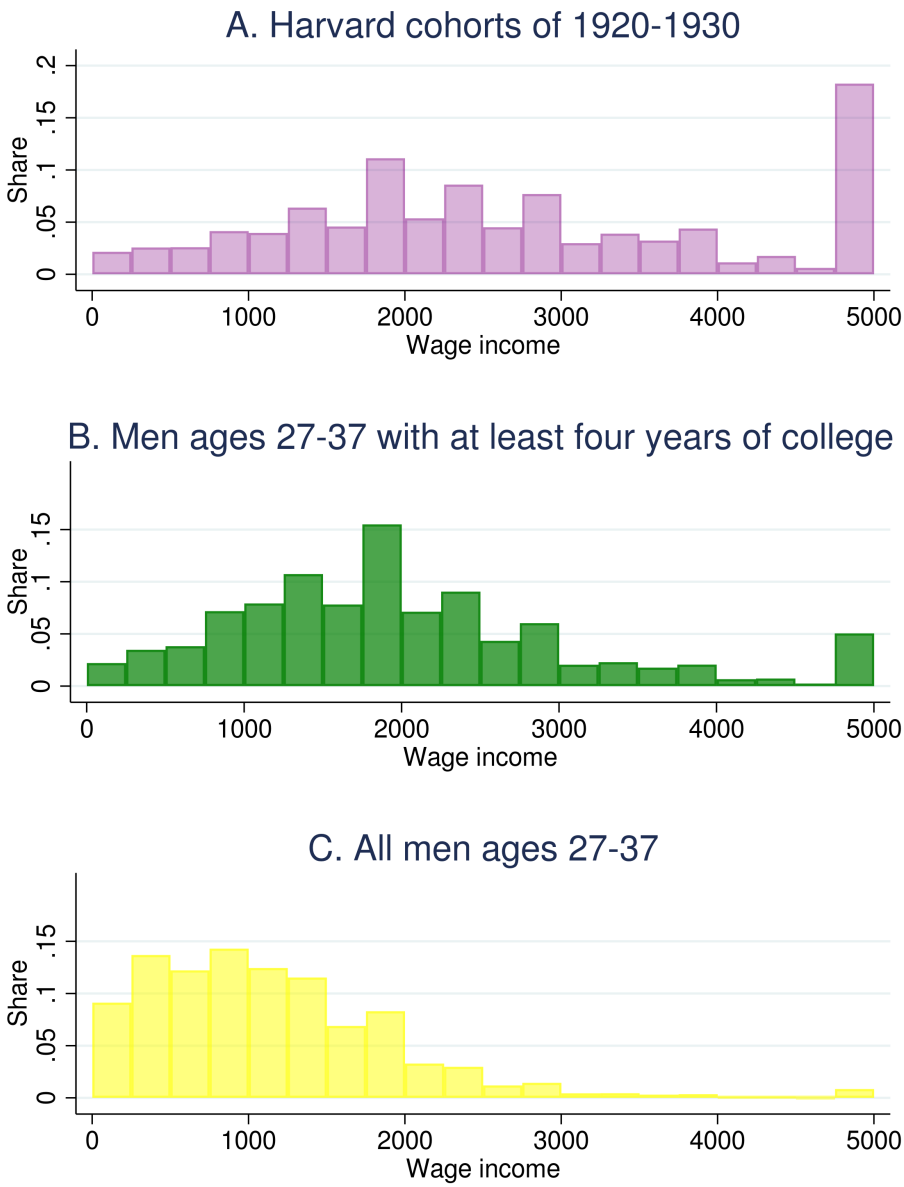
Panel A depicts a student-weighted histogram of randomization block size. Panels B and C depict the range of mean student price (B) and private school share (C) across neighborhoods within each randomization block. Blocks are sorted vertically by maximum neighborhood price (B) or private school share (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. Blocks with fewer than nine students not displayed. In all three panels dashed lines denote 10th and 90th percentiles.

Figure 8: Peer effect estimates by tercile



Estimated effects and 90% confidence intervals of assignment to rooms in the top and middle terciles of the peer neighborhood price distribution, relative to omitted bottom tercile category, for outcomes listed in graph title. Specifications are as in equation 3 but use tercile dummies rather than continuous measures of neighborhood price. See text for outcome definitions.

