# Online Appendix to "Old Boys' Clubs and Upward Mobility Among the Educational Elite"

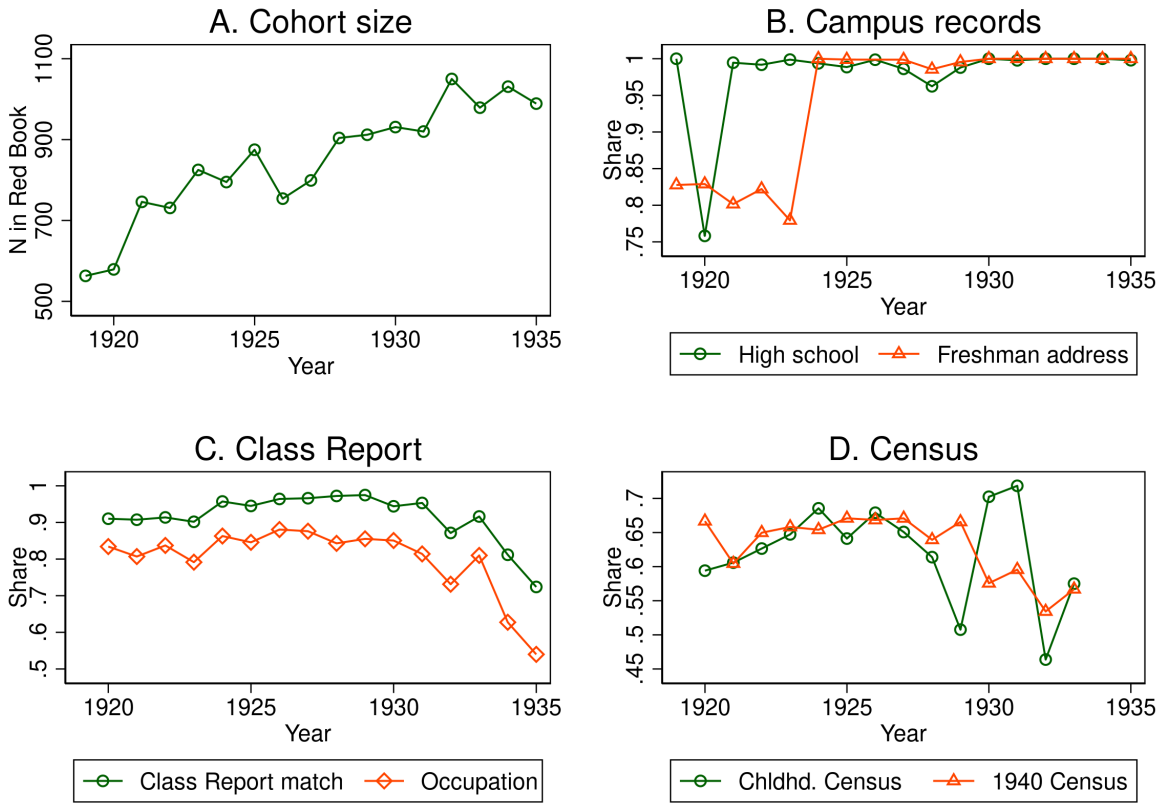
Valerie Michelman
Joseph Price
Seth Zimmerman

August 10, 2021

# A Additional tables and figures

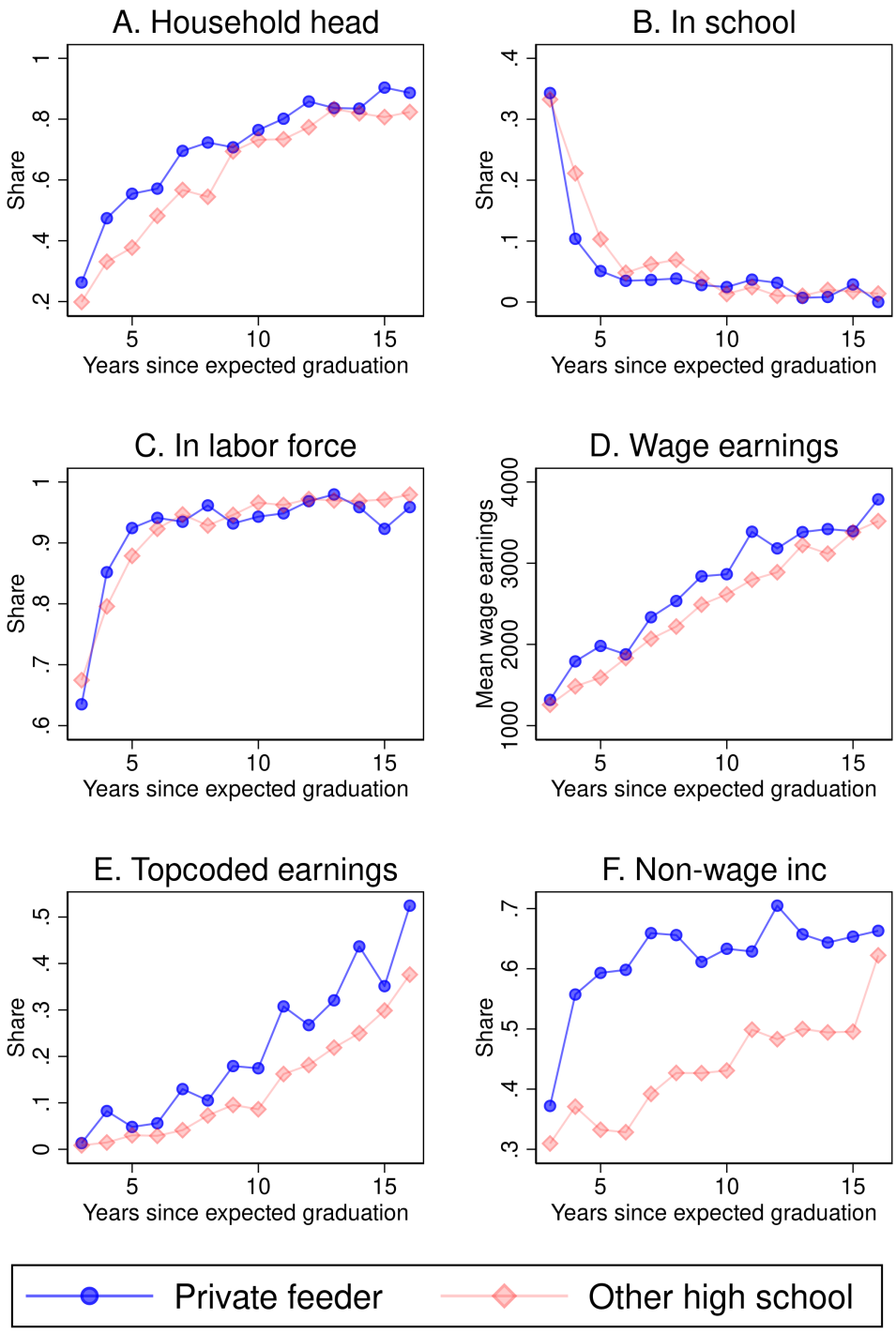
## A.1 Figures

Figure A.1: Data availability by cohort



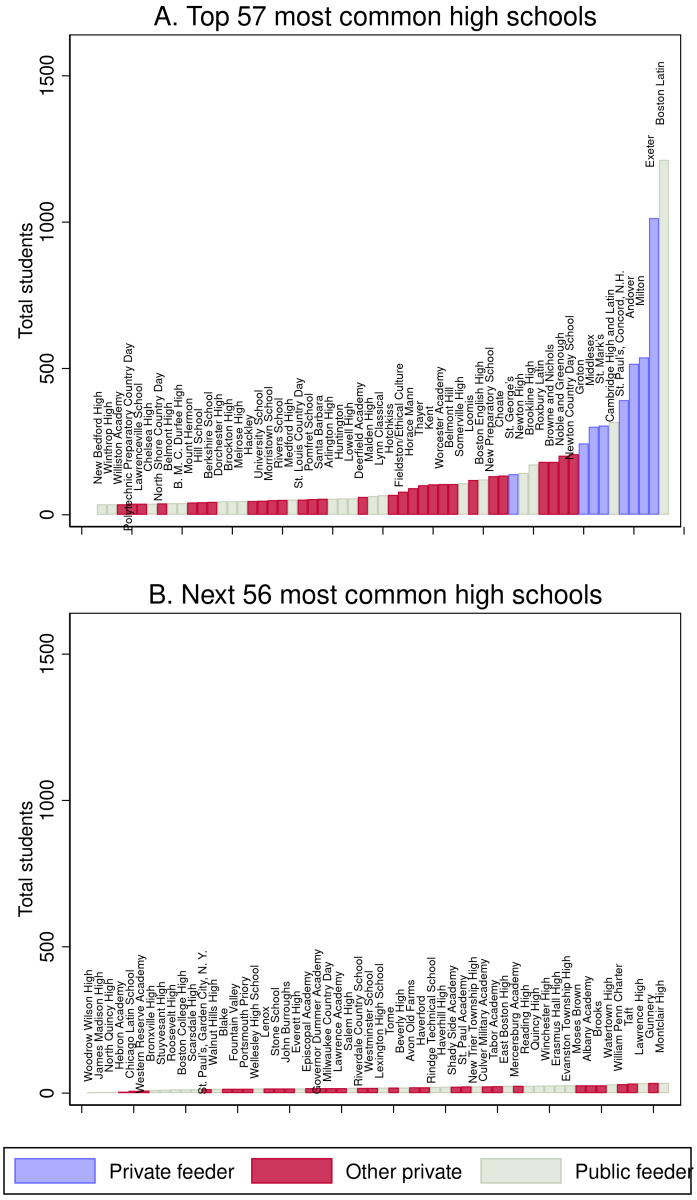

Sample counts and data availability by year entering Harvard. Panel A: Count of students by year. Panel B: Share with data on high school attended and campus address. Panel C: Share matched to 25th Reunion Class Reports and share who are both matched to Class Reports and have non-missing occupation data. 1919 cohort excluded because Class Reports are not available. Panel D: Share matched to childhood and adult Census records (overall, not conditional on Class Report match). We exclude cohorts 1919, 1934, and 1935 from our Census match because Class Reports for those cohorts do not include variables needed for the match. See section 3.3 for details.

Figure A.2: Census outcome profiles by cohort and high school type



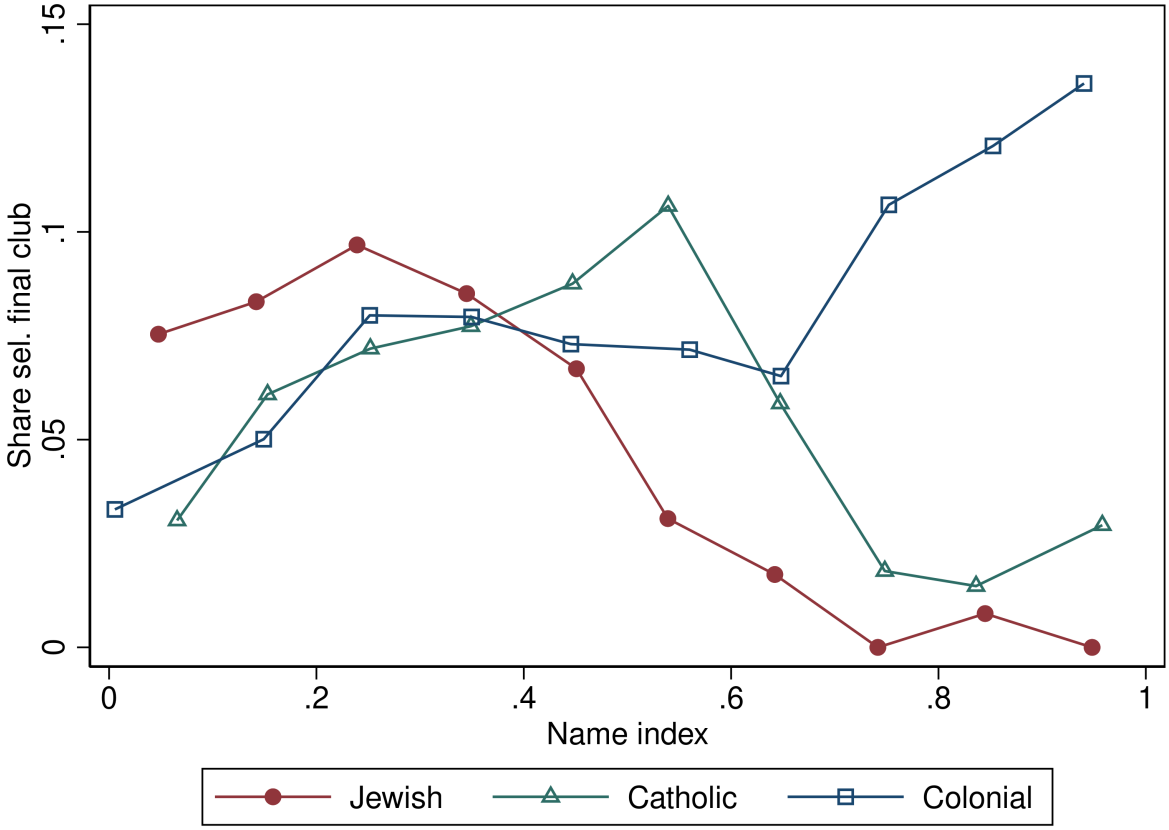

Census outcome profiles presented by cohort. The horizontal axis is years since expected graduation, set equal to 1940 − (entry year + 4). Sample: Harvard students matched to 1940 Census data. See section 3.3 for discussion.

Figure A.3: Counts of Harvard students by high school



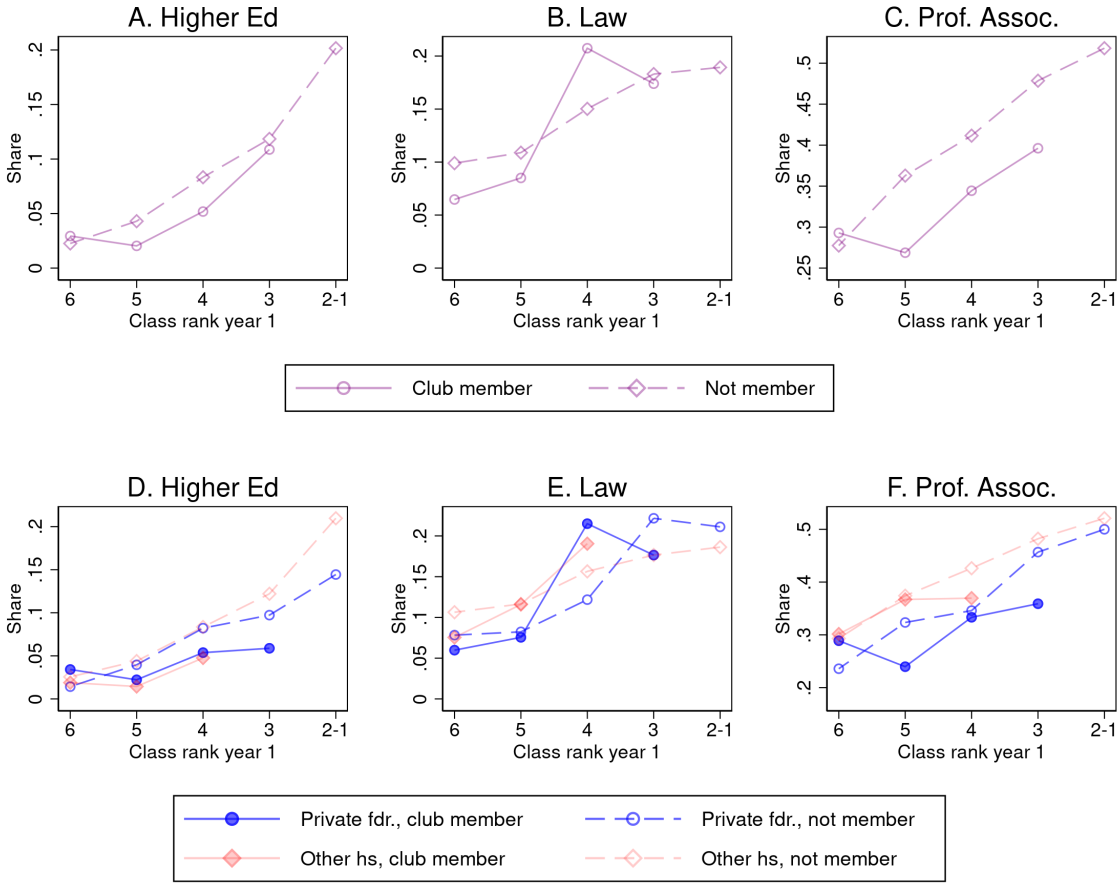

Count of students in Harvard sample universe by high school for each high school that we classify as private feeder, other private, or public feeder. Bar colors denote high school classification. See Section 3.4.1 for school classification details.

Figure A.4: Final club membership by value of name indices



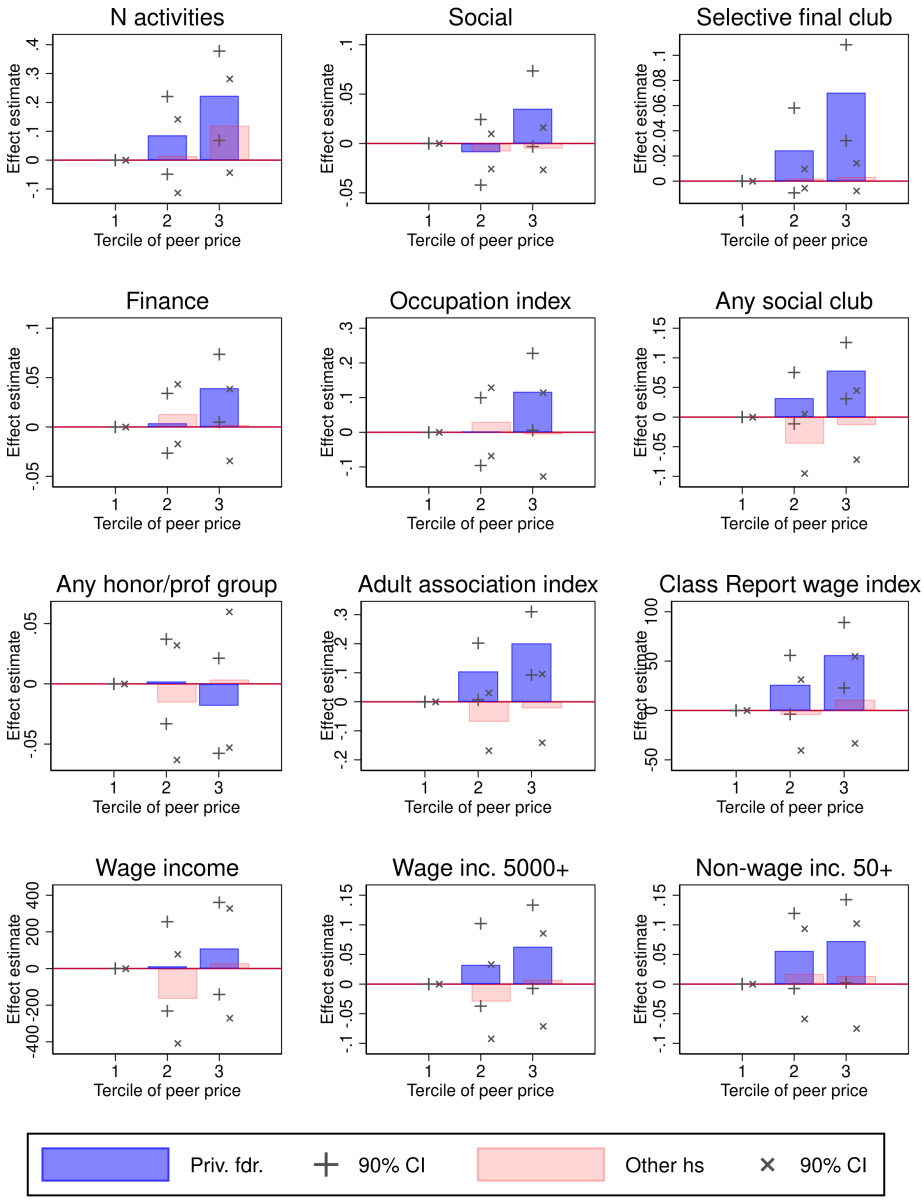

Share of students who report membership in a selective final club by name index value, for different name indices. See section 4.2 for details.

Figure A.5: Additional adult career and social outcomes by academic performance and final club membership



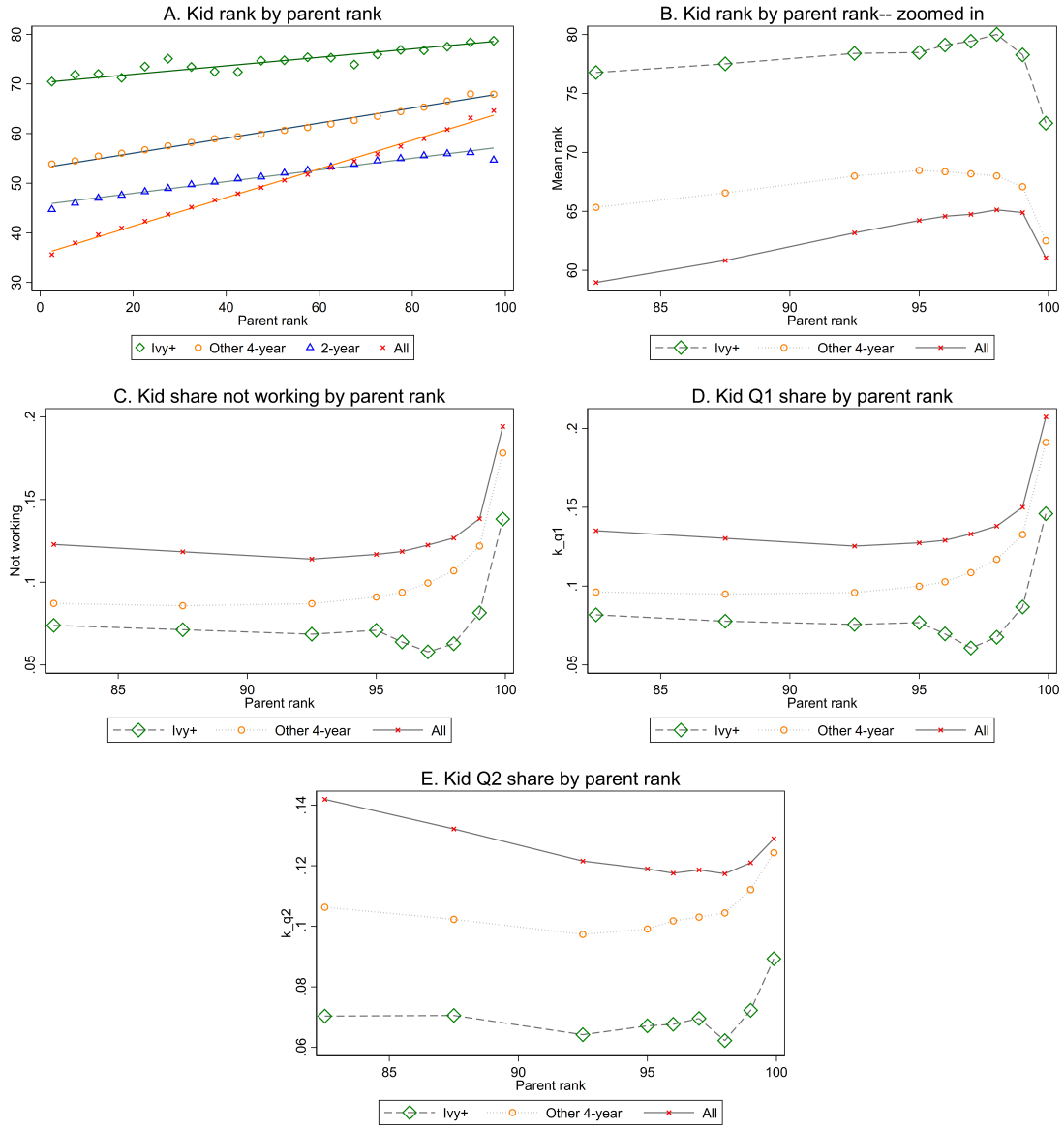

Career outcomes and adult social outcomes by academic performance and membership in a selective final club. Outcome types come from Class Report data and are listed in panel titles. Panels A–C depict adult outcomes for students by freshman academic rank group and selective final club membership. Panels D–F present the same outcomes but also divide students by high school type. We collapse groups 1 and 2 and do not display groups with fewer than 20 students. Sample: students from cohorts 1920–1934 who matched to a Class Report; for occupations we further restrict to students with non-missing occupation data. See section 4.6 for details.

Figure A.6: Peer effect estimates by tercile



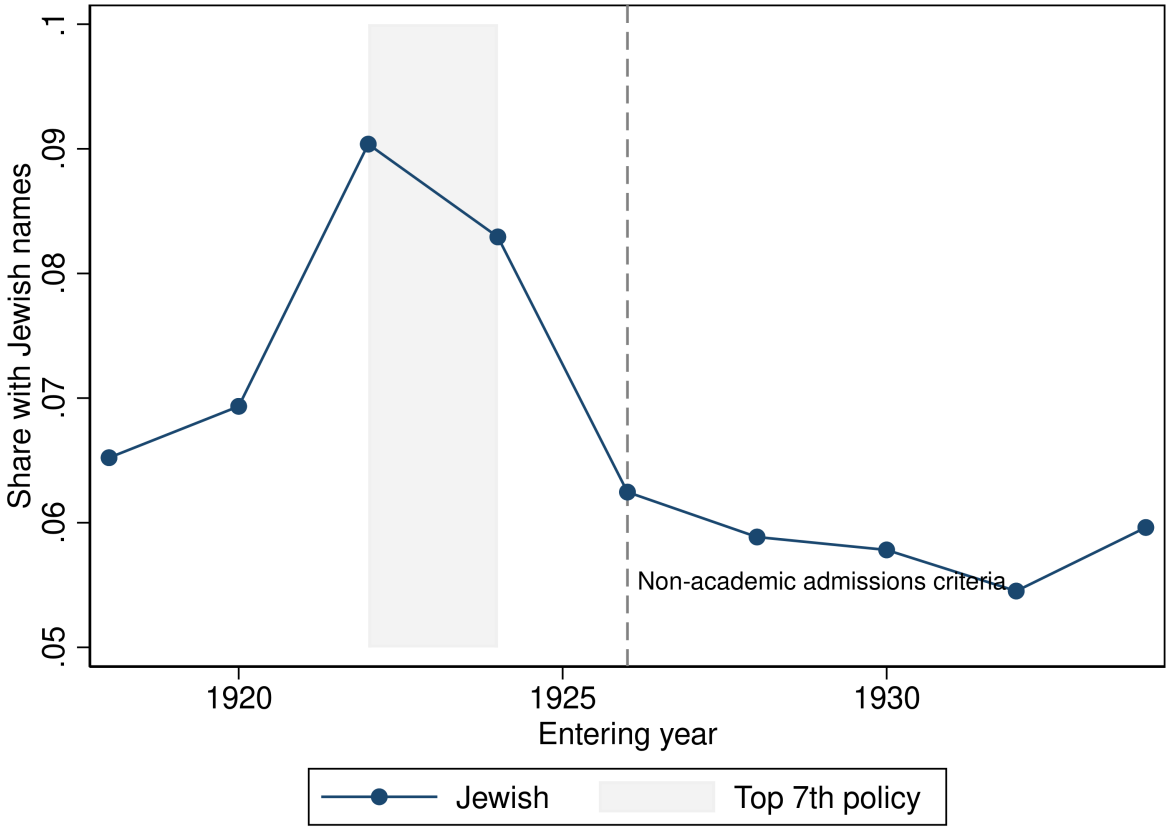

Estimated effects and 90% confidence intervals of assignment to rooms in the top and middle terciles of the peer neighborhood price distribution, relative to omitted bottom tercile category, for outcomes listed in graph title. Specifications are as in equation 2 but use tercile dummies rather than continuous measures of neighborhood price. See text for outcome definitions and section 5.7 for more discussion.

Figure A.7: Additional children's outcomes by parent outcomes for 1980s birth cohorts




Childrens' outcomes by parent income percentile and college type for 1980s birth cohorts. Data are from Chetty et al. (2020). Horizontal axis in all panels is parent income percentile. College types are Ivy+, "Other 4-year" (all other 4-year colleges), "2-year" (2-year colleges), and "All" (all children). Outcomes by panel are as follows. Panel A: mean child rank by parent income percentile across the full distribution of parent income. Each point is a mean within a five percentile bin. This is a public-data replication of Figure III.C in Chetty et al. (2020), but with different group definitions for the most elite schools. We include only Ivy+ schools in the elite category, whereas Chetty et al. (2020) include all schools in the Barron's "most competitive" category. We group non-Ivy+ "most competitive" schools into the "other 4-year" category. Panel B: mean child rank by parent income percentile, zooming in on the top quintile of parent income. Panel C: share of children not working by parent income percentile, zooming in on the top quintile of parent income. Panel D: share of children in bottom income quintile by parent income percentile. Panel E: share of children in second-lowest income quintile by parent income percentile. In panels B through E, points below the 95th percentile are means within centered five percentile bins. Points at the 95th percentile and above are one-percentile bins, with the top percentile split into a 99-99.9 point and a top 0.1% point. See section 6.4 for details.

Figure A.8: The introduction of non-academic admissions criteria




Share of students with characteristically Jewish names (Jewish name index of at least 0.7) by entering class year. Points are averages within two-year bins (e.g., the 1924 bin shows the mean over 1924-25, the 1926 bin for 26-27, and so on). Vertical gray shades denote the period in which Harvard Adopted the "Top 7th" policy, which gave automatic admission to students graduating in the top 7th of their high school class. The vertical dashed line in 1926 corresponds to Harvard's adoption of non-academic admissions criteria for the first time. See section 6.5 for details.

## A.2 Tables

Table A.1: Names with high values of Jewish, Catholic, and Colonial indices

| | Jewish names | | Catholic Names | | Colonial names | |
|---|---|---|---|---|---|---|
| Rank | First | Last | First | Last | First | Last |
| 1 | WOLF | FEINSILVER | SALVATORE | NOBILI | | BLODGET |
| 2 | BARNET | LITVACK | SAVERIO | DIMARZIO | | THAXTER |
| 3 | HIRSH | BORENSTEIN | DANTE | ABRUZZESE | | PLIMPTON |
| 4 | HYMAN | FRADKIN | DOMENIC | MAGLIOZZI | | TILESTON |
| 5 | MANDEL | LIFCHITZ | ETTORE | DIPIETRO | | FROTHINGHAM |
| 6 | MEYER | ISENSTEIN | ANGELO | CIABURRI | | NEWHALL |
| 7 | HYMEN | SHANKMAN | DINO | CAMPOPIANO | | WHEELWRIGHT |
| 8 | GERSHON | LITWACK | ENRICO | BAGLIONE | | DWINNELL |
| 9 | ISIDORE | RUDOFSKY | VIETO | LATORRACA | | BAYLIES |
| 10 | ISRAEL | COOPERSTEIN | COSMO | GALLUCCIO | | ADLINGTON |

The ten common names with highest values of Jewish, Catholic, and Colonial indices. Names are from Harvard student population. Restricted to names with at least 100 occurences in 1920 and 1930 Censuses (combined). Colonial name index is computed only for last names. First and last names are sorted independently– there are Harvard students named "Wolf" and "Feinsilver" but not necessarily "Wolf Feinsilver." See section 3.4.4.

Table A.2: Peer effects on disaggregated business careers

| | All | Private | Non-private | Test |
|---|---|---|---|---|
| *A. Disaggregated business categories* | | | | |
| Accounting | 0.031 | 0.066 | 0.013 | 0.337 |
| | (0.024) | (0.048) | (0.029) | |
| Senior management | 0.036 | 0.058 | 0.014 | 0.551 |
| | (0.032) | (0.061) | (0.040) | |
| Low management | 0.068 | 0.045 | 0.072 | 0.593 |
| | (0.026) | (0.042) | (0.032) | |
| Retail | -0.001 | 0.024 | -0.015 | 0.510 |
| | (0.028) | (0.049) | (0.036) | |
| N | 9343 | 2824 | 6370 | |

Coefficients on peer neighborhood price rank from regressions of form given in equation 2. The first three columns denote samples. Specifications are identical to those reported in Table 7 but with randomization blocks defined only by interactions between per-student price and entry year. Occupancy interactions and controls for high school identifiers are dropped. 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. Reported coefficients are subcategories within the aggregated "business" category in Table 7, Panel B. Standard errors clustered at peer neighborhood level.