Figure 9: Wage income of Harvard students and similarly aged men



Histograms depict the distribution of wage income among those who report non-zero wages in the 1940 census. Wage income of Harvard incoming classes of 1920-1930 are displayed along with the income of similarly aged college graduates (men ages 27 through 37 in 1940 with at least four years of college), and all similarly aged men (27 through 37 in 1940). Wage income is top coded at $5000 in the 1940 census.

Tables

Table 1: Data availability by source

Data type	Share non-missing	Universe	N
A. Match rates for available cohorts			
Red book	1.000	Cohorts 1919-35	14382
Have HS data	0.984	Cohorts 1919-35	14382
School address	0.950	Cohorts 1919-25; 1927-35	13628
Class report	0.893	Cohorts 1920-35	13819
Have occupation	0.778	Cohorts 1920-35	13819
1940 census	0.635	Cohorts 1920-30	8851
Census pre-Harvard	0.609	Cohorts 1920-33	11799
B. Match rates for Class Report records			
Have occupation	0.871	Matched to Class Report	12341
1940 census	0.686	Matched to Class Report	8193
Census pre-Harvard	0.665	Matched to Class Report	10812

Data availibility is presented for each individual data source. The second column presents the share of students in the relevent 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.

Table 2: Harvard sample description: background and college outcomes

	All	Private feeder	All non-private	Randomized
A. Demographics				
Have HS data	0.984	1.000	0.979	0.983
Private HS	0.423	1.000	0.235	0.536
Private feeder	0.241	1.000	0.000	0.313
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	0.229
B. Census childhood household demographics				
Have cen. pre-Harvard	0.609	0.700	0.581	0.640
N servants	0.937	1.881	0.620	1.188
S.E. Eur. gen. 1-2	0.098	0.010	0.127	0.051
Have father occ.	0.689	0.650	0.701	0.680
F. Occ.: Doc.	0.101	0.127	0.094	0.120
F. Occ.: Law.	0.109	0.152	0.097	0.136
C. 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	209.2	234.3	198.1	213.0
Peer neighborhood price	213.9	227.9	207.5	213.9
Peer neighborhood rank	0.486	0.551	0.456	0.485
D. Class rank groups				
Rank not listed	0.209	0.216	0.207	0.209
Rank group 1	0.018	0.008	0.021	0.018
Rank group 2	0.069	0.040	0.078	0.065
Rank group 3	0.155	0.105	0.171	0.152
Rank group 4	0.242	0.217	0.250	0.241
Rank group 5	0.376	0.437	0.357	0.376
Rank group 6	0.140	0.193	0.124	0.148
E. First-year activities				
Have activity	0.526	0.707	0.468	0.605
N activities	1.045	1.671	0.847	1.265
Activity leadership position	0.064	0.134	0.042	0.083
Sports	0.367	0.549	0.309	0.430
Social	0.083	0.197	0.047	0.107
Music	0.133	0.163	0.123	0.156
Other activities	0.174	0.249	0.150	0.211
Activity score	0.000	0.534	-0.170	0.157
F. Upper-year social clubs				
1770/Hasty Pudding	0.153	0.393	0.077	0.209
More sel. final club	0.070	0.214	0.025	0.095
Any final club	0.136	0.369	0.062	0.185
Upper-year club score	-0.000	0.722	-0.229	0.152
N	14382	3466	10916	9342

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 private feeder category. See text for more on high school groups. Randomized: students living in on-campus housing with price and occupancy records in a non-singular randomization block. See Section 5 for more on room randomization. Panel A: demographic variables from Red Books. Panel B: childhood demographics from a census record prior to the first year at Harvard. Cohorts 1919, 1934, and 1935 excluded. 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 acheivement, six is lowest. Panel E: College activity records from Red Books. Panel F: Membership in upper year social clubs. Cohort 1935 excluded. See text for detailed variable definitions.

Table 3: Harvard sample description: adult outcomes

	All	Private feeder	All non-private	Randomized
A. Occupations				
Have occupation	0.748	0.752	0.747	0.746
Finance	0.101	0.177	0.077	0.120
Accounting & RE	0.108	0.118	0.105	0.106
Doctor	0.091	0.060	0.101	0.080
Law	0.122	0.099	0.129	0.124
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	0.104	0.122	0.115
Retail	0.136	0.127	0.139	0.130
Occupation score	0.000	0.277	-0.088	0.065
B. Adult social clubs				
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	0.081	0.138
Fraternal order	0.103	0.078	0.111	0.091
Any honor/prof group	0.361	0.307	0.378	0.355
Prof. Association	0.337	0.278	0.355	0.330
Honor society	0.074	0.064	0.077	0.072
Club PHS score	0.000	0.408	-0.133	0.115
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	2958	2680	2852
Has wage income	0.726	0.709	0.732	0.724
Non-wage inc.50+	0.503	0.647	0.456	0.556
Wage inc. 5000+	0.175	0.247	0.153	0.210
N	14382	3466	10916	9342