Table A.3: Peer effects on short-run outcomes without occupancy controls

| | All | Private | Non-private | Test |
|---|---|---|---|---|
| *A. First-year activities* | | | | |
| Have any activity | 0.083 | 0.100 | 0.063 | 0.564 |
| | (0.031) | (0.053) | (0.039) | |
| N activities | 0.304 | 0.651 | 0.125 | 0.015 |
| | (0.098) | (0.191) | (0.110) | |
| Activity leadership position | 0.036 | 0.088 | 0.007 | 0.043 |
| | (0.017) | (0.037) | (0.018) | |
| Social | 0.062 | 0.200 | -0.016 | 0.000 |
| | (0.021) | (0.048) | (0.019) | |
| Sports | 0.042 | 0.105 | -0.005 | 0.116 |
| | (0.032) | (0.058) | (0.039) | |
| Music | 0.032 | 0.055 | 0.025 | 0.579 |
| | (0.024) | (0.046) | (0.029) | |
| Other activities | 0.038 | -0.008 | 0.071 | 0.158 |
| | (0.028) | (0.049) | (0.032) | |
| First-year activity index | 0.227 | 0.536 | 0.033 | 0.003 |
| | (0.073) | (0.152) | (0.079) | |
| N | 9356 | 2852 | 6390 | |
| *B. Upper-year social clubs* | | | | |
| Selective final club | 0.073 | 0.182 | 0.018 | 0.002 |
| | (0.020) | (0.051) | (0.016) | |
| Less selective final club | -0.016 | -0.060 | -0.002 | 0.182 |
| | (0.018) | (0.040) | (0.020) | |
| Hasty Pudding Inst. 1770 | 0.053 | 0.135 | 0.016 | 0.048 |
| | (0.026) | (0.056) | (0.027) | |
| Upper-year club index | 0.192 | 0.440 | 0.058 | 0.026 |
| | (0.072) | (0.160) | (0.072) | |
| N | 8601 | 2630 | 5864 | |
| *C. First-year academic rank* | | | | |
| Rank group 1 | 0.001 | -0.010 | 0.005 | 0.381 |
| | (0.009) | (0.012) | (0.012) | |
| Rank group 2 | 0.010 | 0.021 | 0.004 | 0.552 |
| | (0.015) | (0.021) | (0.020) | |
| Rank group 3 | -0.003 | 0.006 | -0.010 | 0.720 |
| | (0.022) | (0.036) | (0.027) | |
| Rank group 4 | -0.008 | 0.074 | -0.032 | 0.035 |
| | (0.026) | (0.041) | (0.032) | |
| Rank group 5 | 0.030 | -0.024 | 0.071 | 0.140 |
| | (0.029) | (0.054) | (0.035) | |
| Rank group 6 | 0.028 | 0.010 | 0.016 | 0.890 |
| | (0.020) | (0.039) | (0.024) | |
| Rank listed year 1 | 0.059 | 0.077 | 0.054 | 0.673 |
| | (0.026) | (0.046) | (0.032) | |
| Class rank year 1 | -0.048 | 0.091 | -0.076 | 0.375 |
| | (0.093) | (0.154) | (0.116) | |
| N | 7063 | 2123 | 4832 | |

Coefficients on peer neighborhood price rank from regressions of form given in equation 2. Specifications are identical to those reported in Table 6 but with randomization blocks defined only by interactions between per-student price and entry year. Occupancy interactions and controls for high school identifiers are dropped. 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. "First-year activity index" and "upper-year club index" are the standardized indices of the association between activities and private high school background described in section 3.4.3. Other variables are indicators for a given activity type unless stated otherwise. Panel C describes academic outcomes in the first year. Class rank is a continuous variable from one through six, with one the best and six the worst. The other outcomes are dummies for having grades in the listed rank group and being listed at all. Standard errors clustered at peer neighborhood level.

Table A.4: Peer effects on long-run outcomes without occupancy controls

| | All | Private | Non-private | Test |
|---|---|---|---|---|
| *A. Adult associations* | | | | |
| Any social club | 0.078 | 0.211 | 0.023 | 0.010 |
| | (0.033) | (0.061) | (0.041) | |
| Country club | 0.087 | 0.215 | 0.034 | 0.008 |
| | (0.030) | (0.058) | (0.037) | |
| Gentleman's club | 0.024 | 0.052 | 0.006 | 0.413 |
| | (0.024) | (0.051) | (0.026) | |
| Fraternal order | 0.009 | 0.005 | 0.010 | 0.911 |
| | (0.020) | (0.032) | (0.026) | |
| Any honor/prof group | -0.031 | 0.031 | -0.059 | 0.190 |
| | (0.029) | (0.056) | (0.037) | |
| Prof. Association | -0.039 | 0.019 | -0.071 | 0.179 |
| | (0.029) | (0.055) | (0.037) | |
| Honor society | -0.020 | 0.011 | -0.024 | 0.372 |
| | (0.017) | (0.032) | (0.022) | |
| Adult association index | 0.185 | 0.346 | 0.113 | 0.175 |
| | (0.073) | (0.154) | (0.081) | |
| N | 8193 | 2504 | 5574 | |
| *B. Occupation choice* | | | | |
| Finance | 0.022 | 0.139 | -0.037 | 0.002 |
| | (0.023) | (0.051) | (0.025) | |
| Medicine | 0.002 | -0.008 | 0.016 | 0.548 |
| | (0.021) | (0.031) | (0.027) | |
| Higher ed. | -0.018 | -0.025 | -0.023 | 0.961 |
| | (0.019) | (0.028) | (0.025) | |
| Law | -0.017 | -0.001 | -0.021 | 0.678 |
| | (0.024) | (0.038) | (0.032) | |
| Business | 0.047 | 0.046 | 0.031 | 0.844 |
| | (0.036) | (0.067) | (0.044) | |
| Teach | -0.013 | -0.035 | -0.002 | 0.413 |
| | (0.019) | (0.033) | (0.025) | |
| Government | 0.000 | 0.012 | -0.004 | 0.613 |
| | (0.014) | (0.025) | (0.017) | |
| Art/pub | -0.009 | -0.034 | 0.004 | 0.355 |
| | (0.020) | (0.034) | (0.024) | |
| Occupation index | 0.070 | 0.424 | -0.119 | 0.003 |
| | (0.075) | (0.164) | (0.086) | |
| N | 7067 | 2128 | 4826 | |
| *C. Adult income* | | | | |
| Wage income | 24.8 | -315.9 | 207.1 | 0.189 |
| | (184.1) | (349.5) | (223.1) | |
| Wage inc. 5000+ | 0.052 | 0.001 | 0.096 | 0.385 |
| | (0.050) | (0.097) | (0.059) | |
| Non-wage inc. 50+ | 0.061 | 0.081 | 0.049 | 0.780 |
| | (0.050) | (0.094) | (0.061) | |
| Class Report wage index | 44.9 | 142.5 | -13.0 | 0.005 |
| | (25.9) | (45.2) | (33.1) | |
| N | 2428 | 722 | 1644 | |

Coefficients on peer neighborhood price rank from regressions of form given in equation 2. The first three columns denote samples. Specifications are identical to those reported in Table 7 but with randomization blocks defined only by interactions between per-student price and entry year. Occupancy interactions and controls for high school identifiers are dropped. 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 adult social club and occupation outcomes, respectively. "Adult association index" and "Occupation index" are standardized indices of the association between adult outcomes and private high school background. See section 3.4.3. Other variables are indicators for outcomes of the listed type. The sample in Panels A and B is students matched class reports; Panel B further restricts to students with occupation outcomes. Panel C reports labor market outcomes from the 1940 census. Sample is students matched to Census wage records in the 1920-30 entering cohorts. Standard errors clustered at peer neighborhood level.

Table A.5: Peer effects on short-run outcomes without large high school controls

| | All | Private | Non-private | Test |
|---|---|---|---|---|
| *A. First-year activities* | | | | |
| Have any activity | 0.068 | 0.088 | 0.061 | 0.695 |
| | (0.036) | (0.057) | (0.043) | |
| N activities | 0.279 | 0.717 | 0.090 | 0.014 |
| | (0.118) | (0.226) | (0.127) | |
| Activity leadership position | 0.042 | 0.100 | 0.011 | 0.064 |
| | (0.020) | (0.044) | (0.021) | |
| Social | 0.048 | 0.193 | -0.031 | 0.000 |
| | (0.026) | (0.058) | (0.023) | |
| Sports | 0.024 | 0.088 | -0.010 | 0.206 |
| | (0.037) | (0.064) | (0.043) | |
| Music | 0.050 | 0.114 | 0.028 | 0.165 |
| | (0.027) | (0.053) | (0.033) | |
| Other activities | 0.023 | -0.015 | 0.057 | 0.259 |
| | (0.031) | (0.058) | (0.035) | |
| First-year activity index | 0.199 | 0.527 | 0.012 | 0.008 |
| | (0.089) | (0.174) | (0.090) | |
| N | 9343 | 2824 | 6370 | |
| *B. Upper-year social clubs* | | | | |
| Selective final club | 0.067 | 0.182 | -0.001 | 0.006 |
| | (0.026) | (0.065) | (0.018) | |
| Less selective final club | -0.031 | -0.062 | -0.013 | 0.353 |
| | (0.021) | (0.050) | (0.022) | |
| Hasty Pudding Inst. 1770 | 0.021 | 0.137 | -0.024 | 0.036 |
| | (0.034) | (0.072) | (0.031) | |
| Upper-year club index | 0.126 | 0.443 | -0.040 | 0.030 |
| | (0.098) | (0.211) | (0.082) | |
| N | 8589 | 2603 | 5847 | |
| *C. First-year academic rank* | | | | |
| Rank group 1 | 0.001 | -0.011 | 0.003 | 0.497 |
| | (0.011) | (0.015) | (0.015) | |
| Rank group 2 | 0.018 | 0.015 | 0.015 | 0.989 |
| | (0.016) | (0.025) | (0.021) | |
| Rank group 3 | 0.025 | 0.007 | 0.033 | 0.589 |
| | (0.024) | (0.039) | (0.030) | |
| Rank group 4 | -0.007 | 0.088 | -0.053 | 0.012 |
| | (0.029) | (0.045) | (0.036) | |
| Rank group 5 | 0.019 | -0.063 | 0.069 | 0.064 |
| | (0.032) | (0.060) | (0.039) | |
| Rank group 6 | 0.019 | 0.018 | 0.002 | 0.773 |
| | (0.023) | (0.046) | (0.028) | |
| Rank listed year 1 | 0.075 | 0.053 | 0.069 | 0.798 |
| | (0.028) | (0.050) | (0.035) | |
| Class rank year 1 | 0.038 | 0.093 | 0.041 | 0.810 |
| | (0.106) | (0.179) | (0.129) | |
| N | 7035 | 2085 | 4802 | |

Coefficients on peer neighborhood price rank from regressions of form given in equation 2. Specifications are identical to those reported in Table 6 but with controls for high school identifiers dropped. 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. "First-year activity index" and "upper-year club index" are the standardized indices of the association between activities and private high school background described in section 3.4.3. Other variables are indicators for a given activity type unless stated otherwise. Panel C describes academic outcomes in the first year. Class rank is a continuous variable from one through six, with one the best and six the worst. The other outcomes are dummies for having grades in the listed rank group and being listed at all. Standard errors clustered at peer neighborhood level.

Table A.6: Peer effects on long-run outcomes without large high school controls

| | All | Private | Non-private | Test |
|---|---|---|---|---|
| *A. Adult associations* | | | | |
| Any social club | 0.069 | 0.249 | -0.002 | 0.002 |
| | (0.036) | (0.066) | (0.046) | |
| Country club | 0.073 | 0.224 | 0.004 | 0.004 |
| | (0.032) | (0.064) | (0.040) | |
| Gentleman's club | 0.023 | 0.063 | 0.005 | 0.354 |
| | (0.026) | (0.057) | (0.028) | |
| Fraternal order | 0.001 | 0.015 | -0.000 | 0.755 |
| | (0.022) | (0.037) | (0.029) | |
| Any honor/prof group | -0.012 | 0.040 | -0.024 | 0.382 |
| | (0.032) | (0.059) | (0.041) | |
| Prof. Association | -0.020 | 0.030 | -0.036 | 0.363 |
| | (0.032) | (0.059) | (0.040) | |
| Honor society | -0.013 | 0.029 | -0.018 | 0.267 |
| | (0.019) | (0.035) | (0.024) | |
| Adult association index | 0.146 | 0.356 | 0.041 | 0.104 |
| | (0.079) | (0.172) | (0.087) | |
| N | 8178 | 2476 | 5551 | |
| *B. Occupation choice* | | | | |
| Finance | 0.021 | 0.156 | -0.051 | 0.001 |
| | (0.025) | (0.059) | (0.028) | |
| Medicine | 0.000 | -0.019 | 0.015 | 0.435 |
| | (0.023) | (0.033) | (0.030) | |
| Higher ed. | -0.020 | -0.032 | -0.024 | 0.862 |
| | (0.021) | (0.033) | (0.028) | |
| Law | -0.032 | -0.020 | -0.032 | 0.833 |
| | (0.027) | (0.044) | (0.035) | |
| Business | 0.063 | 0.065 | 0.045 | 0.830 |
| | (0.040) | (0.076) | (0.049) | |
| Teach | -0.016 | -0.024 | -0.005 | 0.684 |
| | (0.021) | (0.037) | (0.027) | |
| Government | 0.005 | 0.013 | 0.008 | 0.901 |
| | (0.015) | (0.031) | (0.018) | |
| Art/pub | -0.006 | -0.019 | 0.006 | 0.552 |
| | (0.021) | (0.034) | (0.027) | |
| Occupation index | 0.074 | 0.501 | -0.162 | 0.001 |
| | (0.083) | (0.183) | (0.094) | |
| N | 7039 | 2094 | 4797 | |
| *C. Adult income* | | | | |
| Wage income | 1.1 | -384.1 | 199.6 | 0.168 |
| | (191.6) | (364.6) | (236.4) | |
| Wage inc. 5000+ | 0.028 | -0.019 | 0.072 | 0.428 |
| | (0.050) | (0.099) | (0.061) | |
| Non-wage inc. 50+ | 0.067 | 0.111 | 0.047 | 0.591 |
| | (0.053) | (0.099) | (0.065) | |
| Class Report wage index | 39.3 | 136.2 | -14.9 | 0.009 |
| | (26.7) | (48.4) | (33.9) | |
| N | 2396 | 697 | 1611 | |

Coefficients on peer neighborhood price rank from regressions of form given in equation 2. Specifications are identical to those reported in Table 7 but with controls for high school identifiers dropped. "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 adult social club and occupation outcomes, respectively. "Adult association index" and "Occupation index" are standardized indices of the association between adult outcomes and private high school background. See section 3.4.3. Other variables are indicators for outcomes of the listed type. The sample in Panels A and B is students matched class reports; Panel B further restricts to students with occupation outcomes. Panel C reports labor market outcomes from the 1940 census. Sample is students matched to Census wage records in the 1920-30 entering cohorts. Standard errors clustered at peer neighborhood level.

Table A.7: Peer effects on short-run outcomes with alternate private high school classifications

| | Private feeder | More expansive | Less expansive | Private non-feeder |
|---|---|---|---|---|
| *A. First-year activities* | | | | |
| Have any activity | 0.080 | 0.087 | 0.054 | 0.069 |
| | (0.057) | (0.053) | (0.060) | (0.071) |
| N activities | 0.648 | 0.571 | 0.630 | -0.043 |
| | (0.222) | (0.206) | (0.231) | (0.208) |
| Activity leadership position | 0.098 | 0.087 | 0.098 | -0.053 |
| | (0.044) | (0.039) | (0.045) | (0.039) |
| Social | 0.183 | 0.124 | 0.176 | -0.031 |
| | (0.057) | (0.053) | (0.059) | (0.044) |
| Sports | 0.085 | 0.070 | 0.072 | -0.029 |
| | (0.064) | (0.058) | (0.066) | (0.070) |
| Music | 0.095 | 0.098 | 0.101 | 0.026 |
| | (0.052) | (0.047) | (0.054) | (0.053) |
| Other activities | -0.027 | -0.049 | -0.040 | -0.062 |
| | (0.058) | (0.053) | (0.059) | (0.059) |
| First-year activity index | 0.494 | 0.396 | 0.477 | -0.089 |
| | (0.171) | (0.158) | (0.180) | (0.159) |
| N | 2824 | 3468 | 2701 | 2341 |
| *B. Academic outcomes* | | | | |
| Rank group 1 | -0.012 | -0.002 | -0.013 | 0.023 |
| | (0.015) | (0.013) | (0.016) | (0.018) |
| Rank group 2 | 0.017 | 0.032 | 0.017 | 0.030 |
| | (0.025) | (0.022) | (0.025) | (0.029) |
| Rank group 3 | 0.006 | 0.016 | -0.005 | 0.002 |
| | (0.039) | (0.036) | (0.040) | (0.042) |
| Rank group 4 | 0.090 | 0.049 | 0.098 | -0.114 |
| | (0.046) | (0.044) | (0.048) | (0.060) |
| Rank group 5 | -0.067 | -0.038 | -0.076 | 0.067 |
| | (0.061) | (0.057) | (0.062) | (0.065) |
| Rank group 6 | 0.013 | 0.016 | 0.003 | 0.040 |
| | (0.046) | (0.043) | (0.048) | (0.048) |
| Rank listed year 1 | 0.047 | 0.072 | 0.025 | 0.048 |
| | (0.051) | (0.046) | (0.051) | (0.063) |
| Class rank year 1 | 0.101 | 0.144 | 0.098 | 0.074 |
| | (0.180) | (0.165) | (0.182) | (0.215) |
| N | 2085 | 2581 | 2000 | 1692 |
| *C. Upper-year social clubs* | | | | |
| Selective final club | 0.167 | 0.135 | 0.169 | 0.014 |
| | (0.056) | (0.049) | (0.059) | (0.041) |
| Less selective final club | -0.075 | -0.056 | -0.079 | -0.003 |
| | (0.050) | (0.045) | (0.052) | (0.047) |
| Hasty Pudding Inst. 1770 | 0.105 | 0.088 | 0.069 | -0.038 |
| | (0.062) | (0.056) | (0.063) | (0.064) |
| Upper-year club index | 0.357 | 0.300 | 0.303 | -0.001 |
| | (0.176) | (0.158) | (0.181) | (0.174) |
| N | 2603 | 3207 | 2491 | 2136 |

Coefficients on peer neighborhood price rank from regressions of form given in equation 2. Specifications are identical to those reported in Table 6 but for alternate high school classifications. "Private feeder" column repeats estimates from Table 6 for private feeder sample. "More expansive" column reports results for the sample of the 13 private schools that sent the most students to Harvard over the period. "Less expansive" reports results for the sample of the seven private schools that sent the most students. "Private non-feeder" reports results for all private schools not part of the main feeder sample. Panel A reports first-year activity outcomes. "First year activity index" and "Upper-year club index" are the standardized private school indices described in section 3.4.3. See text for other variable descriptions. Standard errors clustered at peer neighborhood level.

Table A.8: Peer effects on long-run outcomes with alternate private high school classifications

| | Private feeder | More expansive | Less expansive | Private non-feeder |
|---|---|---|---|---|
| *A. Adult associations* | | | | |
| Any social club | 0.241 | 0.198 | 0.245 | -0.000 |
| | (0.067) | (0.060) | (0.070) | (0.076) |
| Country club | 0.219 | 0.209 | 0.222 | 0.046 |
| | (0.066) | (0.057) | (0.068) | (0.072) |
| Gentleman's club | 0.053 | 0.046 | 0.047 | 0.029 |
| | (0.056) | (0.049) | (0.058) | (0.050) |
| Fraternal order | 0.020 | 0.007 | 0.007 | 0.007 |
| | (0.037) | (0.035) | (0.039) | (0.041) |
| Any honor/prof group | 0.039 | 0.018 | 0.033 | -0.103 |
| | (0.060) | (0.053) | (0.061) | (0.065) |
| Prof. Association | 0.032 | 0.009 | 0.019 | -0.098 |
| | (0.060) | (0.053) | (0.061) | (0.066) |
| Honor society | 0.020 | 0.026 | 0.024 | 0.002 |
| | (0.035) | (0.031) | (0.036) | (0.037) |
| Adult association index | 0.328 | 0.327 | 0.335 | 0.195 |
| | (0.171) | (0.147) | (0.175) | (0.155) |
| N | 2476 | 3038 | 2372 | 2045 |
| *B. Occupation choice* | | | | |
| Finance | 0.144 | 0.105 | 0.149 | -0.106 |
| | (0.058) | (0.051) | (0.059) | (0.051) |
| Medicine | -0.018 | -0.030 | -0.015 | -0.047 |
| | (0.034) | (0.031) | (0.035) | (0.045) |
| Higher ed. | -0.026 | -0.004 | -0.009 | 0.037 |
| | (0.034) | (0.031) | (0.034) | (0.047) |
| Law | -0.009 | -0.013 | -0.027 | -0.007 |
| | (0.044) | (0.038) | (0.046) | (0.055) |
| Business | 0.071 | 0.094 | 0.059 | 0.085 |
| | (0.077) | (0.067) | (0.078) | (0.081) |
| Teach | -0.017 | -0.008 | -0.019 | 0.025 |
| | (0.038) | (0.034) | (0.039) | (0.047) |
| Government | 0.012 | 0.012 | 0.007 | 0.028 |
| | (0.031) | (0.027) | (0.031) | (0.028) |
| Art/pub | -0.017 | -0.039 | -0.013 | -0.004 |
| | (0.035) | (0.031) | (0.036) | (0.037) |
| Occupation index | 0.450 | 0.331 | 0.462 | -0.305 |
| | (0.181) | (0.159) | (0.183) | (0.166) |
| N | 2094 | 2598 | 2006 | 1745 |
| *C. Adult income* | | | | |
| Wage income | -440.5 | -127.7 | -437.5 | 595.1 |
| | (367.8) | (310.2) | (378.0) | (416.4) |
| Wage inc. 5000+ | -0.029 | 0.016 | -0.052 | 0.248 |
| | (0.099) | (0.088) | (0.100) | (0.119) |
| Non-wage inc. 50+ | 0.081 | 0.084 | 0.098 | 0.119 |
| | (0.099) | (0.096) | (0.101) | (0.117) |
| Class Report wage index | 128.8 | 101.2 | 115.2 | -54.8 |
| | (48.1) | (44.1) | (50.5) | (54.8) |
| N | 697 | 881 | 665 | 523 |

Coefficients on peer neighborhood price rank from regressions of form given in equation 2. Specifications are identical to those reported in Table 7 but for alternate high school classifications. "Private feeder" column repeats estimates from Table 6 for private feeder sample. "More expansive" column reports results for the sample of the 13 private schools that sent the most students to Harvard over the period. "Less expansive" reports results for the sample of the seven private schools that sent the most students. "Private non-feeder" reports results for all private schools not part of the main feeder sample. "Adult association index" and "Occupation index" are standardized private school indices described in section 3.4.3. See text for other variable descriptions. Standard errors clustered at peer neighborhood level.

Table A.9: Peer neighborhood effects on short-run outcomes excluding cohorts 1919-1921

| | All | Private | Non-private | Test |
|---|---|---|---|---|
| *A. First-year activities* | | | | |
| Have any activity | 0.074 | 0.081 | 0.066 | 0.846 |
| | (0.037) | (0.063) | (0.046) | |
| N activities | 0.282 | 0.397 | 0.209 | 0.466 |
| | (0.114) | (0.230) | (0.128) | |
| Activity leadership position | 0.036 | 0.104 | 0.003 | 0.047 |
| | (0.021) | (0.047) | (0.022) | |
| Social | 0.048 | 0.157 | -0.010 | 0.006 |
| | (0.025) | (0.058) | (0.022) | |
| Sports | 0.039 | 0.063 | 0.015 | 0.570 |
| | (0.037) | (0.071) | (0.045) | |
| Music | 0.029 | 0.027 | 0.029 | 0.980 |
| | (0.029) | (0.054) | (0.035) | |
| Other activities | 0.045 | 0.008 | 0.074 | 0.334 |
| | (0.033) | (0.061) | (0.038) | |
| First-year activity index | 0.228 | 0.461 | 0.079 | 0.068 |
| | (0.089) | (0.189) | (0.094) | |
| N | 8018 | 2414 | 5473 | |
| *B. Upper-year social clubs* | | | | |
| Selective final club | 0.069 | 0.145 | 0.016 | 0.032 |
| | (0.022) | (0.058) | (0.017) | |
| Less selective final club | -0.024 | -0.055 | -0.008 | 0.424 |
| | (0.022) | (0.055) | (0.023) | |
| Hasty Pudding Inst. 1770 | 0.009 | 0.057 | -0.004 | 0.394 |
| | (0.030) | (0.068) | (0.030) | |
| Upper-year club index | 0.128 | 0.307 | 0.019 | 0.164 |
| | (0.085) | (0.194) | (0.083) | |
| N | 7264 | 2193 | 4950 | |
| *C. First-year academic rank* | | | | |
| Rank group 1 | 0.000 | -0.010 | 0.005 | 0.518 |
| | (0.011) | (0.015) | (0.016) | |
| Rank group 2 | 0.018 | 0.019 | 0.013 | 0.861 |
| | (0.016) | (0.027) | (0.021) | |
| Rank group 3 | 0.020 | 0.008 | 0.027 | 0.715 |
| | (0.026) | (0.042) | (0.032) | |
| Rank group 4 | -0.021 | 0.086 | -0.075 | 0.007 |
| | (0.031) | (0.049) | (0.039) | |
| Rank group 5 | 0.020 | -0.071 | 0.071 | 0.063 |
| | (0.034) | (0.064) | (0.042) | |
| Rank group 6 | 0.030 | -0.002 | 0.027 | 0.586 |
| | (0.022) | (0.045) | (0.027) | |
| Rank listed year 1 | 0.068 | 0.031 | 0.068 | 0.571 |
| | (0.031) | (0.055) | (0.038) | |
| Class rank year 1 | -0.006 | 0.133 | -0.034 | 0.460 |
| | (0.107) | (0.188) | (0.133) | |
| N | 6252 | 1862 | 4264 | |

Coefficients on peer neighborhood price rank from regressions of form given in equation 2. Cohorts 1919, 1920, and 1921 are excluded. The first three columns denote samples. Rows are outcome variables. All specifications include randomization block and dummies for large feeder high schools; see section 5.1 for details."Test" column reports the p-value from a test of the null that the coefficients reported in the private and non-private columns are equal. Panels A and B report first-year activity outcomes and upper-year club outcomes respectively. "First-year activity index" and "upper-year club index" are the standardized indices of the association between activities and private high school background described in section 3.4.3. Other variables are indicators for a given activity type unless stated otherwise. Panel C describes academic outcomes in the first year. Class rank is a continuous variable from one through six, with one the best and six the worst. The other outcomes are dummies for having grades in the listed rank group and being listed at all. 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; students in these blocks are omitted from split sample regressions. Standard errors clustered at peer neighborhood level.

Table A.10: Peer neighborhood effects on long-run outcomes excluding cohorts 1919-1921

| | All | Private | Non-private | Test |
|---|---|---|---|---|
| *A. Adult associations* | | | | |
| Any social club | 0.067 | 0.261 | -0.005 | 0.002 |
| | (0.038) | (0.070) | (0.047) | |
| Country club | 0.082 | 0.238 | 0.018 | 0.007 |
| | (0.034) | (0.069) | (0.042) | |
| Gentleman's club | 0.018 | 0.037 | 0.010 | 0.685 |
| | (0.027) | (0.060) | (0.029) | |
| Fraternal order | -0.001 | 0.023 | -0.009 | 0.519 |
| | (0.023) | (0.039) | (0.030) | |
| Any honor/prof group | -0.016 | 0.048 | -0.026 | 0.365 |
| | (0.035) | (0.067) | (0.044) | |
| Prof. Association | -0.024 | 0.045 | -0.040 | 0.293 |
| | (0.035) | (0.066) | (0.043) | |
| Honor society | -0.013 | 0.026 | -0.015 | 0.378 |
| | (0.021) | (0.039) | (0.026) | |
| Adult association index | 0.147 | 0.300 | 0.072 | 0.264 |
| | (0.082) | (0.182) | (0.091) | |
| N | 7347 | 2235 | 4975 | |
| *B. Occupation choice* | | | | |
| Finance | 0.005 | 0.130 | -0.056 | 0.005 |
| | (0.025) | (0.060) | (0.029) | |
| Medicine | 0.004 | -0.011 | 0.018 | 0.536 |
| | (0.024) | (0.036) | (0.032) | |
| Higher ed. | -0.019 | -0.030 | -0.023 | 0.888 |
| | (0.022) | (0.037) | (0.029) | |
| Law | -0.043 | -0.032 | -0.042 | 0.862 |
| | (0.028) | (0.046) | (0.037) | |
| Business | 0.077 | 0.067 | 0.063 | 0.968 |
| | (0.041) | (0.080) | (0.051) | |
| Teach | -0.020 | -0.019 | -0.010 | 0.867 |
| | (0.023) | (0.040) | (0.029) | |
| Government | 0.006 | 0.004 | 0.010 | 0.887 |
| | (0.016) | (0.033) | (0.019) | |
| Art/pub | 0.000 | -0.010 | 0.014 | 0.586 |
| | (0.022) | (0.036) | (0.029) | |
| Occupation index | 0.034 | 0.429 | -0.171 | 0.004 |
| | (0.085) | (0.190) | (0.097) | |
| N | 6294 | 1877 | 4287 | |
| *C. Adult income* | | | | |
| Wage income | -73.7 | -571.5 | 140.1 | 0.110 |
| | (203.9) | (374.4) | (256.6) | |
| Wage inc. 5000+ | 0.029 | -0.024 | 0.077 | 0.392 |
| | (0.052) | (0.103) | (0.062) | |
| Non-wage inc. 50+ | 0.046 | 0.001 | 0.087 | 0.487 |
| | (0.054) | (0.103) | (0.068) | |
| Class Report wage index | 25.4 | 127.0 | -25.2 | 0.015 |
| | (28.9) | (52.2) | (36.8) | |
| N | 1990 | 592 | 1319 | |

Coefficients on peer neighborhood price rank from regressions of form given in equation 2. Cohorts 1919, 1920, and 1921 are excluded. The first three columns denote samples. Rows are outcome variables. All specifications include randomization block and dummies for large feeder high schools; see section 5.1 for details."Test" column reports the p-value from a test of the null that the coefficients reported in the private and non-private columns are equal. Panels A and B report adult social club and occupation outcomes, respectively. "Adult association index" and "Occupation index" are standardized indices of the association between adult outcomes and private high school background. See section 3.4.3. Other variables are indicators for outcomes of the listed type. The sample in Panels A and B is students matched class reports; Panel B further restricts to students with occupation outcomes. Panel C reports labor market outcomes from the 1940 census. Sample is students matched to Census wage records in the 1920-30 entering cohorts. Standard errors clustered at peer neighborhood level.

Table A.11: Peer neighborhood effects on short-run outcomes with randomization inference

| | All | Private | Non-private | Test |
|---|---|---|---|---|
| *A. First-year activities* | | | | |
| Have any activity | 0.070 | 0.080 | 0.064 | 0.810 |
| | [0.047] | [0.156] | [0.159] | |
| N activities | 0.283 | 0.648 | 0.108 | 0.034 |
| | [0.010] | [0.004] | [0.395] | |
| Activity leadership position | 0.043 | 0.098 | 0.012 | 0.080 |
| | [0.033] | [0.029] | [0.548] | |
| Social | 0.049 | 0.183 | -0.028 | 0.000 |
| | [0.038] | [0.001] | [0.211] | |
| Sports | 0.029 | 0.085 | -0.007 | 0.243 |
| | [0.394] | [0.196] | [0.899] | |
| Music | 0.046 | 0.095 | 0.029 | 0.286 |
| | [0.097] | [0.072] | [0.394] | |
| Other activities | 0.021 | -0.027 | 0.059 | 0.217 |
| | [0.475] | [0.659] | [0.097] | |
| First-year activity index | 0.208 | 0.494 | 0.025 | 0.013 |
| | [0.010] | [0.002] | [0.781] | |
| N | 9343 | 2824 | 6370 | |
| *B. Upper-year social clubs* | | | | |
| Selective final club | 0.065 | 0.167 | 0.001 | 0.008 |
| | [0.002] | [0.005] | [0.955] | |
| Less selective final club | -0.028 | -0.075 | -0.007 | 0.204 |
| | [0.174] | [0.126] | [0.732] | |
| Hasty Pudding Inst. 1770 | 0.019 | 0.105 | -0.018 | 0.083 |
| | [0.518] | [0.088] | [0.568] | |
| Upper-year club index | 0.122 | 0.357 | -0.021 | 0.063 |
| | [0.127] | [0.053] | [0.810] | |
| N | 8589 | 2603 | 5847 | |
| *C. Academic outcomes* | | | | |
| Rank group 1 | -0.000 | -0.012 | 0.003 | 0.390 |
| | [0.991] | [0.346] | [0.808] | |
| Rank group 2 | 0.018 | 0.017 | 0.014 | 0.930 |
| | [0.261] | [0.436] | [0.522] | |
| Rank group 3 | 0.022 | 0.006 | 0.029 | 0.619 |
| | [0.353] | [0.888] | [0.328] | |
| Rank group 4 | -0.008 | 0.090 | -0.055 | 0.019 |
| | [0.759] | [0.063] | [0.127] | |
| Rank group 5 | 0.022 | -0.067 | 0.073 | 0.055 |
| | [0.503] | [0.300] | [0.061] | |
| Rank group 6 | 0.021 | 0.013 | 0.006 | 0.909 |
| | [0.378] | [0.800] | [0.832] | |
| Rank listed year 1 | 0.073 | 0.047 | 0.070 | 0.725 |
| | [0.013] | [0.384] | [0.048] | |
| Class rank year 1 | 0.024 | 0.101 | 0.020 | 0.671 |
| | [0.818] | [0.542] | [0.890] | |
| N | 7035 | 2085 | 4802 | |

Coefficients on peer neighborhood price rank from regressions of form given in equation 2. The first three columns denote samples. Rows are outcome variables. All specifications include randomization block and dummies for large feeder high schools; see section 5.1 for details."Test" column reports the p-value from a test of the null that the coefficients reported in the private and non-private columns are equal. Panels A and B report first-year activity outcomes and upper-year club outcomes respectively. "First-year activity index" and "upper-year club index" are the standardized indices of the association between activities and private high school background described in section 3.4.3. Other variables are indicators for a given activity type unless stated otherwise. Panel C describes academic outcomes in the first year. Class rank is a continuous variable from one through six, with one the best and six the worst. The other outcomes are dummies for having grades in the listed rank group and being listed at all. 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; students in these blocks are omitted from split sample regressions. P-values are presented in brackets below each coefficient. Two-sided p-values computed using randomization inference with 5000 re-randomizations of roommate groups to rooms within blocks. Statistic reported is $\frac{1}{N}\sum_{j=1}^{N} 1[\mid \hat{\beta}_1^j \mid > \mid \hat{\beta}_1 \mid]$, where $N = 5000$ is the number of rerandomizations, $\hat{\beta}_1^j$ is the estimate of $\beta_1$ in the $j$th randomization, and $\hat{\beta}_1$ is the estimate of $\beta_1$ observed in the data.

Table A.12: Peer neighborhood effects on long-run outcomes with randomization inference

| | All | Private | Non-private | Test |
|---|---|---|---|---|
| *A. Adult social and professional organizations* | | | | |
| Any social club | 0.066 | 0.241 | -0.003 | 0.001 |
| | [0.073] | [0.000] | [0.952] | |
| Country club | 0.073 | 0.219 | 0.006 | 0.005 |
| | [0.034] | [0.002] | [0.892] | |
| Gentleman's club | 0.019 | 0.053 | 0.003 | 0.426 |
| | [0.436] | [0.333] | [0.908] | |
| Fraternal order | 0.002 | 0.020 | -0.002 | 0.624 |
| | [0.922] | [0.566] | [0.949] | |
| Any honor/prof group | -0.014 | 0.039 | -0.024 | 0.390 |
| | [0.657] | [0.528] | [0.560] | |
| Prof. Association | -0.021 | 0.032 | -0.037 | 0.339 |
| | [0.506] | [0.599] | [0.372] | |
| Honor society | -0.017 | 0.020 | -0.019 | 0.309 |
| | [0.360] | [0.531] | [0.399] | |
| Adult association index | 0.139 | 0.328 | 0.040 | 0.122 |
| | [0.074] | [0.040] | [0.644] | |
| N | 9343 | 2824 | 6370 | |
| *B. Occupation choice* | | | | |
| Finance | 0.016 | 0.144 | -0.052 | 0.003 |
| | [0.519] | [0.012] | [0.071] | |
| Medicine | 0.001 | -0.018 | 0.017 | 0.457 |
| | [0.935] | [0.630] | [0.541] | |
| Higher ed. | -0.019 | -0.026 | -0.023 | 0.942 |
| | [0.373] | [0.418] | [0.392] | |
| Law | -0.032 | -0.009 | -0.036 | 0.623 |
| | [0.246] | [0.840] | [0.316] | |
| Business | 0.067 | 0.071 | 0.046 | 0.795 |
| | [0.092] | [0.360] | [0.349] | |
| Teach | -0.014 | -0.017 | -0.005 | 0.808 |
| | [0.507] | [0.671] | [0.855] | |
| Government | 0.005 | 0.012 | 0.009 | 0.939 |
| | [0.758] | [0.685] | [0.604] | |
| Art/pub | -0.005 | -0.017 | 0.008 | 0.586 |
| | [0.810] | [0.636] | [0.754] | |
| Occupation index | 0.056 | 0.450 | -0.164 | 0.003 |
| | [0.479] | [0.013] | [0.088] | |
| N | 7039 | 2094 | 4797 | |
| *C. Adult income* | | | | |
| Wage income | -26.1 | -440.5 | 191.0 | 0.136 |
| | [ 0.9] | [ 0.2] | [ 0.4] | |
| Wage inc. 5000+ | 0.020 | -0.029 | 0.073 | 0.379 |
| | [0.698] | [0.778] | [0.222] | |
| Non-wage inc. 50+ | 0.053 | 0.081 | 0.044 | 0.782 |
| | [0.322] | [0.416] | [0.499] | |
| Class Report wage index | 33.585 | 128.824 | -17.155 | 0.019 |
| | [0.212] | [0.011] | [0.618] | |
| N | 2396 | 697 | 1611 | |

Coefficients on peer neighborhood price rank from regressions of form given in equation 2. The first three columns denote samples. Rows are outcome variables. All specifications include randomization block and dummies for large feeder high schools; see section 5.1 for details."Test" column reports the p-value from a test of the null that the coefficients reported in the private and non-private columns are equal. Panels A and B report adult social club and occupation outcomes, respectively. "Adult association index" and "Occupation index" are standardized indices of the association between adult outcomes and private high school background. See section 3.4.3. Other variables are indicators for outcomes of the listed type. The sample in Panels A and B is students matched class reports; Panel B further restricts to students with occupation outcomes. Panel C reports labor market outcomes from the 1940 census. Sample is students matched to Census wage records in the 1920-30 entering cohorts. P-values are presented in brackets below each coefficient. Two-sided p-values computed using randomization inference with 5000 re-randomizations of roommate groups to rooms within blocks. Statistic reported is $\frac{1}{N}\sum_{j=1}^{N} 1[|\ \hat{\beta}_1^j\ | > |\ \hat{\beta}_1\ |]$, where $N = 5000$ is the number of rerandomizations, $\hat{\beta}_1^j$ is the estimate of $\beta_1$ in the $j$th randomization, and $\hat{\beta}_1$ is the estimate of $\beta_1$ observed in the data.

Table A.13: Child outcomes by parent income rank– recent cohorts

| | Ivy+ | Other elite | Other 4yr | 2 year | No college |
|---|---|---|---|---|---|
| *A. Sample counts* | | | | | |
| N students | 52724 | 183973 | 3398406 | 2063764 | 4106026 |
| *B. Distribution of parent income* | | | | | |
| Top 0.1% | 0.0221 | 0.0119 | 0.0016 | 0.0004 | 0.0001 |
| 99-99.9% | 0.1231 | 0.0882 | 0.0161 | 0.0041 | 0.0013 |
| 95-99% | 0.2699 | 0.2304 | 0.0771 | 0.0271 | 0.0075 |
| 90-95% | 0.1386 | 0.1454 | 0.0930 | 0.0485 | 0.0131 |
| 80-90% | 0.1300 | 0.1523 | 0.1615 | 0.1174 | 0.0421 |
| <80% | 0.3163 | 0.3718 | 0.6507 | 0.8025 | 0.9359 |
| *C. Kid outcomes by parent income* | | | | | |
| Mean rank | | | | | |
| Top 0.1% | 0.7248 | 0.6813 | 0.6025 | 0.4982 | 0.3930 |
| 99-99.9% | 0.7827 | 0.7322 | 0.6527 | 0.5306 | 0.4141 |
| 95-99% | 0.7941 | 0.7479 | 0.6719 | 0.5498 | 0.4230 |
| 90-95% | 0.7840 | 0.7501 | 0.6740 | 0.5618 | 0.4487 |
| 80-90% | 0.7721 | 0.7399 | 0.6557 | 0.5574 | 0.4546 |
| <80% | 0.7441 | 0.7115 | 0.5991 | 0.5094 | 0.3693 |
| Mean income | | | | | |
| Top 0.1% | 233 | 143 | 84 | 85 | 44 |
| 99-99.9% | 176 | 114 | 73 | 45 | 34 |
| 95-99% | 161 | 99 | 66 | 41 | 30 |
| 90-95% | 132 | 92 | 61 | 41 | 30 |
| 80-90% | 135 | 87 | 55 | 39 | 29 |
| <80% | 116 | 78 | 46 | 33 | 20 |
| Top 1% share | | | | | |
| Top 0.1% | 0.2567 | 0.1529 | 0.0824 | 0.0416 | 0.0495 |
| 99-99.9% | 0.2450 | 0.1417 | 0.0703 | 0.0246 | 0.0232 |
| 95-99% | 0.2121 | 0.1146 | 0.0466 | 0.0124 | 0.0115 |
| 90-95% | 0.1750 | 0.0937 | 0.0309 | 0.0074 | 0.0063 |
| 80-90% | 0.1629 | 0.0824 | 0.0208 | 0.0044 | 0.0031 |
| <80% | 0.1421 | 0.0656 | 0.0116 | 0.0025 | 0.0009 |
| Share not working | | | | | |
| Top 0.1% | 0.1382 | 0.1557 | 0.1872 | 0.2557 | 0.4027 |
| 99-99.9% | 0.0815 | 0.0983 | 0.1289 | 0.1860 | 0.3471 |
| 95-99% | 0.0630 | 0.0805 | 0.1008 | 0.1516 | 0.3220 |
| 90-95% | 0.0686 | 0.0713 | 0.0884 | 0.1320 | 0.2604 |
| 80-90% | 0.0724 | 0.0740 | 0.0871 | 0.1266 | 0.2297 |
| <80% | 0.0933 | 0.0873 | 0.1024 | 0.1415 | 0.2932 |

Childrens' outcomes by parent income percentile and college type for 1980s birth cohorts. Data are from Chetty et al. (2020). Columns are college types. Panel A gives sample counts. Panel B reports the share of parents with income in different percentile bins. Panel C presents the listed child's outcome for parents with the income given in the row. "Mean rank" is the mean child income percentile. "Mean income" is mean child income in 1000s of USD. "Top 1% share" is the share of children with income in the top 1%. "Share not working" is the share of children not working. See section 6.4 for details.

# B Data construction and additional analysis

## B.1 Data construction details

This appendix describes data sources and match procedures we use to construct the Harvard archival dataset. Unless noted, all materials were accessed through the Harvard Archives.

### B.1.1 Freshman Red Book / Freshman Register

The Freshman Red Books present information about each Harvard class's freshman year in a manner analogous to a yearbook. For the Harvard classes of 1913 through 1952, Freshman Red Books were published at the end of the spring semester of a class's freshmen year.[B.1] In 1953 the Freshmen Red Book was renamed the Freshman Register and publication was moved to the beginning of the fall semester. Starting in 1962, the Freshmen Registers no longer contain campus address or activities. Note that the Freshman Registers are distinct from the Student Council Registers and Official Registers.

In the Red Books, there is a "Class Directory" section with entries for each student containing the student's picture, name, home address, college address, high school, age, and activities the student participated in during his freshman year. See Figure B.1 for an example. Following the student records, the Red Books contain club and activity pages which list freshman membership in campus activities. See Figure B.2 for a club page example. We use the Red Books as our primary source for constructing cohorts of entering Harvard classes and for information on high schools, campus addresses, home addresses, and freshman activities. We have this information for all cohorts in our sample – the entering classes of 1919 through 1935 – with the exception of the entering class of 1926 (the graduating class of 1930) for whom we do not have college addresses because they were not included in the class of 1930 Red Book. We digitized the Red Books using hand transcription.

We clean student home address records into standard address format. We process student campus addresses by cleaning them and then matching to college dorm records so that, in addition to knowing the exact street address or dormitory and room number, we also categorize whether students were living at home, at another address off campus, in a dorm primarily housing freshmen in that year, or in a dorm primarily housing upper classmen or graduate students in that year. The small share of records where students are housed in non-freshman dorms are primarily due to demand for freshman dorms exceeding capacity or the student sharing a room with an older brother.

[B.1] The classes of 1940-1945 have a fall Freshmen Red Book in addition to the standard spring Red Book. The fall Red Books do not contain student activity information.

We categorize student high school records as follows. First, we clean and standardize the free text entries by constructing a key that maps regular expressions description of common string patterns to standardized high school names. We do this for the 67 high schools listed identically by at least 15 students. We supplement this list with 46 high schools needed to distinguish records of the most common schools (e.g., we standardize reports of attending St. Paul's, Garden City, N. Y. in order to distinguish its graduates from students who attended St. Paul's School in New Hampshire). We categorize all standardized school names as either public or private. Among the private schools, we also flag the most elite New England boarding schools as "private feeder" schools. See Table B.1 for a complete list of school codes. We use that key to assign a high school code to each matching student record. 9813 students in our sample attended one of the schools we standardized. 4342 students report a high school other than our set of standardized schools. The majority of these students attended a school that only sent a handful of students to Harvard across our cohorts (e.g., Zanesville High). We include these students as non-private feeder, non-public feeder, and non-private school students.

As discussed in Section 3.4.1, the choice of Andover, Exeter, Groton, Middlesex, Milton, St. George's, St. Paul's, and St. Mark's as our private feeder school population was driven by the high shares of incoming Harvard students from each school and by historical accounts which describe these schools as a signature of high social status. In addition, these schools were characterized by high tuition, making them accessible only to wealthy families. For example, tuition and expenses at Andover and Groton in 1920 ranged from $500 - $1,200 (Levine, 1980), a significant cost compared to average incomes estimated at $1,283 in 1919 (Bowden, 1946) and Harvard tuition and expenses estimated at minimum $650 (Official Register, 1921). We discuss alternative private school classifications in sections 2 and 5.7 of the main text.

Finally, we process student activities records as follows. We clean and standardize activities listed in free text entries by constructing a key that maps regular expressions descriptions of identifiable string patterns to standardized activity names (identifiers). We group activities into a straightforward taxonomy under the guiding principle that activities whose participants may have interacted in a shared activity setting are grouped in the same category. See Table B.2 for the complete code list. Our categorization was exhaustive, meaning that all activites were categorized except for a few rare cases (e.g., activities that were only listed by a couple of students). We also flag if students were described as having a leadership role (e.g., manager, captain, president). Students may participate in multiple activities.

### B.1.2 Student Council Registers

To supplement the freshman student activity participation information we extract from the Red Books, we also digitize the "Registered Clubs" chapter of the Harvard University Registers published by the Student Council of Harvard College. We refer to these as the Student Council

Registers to avoid confusion with the Freshman Registers (a later name for the Red Books) and the Harvard University Official Registers published by the University (described below under 'Official Registers'). We obtained Student Council Registers for school years 1919-1920 through 1923-1924 from the Hathi Trust and Google Books and accessed physical copies of Student Council Registers for school years 1924-1925, 1925-1926, and 1927-1928 in the Harvard Archives.[B.2] Each club record in Student Council Registers lists club officers and members by name and class year. See Figure B.3 for an example. From the "Registered Clubs" section, we only extract the clubs from the "Miscellaneous – Social" subsection which contains records for fraternities and final clubs.[B.3] Club and fraternity membership records extracted from the Student Council Registers list students in all classes, in contrast to the activity reports from the Red Books, which only list Freshmen activity participation. As the typical final club does not admit members until at least sophomore year, we need these additional records in order to track final club membership.

Membership records were digitized by the firm Suntec Digital whose work we then audited for quality. We categorize the clubs into three categories – fraternities, most elite final clubs, and other final clubs. To distinguish the most elite final clubs, we relied on contemporary records naming the most inclusive clubs such as Amory's The Proper Bostonian. See main text for details.

### B.1.3 Class Albums

The Harvard Class Albums edited by the Senior Album Committee are yearbooks produced at the time of graduation. In the Class Albums, each student entry contains the student's photo, name, and brief description of the student's background and time at Harvard. Figure B.4 presents an example. We use these records to extend our series of final club membership from 1928 to 1935 following the end of the Student Council Register series.

### B.1.4 Rank Lists

Rank Lists report class rank based on course grades discretized into six groups.[B.4] These lists were published for freshmen, sophomore, and junior students from 1921-1932 and for freshmen

[B.2] As best we and the librarians at the Harvard Archives can tell, Student Council Registers were either discontinued or not preserved following 1928.

[B.3] It is somewhat surprising and fortuitous for us that final clubs are included in the "Registered Clubs" section as they were not, in fact, formally recognized by the administration of the University. Amory (1947) writes that "[f]or the most part official Harvard prefers to ignore the clubs in the same manner that the columns of the Crimson, undergraduate newspaper, take no notice of their existence. A Harvard professor recently declared that having been a faithful attendance at faculty meetings for twenty-five years he has yet to hear the clubs brought up for discussion."

[B.4] The rank lists were distributed to distinguish academic achievement with more granularity - especially in the middle of the distribution - than other previously available measures such as probation and receiving academic prizes (Annual Report of 1920-21). Lists were distributed publicly to students and their families.

only begining in 1933.[B.5] Students are divided into six groups "according to the average of the grades attained" (Annual Report of 1921-22). See Figure B.5 for the description of the ranks and which students were excluded from the rank list. Students' rank is reported alongside their names and class. See Figure B.6 for an example. Rank Lists were hand transcribed with quality audits.

### B.1.5 Official Register

The Official Registers of Harvard University were a means by which the administration communicated information to entering freshmen. We refer to these as 'Official Registers' to prevent confusion with the Student Council Registers and the Freshman Registers. In these Official Registers, we find descriptions of how freshman students are assigned to dorms. Official Registers from 1922 through 1944 contain near identical language describing how students should submit an application indicating the price range they are willing to pay, if they have special financial hardship, and an optional choice of roommate. The text further describes that, with the exception of trying to reserve the very cheapest rooms for the poorest students and an attempt to make sure that students from the same high school are somewhat spread out, students are assigned to rooms "by lot" within the requested price range. We observe essentially identical text, including the exact phrase "by lot," in each year from 1922 through 1942, except for 1936 where we cannot verify the text because we are missing the relevant page of the Register. The text also refers to application blanks in each year, and describes randomization blocks in the same way. Figure B.7 provides an example of this text from one year.

In addition, we use floor plans and room price menus from the Official Registers of 1920 and 1932. We use the floor plans in order to inform our definition of 'neighborhoods' by the architectural features of the dorm. Price menus allow us to stratify rooms by their per-student price. As we do not have price lists available for every year, we rely on the prices from 1920 and 1932. In 1931, freshmen were relocated from the dorms by the river to dorms in the yard, so we rely on the 1920 records for information on dorms occupied by entering classes 1919-1930 and on the 1932 records for information on dorms occupied by entering classes of 1931 and later. In the rooms for which we observe multiple prices across years, these prices are the same. If prices do change across years in the 1920s, to the extent they preserve relative orderings, our rank-based measures will be unaffected. See Figures B.8 and B.9 for examples of price lists and floor plans. We describe the way we use these data to identify neighborhoods in section B.7

[B.5] In 1933, publication of rank lists for Sophomores and Juniors was discontinued due to concern that producing a metric of course grades alone distorted upperclass students' incentives towards focusing on course grades when departments were defining students' academic performance more broadly to also encompass students' tutorial work, which they preferred to leave ungraded. Freshmen had no tutorial duties, so course grades were considered an appropriate measure and the publication of Freshmen Rank Lists continued (Annual Report of 1933-1934, p. 107-108).

### B.1.6 Miscellaneous Student Housing Records

In the archived records of the Bursar, we located a blank copy of the housing application form entering freshman students were asked to submit in 1922. Figure B.10 displays the application. Students could fill out the form individually or in pairs if they were requesting to live together. The form is consistent with descriptions of the application process in the Official Registers.

### B.1.7 Contemporary descriptions of random assignment

Did students know or talk about the random assignment of first-year rooms? A search of the *Harvard Crimson* for articles on first-year room assignments produced limited results. This makes sense, because the students being assigned first-year rooms are not yet enrolled in Harvard and are neither the writers nor the intended audience of the Crimson.

However, newspaper reports on room assignments to upper-year students indicate that a) assignment of rooms "by lot" was common at Harvard during this period, and b) students understood "by lot" to mean a random draw designed to avoid favoritism. Consider Harvard University (1924). This *Crimson* article describes how room assignments in Yard Dormitories for seniors were determined "with a system based on a draw, and the rooms were assigned strictly by lot, without regard to the prominence of the men nor to the theory of 'congenial entries,' under the cover of which former committees in many cases showed favoritism[.]" The article also notes that this same system had been used in previous years.

The conclusion we draw from the primary- and secondary-source evidence, as well as from our empirical validation of conditional randomization, is that the evidence that rooms were assigned at random conditional on price and occupancy is as strong as one could hope given that we are observing the process 80 to 100 years later.

### B.1.8 Class Reports

The 25th Reunion Class Reports were compiled by Harvard twenty-five years following graduation. Nearly all graduates self-reported the information for the Class Reports. Of the minority of graduates who did not respond, Harvard entered the information it had on file (such as last known address or date and location of death). For graduating classes of 1923 through 1937, these records usually include the graduates' home address, office address, birth place, parents' names, high school, years at Harvard, marriages and children, occupation, offices held, and clubs and associations of which the graduate is a member. This more standardized information is followed by narrative text usually containing some combination of Harvard memories and updates on the graduate's life since graduation. Figure B.11 shows an example of a Class Report record. Starting with the class of 1938 (Red Book year 1934), some previously common fields were frequently omitted, including those relating to parents and place of birth.

We digitize the class reports by photographing the pages with a high resolution camera, running an OCR program to collect all text from the images, and parsing the raw text file to extract standardized information from each record (e.g., the text following "PREPARED AT:" was assigned as a student's high school). A team of BYU research assistants went over each of the OCRed records to fill in missing names and match the individuals to Census records on FamilySearch. Observations that were incomplete or indecipherable due to imperfect OCR were sent to the firm SunTec Digital to be entered by hand.

We categorize occupations listed in the Class Reports by searching for key words associated with occupation categories. Our occupation categories are not mutually exclusive. We do not distinguish between occupation and industry. For example, an indiviual who reports "Lawyer, secretary and resident legal counsel, Brookhav National Laboratory, Associated Universities, Inc" would be categorized in both the "law" category and the "academics/research" category. Starting with common identifying words, we continue to add identifiers until all reported occupations are assigned to at least one category. Table B.3 provides the full list of codes. When we construct private school indices, we include all categories with at least 100 students as potential predictors. This amounts to adding the science and engineering categories to the eleven we discuss in the main text.

We categorize associations and clubs listed in the Class Reports as follows. We clean and standardize clubs and associations listed in free text entries by constructing a key that maps regular expressions descriptions of identfitiable string patterns to standardized club and association names (identifiers). We categorize clubs and associations as social (including subcategories for country clubs, gentlemen's clubs, and fraternal orders), professional, or honorary/political. Table B.4 provides the full list of club codes. We categorize any club listed identically by at least five students across the Class Reports, excluding Harvard-specific clubs.

## B.2 Record linking procedures

### B.2.1 Linking red books and class reports

We linked student records from the Red Books to records from the Class Reports via the following process:

1. First, last, and full (combined first, last, and possibly middle) names, years, and high schools are extracted for each person in both the red books and the class reports.

2. Red book-class report pairs with similar first OR last names (within 3 or 2 positions in an alphabetically sorted list, respectively) AND similar years ($\pm 2$) are identified as candidate matches.

3. The following features are assembled for each candidate match:
    - A first name score that is the Jaro-Winkler distance between the candidates' first names
    - A last name score that is the Jaro-Winkler distance between the candidates' last names
    - A full name score that is the Jaro-Winkler distance between the candidates' full names
    - A year score based on the difference between the birth date in the class report and the estimated birth date from the red book.
    - A high school score that is a normalized Levenshtein distance between the candidates' listed high schools
4. Matches with very bad first name or last name scores are removed.
5. The features for the remaining candidate matches are plugged into a random forest classifier (50 trees, max depth 5, class weights normalized to be balanced), which outputs a match score for each candidate match. These scores are the proportion of trees in the random forest that classify a given candidate pair as a match. The random forest classifier was trained on a data set 690 candidate matches which were marked as a match only if a research assistant was highly confident that the records were for the same person. By being trained on this data set, the classifier learns to estimate the probability that a human would confidently say that the two records belong to the same person.
6. Candidate matches are retained if their score exceeds a threshold (0.4). If a class report entry has more than one Red Book match, the match with the highest score is retained and all others are dropped. It is possible that a given red book entry may still have multiple potential matches indicated in the output file, but this is rare, and the match with a higher score should be preferred.

We assess the performance of links between Red Books and Class Reports using k-fold cross-validation on the manually labeled training data, with k set to 10. We split the data into k folds, then withold one fold at a time and re-fit the model on the other k-1 folds. We compute accuracy, precision, and recall on the witheld fold. Figure B.12 displays precision-recall curves generated from this analysis. At the threshold currently used (0.4) the average precision (percent of candidates predicted to be matches that are actually matches) is about 0.95, and the average recall (percent of true matches correctly predicted to be matches) is essentially 1.0.

#### B.2.2 Linking clubs and honor lists to student records within the Red Books

1. First and last names are extracted from the clubs lists and the main Red Book data. First names are transformed to just first initials because the clubs lists often only include first

initials.

2. For each entry in the clubs lists:

   (a) Candidate matches are selected from the set of those with exactly matching years.

   (b) Candidate matches are filtered down to those with exactly matching last names. The match with the closest first name is kept.

   (c) If there are no exactly matching last names, candidate matches are reduced to those with exactly matching first names, and Jaro Winkler scores are computed for the last names of each candidate match. The candidate with the highest score above a threshold (0.85) is kept; if no candidates meet that threshold, there will be no match.

We evaluate match quality by calculating precision on a random sample of 200 records. See section B.2.4 for a description of this process and Table B.5 for results. We find precision exceeding 98% on our quality check sample.

#### B.2.3 Linking of class rank data to red books and class reports

Entries in the class ranks data with missing information or some invalid value (for example, some entries do not report class years) are dropped.

1. First and last names are extracted from class rank, red book and class report data. First and last initials are also extracted.

2. For each entry in the class rank data:

   (a) Entries missing either a first or a last name are dropped (no match)

   (b) Look for matches with exact last name match and exact first initial (or full name if available) match in the class report for the same year. If more than one match is found, it is ambiguous which one is correct, so we drop the record (no match)

   (c) If still no match found, repeat the above looking in surrounding years ($\pm 1$).

   (d) If still no match found, try again in the original year, but allow inexact first and last name matches (within a damerau-levenshtein distance of 2, i.e. 2 or fewer string edits away)

   (e) Repeat the above 4 steps, this time looking for matches in red book data.

3. We assign each candidate a rough confidence score which we use to select a single match for entries with multiple candidate matches. These scores were constructed to reflect the amount and quality of information used to form the match. Looking for inexact matches

or matches in surrounding years decreases the confidence score, and making use of the full name (rather than just initials) increases the confidence score. The confidence score will also be higher if a match is found both in the Red Books and the Class Reports and those matches agree (as per the Red Book-class report linking described above). If a given Class Report entry or Red Book entry is given as a match to more than one class ranking entry, the match with a lower confidence score is dropped.

We evaluate match quality by calculating precision on a random sample of 200 records. See section B.2.4 for a description of this process and Table B.5 for results. We find precision of 100% on our quality check sample.

#### B.2.4 Report linking accuracy for class rank, Student Council Registers, and Class Albums

When we link records such as class rank data from the Rank Lists or social club members from the Student Council Registers, we only have name and year to rely on. This means that the random forest approach we used to link the Red Books and Class Reports is not feasible given the paucity of features and that we have many ambiguous candidate matches. We err on the side of false negatives rather than false positives when designing the matching schemes. Manually checking a small random subsamples of 200 class rank records, 200 Student Council Register records, and 200 Senior Album records, we report precision for matches with Red Books and Class Reports in Table B.5. All quality checks yield precisions exceeding 98%.

#### B.2.5 Linking class reports and Census records

We combine three methods for linking our Harvard sample with US Census records. First, we link our Harvard sample to profiles in the Family Tree at familysearch.org, using an approach similar to Price et al. (2021) and Feigenbaum and Gross (2020). FamilySearch provides an API that allows researchers to use information about an individual in a record and search if a profile exists for that person on the Family Tree. We extended this method and created a profile for each person in our sample that did not already have an existing profile. The machine learning algorithms employed by FamilySearch provided possible Census hints for all of the profiles in our sample. Research assistants then evaluated these Census hints and attached them to profiles on the Family Tree if they were correct matches.

The FamilySearch Family Tree has very good coverage for the United States, and we were able to identify an existing profile for 59.9% of the Harvard students in our sample. Of those students who already had a profile, we found that 69.5% of the students in our sample were already attached to the 1940 census and 70.5% to at least one census prior to when they entered Harvard.

Second, we use the hand-linking approach employed by Costa et al. (2020) to recover additional matches. As in Costa et al. (2020), we utilize the search tools on Ancestry.com to find Census records of the people in our sample. These search tools allow us to adjust the parameters and features that we use to find possible records. In addition, when a Census record is found on Ancestry, their machine learning algorithm provides links to other records that are likely the same person. Public family trees created by users on Ancestry can also be found, and these will often provide links to Census records that have been hand-linked by users on that platform. We bring all of the Census matches that we find using Ancestry over to the Family Tree on FamilySearch.

Third, for each person in our sample to whom we connect at least one Census record, we then use traditional Census-to-Census linking methods from economics to identify other Census records for the same person. For this step, we use the Census linking method described in Price et al. (2021), which incorporates the linking methods described in Abramitzky et al. (2019). These automated methods help identify additional Census matches that were not already on the Family Tree or suggested by the algorithms used by FamilySearch or Ancestry. Since we wanted to have extremely high levels of both recall and precision for this paper, we used methods that provided the most possible matches and then had research assistants hand check these matches. We attached each of the correct matches to the profiles of our on the Family Tree.

The final result of these three methods is a fully public dataset on the Family Tree that anyone can access using the FamilySearch website or API. As such, our dataset is likely to acquire additional record collections as users on the website attach additional information about these individuals, including military and death records. The extensive hand-linking that we performed for this project could also provide a training set for future projects similar to the training set created by Bailey et al. (2020). Aside from these advantages, the primary reason Family Tree was used to do the linking for our project is that it allowed us to achieve the high levels of recall and precision that come from hand-linking and combine this with the cost savings of automated methods and data already available on the Family Tree.

## B.3 Comparing occupation classifications and alternative wage measures

### B.3.1 Occupation classifications in Class Reports vs. the Census

We use occupation categories constructed from the Class Reports as our primary measure of occupation. As described in section B.1.8, we construct these by applying a categorization scheme to the free text describing occupations in the Class Reports. For students linked to the 1940 Census, the occupations reported in the 1940 Census are available as an alternative measure of occupation. These occupation measures differ in terms of data availability, the number of years post-college when occupation is observed, and the granularity of occupation category definition.

We consider our Class Report occupations to be preferable across each of these dimensions. Class Report occupations capture a student's occupation twenty-five years after graduation for cohorts 1920-1935 and are available for 79% of students in those cohorts, while the 1940 Census reports a students' occupation sixteen to six years after graduation for cohorts 1920-1930[B.6] respectively and is available for 60.1% of students in those cohorts. This advantages Class Report occupations in two ways. First, we observe Class Report occupation for twice as many students (10,752 versus 5,399). Second, we observe Class Report occupations at the same time in students' career trajectories and well into their working years (approximately age 47) while the 1940 census occupations provides a different early-career snapshot for each cohort.

Turning to differences in the granularity of occupation reports between the Class Reports and the Census, Table B.6 presents the modal Census occupation for four illustrative Class Report occupations: finance, senior management, medicine, and law. The Census provides a reasonable approximation of Class Report data for the professions. Medicine and law map closely to the Census occupation codes for, respectively, "Physicians and Surgeons" and "Lawyers and judges." Despite the ten to fifteen year gap in reporting most doctors in our data also show up as doctors in the Census (54%), and most lawyers in our data show up as lawyers in the Census (59%). The weaknesses of Census data are more pronounced for business-oriented careers. The modal occupation for students reporting both finance and senior manager positions is "Managers, officials, and proprietors, not elsewhere classified." The essential issue here is that the Census occupation categories do not offer a very detailed taxonomy of distinctions across elite business and management careers, which are highly relevant for Harvard students.