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 private feeder category. See text for more on high school groups. Randomized: students living in on-campus housing with price and occupancy records in a non-singular randomization block. See Section 5 for more on room randomization. Panel A: Occupation data from 25th Anniversary Class Reports. Cohort 1919 excluded. Panel B: Adult social club and professional association data from 25th Anniversary Class Reports. Cohort 1919 excluded. Panel C: Adult outcomes from the 1940 census. Cohorts 1919 and 1931-35 excluded. See text for detailed variable definitions.

Table 4: Labor market outcomes by academic performance 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. Add less elite final club				
Private feeder	0.004	79	0.032	0.121
	(0.015)	(60)	(0.017)	(0.019)
Class rank	0.001	52	-0.004	0.006
	(0.005)	(20)	(0.005)	(0.006)
Selective final club	-0.006	834	0.246	0.198
	(0.024)	(96)	(0.032)	(0.027)
Less sel. final club	0.064	392	0.088	0.131
	(0.024)	(93)	(0.027)	(0.027)
Sample mean	0.473	2770	0.182	0.511
N	7097	3360	3360	4532
D. 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
E. 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

All models include cohort fixed effects. The sign on class rank groups is reversed so that high numbers correspond with higher academic performance (e.g., the lowest performing group is coded as -6 and the highest performing group is coded -1). Membership in more and less selective final clubs are mutually exclusive. Models A-C include students in cohorts 1920-1930 who are matched with freshman rank records and the 1940 census. Model D includes high school fixed effects and additionally restricts to students from private feeder high schools. Model E. includes family fixed effects and restricts to students who have at least one brother who also attended in cohorts 1920-1930 and is matched with first year class rank and 1940 census records. Robust standard errors in parentheses. Standard errors of model E are clustered at the family level.

Table 5: Test of room randomization

	Year FEs		Price * Year FEs		Block FEs	
A. Balance test						
Room price	193.1	(3.6)	–	–	–	–
Private HS	0.389	(0.024)	-0.025	(0.038)	-0.048	(0.030)
Private feeder	0.289	(0.021)	0.018	(0.034)	–	–
Have HS code	0.292	(0.022)	-0.020	(0.033)	-0.024	(0.030)
From MA	-0.049	(0.022)	0.008	(0.035)	0.027	(0.034)
From NY	0.074	(0.017)	-0.030	(0.028)	-0.039	(0.029)
Have harvard father	0.067	(0.011)	0.000	(0.019)	0.009	(0.022)
Have harvard brother	0.114	(0.018)	0.020	(0.028)	0.007	(0.031)
S.E. Eur. gen. 1-2	-0.050	(0.014)	0.002	(0.018)	0.006	(0.019)
Have father occ.	-0.065	(0.028)	-0.001	(0.048)	-0.027	(0.052)
F. Occ.: Doc.	0.045	(0.024)	0.045	(0.040)	0.073	(0.045)
F. Occ.: Law.	0.084	(0.025)	-0.031	(0.045)	-0.038	(0.052)
Joint balance test [p-value]	[0.000]		[0.900]		[0.505]	
B. Links to other data sources						
Class report link	0.003	(0.011)	-0.009	(0.017)	-0.002	(0.019)
Have occupation	0.000	(0.015)	-0.014	(0.024)	0.003	(0.027)
Cen 1940	0.028	(0.014)	0.025	(0.021)	0.030	(0.025)
C. First stage						
Private HS peer share	0.374	(0.024)	0.286	(0.030)	0.261	(0.032)
Feeder HS peer share	0.273	(0.021)	0.213	(0.027)	0.192	(0.031)
Legacy share	0.065	(0.012)	0.063	(0.015)	0.063	(0.018)
E Eur. peer share	-0.047	(0.014)	-0.047	(0.015)	-0.047	(0.016)
Share nbd price above median	1.146	(0.017)	0.775	(0.018)	0.758	(0.018)
Share nbd price 90th pctl	0.307	(0.016)	0.426	(0.016)	0.461	(0.018)
D. Endogenous peer outcomes						
Peer mean acad. rank	0.390	(0.057)	0.243	(0.069)	0.136	(0.073)
Peer activity index	0.303	(0.049)	0.289	(0.062)	0.281	(0.069)
Peer mean activity count	0.291	(0.065)	0.267	(0.081)	0.306	(0.090)
Peer mean social act.	0.055	(0.014)	0.059	(0.018)	0.053	(0.020)
Peer more sel. final	0.135	(0.014)	0.122	(0.019)	0.106	(0.021)
Peer occ. index	0.397	(0.051)	0.494	(0.066)	0.437	(0.074)
Peer mean finance	0.115	(0.015)	0.146	(0.020)	0.133	(0.023)
Peer mean wage inc.	373.2	(95.9)	416.5	(128.8)	347.7	(139.8)
N	9342		9342		9342	