A corrollary to the point that Census occupation measures do not describe elite business careers very well is that measures of income based on Census occupations do not perform well in our setting. Some studies of historical Census data use occupation wage scores or occupation prestige scores from the 1950 Census to describe individual outcomes. Table B.6 compares the scores for the modal Census occupation for each Class Report occupation to observed earnings for Harvard students in the Class Report occupation who also report the modal Census occupation. What we see is that students reporting finance and senior management occupations in the Class Reports (and who the Census places in the "Managers, officials, and proprietors, not elsewhere classified" category) have very high incomes. The average value is this group is close to $3,700, and 40 percent have topcoded income. This is not captured by the occupation wage score for this group, which is middling– about half of the score for doctors and two-thirds the score for lawyers– despite the fact observed earnings for students in finance and top management are higher than for either of the two professional careers. Harvard graduates with successful careers

[B.6] As with the rest of the analysis with labor market outcomes in the 1940 Census, we do not include entering cohorts after 1931 as their labor force participation rates have not convereged by 1940 and we are systematically less likely to observe the profession and income of future doctors and other professions that require years of graduate preparation.

in business and finance are, unsurprisingly, quite rich. The Census occupation categories do not capture this.

Table B.7 shows that the limits of Census occupation categories are quantitatively important. It reports income information about the Census occupations most common in our sample (panel A) and for all men from the same birth cohorts in the 1940 Census (panel B). Twenty percent of Harvard students with a Census occupation are in the broad category "Managers, officials, and proprietors (not elsewhere classified)" and over half of students with a Census occupation are in one of the five most frequent occupations. The fact that a handful of occupations covers a large portion of the sample suggests that Census occupation-level measures such as wage-score and occupational prestige may poorly reflect between student variation in our sample by discarding all variation within Census occupation categories. One way to quantify this is through the correlation between observed earnings and Census occupation scores. In our Harvard sample, this correlation is 0.15, less than one third the value of 0.49 that we obtain for similarly-aged men in the 1940 Census population. There are many applications where Census occupation scores are a reasonable measure of career outcomes, but the highly-selected Harvard sample does not appear to be one of them.

The final point to make here is that limitations of the Census are asymmetric across the groups of Harvard students we are particularly interested in: private feeder students and final club members. For example, private feeder students are more likely to be in finance and/or in an upper management position which is likely to show up in Census as "managers, officials, and proprietors (n.e.c.)". Over a third of our Harvard students are in that or a similarly broadly-defined Census occupation. Comparing panels A and B of Table B.7 shows that while Harvard students report higher average earnings than the general population across all occupations, the discrepencies are larger for broadly-defined occupations. Harvard students in the "managers, officials, and proprietors (n.e.c.)" occupation are six times as likely to have top-coded wage income as similarly-aged men on average. In contrast, Harvard students categorized as lawyers and judges are only twice as likely as similarly-aged men to earn top wage income, and Harvard physicians and surgeons are equally likely. From these comparisions, we would expect occupation-level measures to perform particularly poorly for the more than a third of students in broadly-defined Census occupations, which would introduce bias unevenly along key dimensions of interest for this study.

In section 5.6 we analyze room randomization using an alternate wage index that addresses most of these limitations by using Class Report records of social and occupational outcomes in concert with Census reports of wage earnings. This amounts to computing predicted wage earnings based on coarse occupation and adult social assocations. This has two main advantages over Census-only wage indices. The first is that it allows us to exploit the fine-grained variation within elite business careers that shows up in the Class Reports but not the Census. The second

is that we can construct this index for many more students.

See section B.6.2 below for a discussion of how the use of alternate measures of income and wealth affects the analysis of Census outcomes in section 4.

### B.4 Computing private school indices

This section presents more detail on the construction of private school indices for activities, occupations, and adult social categories first described in section 3.4.3. To summarize how changes in first-year activities, occupation, and adult club participation relate to integration across social groups, we collapse activities (and occupations, and adult social categories) into standardized linear indices capturing how characteristic the set of outcomes is of private school students. We construct these indices by regressing an indicator for private feeder school background on a set of indicators for participation in different first-year activities (occupations, adult social clubs) and cohort fixed effects. We compute predicted values, and then standardize (mean zero, standard deviation one) to facilitate cross-outcome comparisons. To avoid using own schooling background as an input to prediction, we estimate this specification for each cohort $c^*$ using cohorts $c \neq c^*$. We use a Lasso to select coefficients and use EBIC for model selection. Table B.8 reports coefficients estimated using data from all cohorts, and Figure B.13 displays the distributions of predicted values.

### B.5 Availability of cultural name indices over time

We compute Jewish, Catholic, and Colonial name indices for all Harvard students with names that we also observe in the combined 1920/1930 population Census data. Table B.9 reports the share of students for whom we compute indices, split by first and last name. In our main of students entering Harvard in the 1920s and 1930s, we compute indices for essentially all names– 99.7% of first names and 98.7% of last names. The high rates here make sense given that we expect students in Harvard during this period to show up in the population Census records themselves. Non-matched names reflect a combination of typographical errors and students not captured in Census data. We treat students with unmatched names as missing data in our main sample analyses. We compute index values for almost all students in both private feeder and public feeder high schools. Note that the public feeder definition used in this table is the extended definition from the long-run sample.

In the full long run sample, which runs through the graduating class of 2015, we are able to compute indices for a somewhat lower but still high share of students. 96.4% for first names and 94.6% for last names. The decline relative to the early sample makes sense given that the students in our data move progressively farther from the 1920/1930 Census population as cohorts progress. We treat missing records with missing index value as non-Jewish/non-Colonial in our

long-run analysis. See section B.9 for details.

## B.6 Additional descriptive analyses

### B.6.1 Club selection regression analyses

We explore how student background interacts with academic and social performance to determine who advances through the Harvard club system by estimating regressions with Hasty Pudding and final club membership as the outcomes of interest. Table B.11 reports our findings. Of the pre-college characteristics, the most important predictor of Hasty Pudding membership is attending a private feeder school, which raises membership probability by 23.7 percentage points. Being a Harvard legacy (9.0pp) and having a Colonial name (5.3pp) also predict membership, while Jewish names ($-3.9$pp) predict non-membership. First-year social leaders (team captains, club presidents, and social organizers) have membership rates that are 32.5pp higher. Low academic rank raises rates by 4.6pp and high academic rank *reduces* rates ($-3.3$pp). Looking at interactions between pre-college characteristics and first-year activities, we see that private feeder students who are also social leaders get a boost, while private feeder students who have grades that are anything but in the middle of the rank distribution are less likely to join Hasty Pudding.[B.7]

Similar patterns persist when the outcome is membership in selective final clubs, with the key difference being that having both a pre-college background at a private feeder school and social leadership activities early in college is more important for clearing the steeper hurdle. These results are reported in column 2 of Table B.11. Column 3 of Table B.11 restricts the sample to Hasty Pudding members. What emerges here is a large, negative correlation between final club membership and Jewish and Catholic names. The relatively few such students who make it into Hasty Pudding are much less likely than other Hasty members to make the jump to a final club.

### B.6.2 Alternate income definitions

One limitation of earnings outcomes in the 1940 Census is that they do not include unearned income. To understand how this may affect the analysis in section 4, we consider proxies for income as well as exercises that impute income for occupations where unearned income may be particularly important.

Figure B.14 presents a version of Figure 1 with alternate outcome measures. The first outcome measure (Panels A and C) is an alternative measure of wage income where we adjust the

[B.7] The social benefits of middling grades follows contemporary accounts. Amory (1947) describes how Harvard "social taboos" include "grades above C." At the same time, F. Scott Fitzgerald's semi-autobiographical novel about social climbing at Princeton describes how "[i]f you don't pass [math class], you're the world's worst goopher. Your stock will go down like an elevator at the club and on the campus" (Fitzgerald, 1920).

wage incomes reported in the 1940 census by imputing top wages ($5,000) for all doctors and lawyers. We present this measure to address concerns that law and medicine may underreport earned income relative to unearned business income, and that these occupations may be disproportionately represented among non-club members, and non-private feeder students. The second is alternative measure (Panels B and D) is predicted earnings based on home values. We construct this measure for students who report being head of their household and owning their own home by predicting earnings on cohort indicators and a third-degree polynomial of home values. We present the predicted income measure as a means of stepping back from the limitations imposed by top coding of wage income and only an indicator for non-wage income. Among home-owners, housing value provides an alternative measure of economic well-being. One drawback from this approach is that home values may also reflect inherited wealth. Using these alternate measures does not affect the conclusions one draws from Figure 1.

#### B.6.3 Sample description for within-family analysis

This section presents a more detailed description of the sample used for the within-family analysis in section 4.5. Families that send multiple sons to Harvard tend to be wealthier than other Harvard families. Tables B.12 and B.13 present descriptive statistics for 1) our full sample of students, 2) the sample of students for whom data on wages and academic class rank is available (i.e., the sample used in our non-family descriptive analyses), 3) the sample of students in multi-brother families who also have wage and rank data (i.e., the sample used in our family fixed effect analysis), and 4) the subset of group 3 for whom one there is within-family variation in selective final club membership. We see that 37% of students in the brothers sample went to private feeder schools, compared to 24% in wages and ranks sample overall. Cross-group differences in legacy status, and immigrant status all similarly indicate that the brothers' sample is higher-SES than the full sample.

There are relatively few families where the brothers have different final club outcomes. 8% of brothers fit this description, or 23 in total from 11 families. The rightmost columns of Tables B.12 and B.13 show that this group is from higher-status backgrounds than brothers over all. For example, 78% went to private feeder schools. The small number of brother pairs who differ in terms of final club membership helps explain the imprecision of the within-family estimates presented in Panel D of Table 4. Note that students for whom final club membership does not vary within family help identify the cohort effects and grade effects.

#### B.6.4 Missing class rank data

The regression analysis in Table 4 uses data on the sample of students with non-missing data on first-year academic rank. The choice to include or exclude students with missing rank data

does not affect our findings. Table B.14 reports alternate estimates of Panel B of Table 4 that include students with missing rank data. The "no academic rank" coefficient we report reflects the differences in outcomes for students with no academic rank relative to students in the bottom rank group. Estimates of the final club, private school, and academic rank premia are unaffected across all outcomes.

### B.6.5 OLS specifications with third-year grades

This section discusses how the descriptive analysis in section 4.5 changes when we use other measures of academic performance. Our analysis of academic achievement in the main text focuses on first year class rank, because these data are available for all cohorts of Harvard students in our data. This appendix repeats the analysis for third-year class rank. These are the most advanced students for whom we observe class rank. We only observe these data through calendar year 1930.

We first show that first-year and third-year class rank are strongly correlated. The upper panel of Figure B.15 plots the distribution of third year rank group by each value of first-year rank group. Within each first year rank group except the sixth, the modal third-year rank group is the same as the first year group. The lower panel of Figure B.15 plots the mean and interquartile range of the third year rank group by the first year group.

Figure B.16 presents a version of Figure 1 with third year class rank. Table B.15 presents an alternate version of Table 4 that uses third-year class rank rather than first year class rank, and also includes some additional outcome measures. The first four columns repeat Table 4 but using third year class rank. Sample sizes are slightly smaller because class rank is missing more often in the third year than the first, even in cohorts for which both measures are available for some students. This change in specification does not affect the conclusions we draw about the size of the final club membership earnings premium, which remains large across all specifications. The main difference we see here is that the class rank premium rises somewhat and becomes more precisely estimated. For example, in Panel C of Table B.15, it is equal to \$94 (SE=\$20), compared to a \$50 estimate (SE=\$19) in Panel C of Table 4. The larger effects for third-year grades are still small relative to the final club membership premium. Continuing to focus on Panel C of Table B.15, for example, we see that the return to a one-group increase in rank is equal to one-eighth the return to membership in a selective final club when the outcome is earned income (column 2) and about 1/16th when the outcome is topcoding (column 3). There are only six rank groups, so for both of these outcomes the predicted effect of a move from the bottom rank group to the top rank group is substantially smaller than the effect of final club membership.

We continue our analysis of alternate income variables in the rightmost columns of Table B.15. In column 5, we display specifications that impute maximum income values for doctors and lawyers, as above. The second shows specifications where in addition to imputing outcome for

doctors and lawyers, we follow Goldin and Katz (2008) and multiply all topcoded values by 1.4. These adjustments tend to raise the return to class rank. The first procedure tends to lower the return to final club membership relative to the column 2 baseline, while the second procedure tends to raise it. Regardless of which procedure we use, the return to final club membership remains very large relative to the return to improved academic standing across all specifications. In section B.10.1 we use estimates from columns 5 and 6 of Panel C in this table to compare our findings to Goldin and Katz (2008). Note that the standard deviation of class rank in the sample used Column 2 of Panel C is 1.26.

### B.7 Defining residential peer groups

We define residential peer groups using archival floor plans for dorms housing freshman students. Our aim is to identify peer groups that form as a result of residential proximity, i.e., the loci for casual interactions between residential peers. To define a peer group, we apply two rules. The first rule is that for two rooms to be part of the same residential peer group, a student must be able to get from one to the other without going outside. Most Harvard dorms are organized by entryways (a group of rooms surrounding a stairwell that can be accessed by a common exterior door), so this rule means that residential peer groups are composed of students within the same stairwell in the same dorm. The next rule is that floors (or hallways) are important social groupings. This follows Marmaros and Sacerdote (2006), who show that being members of the same hall more than doubles interactions among college students. Not all Harvard dorms have hallways– some entryways are vertically oriented and consist of one or two rooms around a staircase entrance on each floor– but, for those that do, we split peer groups by floor. Applying these two rules results in almost all of our peer groups being defined as either the entire entryway of a dorm or the floor within an entryway of a dorm. (Exceptions to this rule are for two dorms with layouts that do not use the entryway model, described below.)

For clarity, we present examples of residential living arrangments that give rise to each main type of peer group. Figure B.9 presents the floor plans of two dorms, Gore and Grays. Gore is composed of entryways A through E. Within each entryway, multiple rooms are connected by shared hallway space. For example, entryway B in the northwest (upper left) corner of the building contains rooms B21 through B25 on the second floor. These rooms are connected by a hallway around the stairs and past B24 and B25. We consider these groupings of rooms on a single floor within an entryway analogous to the floor (hallway) peer groups in Marmaros and Sacerdote (2006), and define residential peer groups in Gore and similar dorms as floors within entryways. Grays, on the other hand, has only two rooms per floor within an entryway, and no common space. Due to the absence of floor-level common space, we define residential peer groups in Grays and similar dorms at the entryway level.

Looking across all dorms, we find that Hollis, Holworthy, Lionel, Mower, Stroughton, Straus,

and Wigglesworth have similar layouts to Grays and define peer groups at the entryway level. Massachusetts, Matthews, McKinlock, Smith, Standish, Thayer, and Weld are similar to Gore, and we define peer groups at the entryway by floor level. Finally, Harvard Union and Shepherd are unusual in that they do not have multiple entryways. We define peer groups in those dorms as the entire dorm for Harvard Union and each floor within the dorm for Shepherd.

To understand how peer effects change with group definition and group size, we consider two alternate classifications based on residential proximity. The first is the *entryway*. This is a large grouping; it includes entire dorms when they have a single entrance, as is the case for Shepherd. The second is a smaller grouping that always defines peer groups as floor within entryway, regardless of the dorm's architectural features. We call this group *nearest neighbors*.

Figure B.17 displays histograms of group sizes for our main peer neighborhood definition (Panel A), for nearest neighborhors (Panel B), and for entryways (Panel C). Nearest neighbor groups are weakly smaller than our main peer neighborhood definition. Entryways are dramatically larger. The median size of a peer neighborhood is 8. For nearest neighbors it is 7, and for entryways it is 28, with the largest entryway groups containing 61 students.

Given the large differences in group sizes, direct comparisons of the effect of neighborhood rank across different spatial units are not sensible, especially for the larger entryway groups. Cross-group variation in mean group price is 44% as large for the entryway group as in our main measure. This means that a given change in rank within large groups corresponds to a much smaller change in mean price. For example, moving from the 10th percentile to the 90th percentile in the distribution of peer group means is associated with an $188 increase in mean price, averaging over years. For the larger entryway groups, this value is $126 . To make comparing magnitudes of regression estimates more straightforward, we change the units for our nearest-neighbor and entryway group definitions so that they correspond to where a group would place in the peer neighborhood rank in the same year. That is, we compute the mean price within a designated spatial unit, and then assign it a rank based on where it would fall in the peer neighborhood distribution by our main definition.

We use these modified rank measures to re-estimate equation 2 for alternate spatial groupings. *We view these specifications not as robustness tests on our main results, but as ways to explore how residential peer effects operate in larger and smaller groups.* This is for three reasons. First, the effects are conceptually different: social multiplier effects may depend on group size, as may students' ability to sort *within* spatial groups (Glaeser et al., 2003; Carrell et al., 2013). Second, we are less confident in the internal validity of the alternate measures as reflections of the constraints the lived environment imposes on social interaction than we are for our main measure. We view both as being worse approximations of the component of the peer environment that varies with residential room assignment than our main definition; they are either too inclusive or not inclusive enough. Third, for the larger entryway definition, the variation in peer environment

induced by randomization is more limited than in our main definition.

We report our findings in Tables B.16 through B.19. Tables B.16 and B.17 use the larger peer group definition, while Tables B.18 and B.19 use the smaller definition. Our central finding is that we see comparable estimates of peer effects at our alternative levels of aggregation. We leave the interpretation of this similarity for future work.

## B.8 Mechanisms in the quasi-experimental design

This subsection provides details on the analysis of mechanisms and heterogeneous effects summarized in section 5.8.

### B.8.1 Academic achievement and the opportunity cost of finance careers

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

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

Results reported in Online Appendix Table B.20 show that the career and social shifts we observe are driven by academic high-achievers. Assignment to higher-status residential peers raises the rate at which students both have high grades and are members of final clubs in college. In the long run the joint high grades and finance outcomes shifts up, while the joint occurence of high grades and medical or higher education careers shifts down (i.e., the academic high-achievers switch from medicine and higher education to finance).

The rates at which students hit these career and social milestones and have low grades are not affected by residential assignment.

### B.8.2 The Great Depression

Social ties may be particularly important when the economy is bad early in one's career (Kramarz and Skans, 2014). Harvard students in roughly the latter half of our sample (classes of 1930 and later) graduated into the teeth of the Great Depression. To understand whether peer ties were more valuable for Depression-exposed students than others, we estimate versions of equation 2 that split the sample by whether students graduated into the Depression or not. As reported in Online Appendix Table B.22, we find no evidence of heterogeneous effects on joining a final club while in college, adult social outcomes, or adult occupation choice. As reported in Online Appendix Table B.21, descriptive regressions of career outcomes from Class Report data suggest that final club membership is somewhat more closely correlated with private-feeder typical long-run outcomes for Depression cohorts than for other cohorts.[B.8] Final club membership may matter somewhat more for Depression cohorts, but the effects of residential peers do not appear to be a Depression-specific phenomenon.

### B.8.3 Major choice and career intent

School peers may shape long run outcomes by changing students' academic specialization in college, by changing their career goals at the time of college completion, or by changing how their careers play out conditional on specialization and goals. To understand where the effects we see come from, we use data on student major (available for entering cohorts from 1927 on) and occupational intent at the time of graduation (available for entering cohorts from 1924 on). We divide majors into four broad categories: Humanities, STEM, Economics, and other social sciences. Occupational intent comes from reported career choices in senior yearbooks. Students list things like "banking" or "law." We code stated intent using our main occupation classification scheme. As in our main analysis, we use the major and intent data to generate standardized indices of private-school typicality.

Online Appendix Table B.23 describes the sample of students for whom major data is available, and the characteristics of students choosing different kinds of majors. We observe major for 83% of students in cohorts where it is systematically reported. The subsample of students for whom major data is available closely resemble the sample population in terms of their pre-college characteristics, college activities, and long-run outcomes. The same is true for the 63% of students in 1924 and later cohorts who report data on career intent. The four major types we consider differ in terms of the students who sort into them. Humanities majors have the largest share of private feeder students (31.1%), with Economics and other social sciences in the middle,

[B.8] We do not consider Census wage outcomes in the descriptive specifications because we cannot differentiate differential effects of the Depression on final club members from differential experience-earnings profiles using earnings data from a single calendar year. This problem is not present in Class Report data, where outcomes are observed at the same experience level for all students (in different calendar years).

followed by STEM (16.5%). Career intent differs across majors in intuitive ways. For example, 14.8% of economics give finance as their intended career path, and 49.5% list business, while only 1.6% list Medicine. 38.1% of STEM majors intend to pursue medical careers, and only 0.7% list finance. Following students forward, we see that 13.7% of Economics majors work in finance 25 years later, compared to 2.9% of STEM majors. 27.0% of STEM majors work in medicine, but only 1.7% of Economics majors.

We study how major choice, grades, and social success in college interact to affect long run outcomes by regressing labor market outcomes on indicators for each major (taking STEM as the omitted category), high school type descriptors, first year class rank, and final club membership. All specifications also include cohort effects. Online Appendix Table B.24 reports the results. We find that the return to studying economics relative to STEM is nearly as large as the final club wage premium. The effect of majoring in economics on pursuing a finance career is about two thirds as large as the effect of final club membership.

The large return to majoring economics is consistent with recent studies at Harvard (Goldin and Katz, 2008) and elsewhere (Bleemer and Mehta, Forthcoming). A difference between our findings and those focusing on recent cohorts (Altonji et al., 2016) is that STEM fields are the lowest-earning in our data. The returns to studying economics do not change when we add controls for final club participation, and the returns to final club participation remain large when we control for major. Further, there is no interaction between having attended a private feeder high school and the effect of the economics major. In short, majoring in economics is not a mediator of the (large, positive) effects of final club membership on long-run labor market outcomes; it appears to offer a separate path to high-paying finance career tracks.

We then place major choice and career intent on the left side of the quasi experimental peer effects specification in equation 2. Results are reported in Online Appendix Table B.25. As in Sacerdote (2001), we find no effects of residential peers on major choice. Assignment to high-status peers does not shift private or public-school students towards private-school typical majors, and shifts in individual major categories are statistically insignificant, with the exception of a modest shift towards double-majoring. We find modest evidence of a shift towards private school typical occupational intent, and in indices that combine the major and intent data.

We construct an additional index that takes the finance career outcome as the dependent variable and bases predictions on the combined major and intent data. This index is non-standardized; it thus captures the shift towards finance careers that we would predict based on changes in major and career intent at graduation. For private feeder students, the changes in major and occupational intent can explain about one tenth of the observed shift towards finance careers.

We conclude that high-status residential peers shape long-run outcomes mainly by shaping career paths conditional on career intent and academic specialization, not by shifting broad ca-

reer goals or study plans while in college. At the same time, majoring in economics appears to offer a different path towards the finance career track than the "social" path from selective final clubs.

#### B.8.4 The marriage market

In addition to shaping career and social outcomes for high-status Harvard students, exposure to high-status residential peers may also shape outcomes in the marriage market. Historical accounts of Harvard social clubs such as Amory (1947) emphasize their importance for marriage, while studies in other settings emphasize the importance of marriage for the intergenerational transmission of social status (Ager et al., 2019).

To test this hypothesis, we use data on marriage outcomes from the Class Reports. These records report the wife's birth (i.e., maiden) name, which we classify according to our Colonial name score. As reported in Online Appendix Table B.26, we identify 74% of students as married in the sample to students merged to Class Report records. Private feeder and other students marry at similar rates, but Hasty Pudding and selective final club members are more to report being married, with rates for final club members reaching 82%.

We obtain spouse birth name scores for essentially all married individuals— more than 97%. Unscored names are those we could not match to population records and likely reflect typographical errors in the Class Reports or our transcription of the reports. We compute Jewish spouse and Catholic spouse indicators using spouse first and birth names, and Colonial indices using spouse birth name. We define a "Colonial marriage" as one where the Harvard student is married, has a Colonial name themselves, and is married to a wife with a Colonial birth name.

Conditional on being married, spouse attributes vary across groups. 12.1% of students have spouses with Jewish (birth) names, which we define as all last names with a Jewish index value above 0.7 (we maintain the 0.7 cutoff across each index type). 5.1% of private feeder students and 4.5% of final club members have spouses with Jewish names. In contrast, 24.8% of all students have spouses with Colonial names, with much higher rates for private feeder students (32.5%) and final club members (41.0%).

We also find strong evidence of assortative matching on cultural background. As reported in Online Appendix Figure B.18, Harvard students with high values of own-name indices are much more likely to marry spouses from that background. For example, roughly three-fourths of students with the highest values of the Jewish name index marry a spouse we classify as having a Jewish name.

To understand the relationship between marriage and career outcomes, we first present descriptive regressions of career and social outcomes on own attributes and attributes of the spouse and marriage. We report these findings in Online Appendix Table B.27. We find that marriage in general is not predictive of embarking on a finance career, but that marriage to a Colonial spouse

is. The effect is about one sixth the size of the effect of final club membership. There is no additional effect of being in a Colonial marriage— i.e., no interaction effect for both the husband and wife having Colonial names. In contrast, being married is strongly predictive of participation in adult social activities, with an effect nearly as large as the effect of final club membership. Marriage to a Colonial spouse also predicts adult social club membership, and in particular adult country club membership. The Colonial spouse effect is about one seventh the size of the final club effect. Marriage, career, and adult social outcomes are in all likelihood jointly determined so we interpret these regressions cautiously.

We then place marriage outcomes on the left side of our quasi-experimental peer effects specifications. We report these findings in Online Appendix Table B.28. In addition to our standard splits by high school type, we also split by whether the Harvard student had a Colonial name. This allows us to focus directly on assortativeness. The first row of Table B.28 reports effects for the final club membership outcome. We display these to show heterogeneous effects for Harvard students with and without Colonial names.

In the full sample, random assignment to high-status peers does not change Harvard students' chances of getting married at all, but does shift students towards Colonial spouses, and lead to higher rates of assortative Colonial marriages. Splitting the sample, we see that the net-zero effect on marriage rates emerges from large positive effects for private feeder students and especially students with Colonial names, coupled with modest negative effects for other groups. The overall Colonial spouse and Colonial marriage effects are likewise driven by larger effects for private feeder and Colonial Harvard students. Though the observed effects are large relative to sample means, scaling the effects by the descriptive relationship between marriage outcomes and career or social outcomes reported in Table 7 suggests that they can account for only a small part of the relationship between peer attributes and long-run outcomes reported in our main peer effects results.

We conclude that assignment to high-status residential peers shifts students towards high-status spouses and increases marital assortativeness. This is consistent with the idea that interactions with high-status peers at college shape important life outcomes in the long run.

## B.9 Long-run time series

### B.9.1 Construction

This section describes the construction of the 1924-1990 long-run series used in section 6.1, and discusses the representativeness and consistency over time of the data sources we use. We make six main points. First, changes in the content and timing of Red Book publications mean that we cannot track first-year social activities over the long run. Second, we observe high school background over the full time series. Third, we classify student gender on the basis of their first

and middle names over the full time period, and these classifications closely track published statistics in the HEGIS/IPEDS in years when both are available. Fourth, we classify student race/ethnicity by combining surnames with visual codes. These match published aggregates for the share of Asian students over the full period. They match published shares of URM students through the 1980s and but somewhat undercount URM students relative to published aggregates in the 1990s and later. Fifth, we are able to observe students' Latin honors at graduation over the full period, and the rates of honors graduation that we observe track published aggregate statistics. Sixth, our coverage of occupation declines somewhat over time, but we see no evidence of differential attrition on the basis of high school type, and the values we obtain are similar to those observed in surveys of Harvard students. Our overall assessment is that the long-run series we assemble provides reliable evidence on trends in academic and occupational attainment, and differences in these outcomes by high school type.

We construct the long run series by extending the two main sources of Harvard records used in the main analysis: Freshman Red Books and 25th Reunion Class Reports. See section 3 for an extended discussion of these sources. We digitize Red Book records through the entering class of 2019. We have Class Report records for every graduating class through 1940 and then at five year intervals starting in 1945 and continuing through the graduating class of 1990. At the time of digitization, 1990 was the last 5-year interval for which 25-year Class Report outcomes were available. As in our main analysis, we define our sample universe using the Red Book records. Panel A of Figure B.19 shows how Red Book cohort sizes evolve across class years. Red Book counts closely match cohort sizes reported in HEGIS/IPEDS data when these series first become available in the late 1960s. This is consistent with the observation that our pre-war Red Book sample universe matches the Harvard class size from that era. Red Book counts hold steady at 85-90% of IPEDS levels in the 1980s and 1990s. They then drop briefly in the early 2000s before rebounding to match IPEDS counts almost exactly in the 2010s. In short, the Red Books appear to include the vast majority of Harvard enrollment over the full century.

Ideally, our long run analysis would have included data on first-year social activities. However, changes in the information included in the Red Books makes doing so impossible. Prior to 1940, Red Books were published in the spring, at the end of students' first year. These spring books are the main source for social activity data in our analysis of data from the 1920s and 1930s. Beginning in 1940, Harvard published Red Books in both the fall and the spring. The fall Red Book lists students' high schools and hometowns, but not college activities, while the spring Red Books continue to list college activities but with decreasing detail. In 1950, the spring Red Books cease publication, leaving only the fall Red Books. The fall Red Book series continues through the present. Our analysis uses spring Red Books through the end of their publication in 1949, then switches to the fall books for the remaining years.

The one piece of information on social outcomes we do have for the full time series is a list

of members of the A.D. club for each class from 1837 through 2015. The A.D. club is one of the selective final clubs we discuss in section 3. We obtain this data from the 2015 edition of the A.D. club catalog (A.D. Club, 2015). We would have liked to obtain similar records for each final club. However, unlike the other Harvard records we use for this paper, club catalogs are not available through the Harvard archives, or any of the other major archives we searched. We located the A.D. club catalog in a non-archival collection. The document, which is distributed to members, contains a history of the club, officer and membership lists for each class year, and cross-indices of members by family name and current location (if applicable). Figure B.20 shows an example record page. Comparisons of the membership lists reported in the club catalog to those in our archival data suggests the lists are accurate.

Women were admitted to Harvard as undergraduates starting in the early 1970s. This change is visible in the cohort size data; class size rises to partially accommodate the addition of female students. We identify the gender of Harvard students using data on name frequency by gender and birth year from the Social Security Administration (2021). We first search for a perfect match of the first name in the cohorts close to the expected birth year of a student. If we find a unique perfect match, we assign the corresponding gender. If we find multiple perfect matches, i.e. the name has been given to both male and female babies, we proceed by searching for perfect matches on the middle name if recorded in our data. If there is unique perfect match for the middle name, we assign the corresponding gender. Otherwise, we stop and assign the gender for which the first name frequency is higher. If no perfect match can be found, we calculate the Damerau-Levenshtein distances to all candidate names. We require the set of likely matches to have a distance of less than three. We then assign the gender for which the total frequency among these likely matches is the highest. Panel B of Figure B.19 shows that this procedure works well. The share of students we identify as male is almost exactly one prior to the entering cohort in 1971; it declines slowly, reaching 50% by the early 2000s. Gender shares in our data are almost exactly equal to those reported in HEGIS/IPEDS data in the years records of both types are available.

We identify race/ethnicity in the long run data using a combination of surnames and visual codes. For the visual codes, a team of research assistants inspected the Red Book photographs and coded students as white, Black, Latino, Asian, or other non-white. They did this for each class year through 1955 and then in three-year blocks surrounding the digitized Class Reports in later years. For example, they coded the 1959, 1960, and 1961 Red Books for the 1960 report, the 1964, 1965, and 1966 blocks for the 1965 report, and so on. They continued to follow this pattern even after the last available Class Report in 1990. We report long-run race/ethnicity statistics only in the years for which we have visual codes. We augment the visual codes with surname indices. For the name codes, we compute our Jewish and Colonial name indices (using surnames only) for all class years and apply the 0.7 cutoff for binary classification. In contrast

to our main-text analysis, we classify the small share of students with missing Jewish (Colonial) name indices as non-Jewish (non-Colonial). See section B.5 for details. We define a surname as characteristically Black, Hispanic, or Asian if the 2010 Census Surname Table reports that at least 80% of individuals with the name have the given racial/ethnic background. Students whose last names do not match to the Census surname table are coded as not having characteristically Black, Hispanic, or Asian names. We categorize students as Black, Hispanic, or Asian if either the visual or surname coding scheme identifies them as such. In our main analysis, we aggregate Black and Hispanic students into a single under-represented minority (URM) category. We split white students into three groups: Colonial surnames, Jewish surnames, and other white students.

We aggregate the codes into the mutually exclusive race/ethnicity categorization scheme used in the main text. For individuals with multiple codes, the order of precedence is Asian, URM, Jewish, Colonial, other white. The small (<1%) share of students for whom we could not obtain visual codes or whom we classify as other non-white (not Asian, Black, or Hispanic) are omitted from reported race/ethnicity splits.

Panel C of Figure B.19 compares the race/ethnicity shares we compute in our data to those reported in the IPEDS. IPEDS reports begin in 1980. An important contribution we make here is to show how the shares of non-white students at Harvard evolved before that. There were very few such students before the mid-1960s, when shares of Black students began to rise. Our race codes give a near-exact match to IPEDS reports of the share of Asian students. We track IPEDS share of Black and Hispanic students through the 1980s. In the 1990s and later, IPEDS reports of Black and Hispanic shares continue to rise while the shares we compute remain fairly steady at 1980s levels. IPEDS data are based on self-reports; the contrast between the indices we create and these reports may be an interesting topic for future work. In any case, note that the last entering year considered in our long-run analysis of grades and career outcomes is 1986, before our counts diverge from IPEDS values.

Panel D of Figure B.19 shows that the share of students graduating with any kind of honors designation rose from 29% in the mid 1920s to 77% in the late 1980s, with the fastest expansion coming in the 1960s. To smooth over smaller samples in the Class Report data, each point on this graph and those that follow displays means over graduating classes within 2.5 years of the centered value. For example, the 1982.5 data point is an average of the 1980 and 1985 class years. Our findings closely match aggregate data from Harvard administrative sources reported in Healy (2001); we display the Healy data on the graph as well. Because most students graduate with honors by the late 20th century, we use a definition of academic high achievement that includes only students awarded a *magna* or *summa cum laude* degree.[B.9] By this definition, the share of academic high achievers rises from 8% in the mid 1920s to 33% in the late 1980s.

Match rates from Red Book data to Class Report data remain steady over time. The upper

[B.9] Healy (2001) attributes the increase in honors degrees to grade inflation during the Vietnam War.

line in Panel E of Figure B.19 shows the share of Red Book records linked to Class Report data. We match roughly 80% of Red Book records to Class Report data over the full time series.

The detail of Class Report records falls somewhat over time, particularly after 1980. The lower two lines in Panel E of Figure B.19 show the share of Red Book records a) linked to a Class Report record that includes the student's occupation, and b) linked to a Class Report record that includes the student's educational history. The rates at which students report occupation conditional on Class Report match fall from roughly 90% in our main analysis period to roughly 70% from 1940 through 1970 to roughly 53% from 1980 on. The rates at which students report their educational history are steady from the 1920s through the 1970s at roughly 72% before falling to roughly 58% from 1980 on.

Three features of the data mitigate concerns about reduction in data quality in the later part of the series. First, measures of academic achievement that we observe in our data closely track those published in aggregate statistics, as shown in Panel D of Figure B.19. The sample of students for whom educational histories are available in Class Report data is representative of the broader Harvard population.

Second, rates of graduate degree receipt we observe are also similar to those reported in previously published work. The degrees we use for career classification are MDs, JDs, and MBAs. Goldin and Katz (2008) report the rates at which Harvard students receive these degrees. Their focus is on Harvard graduating classes within four-year windows around 1970, 1980, and 1990. There are many reasons that rates in our data might differ from rates in Goldin and Katz (2008), including differences in the graduating class years we analyze and differences in the timing of survey reporting in the life course.[B.10] Nevertheless, our findings are quite similar. For MD degrees, GK report rates of degree receipt among male students in classes near 1970, 1980, and 1990 as 18.5%, 18.0%, and 14.3%, respectively. For us, these values are 17.1%, 19.5%, and 13.6%. For JD degrees, GK report values of 24.8%, 24.1%, and 20.4%. We report values of 22.9%, 21.6%, and 20.7%. For MBA degrees, GK report values of 10.9%, 20%, and 19.4%, while we report values of 9.4%, 18.4%, and 15.1%.

Third, there is little evidence of differential change in match rates to occupation records. Panel F of Figure B.19 display match rates to occupational data for students in the Red Books, splitting by high school type. Match rates are similar for private feeder and public feeder students over the full period. These are the two groups for whom we draw cross-time comparisons in section 6.1. The observation that coverage declines similarly for these two groups makes us less worried that changes in selection into the data over time could affect the cross-group comparisons we draw.

Our overall conclusion is that data we have provide a credible picture of long-run changes in

[B.10] Goldin and Katz (2008) conduct all surveys in 2006; for us, all degree reports are from 25 years following graduation.

academic and occupational outcomes by group.

### B.9.2 Changes in academic and career outcomes for men

Our main text analysis of long-run patterns in academic and career outcomes uses data for all Harvard students. Online Appendix Figure B.21 replicates our main long-run trends figure in the sample of male students. We observe the same patterns— persistent differences in academic performance by high school type and race/ethnicity, convergence in finance careers, divergence in MBAs— as in our main analysis.

### B.9.3 Job classification in the extended time series

**Overview.** Classifying occupations over the long run creates additional challenges relative to our main classification scheme because the way Harvard graduates describe their jobs may change over time. We address this issue in four ways.

1. We focus on a narrower set of careers than in our main analysis. Specifically, we limit our long-run analysis to finance, higher education, medicine, and law.

2. For medicine and law, we use descriptors of graduate education that map closely to career choices and are easy to code consistently over time, rather than the text descriptors of job title or occupation.

3. For all four career types, we expand the set of job descriptors we consider to cover new ways of describing these careers that emerge over time. In addition, we introduce regular expression representations of the existing identifiers, to account for spelling mistakes more flexibly.

4. For all four career types, we consider alternate classification approaches and reach substantively similar conclusions to those presented in the main analysis.

These supplementary analyses reinforce the argument in the main text that differences in career outcomes by high school type persist over the the 20th century, by showing that the conclusion holds across a variety of measures of career outcomes.

**Extended coding of careers.** Table B.29 shows the set of strings we add to those we code as finance, legal, higher education and medical careers for the extended analysis. These strings tend to pick up job title or detailed occupation descriptions, for example "fixed income." We also consider alternate specifications for finance careers that add names of large investment banks (current and historical) to list of strings we consider, to pick up people who do not list an occupation but do list an employer. We do not use the firm names in the analysis in the main text.

Table B.29 also displays the list of banks we consider in the row referring to the subcategory "Firms (ext. definition)".

**Alternate classification approaches.** We consider alternate classification approaches for each of our four career categories. For finance careers, we consider an augmented classification approach that adds names of major investment banks to our baseline scheme. Figure B.22 compares results from this exercise with those from our main analysis. Including firm names raises our estimate of the rate at which students pursue finance careers slightly from the late 1970s and onward. Our central conclusion that there is a large gap by high school type that converges by the end of the period is unchanged.

For medical and legal careers, we compare our main measures based on MD and JD/LLB receipt with measures based on reported job outcomes. Figure B.23 shows how the share of students reporting MD and JD/LLB degrees relates to the share of students reporting medical and legal careers in the sample of students who report both career and degree information in the Class Report. We do not expect these measures to be identical; degree and career outcomes need not correspond precisely. For medical degrees, measures of medical careers based on job description are slightly higher than those based on degree receipt prior to 1950; after 1950 the measures are nearly identical. For law careers, measures are near identical prior to 1970. After 1970, measured rates of law careers in job descriptions fall off relative to JD receipt. Inspection of records suggests that this reflects a combination of people who receive law degrees but work outside of the law field, and the greater difficulty in classifying workers who list only their firm and not occupation descriptors in the more recent Class Reports.

Figure B.24 reports long-run trends by high school type for the alternate measure of career attainment. The "doctor" panel shows the rates at which students' job description fall into the medical doctor/health career category given in Table B.3. While public high school students are more likely to pursue medical careers throughout the period, we see more evidence of convergence by high school type for this measure than for MD degrees. This may reflect a higher propensity for public feeder students with MD degrees to work outside of the medical field, for example in research positions. For law, our results are essentially the same as for the JD measure, with little evidence of a gap by high school type. For PhD degrees, we see a similar overall pattern as we did for higher education careers in the main text, with similar rates for private and public feeder students in the 1920s and 1930s, but a large gap by the end of the period. In the 1980s and 1990s, the gap in PhD receipt by high school type is larger than the gap we observe in higher education careers.

### B.9.4 High school classification in the extended time series

Our long-run analysis extends the set of high schools in the "public feeder" and ''other private" categories. We add schools to these categories by examining the set of high schools listed at

least ten times across in the cohorts 1925, 1930, 1935,..., 2010. We categorize each such school not already included in our main-sample categorization scheme. The only exceptions to this rule are schools with common names, such as "Jesuit" or "Classical" that are difficult to consistently differentiate from one another. Through this process we add 18 public feeder schools and 72 private schools to the public feeder and other private categories. See Table B.30 for a list of these schools.

## B.10 Additional cross-time comparisons

This section describes the empirical exercises underlying the results discussed in section 6.5.

### B.10.1 The return to academic success

We observe a final club membership premium that is much larger than the academic success premium. Comparing our findings to Goldin and Katz (2008) suggests that the returns to academic performance at Harvard may have risen over time, but that the social success premium we observe is large even compared to returns to academic success for relatively recent cohorts. Goldin and Katz (2008) find that a one standard deviation increase in college grades raises earnings by 15 to 20%.[B.11] Converting our estimates to standard deviation units, the largest effects we observe across multiple specifications correspond to a 5 to 7% premium. We report these results in columns 5 and 6 of Panel C in Online Appendix Table B.15, which focuses on third-year rather than first year class rank. In particular, column 6 adopts the Goldin and Katz approach of scaling topcoded earnings values by 1.4. See Online Appendix B.6.5 for details. The difference between our estimates of grade effects and the Goldin and Katz estimates may reflect differences in measurement (they use survey and administrative reports of income and cumulative GPA; we use census income data and year-specific class rank). The 29% final club premium that we estimate in our main specifications is equivalent to about a 1.5 to 2$\sigma$ change in grades using the Goldin and Katz estimate, and a 4$\sigma$ change using our largest in-sample estimate.

### B.10.2 Sports participation

The returns to athletic participation we observe in our data are similar to reported for more recent cohorts, and appear to be mediated in part by social success. Shulman and Bowen (2011) study the academic and labor market correlates of sports participation using College & Beyond survey data for the class of 1976. They find that male college athletes earn about 13% more than non-athletes, despite having class ranks between five and ten percentiles lower than other students (the latter statistic controls for baseline covariates like test scores). They also find that

[B.11] These coefficients are estimated but not reported in the paper. We thank Larry Katz for providing these results.

athletes choose very different careers than other students. Shulman and Bowen report that 24% of athletes are executives, 14% work in financial services, compared to 19% and 10%, respectively, for non-athletes.

To compare what we observe to Shulman and Bowen (2011), we use our data on first-year activities to identify individuals who participate in intercollegiate (as opposed to intramural) sports. The different levels of competition available to Harvard students during this period do not always have clear parallels to those in modern college athletics. Our guiding principle is to include students on teams that compete for Harvard against other universities. Online Appendix Table B.31 reports the shares of students participating in different types of sports, split by high school type. 33% of students in what we term "schoolwide" sports, with higher rates for private feeder students. The most popular sports are Crew, Track, and Football.

Simple descriptive statistics in our data are similar to those in Shulman and Bowen (2011). Intercollegiate athletes earn about 12% more than non-athletes in our sample. 24% of athletes are in the "high management" occupation, and 14% work in finance, compared to 20% and 8%, respectively, for non-athletes. We do not observe pre-college test scores or precise class rank. However, if we assign each student the class rank equal to the median rank within their coarse rank group, we find that athletes are ranked an average of 7 percentiles lower than non-athletes.

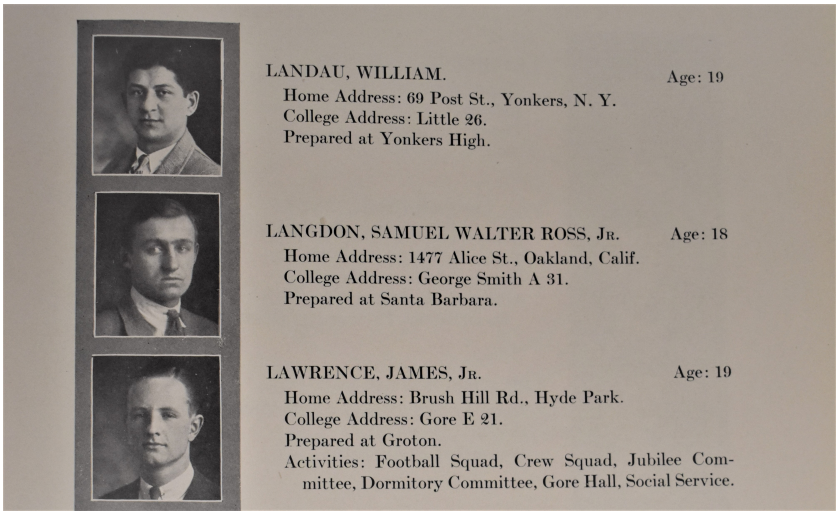
To shed light on how social success mediates the effect of athletic participation, we regress labor market outcomes on sports participation and student covariates, with and without controls for college social activities. Results are reported in Online Appendix Table B.32. Sports participation predicts worse academic performance and an increased likelihood of final club participation, even conditional on high school type. Sports also predict higher earnings and an increased likelihood of going into finance, conditional on high school type. Adding controls for social success— defined here as final club membership, Hasty Pudding membership, and participation in a first-year social committee– reduces the coefficient on sports participation by about 50%. Sport-specific regressions in columns 7 and 8 suggest that cross-sport heterogeneity may be substantial; point estimates are largest for basketball, track, and football.

### B.10.3 Field of study

See section B.8.3 above for a discussion of field-specific premia over time.

## B.11 Figures

Figure B.1: Red book information on freshmen residence and activities



LANDAU, WILLIAM. Age: 19
Home Address: 69 Post St., Yonkers, N. Y.
College Address: Little 26.
Prepared at Yonkers High.

LANGDON, SAMUEL WALTER ROSS, Jr. Age: 18
Home Address: 1477 Alice St., Oakland, Calif.
College Address: George Smith A 31.
Prepared at Santa Barbara.

LAWRENCE, JAMES, Jr. Age: 19
Home Address: Brush Hill Rd., Hyde Park.
College Address: Gore E 21.
Prepared at Groton.
Activities: Football Squad, Crew Squad, Jubilee Committee, Dormitory Committee, Gore Hall, Social Service.

Missing dorm information for entering class of 1926. See Section B.1.1 for more details on the Red Books.

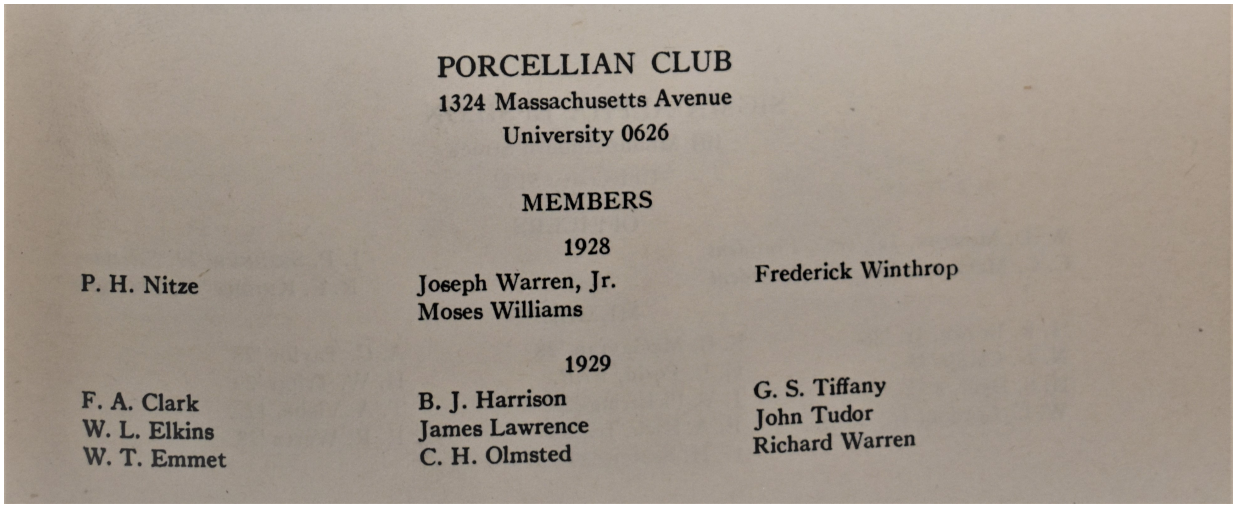
Figure B.2: Red book lists of freshmen activity participants

Crew O—The First Crew

| *Name* | *Position* | *Age* | *Height* | *Weight* | *School* |
|---|---|---|---|---|---|
| J. Lawrence | B. | 18 | 6′¾″ | 166 | Groton |
| E. Hamlen | 2 | 19 | 6′1″ | 173 | Groton |
| W. T. Emmet | 3 | 19 | 6′3¾″ | 176 | St. Paul's |
| F. A. Clark | 4 | 20 | 6′5″ | 198 | St. Mark's |
| G. Murchie | 5 | 19 | 6′4¾″ | 189 | Kent |
| B. J. Harrison | 6 | 19 | 6′4½″ | 186 | St. Paul's |
| G. N. Saum | 7 | 19 | 5′10½″ | 176 | Watseka High |
| C. McK. Norton | S. | 19 | 6′¼″ | 167 | Groton |
| C. I. Neiman | Cox | 17 | 5′ | 94 | Boston Latin |

See Section B.1.1 for more details on the Red Books.

Figure B.3: Student council register information including club membership



PORCELLIAN CLUB
1324 Massachusetts Avenue
University 0626

MEMBERS

1928

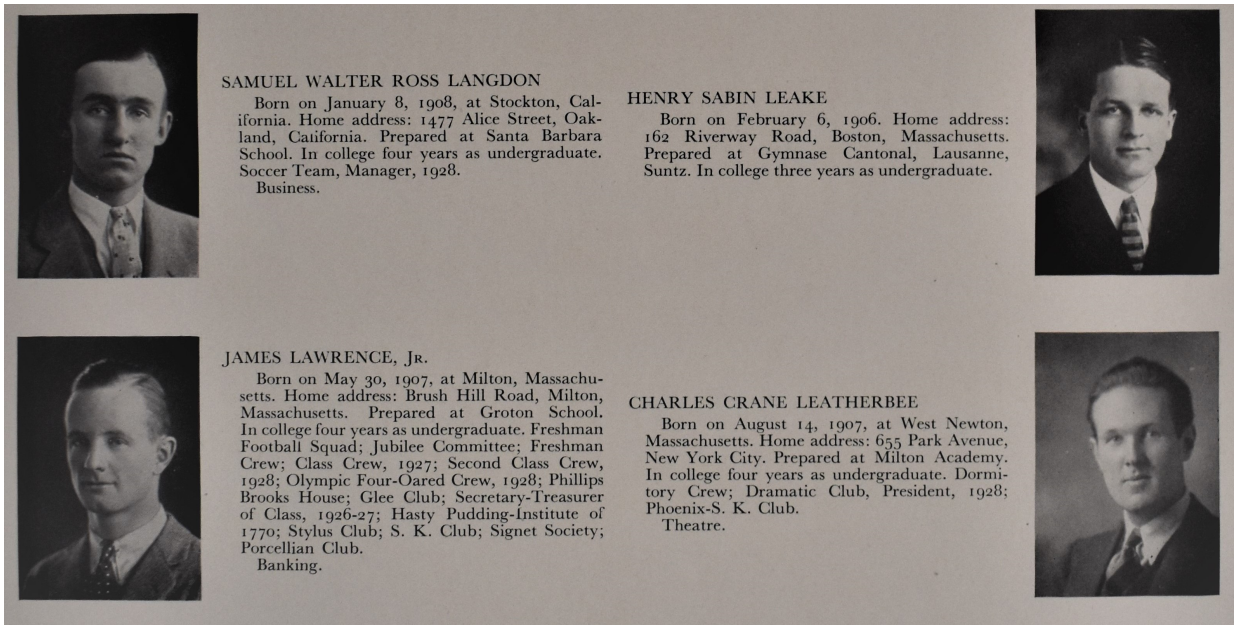
P. H. Nitze
Joseph Warren, Jr.
Moses Williams
Frederick Winthrop

1929

F. A. Clark
W. L. Elkins
W. T. Emmet
B. J. Harrison
James Lawrence
C. H. Olmsted
G. S. Tiffany
John Tudor
Richard Warren

Missing after school year 1927-28. See Section B.1.2 for more details on the Student Council Registers.

Figure B.4: Class albums, includes club membership





SAMUEL WALTER ROSS LANGDON

Born on January 8, 1908, at Stockton, California. Home address: 1477 Alice Street, Oakland, California. Prepared at Santa Barbara School. In college four years as undergraduate. Soccer Team, Manager, 1928.
Business.

HENRY SABIN LEAKE

Born on February 6, 1906. Home address: 162 Riverway Road, Boston, Massachusetts. Prepared at Gymnase Cantonal, Lausanne, Suntz. In college three years as undergraduate.





JAMES LAWRENCE, Jr.

Born on May 30, 1907, at Milton, Massachusetts. Home address: Brush Hill Road, Milton, Massachusetts. Prepared at Groton School. In college four years as undergraduate. Freshman Football Squad; Jubilee Committee; Freshman Crew; Class Crew, 1927; Second Class Crew, 1928; Olympic Four-Oared Crew, 1928; Phillips Brooks House; Glee Club; Secretary-Treasurer of Class, 1926-27; Hasty Pudding-Institute of 1770; Stylus Club; S. K. Club; Signet Society; Porcellian Club.
Banking.

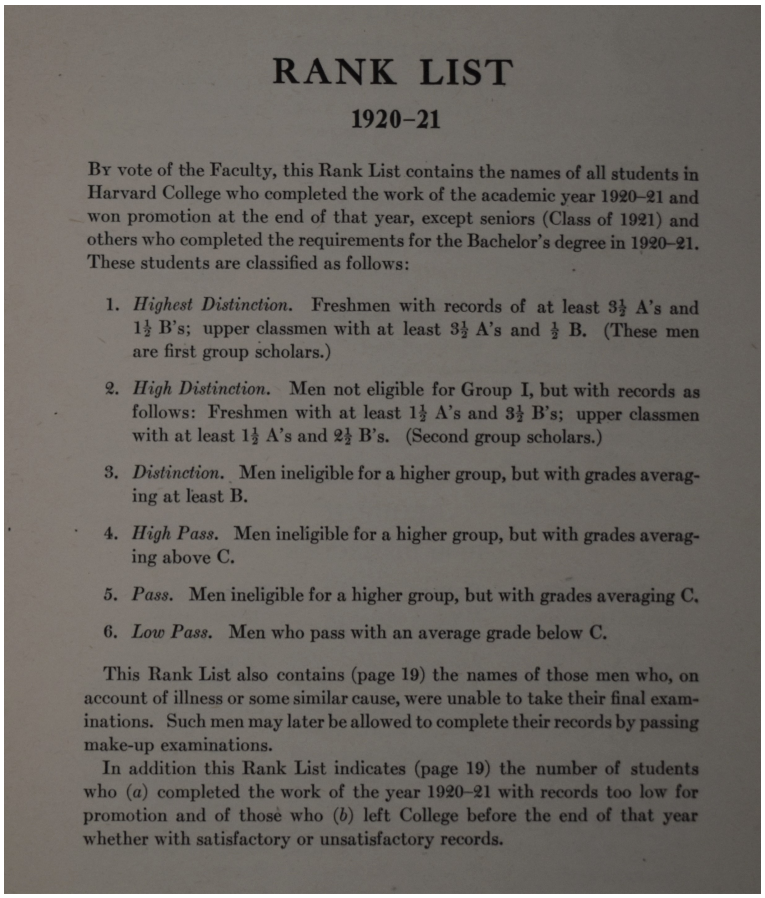
CHARLES CRANE LEATHERBEE

Born on August 14, 1907, at West Newton, Massachusetts. Home address: 655 Park Avenue, New York City. Prepared at Milton Academy. In college four years as undergraduate. Dormitory Crew; Dramatic Club, President, 1928; Phoenix-S. K. Club.
Theatre.



Missing for students who do not persist to graduation. Used to supplement Student Council Register information for years after the Register was discontinued. See Section B.1.3 for more details on the Class Albums.

Figure B.5: Rank group lists explanation of ranks and inclusion



# RANK LIST

## 1920–21

By vote of the Faculty, this Rank List contains the names of all students in Harvard College who completed the work of the academic year 1920–21 and won promotion at the end of that year, except seniors (Class of 1921) and others who completed the requirements for the Bachelor's degree in 1920–21. These students are classified as follows:

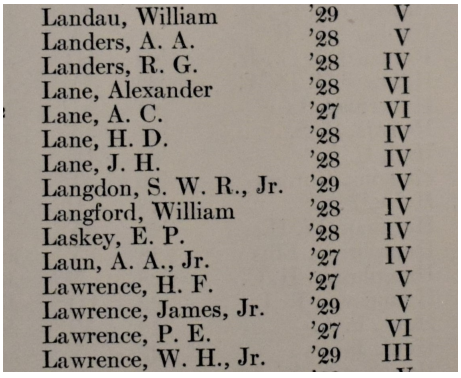
1. *Highest Distinction.* Freshmen with records of at least 3½ A's and 1½ B's; upper classmen with at least 3½ A's and ½ B. (These men are first group scholars.)
2. *High Distinction.* Men not eligible for Group I, but with records as follows: Freshmen with at least 1½ A's and 3½ B's; upper classmen with at least 1½ A's and 2½ B's. (Second group scholars.)
3. *Distinction.* Men ineligible for a higher group, but with grades averaging at least B.
4. *High Pass.* Men ineligible for a higher group, but with grades averaging above C.
5. *Pass.* Men ineligible for a higher group, but with grades averaging C.
6. *Low Pass.* Men who pass with an average grade below C.

This Rank List also contains (page 19) the names of those men who, on account of illness or some similar cause, were unable to take their final examinations. Such men may later be allowed to complete their records by passing make-up examinations.

In addition this Rank List indicates (page 19) the number of students who (*a*) completed the work of the year 1920–21 with records too low for promotion and of those who (*b*) left College before the end of that year whether with satisfactory or unsatisfactory records.

Available for all non-graduating students school years 1920-21 through 1930-31. Available for Freshmen only starting in school year 1931-32. See Section B.1.4 for more details on the rank group lists.

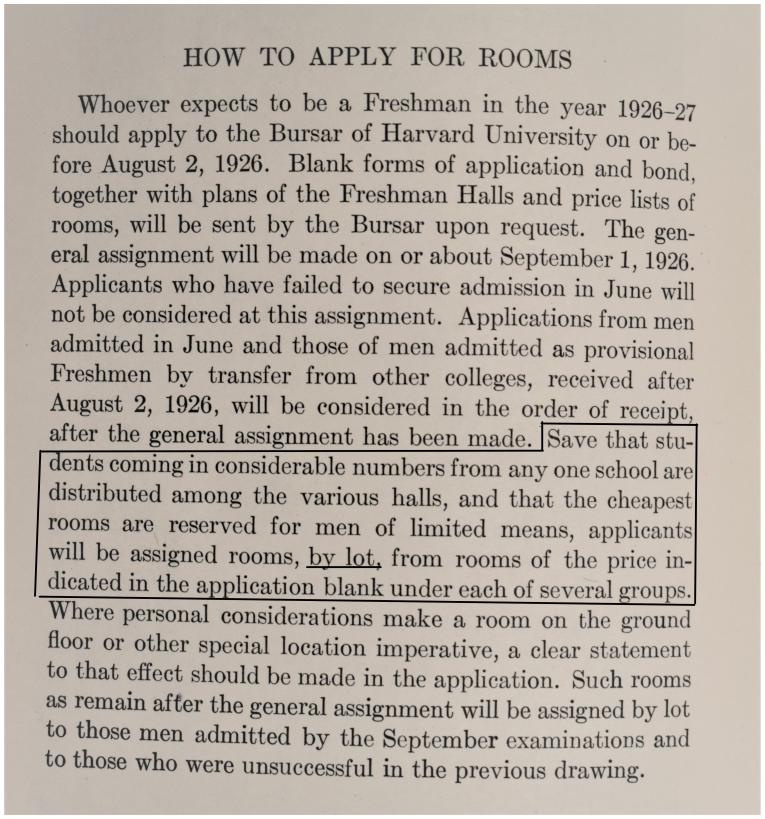
Figure B.6: Rank group lists



| | | |
|---|---|---|
| Landau, William | '29 | V |
| Landers, A. A. | '28 | V |
| Landers, R. G. | '28 | IV |
| Lane, Alexander | '28 | VI |
| Lane, A. C. | '27 | VI |
| Lane, H. D. | '28 | IV |
| Lane, J. H. | '28 | IV |
| Langdon, S. W. R., Jr. | '29 | V |
| Langford, William | '28 | IV |
| Laskey, E. P. | '28 | IV |
| Laun, A. A., Jr. | '27 | IV |
| Lawrence, H. F. | '27 | V |
| Lawrence, James, Jr. | '29 | V |
| Lawrence, P. E. | '27 | VI |
| Lawrence, W. H., Jr. | '29 | III |

Available for all non-graduating students school years 1920-21 through 1930-31. Available for Freshmen only starting in school year 1931-32. See Section B.1.4 for more details on the rank group lists.

Figure B.7: Room assignment procedure 1922-1941

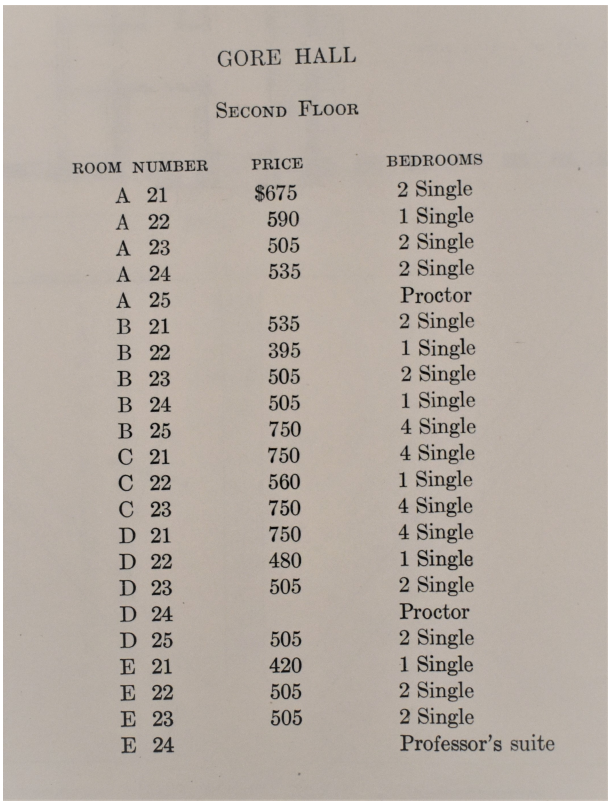


HOW TO APPLY FOR ROOMS

Whoever expects to be a Freshman in the year 1926–27 should apply to the Bursar of Harvard University on or before August 2, 1926. Blank forms of application and bond, together with plans of the Freshman Halls and price lists of rooms, will be sent by the Bursar upon request. The general assignment will be made on or about September 1, 1926. Applicants who have failed to secure admission in June will not be considered at this assignment. Applications from men admitted in June and those of men admitted as provisional Freshmen by transfer from other colleges, received after August 2, 1926, will be considered in the order of receipt, after the general assignment has been made. Save that students coming in considerable numbers from any one school are distributed among the various halls, and that the cheapest rooms are reserved for men of limited means, applicants will be assigned rooms, by lot, from rooms of the price indicated in the application blank under each of several groups. Where personal considerations make a room on the ground floor or other special location imperative, a clear statement to that effect should be made in the application. Such rooms as remain after the general assignment will be assigned by lot to those men admitted by the September examinations and to those who were unsuccessful in the previous drawing.

Emphasis added. Prior to 1922 and after 1944 students' ranked preferences were taken into account rooms assigned taking ranked preferences into account. Procedure from 1942-1944 unknown. See Section B.1.5 for more details on freshman housing.