Coefficients on peer neighborhood price rank from regressions of form given in equation 3. Sample contains 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 different control sets as follows. Year: controls are year dummies only. Price X year: controls for all interactions of per-occupant price and year dummies. Block FEs: controls for all interactions per-occupant price, year dummies, and room occupancy, as well as separate linear 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 null hypothesis that all coefficients in Panel A are zero. Panel B: Matches to later-life data sources. Panel C: Peer neighborhood attributes include price composition which is fixed before randomization and peer student characteristics which is determined by randomization. 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 activity	0.069	0.083	0.055	0.678
	(0.033)	(0.055)	(0.041)	
N activities	0.293	0.684	0.098	0.015
	(0.106)	(0.216)	(0.118)	
Activity leadership position	0.044	0.096	0.015	0.080
	(0.019)	(0.043)	(0.020)	
Social	0.051	0.182	-0.027	0.000
	(0.024)	(0.056)	(0.022)	
Sports	0.024	0.082	-0.014	0.200
	(0.034)	(0.062)	(0.042)	
Music	0.049	0.098	0.028	0.243
	(0.026)	(0.051)	(0.032)	
Other activities	0.025	-0.008	0.056	0.310
	(0.030)	(0.057)	(0.034)	
Activity score	0.207	0.490	0.021	0.011
	(0.079)	(0.165)	(0.084)	
N	9342	2828	6366	
B. Upper-year social clubs				
More sel. final club	0.057	0.147	-0.003	0.009
	(0.021)	(0.055)	(0.017)	
Less sel. final club	-0.021	-0.051	-0.005	0.373
	(0.020)	(0.049)	(0.021)	
1770/Hasty Pudding	0.017	0.103	-0.022	0.061
	(0.028)	(0.061)	(0.029)	
Upper-year club score	0.116	0.352	-0.033	0.038
	(0.076)	(0.172)	(0.076)	
N	8588	2606	5844	
C. Academic outcomes				
Rank group 1	-0.002	-0.010	0.004	0.495
	(0.010)	(0.015)	(0.014)	
Rank group 2	0.014	0.016	0.011	0.882
	(0.015)	(0.024)	(0.021)	
Rank group 3	0.029	0.021	0.032	0.819
	(0.023)	(0.038)	(0.029)	
Rank group 4	-0.001	0.088	-0.047	0.014
	(0.029)	(0.045)	(0.035)	
Rank group 5	0.012	-0.076	0.060	0.050
	(0.031)	(0.059)	(0.038)	
Rank group 6	0.028	0.026	0.010	0.769
	(0.022)	(0.046)	(0.027)	
Rank listed	0.079	0.065	0.071	0.918
	(0.028)	(0.050)	(0.035)	
Class rank year 1	0.026	0.114	0.029	0.685
	(0.102)	(0.175)	(0.127)	
N	7017	2085	4787	