Figure B.8: Price lists



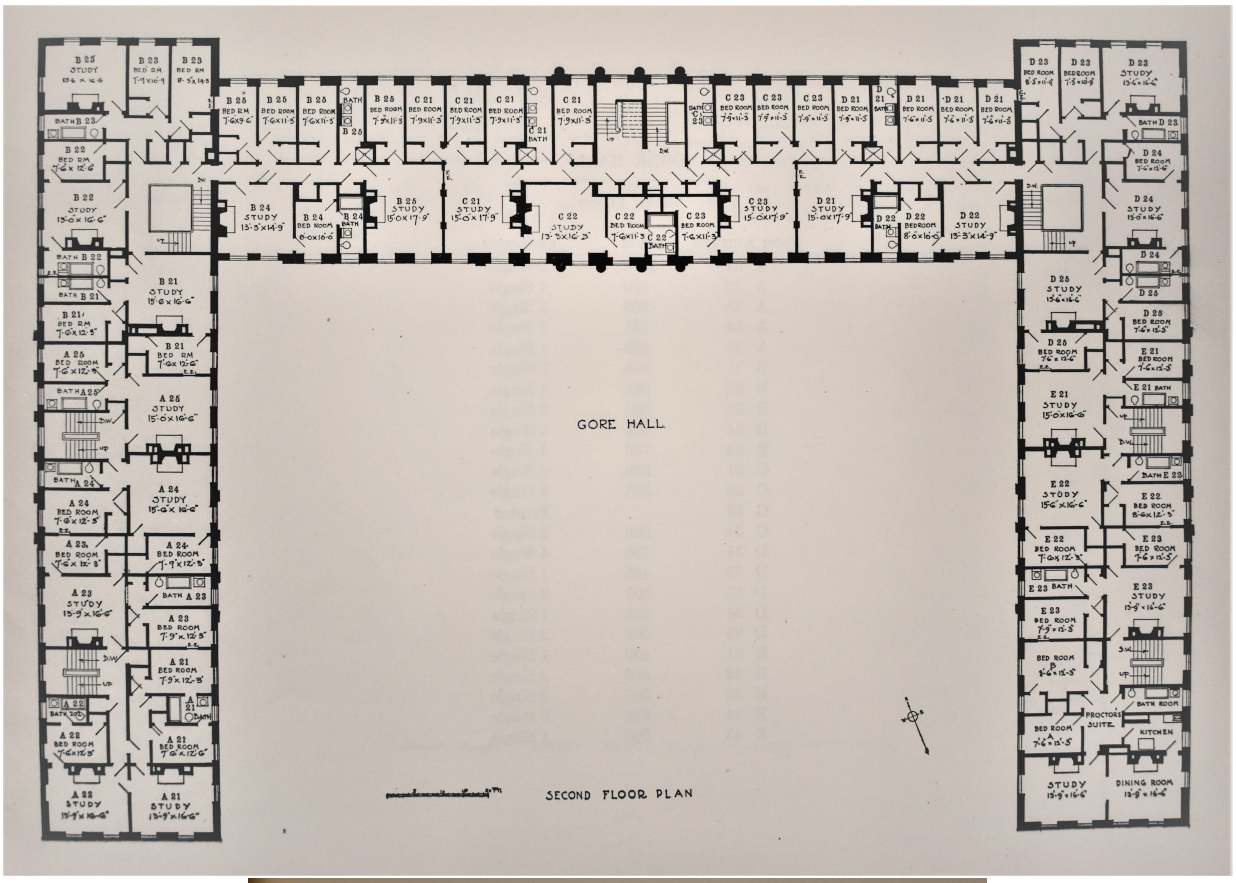
GORE HALL

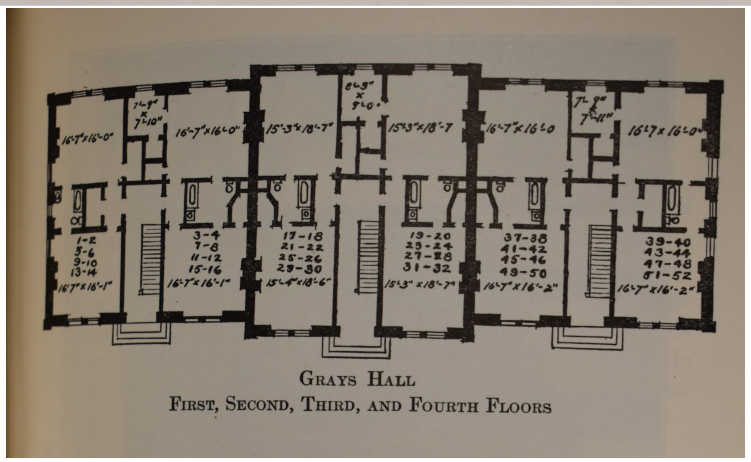
SECOND FLOOR

| ROOM NUMBER | PRICE | BEDROOMS |
|---|---|---|
| A 21 | $675 | 2 Single |
| A 22 | 590 | 1 Single |
| A 23 | 505 | 2 Single |
| A 24 | 535 | 2 Single |
| A 25 | | Proctor |
| B 21 | 535 | 2 Single |
| B 22 | 395 | 1 Single |
| B 23 | 505 | 2 Single |
| B 24 | 505 | 1 Single |
| B 25 | 750 | 4 Single |
| C 21 | 750 | 4 Single |
| C 22 | 560 | 1 Single |
| C 23 | 750 | 4 Single |
| D 21 | 750 | 4 Single |
| D 22 | 480 | 1 Single |
| D 23 | 505 | 2 Single |
| D 24 | | Proctor |
| D 25 | 505 | 2 Single |
| E 21 | 420 | 1 Single |
| E 22 | 505 | 2 Single |
| E 23 | 505 | 2 Single |
| E 24 | | Professor's suite |

Available for 1920, 1932-1941. See Section B.1.5 for more details on freshman housing.

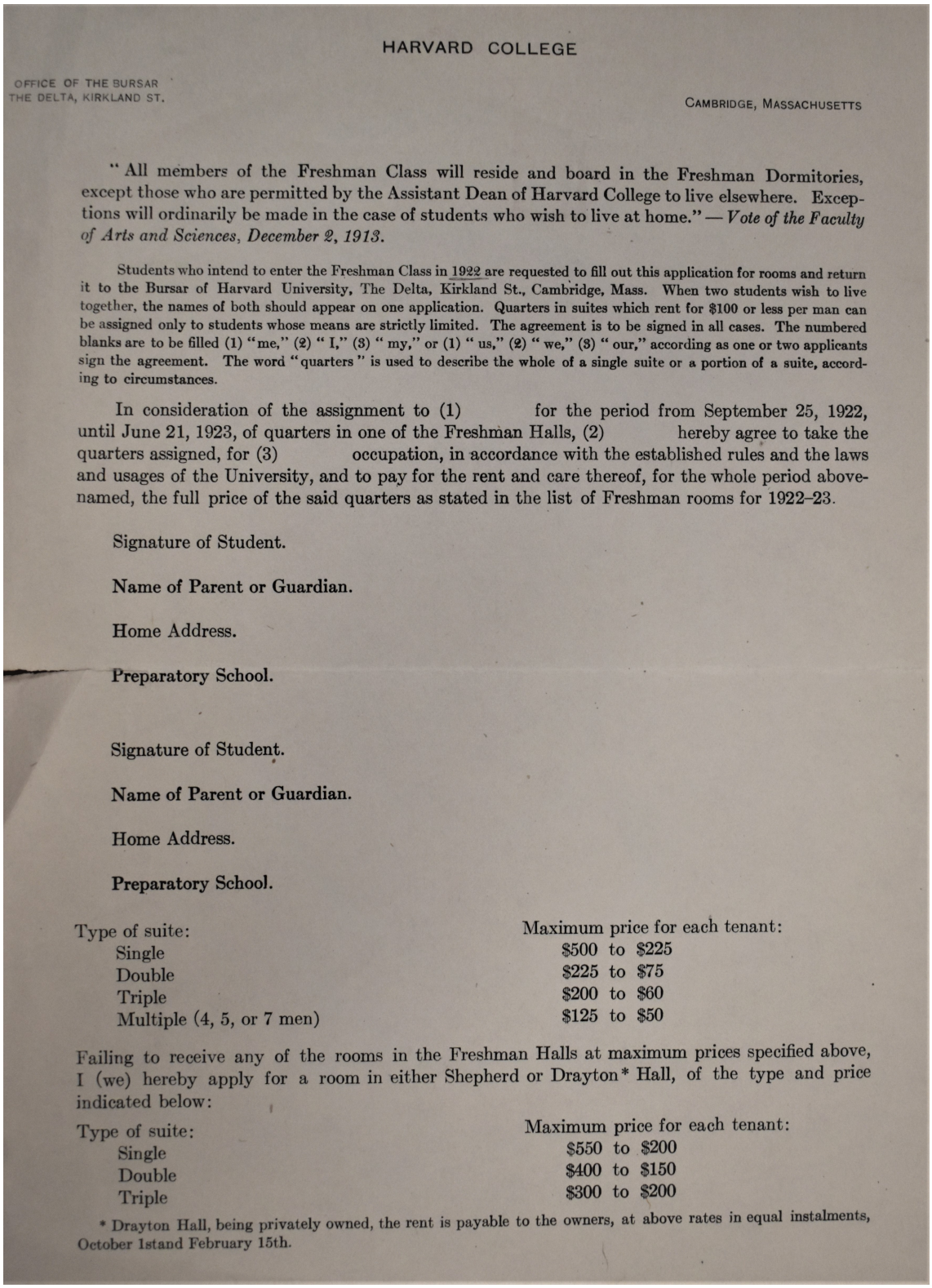
Figure B.9: Floor plan




Available for 1920, 1932-1941. The top image depicts Gore as an example of a dorm where the combination of floor and entryway (denoted by letter) is our preferred peer group definition. The bottom image depicts Grays as an example of a dorm where entryway is our preferred peer group definition. See Section B.1.5 for more details on freshman housing and Section B.7 for details on how we defined peer groups.

Figure B.10: Freshmen housing application blank



HARVARD COLLEGE

OFFICE OF THE BURSAR
THE DELTA, KIRKLAND ST.

CAMBRIDGE, MASSACHUSETTS

"All members of the Freshman Class will reside and board in the Freshman Dormitories, except those who are permitted by the Assistant Dean of Harvard College to live elsewhere. Exceptions will ordinarily be made in the case of students who wish to live at home." — *Vote of the Faculty of Arts and Sciences, December 2, 1913.*

Students who intend to enter the Freshman Class in 1922 are requested to fill out this application for rooms and return it to the Bursar of Harvard University, The Delta, Kirkland St., Cambridge, Mass. When two students wish to live together, the names of both should appear on one application. Quarters in suites which rent for $100 or less per man can be assigned only to students whose means are strictly limited. The agreement is to be signed in all cases. The numbered blanks are to be filled (1) "me," (2) " I," (3) " my," or (1) " us," (2) " we," (3) " our," according as one or two applicants sign the agreement. The word "quarters " is used to describe the whole of a single suite or a portion of a suite, according to circumstances.

In consideration of the assignment to (1) for the period from September 25, 1922, until June 21, 1923, of quarters in one of the Freshman Halls, (2) hereby agree to take the quarters assigned, for (3) occupation, in accordance with the established rules and the laws and usages of the University, and to pay for the rent and care thereof, for the whole period above-named, the full price of the said quarters as stated in the list of Freshman rooms for 1922–23.

Signature of Student.

Name of Parent or Guardian.

Home Address.

Preparatory School.

Signature of Student.

Name of Parent or Guardian.

Home Address.

Preparatory School.

| Type of suite: | Maximum price for each tenant: |
|---|---|
| Single | $500 to $225 |
| Double | $225 to $75 |
| Triple | $200 to $60 |
| Multiple (4, 5, or 7 men) | $125 to $50 |

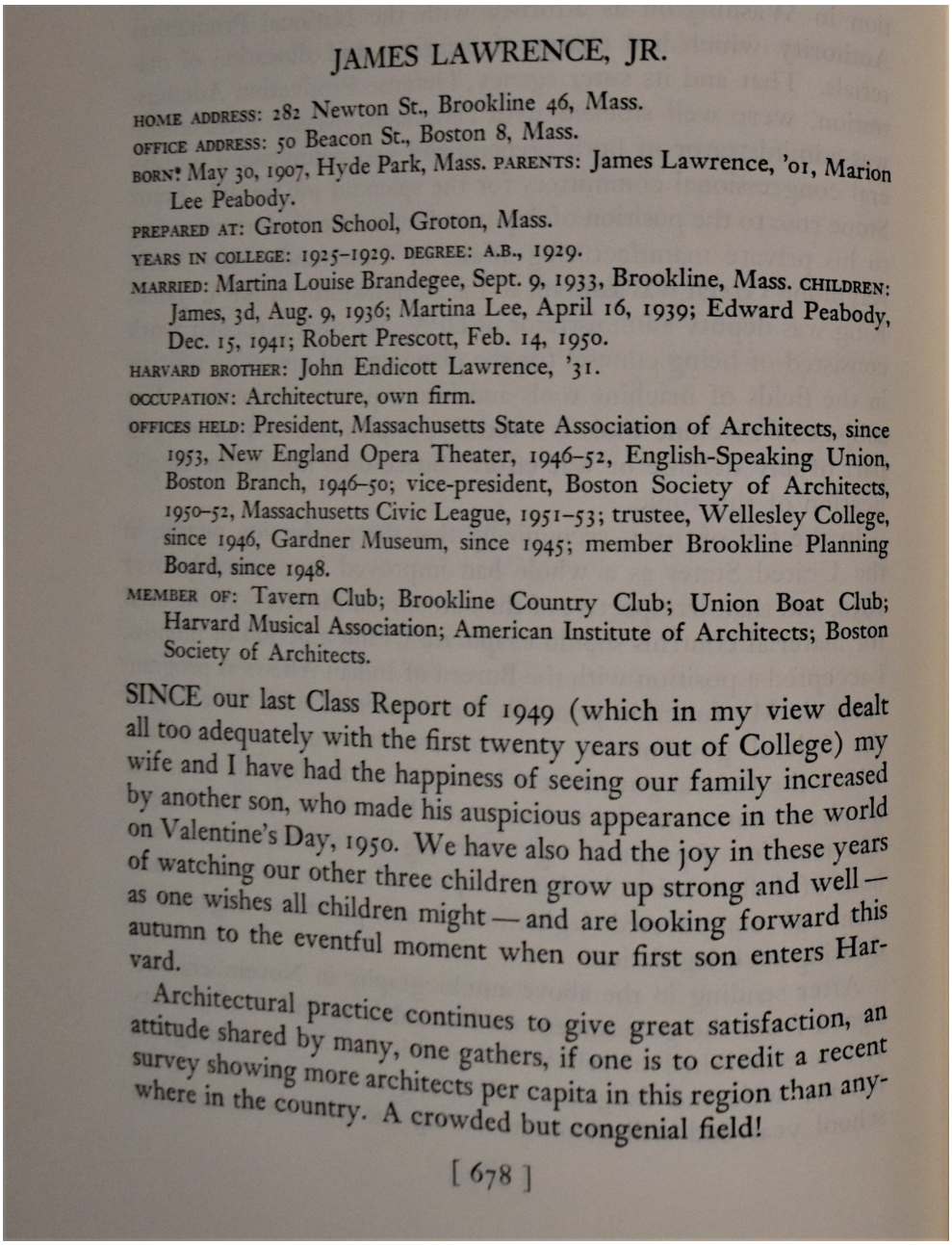
Failing to receive any of the rooms in the Freshman Halls at maximum prices specified above, I (we) hereby apply for a room in either Shepherd or Drayton* Hall, of the type and price indicated below:

| Type of suite: | Maximum price for each tenant: |
|---|---|
| Single | $550 to $200 |
| Double | $400 to $150 |
| Triple | $300 to $200 |

* Drayton Hall, being privately owned, the rent is payable to the owners, at above rates in equal instalments, October 1stand February 15th.

See Sections B.1.5 and B.1.6 for more details on freshman housing and the housing application process.

Figure B.11: 25th Reunion Class Report



## JAMES LAWRENCE, JR.

HOME ADDRESS: 282 Newton St., Brookline 46, Mass.
OFFICE ADDRESS: 50 Beacon St., Boston 8, Mass.
BORN: May 30, 1907, Hyde Park, Mass. PARENTS: James Lawrence, '01, Marion Lee Peabody.
PREPARED AT: Groton School, Groton, Mass.
YEARS IN COLLEGE: 1925–1929. DEGREE: A.B., 1929.
MARRIED: Martina Louise Brandegee, Sept. 9, 1933, Brookline, Mass. CHILDREN: James, 3d, Aug. 9, 1936; Martina Lee, April 16, 1939; Edward Peabody, Dec. 15, 1941; Robert Prescott, Feb. 14, 1950.
HARVARD BROTHER: John Endicott Lawrence, '31.
OCCUPATION: Architecture, own firm.
OFFICES HELD: President, Massachusetts State Association of Architects, since 1953, New England Opera Theater, 1946–52, English-Speaking Union, Boston Branch, 1946–50; vice-president, Boston Society of Architects, 1950–52, Massachusetts Civic League, 1951–53; trustee, Wellesley College, since 1946, Gardner Museum, since 1945; member Brookline Planning Board, since 1948.
MEMBER OF: Tavern Club; Brookline Country Club; Union Boat Club; Harvard Musical Association; American Institute of Architects; Boston Society of Architects.

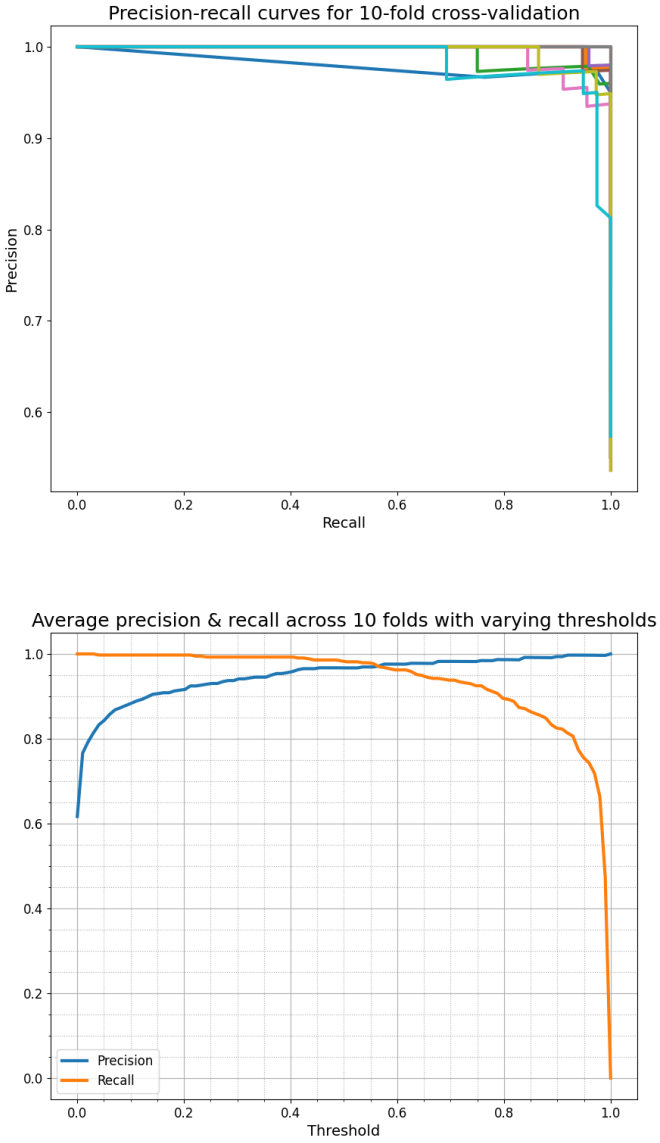
SINCE our last Class Report of 1949 (which in my view dealt all too adequately with the first twenty years out of College) my wife and I have had the happiness of seeing our family increased by another son, who made his auspicious appearance in the world on Valentine's Day, 1950. We have also had the joy in these years of watching our other three children grow up strong and well — as one wishes all children might — and are looking forward this autumn to the eventful moment when our first son enters Harvard.

Architectural practice continues to give great satisfaction, an attitude shared by many, one gathers, if one is to credit a recent survey showing more architects per capita in this region than anywhere in the country. A crowded but congenial field!

[ 678 ]

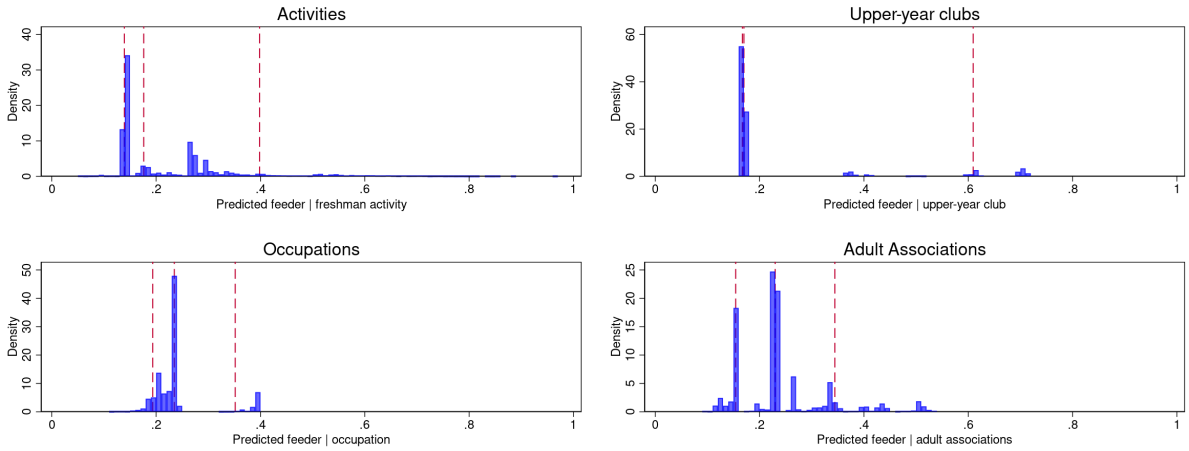
Classes 1923-1950 digitized. Includes birthplace through graduating class of 1937 (corresponding with Freshmen entering in fall 1933). See Section B.1.8 for more details on the Class Reports.

Figure B.12: Cross-validation for linking Red Books to Class Reports



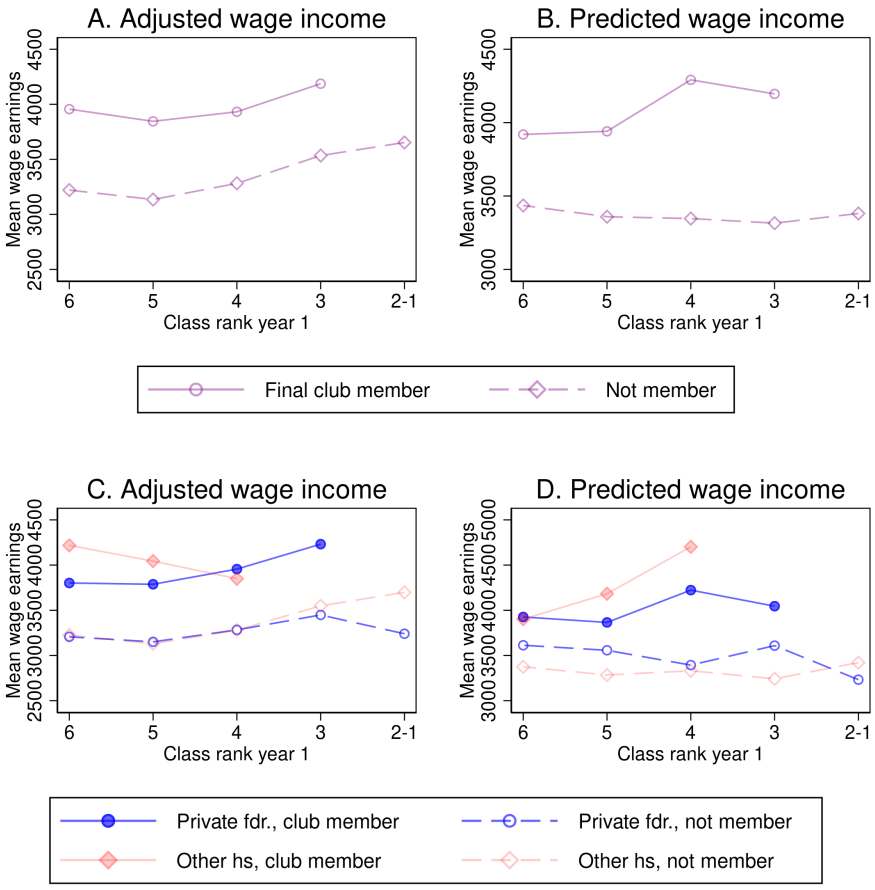

Precision and recall calculated for a 10-fold cross validation exercise on the manually labeled traing data set of 690 observations. The top figure presents the precision/recall curve for each of 10 folds. The bottom figure presents a precision curve and a recall curve by our choice of threshold (minimum similarity score at which we accept a candidate match), averaging over the results from the 10 folds. We use matches above a threshold of 0.4 in our analysis. For more details on the matching of Red Book and Class Report records, see description in Section B.2.1.

Figure B.13: Occupation private school indices



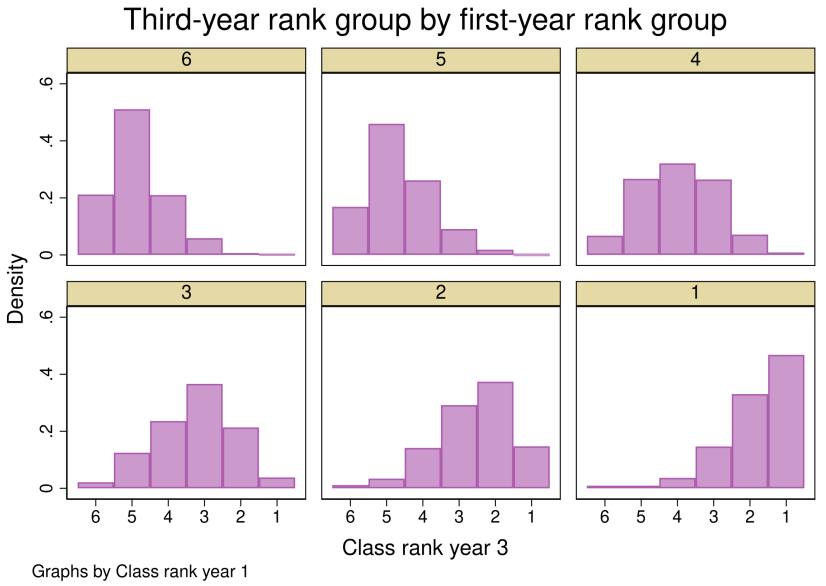

Histograms of activity private high school index (left panel) and occupation private high school index (right panel). Indices estimated using Lasso specifications as described in section 3.4.3. Predictions for each entering cohort are based on students in other entering cohorts. See Table B.8 for estimated coefficients from Lasso procedure in full sample. These distributions are of predicted values, prior to standardization.

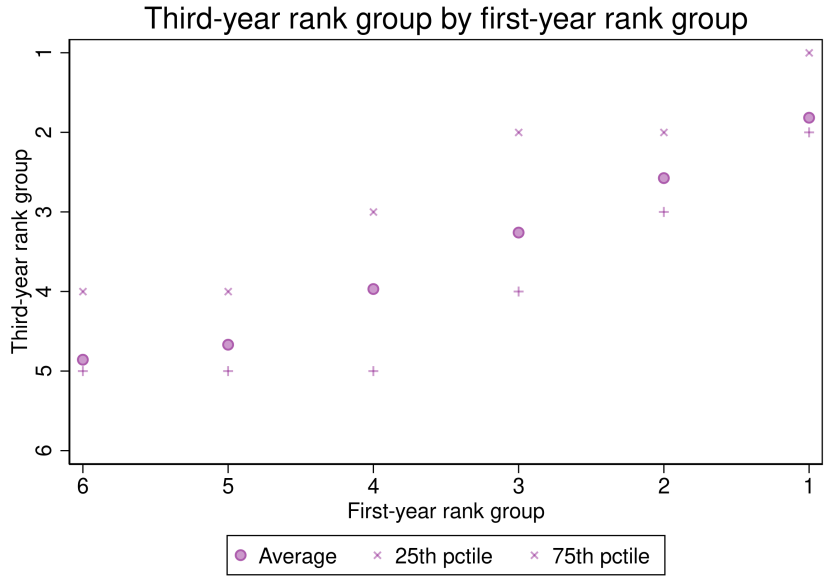
Figure B.14: Alternative earnings by academic performance and final club membership




Panels A-B depict earnings by freshman academic rank group and selective final club membership. Panels C-D 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 C (and owned own home for Panels B and D).

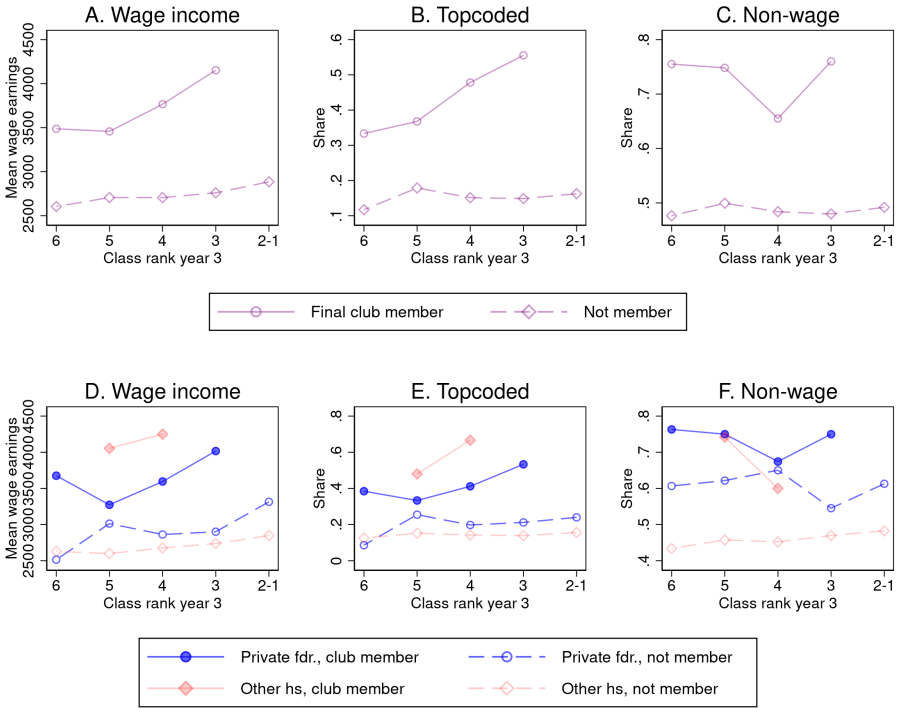
Figure B.15: Academic rank in years 1 and 3



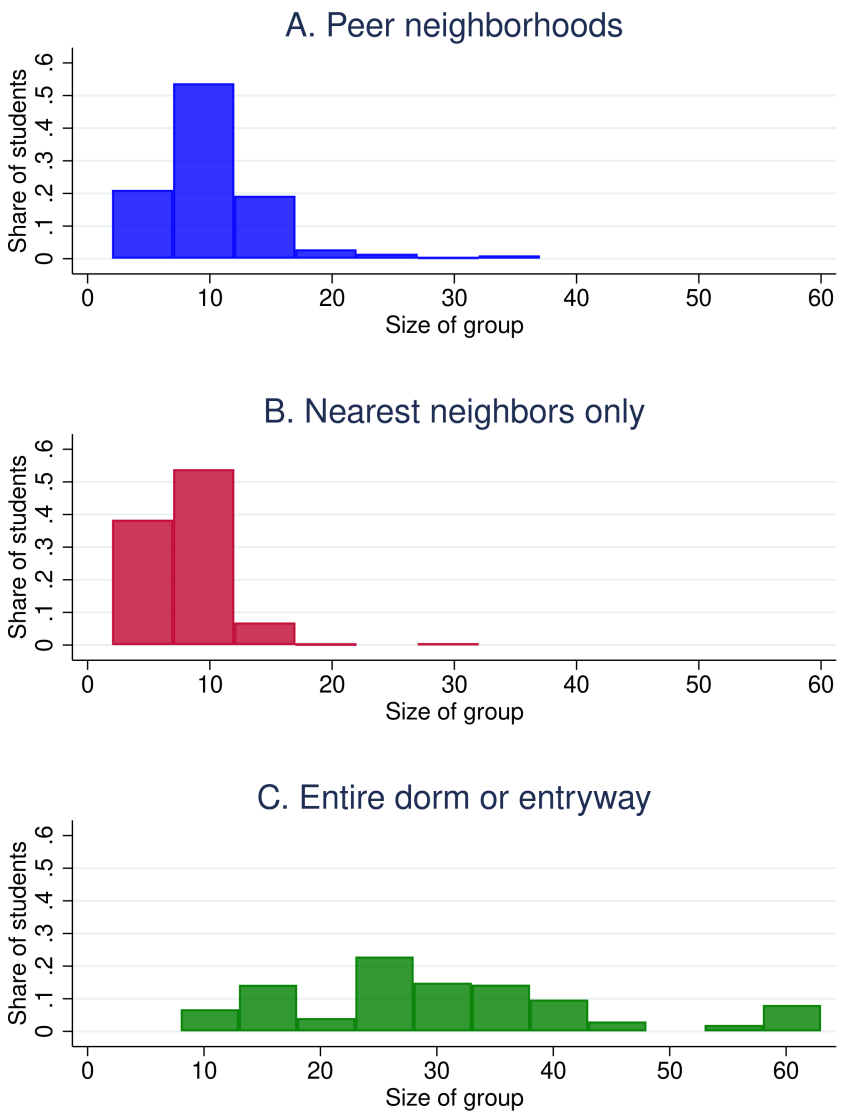

Figures present the within-student relationship between first- and third- year rank groups among students for whom both measures are available. See Section B.1.4 for more details on the rank group lists.