Coefficients on peer neighborhood price rank from regressions of form given in equation 3. All columns include randomization block and large feeder dummies; see section 5.1 for details. The first three columns denote samples. Rows are outcome variables. "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 respectively. 'Activity score" and 'Upper-year club score" are the standardized scores described in section 3.4.4. See text for other variable descriptions. Panel C describes academic outcomes. 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 at least as high as the listed rank group. Private and non-private column sample sizes add up to slightly less than "all" column sample size because there is no variation in peer neighborhood assignment for private-feeder or non-private-feeder students only within some randomization blocks; such students 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. Occupation choice				
Finance	0.018	0.136	-0.042	0.005
	(0.025)	(0.058)	(0.027)	
Accounting & RE	0.019	0.042	0.001	0.446
	(0.024)	(0.048)	(0.029)	
Doctor	0.012	-0.021	0.026	0.275
	(0.022)	(0.033)	(0.028)	
Law	-0.023	-0.005	-0.030	0.649
	(0.026)	(0.044)	(0.034)	
Higher ed.	-0.014	-0.026	-0.017	0.846
	(0.021)	(0.033)	(0.027)	
Teach	-0.016	-0.019	-0.011	0.849
	(0.021)	(0.037)	(0.026)	
Government	0.006	0.021	0.003	0.607
	(0.014)	(0.030)	(0.017)	
Art/pub	-0.012	-0.012	-0.003	0.823
	(0.021)	(0.034)	(0.027)	
Senior management	0.031	0.057	0.012	0.536
	(0.032)	(0.060)	(0.040)	
Low management	0.066	0.048	0.069	0.678
	(0.026)	(0.041)	(0.032)	
Retail	0.006	0.024	-0.004	0.641
	(0.028)	(0.048)	(0.036)	
Occupation score	0.039	0.423	-0.156	0.004
	(0.081)	(0.180)	(0.091)	
N	6922	2075	4707	
B. Adult social and professional organizations				
Any social club	0.066	0.214	0.002	0.008
	(0.035)	(0.066)	(0.044)	
Country club	0.068	0.193	0.010	0.016
	(0.032)	(0.065)	(0.039)	
Gentleman's club	0.017	0.050	-0.001	0.412
	(0.026)	(0.056)	(0.028)	
Fraternal order	0.004	0.033	-0.011	0.341
	(0.022)	(0.036)	(0.028)	
Any honor/prof group	-0.008	0.036	-0.021	0.428
	(0.031)	(0.058)	(0.039)	
Prof. Association	-0.018	0.034	-0.038	0.312
	(0.030)	(0.058)	(0.039)	
Honor society	-0.007	0.026	-0.012	0.363
	(0.018)	(0.034)	(0.023)	
Club PHS score	0.128	0.290	0.041	0.190
	(0.076)	(0.170)	(0.084)	
N	6922	2075	4707	
C. Adult income				
Wage income	-0.0	-350.2	211.4	0.172
	(183.9)	(360.0)	(224.2)	
Wage inc. 5000+	0.018	-0.005	0.061	0.556
	(0.049)	(0.097)	(0.061)	
Non-wage inc.50+	0.064	0.141	0.032	0.337
	(0.052)	(0.096)	(0.063)	
N	2360	683	1587	

Coefficients on peer neighborhood price rank from regressions of form given in equation 3. All columns include randomization block and large feeder dummies; see section 5.1 for details. The first three columns denote samples. Rows are outcome variables. "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 25-year post-graduation occupation and social club outcomes respectively. 'Occupation and Club PHS score" are the standardized scores described in section 3.4.4. See text for other variable descriptions. Panel C reports earnings outcomes from the 1940 census. Effects for census outcomes are estimated on cohorts 1920-1930. Private and non-private column sample sizes add up to slightly less than "all" column sample size because there is no variation in peer neighborhood assignment for private-feeder or non-private-feeder students only within some randomization blocks; such students are omitted from split sample regressions. Standard errors clustered at peer neighborhood level.

Table 8: Peer neighborhood effects on selected outcomes interacted with grades for private feeder students

	Rank group 1-4		Rank group 5-6 or unlisted	
A. Upper-year social clubs				
1770/Hasty Pudding	0.089	(0.049)	-0.002	(0.057)
More sel. final club	0.118	(0.038)	0.045	(0.050)
N	2460		2460	
B. Occupation choice				
Finance	0.112	(0.034)	0.024	(0.048)
Doctor	-0.041	(0.021)	0.020	(0.026)
Law	-0.002	(0.032)	-0.003	(0.032)
Higher ed.	-0.050	(0.026)	0.024	(0.019)
N	2075		2075	
C. Adult social and professional organizations				
Any social club	0.159	(0.045)	0.055	(0.061)
Prof. Association	0.061	(0.042)	-0.026	(0.052)
N	2682		2682	

Coefficients on peer neighborhood price rank from regressions of form given in equation 3. All columns include randomization block and large feeder dummies; see section 5.1 for details. Rows are outcome variables interacted with either an indicator for high grades (first column) or low grades (second column). See section 6.1 for description of outcome variables. Private feeder students only. Standard errors clustered at peer neighborhood level are in the columns to the right of coefficients.