Figure B.16: Earnings by third-year academic performance and final club membership



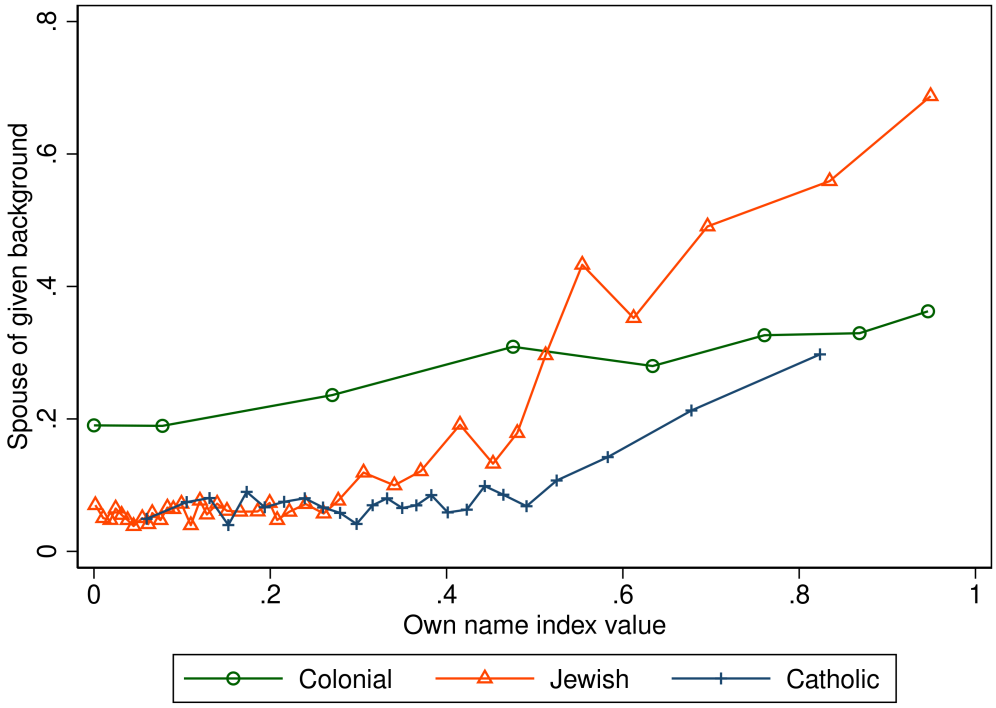

Panels A-C depict earnings by third year academic rank group and selective final club membership. Panels D-F present the same but also divide students by high school type. Rank groups 1 and 2 are collapsed and groups with fewer than 20 students not displayed. Includes students from cohorts 1920-1930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

Figure B.17: Histograms of group size for alternate spatial groupings



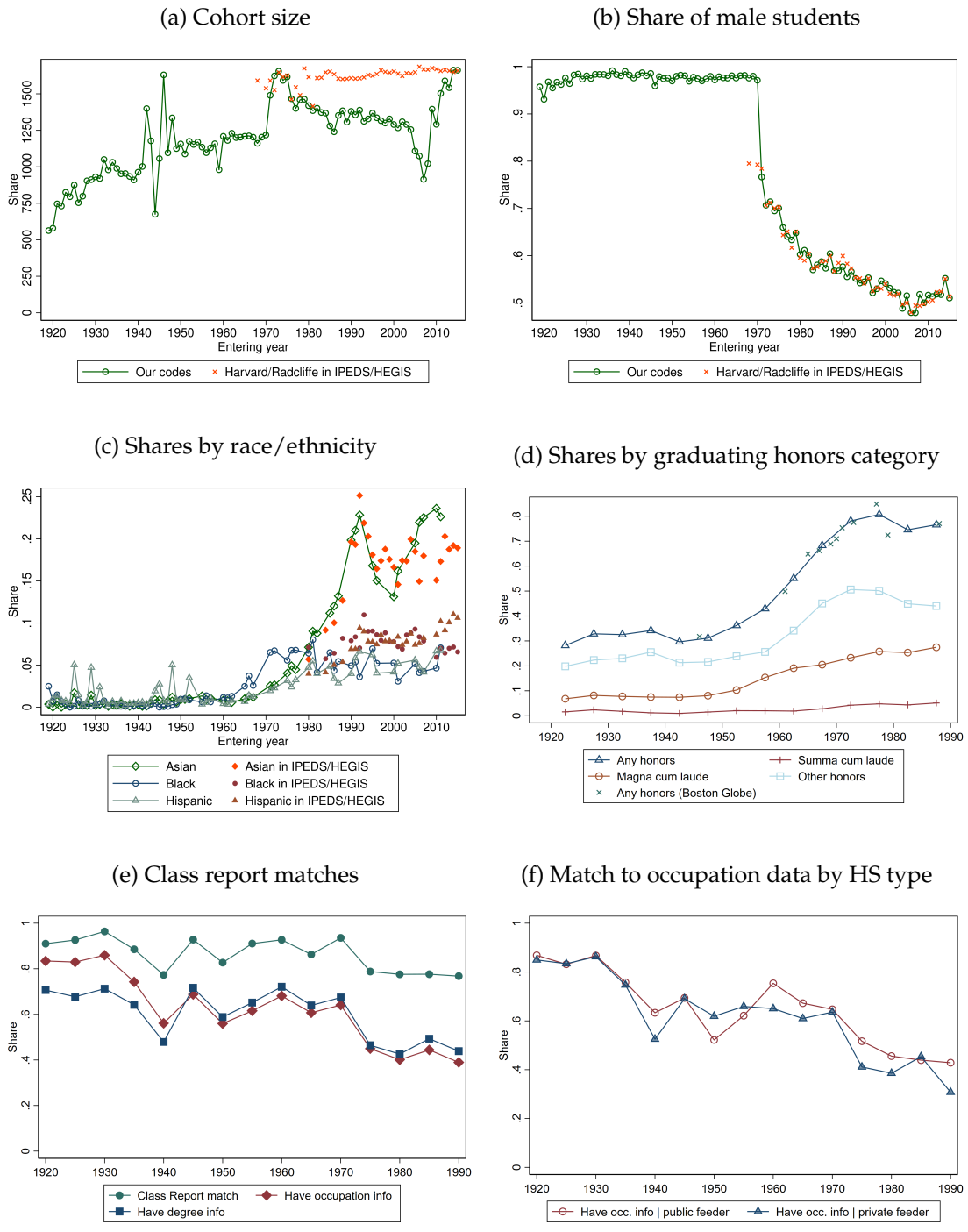

Histograms show the number of students who have different sized peer neighborhoods (Panel A; our main definition), nearest neighbor groups (Panel B), and larger groups of entire dorms (for small dorms) or entire entryway (Panel C). See B.7 for details.

Figure B.18: Own vs spouse name index



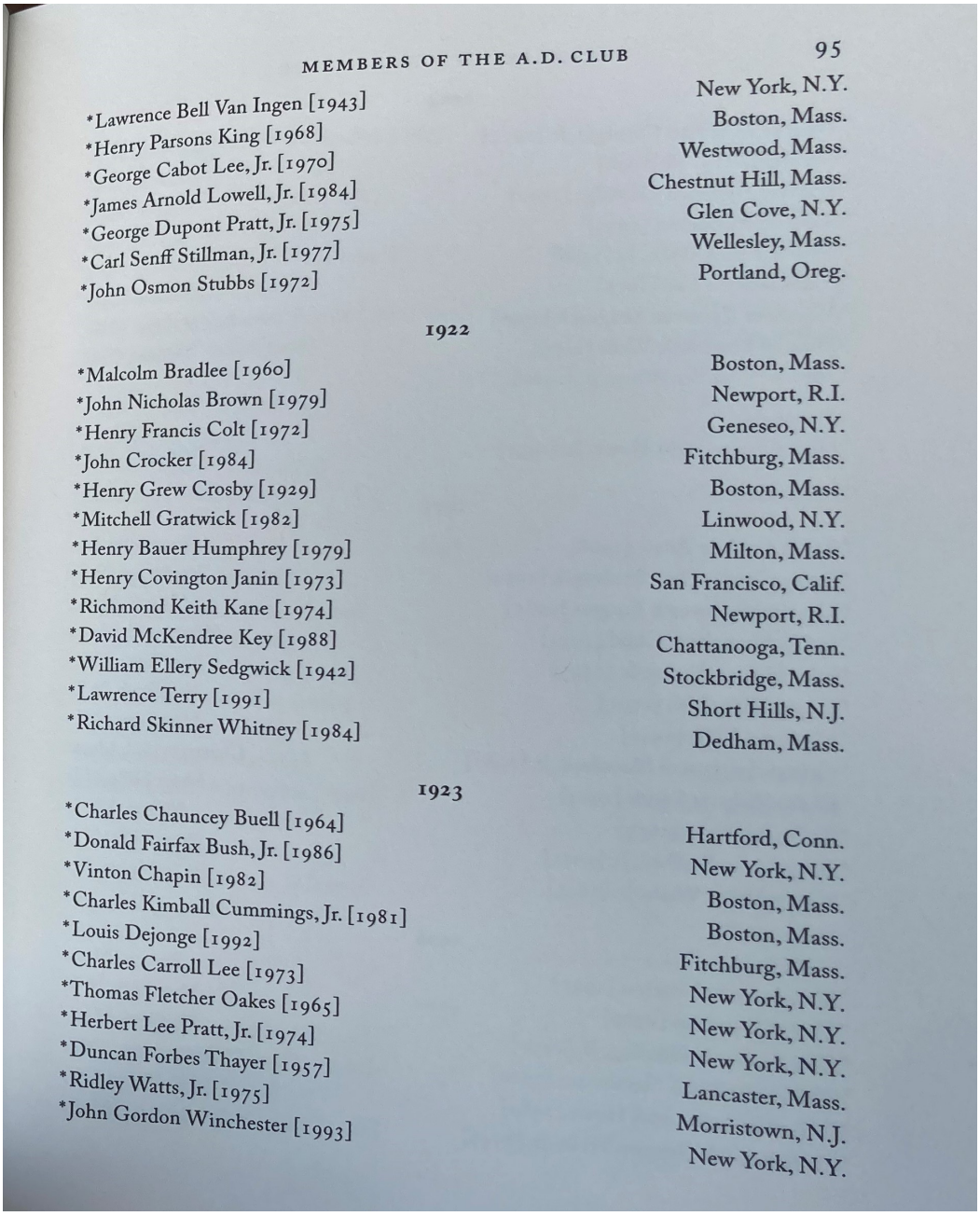

Mean value of spouse name index by quantile bins of own name index. Spouse index type is the same as own index type; e.g. the Colonial series displays the mean value of the spouse Colonial index for each bin of own Colonial index. Sample: married Harvard students. Bin placements and counts determined using Cattaneo et al. (2019). See section B.8.4 for details.

Figure B.19: Long run data construction




Panel A: Red Book cohort size. Panel B: share of male students in Red Books. Panel C: shares by race/ethnicity in Red Books. Panel D shares of students graduating with different honors designations. The Boston Globe series is from Healy (2001). Points are means of data within 2.5 years on either side of the centered value. Panel E: match rates from Red Book data to Class Report records, and to Class Report records that include either occupation data or educational history data. Panel F: Merge rates to occupation data by high school type. For Panels E and F, 1940 and earlier data are averages of centered 5-year bins. 1945 and later data are single-year means.

Figure B.20: A.D. club catalog



MEMBERS OF THE A.D. CLUB 95

*Lawrence Bell Van Ingen [1943] New York, N.Y.
*Henry Parsons King [1968] Boston, Mass.
*George Cabot Lee, Jr. [1970] Westwood, Mass.
*James Arnold Lowell, Jr. [1984] Chestnut Hill, Mass.
*George Dupont Pratt, Jr. [1975] Glen Cove, N.Y.
*Carl Senff Stillman, Jr. [1977] Wellesley, Mass.
*John Osmon Stubbs [1972] Portland, Oreg.

**1922**

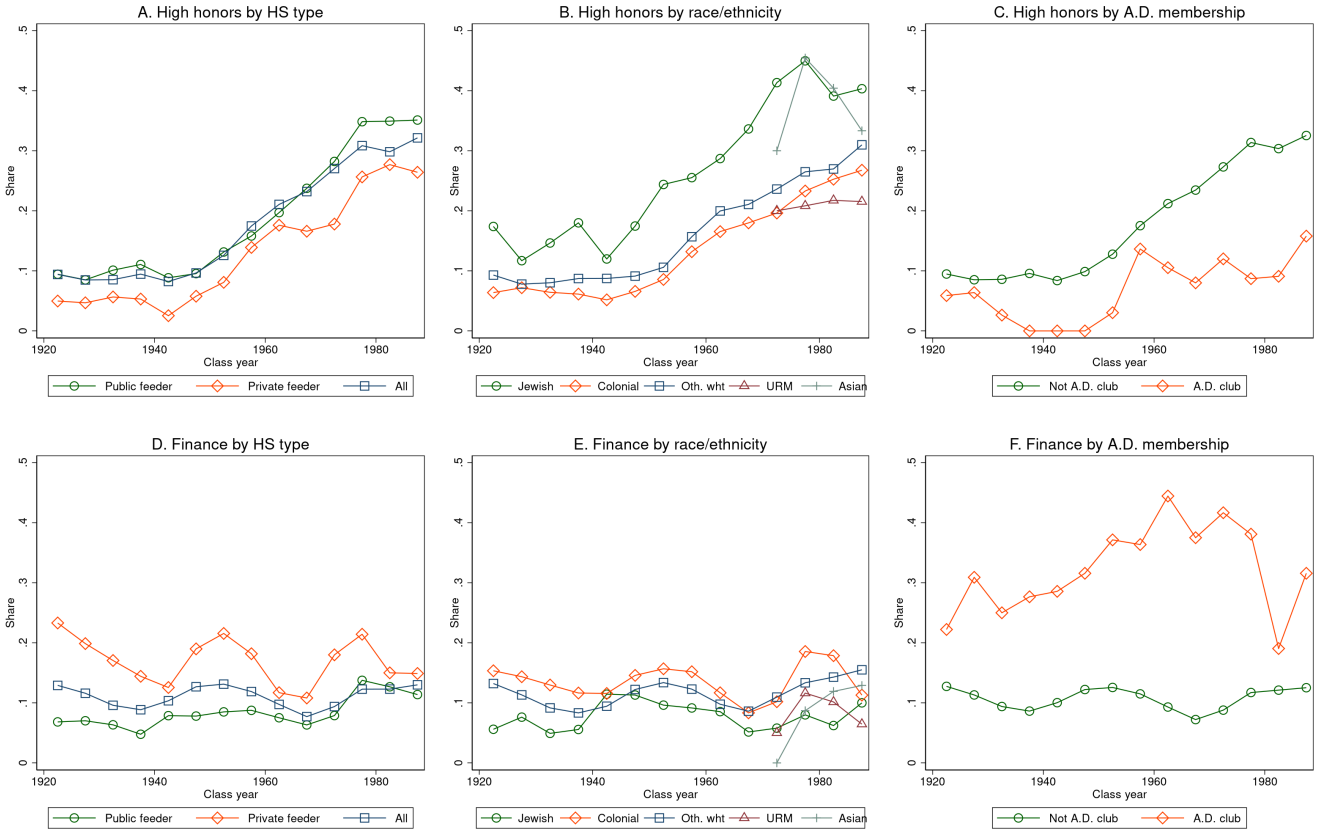
*Malcolm Bradlee [1960] Boston, Mass.
*John Nicholas Brown [1979] Newport, R.I.
*Henry Francis Colt [1972] Geneseo, N.Y.
*John Crocker [1984] Fitchburg, Mass.
*Henry Grew Crosby [1929] Boston, Mass.
*Mitchell Gratwick [1982] Linwood, N.Y.
*Henry Bauer Humphrey [1979] Milton, Mass.
*Henry Covington Janin [1973] San Francisco, Calif.
*Richmond Keith Kane [1974] Newport, R.I.
*David McKendree Key [1988] Chattanooga, Tenn.
*William Ellery Sedgwick [1942] Stockbridge, Mass.
*Lawrence Terry [1991] Short Hills, N.J.
*Richard Skinner Whitney [1984] Dedham, Mass.

**1923**

*Charles Chauncey Buell [1964] Hartford, Conn.
*Donald Fairfax Bush, Jr. [1986] New York, N.Y.
*Vinton Chapin [1982] Boston, Mass.
*Charles Kimball Cummings, Jr. [1981] Boston, Mass.
*Louis Dejonge [1992] Fitchburg, Mass.
*Charles Carroll Lee [1973] New York, N.Y.
*Thomas Fletcher Oakes [1965] New York, N.Y.
*Herbert Lee Pratt, Jr. [1974] New York, N.Y.
*Duncan Forbes Thayer [1957] Lancaster, Mass.
*Ridley Watts, Jr. [1975] Morristown, N.J.
*John Gordon Winchester [1993] New York, N.Y.

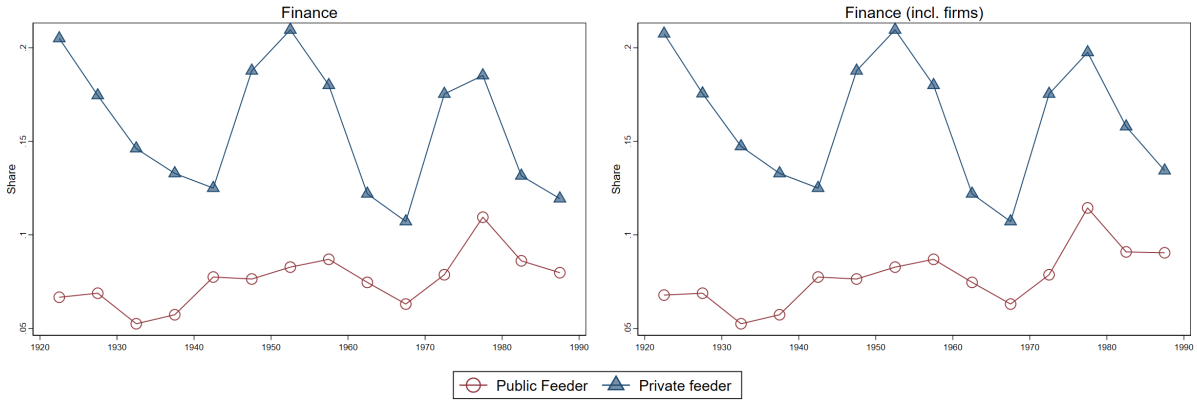
Example page from A.D. Club (2015). Bold numbers are class years. Years in brackets are years of death.

Figure B.21: Grades and career outcomes over time– men only



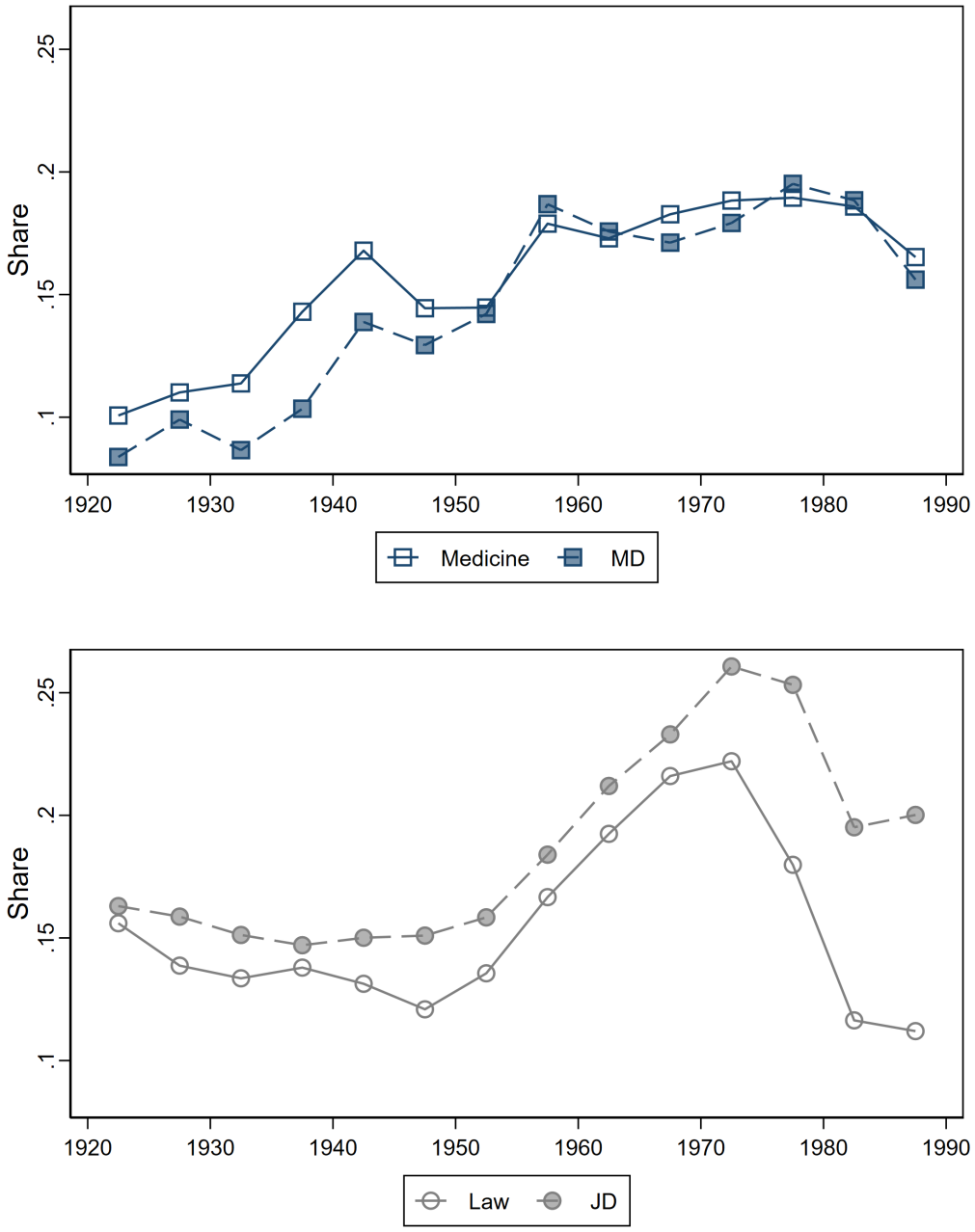

Share of students in different groups graduating with high honors (Panels A-C) or working in finance 25 years after graduation (D-F). Sample is male Harvard students. Horizontal axis in all panels is graduating class year. "High honors" is defined as magna or summa cum laude. Academic and career outcomes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

Figure B.22: Main vs. alternative finance occupation measures



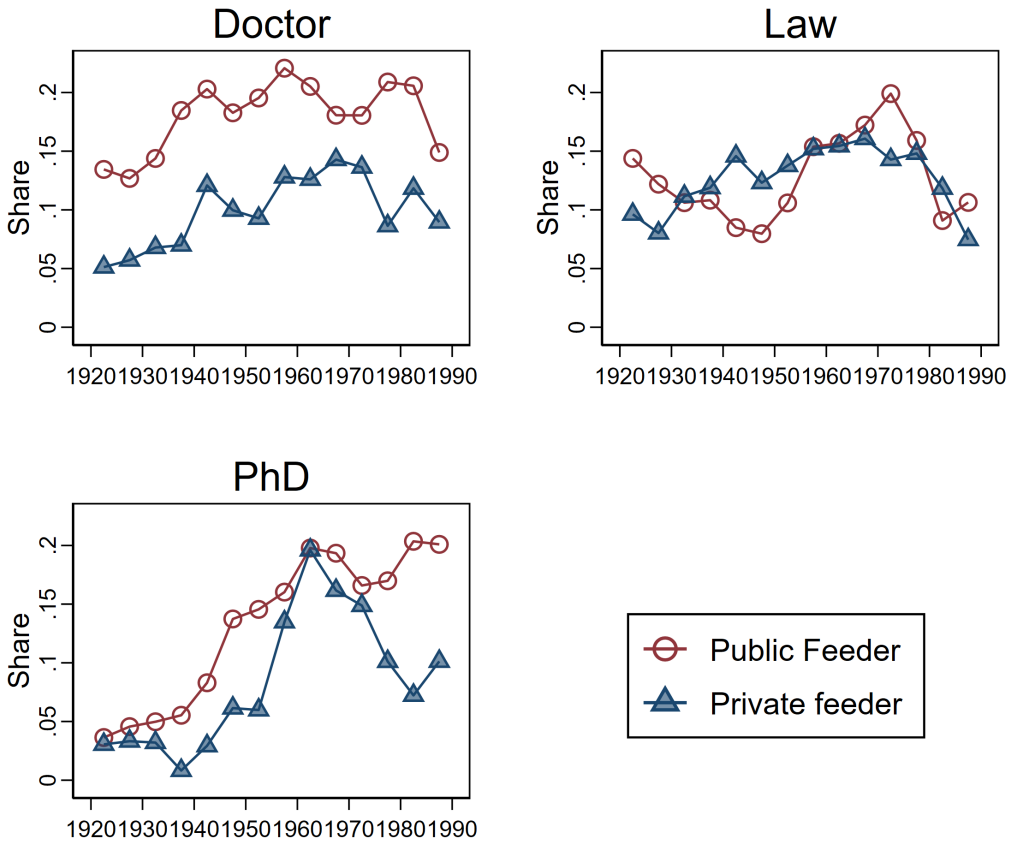

Share of students in finance careers by high school type, using main and supplemental definition. Supplemental definition uses names of major firms. Horizontal axis is graduating class year.. Sample is male students. Both panels use data from Class Reports that is available for each graduating class from 1924 through 1939 and then at five year intervals starting in 1940. Points display means over all years within 2.5 years on either side of the centered value. For example, the 1982.5 datapoint is an average of 1980 and 1985 class years, and the 1987.5 datapoint is an average of 1985 and 1990 class years.

Figure B.23: Trends in career classifications as measured by degree rececipt and job description




Share of students with listed degree/career outcome. Horizontal axis is graduating class year. Sample is students (male and female) with both degree and career information on Class Reports. Both panels use data from Class Reports that is available for each graduating class from 1924 through 1939 and then at five year intervals starting in 1940. Points display means over all years within 2.5 years on either side of the centered value. For example, the 1982.5 datapoint is an average of 1980 and 1985 class years, and the 1987.5 datapoint is an average of 1985 and 1990 class years.

Figure B.24: Long-run trends in alternate career outcomes by high school type




Share of students with listed degree/career outcome by high school type. Horizontal axis is graduating class year. Sample is male students. "Doctor" and "Law" measures are based on job text. "PhD" is students receiving PhD degree. Both panels use data from Class Reports that is available for each graduating class from 1924 through 1939 and then at five year intervals starting in 1940. Points display means over all years within 2.5 years on either side of the centered value. For example, the 1982.5 datapoint is an average of 1980 and 1985 class years, and the 1987.5 datapoint is an average of 1985 and 1990 class years.

## B.12 Tables

Table B.1: Codes for common high schools

| School Name | Private Feeder | Private | Public feeder |
|---|---|---|---|
| Albany Academy | 0 | 1 | 0 |
| Andover | 1 | 1 | 0 |
| Arlington High | 0 | 0 | 1 |
| Avon Old Farms | 0 | 1 | 0 |
| B. M. C. Durfee High | 0 | 0 | 1 |
| Belmont High | 0 | 0 | 1 |
| Belmont Hill | 0 | 1 | 0 |
| Berkshire School | 0 | 1 | 0 |
| Beverly High | 0 | 0 | 1 |
| Blake | 0 | 1 | 0 |
| Boston College High | 0 | 0 | 1 |
| Boston English High | 0 | 0 | 1 |
| Boston Latin | 0 | 0 | 1 |
| Brockton High | 0 | 0 | 1 |
| Bronx High School of Science | 0 | 0 | 1 |
| Bronxville High | 0 | 0 | 1 |
| Brookline High | 0 | 0 | 1 |
| Brooks | 0 | 1 | 0 |
| Browne and Nichols | 0 | 1 | 0 |
| Cambridge High and Latin | 0 | 0 | 1 |
| Chelsea High | 0 | 0 | 1 |
| Chicago Latin School | 0 | 1 | 0 |
| Choate | 0 | 1 | 0 |
| Culver Military Academy | 0 | 1 | 0 |
| Deerfield Academy | 0 | 1 | 0 |
| Dorchester High | 0 | 0 | 1 |
| East Boston High | 0 | 0 | 1 |
| Episcopal Academy | 0 | 1 | 0 |
| Erasmus Hall High | 0 | 0 | 1 |
| Evanston Township High | 0 | 0 | 1 |
| Everett High | 0 | 0 | 1 |
| Exeter | 1 | 1 | 0 |
| Fieldston/Ethical Culture | 0 | 1 | 0 |
| Fountain Valley | 0 | 1 | 0 |
| Governor Dummer Academy | 0 | 1 | 0 |
| Groton | 1 | 1 | 0 |
| Gunnery | 0 | 1 | 0 |
| Hackley | 0 | 1 | 0 |
| Haverford | 0 | 1 | 0 |
| Haverhill High | 0 | 0 | 1 |
| Hebron Academy | 0 | 1 | 0 |
| Hill School | 0 | 1 | 0 |
| Horace Mann | 0 | 1 | 0 |
| Hotchkiss | 0 | 1 | 0 |
| Huntington | 0 | 0 | 1 |
| James Madison High | 0 | 0 | 1 |

Continued on next page

**Table B.1 – continued from previous page**

| School Name | Private Feeder | Private | Public feeder |
|---|---|---|---|
| John Burroughs | 0 | 1 | 0 |
| Kent | 0 | 1 | 0 |
| Lawrence Academy | 0 | 1 | 0 |
| Lawrence High | 0 | 0 | 1 |
| Lawrenceville School | 0 | 1 | 0 |
| Lenox | 0 | 1 | 0 |
| Lexington High School | 0 | 0 | 1 |
| Loomis | 0 | 1 | 0 |
| Lowell High | 0 | 0 | 1 |
| Lynn Classical | 0 | 0 | 1 |
| Malden High | 0 | 0 | 1 |
| Medford High | 0 | 0 | 1 |
| Melrose High | 0 | 0 | 1 |
| Mercersburg Academy | 0 | 1 | 0 |
| Middlesex | 1 | 1 | 0 |
| Milton | 1 | 1 | 0 |
| Milwaukee Country Day | 0 | 1 | 0 |
| Montclair High | 0 | 0 | 1 |
| Morristown School | 0 | 1 | 0 |
| Moses Brown | 0 | 1 | 0 |
| Mount Hermon | 0 | 1 | 0 |
| New Bedford High | 0 | 0 | 1 |
| New Preparatory School | 0 | 1 | 0 |
| New Trier Township High | 0 | 0 | 1 |
| Newton Country Day School | 0 | 1 | 0 |
| Newton High | 0 | 0 | 1 |
| Noble and Greenough | 0 | 1 | 0 |
| North Quincy High | 0 | 0 | 1 |
| North Shore Country Day | 0 | 1 | 0 |
| Polytechnic Preparatory Country Day | 0 | 1 | 0 |
| Pomfret School | 0 | 1 | 0 |
| Portsmouth Priory | 0 | 1 | 0 |
| Quincy High | 0 | 0 | 1 |
| Reading High | 0 | 0 | 1 |
| Rindge Technical School | 0 | 0 | 1 |
| Riverdale Country School | 0 | 1 | 0 |
| Rivers School | 0 | 1 | 0 |
| Roosevelt High | 0 | 0 | 1 |
| Roxbury Latin | 0 | 1 | 0 |
| Salem High | 0 | 0 | 1 |
| Santa Barbara | 0 | 1 | 0 |
| Scarsdale High | 0 | 0 | 1 |
| Shady Side Academy | 0 | 1 | 0 |
| Somerville High | 0 | 0 | 1 |
| St. George's | 1 | 1 | 0 |
| St. Louis Country Day | 0 | 1 | 0 |
| St. Mark's | 1 | 1 | 0 |
| St. Paul Academy | 0 | 1 | 0 |
| St. Paul's, Concord, N.H. | 1 | 1 | 0 |

Continued on next page

**Table B.1 – continued from previous page**

| School Name | Private Feeder | Private | Public feeder |
|---|---|---|---|
| St. Paul's, Garden City, N. Y. | 0 | 1 | 0 |
| Stone School | 0 | 1 | 0 |
| Stuyvesant High | 0 | 0 | 1 |
| Tabor Academy | 0 | 1 | 0 |
| Taft | 0 | 1 | 0 |
| Thayer | 0 | 1 | 0 |
| Tome | 0 | 1 | 0 |
| University School | 0 | 1 | 0 |
| Volkmann | 0 | 1 | 0 |
| Walnut Hills High | 0 | 0 | 1 |
| Watertown High | 0 | 0 | 1 |
| Wellesley High School | 0 | 0 | 1 |
| Western Reserve Academy | 0 | 1 | 0 |
| Westminster School | 0 | 1 | 0 |
| William Penn Charter | 0 | 1 | 0 |
| Williston Academy | 0 | 1 | 0 |
| Winchester High | 0 | 0 | 1 |
| Winthrop High | 0 | 0 | 1 |
| Woodrow Wilson High | 0 | 0 | 1 |
| Worcester Academy | 0 | 1 | 0 |

Table B.2: Codes for activities listed in red books

| Category | Subcategory | Identifiers |
|---|---|---|
| Dormitory Committees | | Smith Hall Dormitory Committee, Standish Dormitory Committee, Gore Hall Dormitory Committee, Dormitory Committee, McKinlock Hall Dormitory Committee, Dudley House Committee |
| Sports Teams | University Teams | baseball squad, soccer squad, track team, lacrosse squad, football squad, basketball squad, Interclass Basketball Team, cross country team, Golf Team, Hockey, Swimming Team, Wrestling Team, Tennis Team, Relay Team, Squash Raquets and Squash Tennis Champion, Freshmen Squash Tournament, Interclass Squash Team, Gym Team, University Fencing Squad, Freshman Golf Tournament, Polo Team, Freshman cheerleader, Handball, Fall Wherry Champtionship, rowing, Boxing, Freshman Football, Ski Team, Rugby, Association Football Team, 2d Football Team, Interclass Football, Sailing |
| | Dormitory Teams | Dormitory Basketball, Dormitory Football, Interdormitory Hockey, Dormitory Crew, Dormitory Hockey, Dormitory Squash Team, Dormitory Tennis, Dormitory Cross Country, Interhall Basketball, Interhall/Interdormitory athletics |
| | Intramural Teams | Intramural Football, Intramural Basketball, Intramural Boxing, Intramural Crew, Intramural Squash, Intramural Swimming |
| Musical Groups | University | University Orchestra, University Band, University Musical Clubs, University Instrumental Clubs, The College Choir, University Jazz Band, Appleton Chapel Choir, Gold Coast Orchestra, University Glee Club, Banjo Club, Mandolin Club, Pierian Sodality, Regimental Band, Concert Audience |
| | Freshman | freshman banjo club, freshman mandolin club, freshman glee club, Freshman Orchestra, Freshman Musical Clubs, Freshman Instrumental Clubs, Freshman Vocal Club |
| Red Book | | Red Book |
| Foreign Clubs | | International Council, Foreign Student Committee, |
| | Language Clubs | Cercle Francais, Circulo Espafiol, Circolo Italiano, Deutscher Verein, Russian Club |
| Outdoor Clubs | | Gun Club, Harvard Flying Club, Harvard Mountaineering Club |
| Politics And Debate Clubs | Debate | Debating Team, Freshman Debating Council, Union Debating Society |
| | Politics | Harvard Chapter of League for industrial Democracy, Harvard Chapter of the Fellowship of Youth For Peace, Harvard Socialist Club, John Marshall Law club, Democratic Club, Liberal Club, Harvard Peace Society |
| | Other | Speakers committee, Discussion Club, Freshman Discussion Club |
| | Politics | Harvard Young Republican Club, Conservative League<br>World Federalists |
| | Politics | Free Enterprise Society |
| Stem Clubs | | Harvard Engineering Society, Boylston Chemical Club, Mathematics Club |
| Drama Clubs | | Freshman Players, Dramatic Club, 47 Workshop, Freshman Vaudeville, Stage crew, Theatre Workshop |

Continued on next page

**Table B.2 – continued from previous page**

| Category | Subcategory | Identifiers |
|---|---|---|
| Newspaper And Literary Clubs | | harvard crimson, The Advocate, Harvard Magazine, Literary Society, Journal, Harvard Lampoon, Harvard Monthly, Harvard Guardian, Service News, Yardling |
| Social Committees | | finance committee, jubilee committee, Dinner Committee, entertainment committee, Smoker Committee, Regatta Committee, Visiting Teams Committee, Freshman library committee, Delegations Committee, Fall Tea Dance Committee, Confidential Guide Committee, Reviews Committee, Union Committee, Union Dance Committee, Winter Informal, Freshman Affairs Committee, Christmas Dance Committee, Dance Committee |
| Jewish Clubs | | Harvard Menorah, Avukah Club |
| Other Organizations | | Executive Board, Student Government, Chess Club, Endowment Fund, Hoover Fund, Social, Social Service, Register Board, Wireless club, Lowell House, H. U. B. Club, Phillips Brooks House Association, John Barnard Associates, Harvard Thomas-For-President Club, Harvard Memorial Society, Student Union, Stamp Club, Price Greenleaf Aid, Yacht Club, Photography Club, Outing Club, Aeronautical Society, Ornithological Club, WHRV - Harvard Radio Voice, Cosmopolitan Club, Wake, Undergraduate Faculty, Prospect Union, Railroading Association, American Civilization Group, Pre-Medical Society, Bridge Club, Film Society, [Major/Minor] Numerals, Social Relations Society |
| | High School Alumni Club | St. Paul's Club, Andover club, Harvard Choate Club, Harvard Classical Club |
| | Geographical | Western Club, Southern Club, Canadian Club |
| | Christian Groups | Christian Club, Catholic Club, Brotherhood of St. Andrew, Usher in Memorial Church, Usher of University Church |
| | War Effort | Student Defence League, Air Raid Precautions, Postwar Council |
| Military | Not deployed | R. O. T. C, U. S. Naval Reserve, Caisson Club, Harvard Naval Unit, Regiment |
| | Service | Infantry, Ambulance, Air Service, State Militia, Navy, Coast Artillery Corps, Royal Air Force, Marine Section, Artillery, British Ambulance, Machine Gun Battalion, Canadian Field Artillery, Signal Corps, U. S. Army, U.S. Coast Guard, other |

Table B.3: Codes for occupations listed in class reports

| Category | Subcategory | Identifiers |
|---|---|---|
| Senior Management | | director, president, executive, chairman |
| Low Management | | manager, supervisor, administration |
| Finance | investment | securities, financial, investments, finance, trust, brokerage, broker, stock, investment, investors, trustee, stockbroker |
| | banking | banking, banker, bankers, bank |
| Accounting/Real Estate | treasury | treasurer, treasury |
| | accountancy | accounting, accountant, accountancy, bookkeeper, auditor, tax |
| | real estate | real, estate |
| | insurance | insurance, actuary |
| Lawyer | law | lawyer, attorney, law, judge, legal, counsel |
| Retail | business | business, owner |
| | sales | sales, salesman, selling, retail, merchandise, mercantile, merchant, store, stores, storekeeper, bookseller, retailer |
| | advertising | advertising, marketing |
| | hr | relations, personnel |
| | secretarial | secretary |
| | consulting | consultant, consulting |
| Doctor | doctor | physician, surgeon, doctor, medicine, anaesthesiologist, ophthalmologist, hospital, obstetrics, obstetrician, gynecologist, urologist, surgery, orthopedic, neurosurgeon, pediatrician |
| | psychiatry | psychiatrist, neuropsychiatrist, psychiatry, psychology, psychologist, psychoanalyst |
| | dentistry | dentist, orthodontist, dentistry |
| | pharmacy | druggist, pharmacist |
| Engineer | engineering | engineering, engineer, inventor |
| | design | designer, draftsman |
| Higher Education | professor | professor, university, college |
| Scientist | research | research, laboratory |
| | chemistry | chemist, chemistry, chemical |
| | other science | physical, botanist, biology, mineralogist, science, geologist, zoologist, metallurgist, physicist |
| | anthropology/history | anthropologist, historian |
| | economics | economics, economist, economic |
| | mathematics | mathematics, statistician |
| | library | library, librarian |
| Teacher | education | teacher, principal, school, admissions, lecturer, education, educational, teaching, tutor, educator, instructor, schoolmaster |
| Art/Publishing | publishing | publishing, copy, book, editor, proofreading, publisher |
| | writing | author, writer, literature, novelist |
| | journalism | journalist, newspaperman, newspaper, journalism, reporter, news |
| | art | painter, artist, arts, art, painting, curator |
| | music | composer, conductor, music, organist, musician |
| | entertainment | actor, theatre, theatres, theatrical, dramatist, entertainer |
| | photography | photography, photographer |
| Agriculture | agriculture | farmer, agriculture, ranching, farm, grower, dairy, farming, rancher |
| Government | government | government, state, bureau |

Continued on next page

**Table B.3 – continued from previous page**

| Category | Subcategory | Identifiers |
|---|---|---|
| | politics | political, senator, politics |
| Manufacturing | manufacturing | manufacturers, manufacturer, manufacturing, factory |
| Mining | mining | mining, coal, mine |
| | oil | petroleum, oil, gas |
| Construction/Architecture | construction | construction |
| | architecture | architect, architecture, architects |
| Religious | religious | rabbi, clergyman, priesthood, church, priest, ministry, bishop, episcopal, vicar, minister |
| Military | military | army, naval, navy, commander, military |
| Aviation | aviation | aviation, air, aircraft, airplane |
| Retired | retired | retired |

Table B.4: Codes for common adult clubs and associations

| Name | Social, any | Country Club | Gent. Club | Frat. Order | Honorary Politic. | Professional |
|---|---|---|---|---|---|---|
| Adventurers Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Agawam Hunt Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Algonquin club | 1 | 0 | 1 | 0 | 0 | 0 |
| Alpha Chi Sigma | 0 | 0 | 0 | 0 | 1 | 1 |
| Alpha Omega Alpha | 0 | 0 | 0 | 0 | 1 | 0 |
| American Academy of Arts and Sciences | 0 | 0 | 0 | 0 | 1 | 1 |
| American Academy of Pediatrics | 0 | 0 | 0 | 0 | 0 | 1 |
| American Anthropological Association | 0 | 0 | 0 | 0 | 0 | 1 |
| American Association for the Advancement of Science | 0 | 0 | 0 | 0 | 0 | 1 |
| American Association of University Professors | 0 | 0 | 0 | 0 | 0 | 1 |
| American Bar Association | 0 | 0 | 0 | 0 | 0 | 1 |
| American Board of Internal Medicine | 0 | 0 | 0 | 0 | 0 | 1 |
| American Board of Surgery | 0 | 0 | 0 | 0 | 0 | 1 |
| American Chemical Society | 0 | 0 | 0 | 0 | 0 | 1 |
| American Civil Liberties Union | 0 | 0 | 0 | 0 | 1 | 1 |
| American College of Physicians (fellow) | 0 | 0 | 0 | 0 | 0 | 1 |
| American College of Surgeons (fellow) | 0 | 0 | 0 | 0 | 0 | 1 |
| American Dental Association | 0 | 0 | 0 | 0 | 0 | 1 |
| American Economic Association | 0 | 0 | 0 | 0 | 0 | 1 |
| American Federation for clinical research | 0 | 0 | 0 | 0 | 0 | 1 |
| American Geophysical Union | 0 | 0 | 0 | 0 | 0 | 1 |
| American Heart Association | 0 | 0 | 0 | 0 | 0 | 1 |
| American Historical Association | 0 | 0 | 0 | 0 | 0 | 1 |
| American Institute of Accountant | 0 | 0 | 0 | 0 | 0 | 1 |
| American Institute of Architects | 0 | 0 | 0 | 0 | 0 | 1 |
| American Institute of Certified Public Accountants | 0 | 0 | 0 | 0 | 0 | 1 |
| American Institute of Electrical Engineers | 0 | 0 | 0 | 0 | 0 | 1 |
| American Institute of Mining and Metallurgical Engineers | 0 | 0 | 0 | 0 | 0 | 1 |
| American Jewish Committee | 0 | 0 | 0 | 0 | 1 | 0 |
| American Jewish Congress | 0 | 0 | 0 | 0 | 1 | 0 |
| American Judicature Society | 0 | 0 | 0 | 0 | 0 | 1 |
| American Law Institute | 0 | 0 | 0 | 0 | 0 | 1 |
| American Library Association | 0 | 0 | 0 | 0 | 0 | 1 |
| American Management Association | 0 | 0 | 0 | 0 | 0 | 1 |
| American Marketing Association | 0 | 0 | 0 | 0 | 0 | 1 |
| American Mathematical Society | 0 | 0 | 0 | 0 | 0 | 1 |
| American Medical Association | 0 | 0 | 0 | 0 | 0 | 1 |
| American Medical Society | 0 | 0 | 0 | 0 | 0 | 1 |
| American Meteorological Society | 0 | 0 | 0 | 0 | 0 | 1 |
| American Philosophical Association | 0 | 0 | 0 | 0 | 0 | 1 |
| American Physical Society | 0 | 0 | 0 | 0 | 0 | 1 |
| American Physiological Society | 0 | 0 | 0 | 0 | 0 | 1 |
| American Political Science Association | 0 | 0 | 0 | 0 | 0 | 1 |
| American Psychiatric Association | 0 | 0 | 0 | 0 | 0 | 1 |
| American Psychological Association | 0 | 0 | 0 | 0 | 0 | 1 |
| American Public Health Association | 0 | 0 | 0 | 0 | 0 | 1 |

Continued on next page

**Table B.4 – continued from previous page**

| Name | Social, any | Country Club | Gent. Club | Frat. Order | Honorary Politic. | Professional |
|---|---|---|---|---|---|---|
| American Society for Metals | 0 | 0 | 0 | 0 | 0 | 1 |
| American Society of Civil Engineers | 0 | 0 | 0 | 0 | 0 | 1 |
| American Society of Mechanical Engineers | 0 | 0 | 0 | 0 | 0 | 1 |
| American Statistical Association | 0 | 0 | 0 | 0 | 0 | 1 |
| American Trudeau Society | 0 | 0 | 0 | 0 | 0 | 1 |
| Anglers Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Anteaters Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Argyle Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Arkwright Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Association of Ex-Members of Squadron A | 1 | 0 | 1 | 0 | 0 | 0 |
| Association of Harvard Chemists | 0 | 0 | 0 | 0 | 0 | 1 |
| Association of the Bar of the City of New York | 0 | 0 | 0 | 0 | 0 | 1 |
| Athenaeum Club | 1 | 0 | 1 | 0 | 0 | 0 |
| B'nai B'rith | 1 | 0 | 0 | 1 | 0 | 0 |
| Belmont Country Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Boston Bar Association | 0 | 0 | 0 | 0 | 0 | 1 |
| Boston Club of New Orleand | 1 | 0 | 1 | 0 | 0 | 0 |
| Boston Surgical Society | 0 | 0 | 0 | 0 | 0 | 1 |
| Boston Yacht Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Brae Burn Country Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Broad Street Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Buffalo Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Burlingame Country Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Cape Fear Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Capitol Hill club | 1 | 0 | 1 | 0 | 1 | 0 |
| Cavendish Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Century Association | 1 | 0 | 1 | 0 | 0 | 0 |
| Chevy Chase Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Chicago Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Cincinnati Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Cincinnati Country Club | 1 | 1 | 0 | 0 | 0 | 0 |
| City Club Coporation (The Lunch Club) | 1 | 0 | 1 | 0 | 0 | 0 |
| City Midday Club | 1 | 0 | 1 | 0 | 0 | 0 |
| City Tavern Association | 1 | 0 | 1 | 0 | 0 | 0 |
| Cliff Dwellers Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Cloud Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Clover Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Club of Odd Volumes | 1 | 0 | 1 | 0 | 0 | 0 |
| Coffee House Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Cohasset Golf Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Cohasset Yacht Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Cold Spring Harbor Beach Club | 1 | 1 | 0 | 0 | 0 | 0 |
| College Art Association | 0 | 0 | 0 | 0 | 0 | 1 |
| Colony Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Columbia Country Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Columbus Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Concord Country Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Cosmos Club | 1 | 0 | 1 | 0 | 0 | 0 |

Continued on next page

**Table B.4 – continued from previous page**

| Name | Social, any | Country Club | Gent. Club | Frat. Order | Honorary Politic. | Professional |
|---|---|---|---|---|---|---|
| Creek Club , Locust Valley | 1 | 1 | 0 | 0 | 0 | 0 |
| Crusing Club of America | 1 | 1 | 0 | 0 | 0 | 0 |
| Cumberland Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Dedham Country and Polo Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Denver Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Des Moines Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Detroit Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Down Town Association (NYC) | 1 | 0 | 1 | 0 | 0 | 0 |
| Down Town Club (Boston) | 1 | 0 | 1 | 0 | 0 | 0 |
| Downtown athletic club (NYC) | 1 | 0 | 1 | 0 | 0 | 0 |
| Dublin Lake Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Duquensne Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Dutch Treat Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Duxbury Yacht Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Eastern Yacht Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Elks | 1 | 0 | 0 | 1 | 0 | 0 |
| Essex Club, Newark | 1 | 0 | 1 | 0 | 0 | 0 |
| Essex County Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Examiner Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Explorers Club | 0 | 0 | 0 | 0 | 0 | 1 |
| Fay Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Fort Orange Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Fort Schuyler Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Fraternal Order of Eagles | 1 | 0 | 0 | 1 | 0 | 0 |
| Genesee Valley Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Graduates Club of New Haven | 1 | 0 | 1 | 0 | 0 | 0 |
| Grand Street Boys Association | 1 | 0 | 1 | 0 | 0 | 0 |
| Harmonie club | 1 | 0 | 1 | 0 | 0 | 0 |
| Hartford Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Harvey Society | 0 | 0 | 0 | 0 | 0 | 1 |
| Home Club of Meriden | 1 | 0 | 1 | 0 | 0 | 0 |
| Hope Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Houston Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Idlewild Club | 1 | 0 | 1 | 0 | 0 | 0 |
| India House | 1 | 0 | 1 | 0 | 0 | 0 |
| Institute of Radio Engineers | 0 | 0 | 0 | 0 | 0 | 1 |
| Interdependent Order of Odd Fellows | 1 | 0 | 0 | 1 | 0 | 0 |
| Jonathan Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Kittansett Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Kiwanis [any] | 1 | 0 | 0 | 1 | 0 | 0 |
| Knickerbocker club | 1 | 0 | 1 | 0 | 0 | 0 |
| Knights of Columbus [any] | 1 | 0 | 0 | 1 | 0 | 0 |
| Knights of Pythias [any] | 1 | 0 | 0 | 1 | 0 | 0 |
| Links Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Lions Club [any] | 1 | 0 | 0 | 1 | 0 | 0 |
| Loyal Order of Moose | 1 | 0 | 0 | 1 | 0 | 0 |
| Madison Square Garden Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Manchester Yacht Club | 1 | 1 | 0 | 0 | 0 | 0 |

Continued on next page

**Table B.4 – continued from previous page**

| Name | Social, any | Country Club | Gent. Club | Frat. Order | Honorary Politic. | Professional |
|---|---|---|---|---|---|---|
| Manhattan Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Maryland Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Masons [any] | 1 | 0 | 0 | 1 | 0 | 0 |
| Massachusetts Bar Association | 0 | 0 | 0 | 0 | 0 | 1 |
| Massachusetts Medical Society | 0 | 0 | 0 | 0 | 0 | 1 |
| Mathematical Association of America | 0 | 0 | 0 | 0 | 0 | 1 |
| Merchants Club (NYC or Baltimore) | 1 | 0 | 1 | 0 | 0 | 0 |
| Merion Cricket Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Metropolitan club | 1 | 0 | 1 | 0 | 0 | 0 |
| Milton-Hoosic Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Minneapolis Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Modern Language Association [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| Myopia Hunt Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Nation Council of Teachers of English | 0 | 0 | 0 | 0 | 0 | 1 |
| National Association of Cost Accountants | 0 | 0 | 0 | 0 | 0 | 1 |
| National Education Association | 0 | 0 | 0 | 0 | 0 | 1 |
| National Geographic Society | 0 | 0 | 0 | 0 | 0 | 1 |
| National Press Club | 1 | 0 | 1 | 0 | 0 | 1 |
| New England Pediatric Society | 0 | 0 | 0 | 0 | 0 | 1 |
| New York Academy of Sciences | 0 | 0 | 0 | 0 | 0 | 1 |
| New York County Lawyers Association | 0 | 0 | 0 | 0 | 0 | 1 |
| New York Society of Security Analysts | 0 | 0 | 0 | 0 | 0 | 1 |
| New York State Bar Association | 0 | 0 | 0 | 0 | 0 | 1 |
| New York Yacht Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Newcomen Society [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| Noonday Club of St Louis | 1 | 0 | 1 | 0 | 0 | 0 |
| Omicron Delta Kappa | 0 | 0 | 0 | 0 | 1 | 0 |
| PEN club | 0 | 0 | 0 | 0 | 0 | 1 |
| Pacific Club (either nantucket or honolulu) | 1 | 0 | 1 | 0 | 0 | 0 |
| Pacific Union Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Pendennis Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Phi Beta Kappa | 0 | 0 | 0 | 0 | 1 | 0 |
| Phi Delta Kappa | 0 | 0 | 0 | 0 | 1 | 0 |
| Phi Kappa Phi | 0 | 0 | 0 | 0 | 1 | 0 |
| Philadelphia Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Pilgrims Society | 1 | 0 | 1 | 0 | 0 | 0 |
| Piping Rock Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Pittsburgh Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Public Relations Society of America | 0 | 0 | 0 | 0 | 0 | 1 |
| Quadrangle Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Quequechan club | 1 | 0 | 1 | 0 | 0 | 0 |
| Quiet birdmen | 1 | 0 | 1 | 0 | 0 | 0 |
| Racquet and Tennis Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Rainier Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Rittenhouse Club | 1 | 0 | 1 | 0 | 0 | 0 |
| River club (NYC) | 1 | 0 | 1 | 0 | 0 | 0 |
| Rockefeller Center Luncheon Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Rotary Club [any] | 1 | 0 | 0 | 1 | 0 | 0 |

Continued on next page

**Table B.4 – continued from previous page**

| Name | Social, any | Country Club | Gent. Club | Frat. Order | Honorary Politic. | Professional |
|---|---|---|---|---|---|---|
| Saturn Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Savile Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Sigma Xi | 0 | 0 | 0 | 0 | 1 | 0 |
| Society in Dedham for Apprehending Horse Theives | 1 | 0 | 1 | 0 | 0 | 0 |
| Somerset Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Sons of the Revolution | 1 | 0 | 0 | 1 | 0 | 0 |
| St. Botolph Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Standard Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Tau Beta Pi (engineer honor society) | 0 | 0 | 0 | 0 | 1 | 1 |
| Tavern Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Tennis and Racquet Club of Boston | 1 | 0 | 1 | 0 | 0 | 0 |
| The Brook Club | 1 | 0 | 1 | 0 | 0 | 0 |
| The Country Club, Brookline | 1 | 1 | 0 | 0 | 0 | 0 |
| The National Grange of the Order of the Patrons of Husbandry | 1 | 0 | 0 | 1 | 0 | 0 |
| The Recess | 1 | 0 | 1 | 0 | 0 | 0 |
| The Town Club of Scarsdale | 1 | 0 | 1 | 0 | 0 | 0 |
| The Travellers Club, Paris | 1 | 0 | 1 | 0 | 0 | 0 |
| Toledo Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Tuxedo Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Union Boat Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Union Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Wellesley Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Wellesley County Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Weston Golf Club | 1 | 1 | 0 | 0 | 0 | 0 |
| White Hall Club | 1 | 0 | 1 | 0 | 0 | 0 |
| White's Club London | 1 | 0 | 1 | 0 | 0 | 0 |
| Winchester Country Club | 1 | 1 | 0 | 0 | 0 | 0 |
| Worcester Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Yorick Club | 1 | 0 | 1 | 0 | 0 | 0 |
| Zionist Organization of America | 0 | 0 | 0 | 0 | 1 | 0 |
| accounting association [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| agriculture [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| american academy of [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| american club abroad [any] | 1 | 0 | 1 | 0 | 0 | 0 |
| arts professional association [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| athletic club [any] | 1 | 1 | 0 | 0 | 0 | 0 |
| bar or legal assocaition [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| beach club [any] | 1 | 1 | 0 | 0 | 0 | 0 |
| bond club [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| business honor society | 0 | 0 | 0 | 0 | 1 | 1 |
| century club [any] | 1 | 0 | 1 | 0 | 0 | 0 |
| chamber of commerce [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| city club [any] | 1 | 0 | 1 | 0 | 0 | 0 |
| commercial club [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| cotillion clubs [any] | 1 | 0 | 1 | 0 | 0 | 0 |
| country club [any] | 1 | 1 | 0 | 0 | 0 | 0 |

Continued on next page

**Table B.4 – continued from previous page**

| Name | Social, any | Country Club | Gent. Club | Frat. Order | Honorary Politic. | Professional |
|---|---|---|---|---|---|---|
| day of the week club -cannot distinguish among cities [any] | 1 | 0 | 1 | 0 | 0 | 0 |
| education professional [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| elite book club [any] | 1 | 0 | 1 | 0 | 0 | 0 |
| engineering professional association [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| finance professional association [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| insurance association [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| legal honor society | 0 | 0 | 0 | 0 | 1 | 1 |
| leisure animals [any] | 1 | 1 | 0 | 0 | 0 | 0 |
| management association | 0 | 0 | 0 | 0 | 0 | 1 |
| medical [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| misc industry and commerce assocation | 0 | 0 | 0 | 0 | 0 | 1 |
| officers club [any] | 1 | 0 | 1 | 0 | 0 | 0 |
| other professions | 0 | 0 | 0 | 0 | 0 | 1 |
| political club [any] | 0 | 0 | 0 | 0 | 1 | 0 |
| press club or jounalist professional association [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| public administration [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| science [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| shriners [any] | 1 | 0 | 0 | 1 | 0 | 0 |
| social science [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| technical associations [any] | 0 | 0 | 0 | 0 | 0 | 1 |
| union league club [any] | 1 | 0 | 1 | 0 | 0 | 0 |
| yacht club [any] | 1 | 1 | 0 | 0 | 0 | 0 |
| zionist organization [any] | 0 | 0 | 0 | 0 | 1 | 0 |

Table B.5: Precision of matches from data sources with only name and year

| Data source | N matched with Red Books | Red Book precision | N matched with Class Reports | Class Report precision |
|---|---|---|---|---|
| Rank Lists | 176 | 100% | 182 | 100% |
| Student Council Registers | 192 | 98.4% | 164 | 98.2% |
| Class Albums | 189 | 98.9% | 190 | 99.5% |

Quality checks for matches relying only on names and years done on random samples of 200 records from the data source being matched to Red Book and Class Report records.

## Table B.6: Most common Census occupation by Class Report occupation

| Class report occ. | Has wage inc. | Wage inc. | Wage inc. 5k+ | Has non-wage inc. | Modal census occ. | Occ. score | N | Share |
|---|---|---|---|---|---|---|---|---|
| Finance | 0.75 | 3637 | 0.39 | 0.64 | Managers, officials, and proprietors (n.e.c.) | 4200 | 246 | 0.30 |
| Senior management | 0.79 | 3703 | 0.40 | 0.56 | Managers, officials, and proprietors (n.e.c.) | 4200 | 385 | 0.26 |
| Law | 0.48 | 3217 | 0.25 | 0.68 | Lawyers and judges | 6200 | 545 | 0.59 |
| Medicine | 0.34 | 1921 | 0.11 | 0.74 | Physicians and surgeons | 8000 | 356 | 0.55 |

Income measures of individuals in our main occupation groups (finance, upper-management, law, medicine) from the Class Reports who also report the modal Census occupation (using 1950 occupation classificiations) within Class Report occupation. Inlcudes cohorts 1920-1930. Occupation scores are the median total income reported by in the 1950 Census of all individuals in that occupation rounded to the nearest hundred dollars.

## Table B.7: Most common Census occupations

| Census occ. | N | Has wage inc. | Wage inc. | Wage inc. 5k+ | Has non-wage inc. |
|---|---|---|---|---|---|
| *A. Harvard cohorts 1920-1930* | | | | | |
| Managers, officials, and proprietors (n.e.c.) | 1100 | 0.75 | 3476 | 0.35 | 0.56 |
| Lawyers and judges | 688 | 0.50 | 3228 | 0.26 | 0.67 |
| Physicians and surgeons | 385 | 0.36 | 1958 | 0.12 | 0.73 |
| Salesmen and sales clerks (n.e.c.) | 370 | 0.91 | 2449 | 0.13 | 0.36 |
| Teachers (n.e.c.) | 343 | 0.95 | 2247 | 0.04 | 0.43 |
| Professional, technical and kindred workers (n.e.c.) | 275 | 0.88 | 2750 | 0.12 | 0.40 |
| Clerical and kindred workers (n.e.c.) | 253 | 0.91 | 2069 | 0.04 | 0.32 |
| College professors | 201 | 0.97 | 2609 | 0.09 | 0.39 |
| All occupations | 5707 | 0.73 | 2747 | 0.17 | 0.50 |
| *B. All men ages 27-37* | | | | | |
| Managers, officials, and proprietors (n.e.c.) | 722984 | 0.54 | 1997 | 0.06 | 0.49 |
| Lawyers and judges | 63532 | 0.39 | 2639 | 0.14 | 0.66 |
| Physicians and surgeons | 52091 | 0.37 | 2074 | 0.11 | 0.66 |
| Salesmen and sales clerks (n.e.c.) | 560366 | 0.90 | 1490 | 0.01 | 0.15 |
| Teachers (n.e.c.) | 123466 | 0.95 | 1487 | 0.00 | 0.21 |
| Professional, technical and kindred workers (n.e.c.) | 78307 | 0.89 | 1825 | 0.02 | 0.19 |
| Clerical and kindred workers (n.e.c.) | 424290 | 0.95 | 1406 | 0.00 | 0.09 |
| College professors | 17977 | 0.90 | 2185 | 0.03 | 0.29 |
| All occupations | 11161415 | 0.76 | 1167 | 0.01 | 0.25 |

Income measures for Harvard cohorts 1920-1930 (panel A) and similarly-aged men (panel B) with most common Census occupations (using 1950 occupation classificiations) reported by Harvard cohorts 1920-1930.

Table B.8: Coefficient estimates from Lasso estimation

| Activities | | UY Clubs | | Occupations | | Associations | |
|---|---|---|---|---|---|---|---|
| Have any activity | 0.011 | Selective final club | 0.093 | Low management | -0.011 | Country club | 0.110 |
| Activity count | 0.023 | Hasty Pudding Inst. 1770 | 0.204 | Finance | 0.158 | Gentleman's club | 0.171 |
| Activity leadership position | 0.069 | Any final club | 0.236 | Medicine | -0.049 | Fraternal order | -0.033 |
| Social | 0.216 | | | Law | -0.031 | Prof. Association | -0.075 |
| Sports | 0.089 | | | Higher ed. | -0.029 | Honor society | -0.004 |
| Redbook | 0.041 | | | Science | -0.032 | | |
| Dorm committee | 0.039 | | | | | | |
| Language club | 0.116 | | | | | | |
| Politics club | -0.080 | | | | | | |
| N | 14383 | | 13394 | | 10970 | | 12597 |

Estimated coefficients from Lasso regressions of private feeder indicators on freshman activity category indicators, upper-year social clubs, career type indicators, and adult associations. Model selection conducted based on EBIC. Sample: students in all entering cohorts. Indices estimated using Lasso specifications as described in section 3.4.3.

Table B.9: Shares with non-missing first/last name indices in main and long-run samples

| | All | Private feeder | Public feeder |
|---|---|---|---|
| *A. Main sample (1923-1939)* | | | |
| Have FN culture indices | 0.997 | 0.997 | 0.998 |
| Have LN culture indices | 0.990 | 0.993 | 0.986 |
| *B. LR sample (1923-2015)* | | | |
| Have FN culture indices | 0.965 | 0.983 | 0.972 |
| Have LN culture indices | 0.949 | 0.974 | 0.959 |
| N (full sample) | 14383 | 3441 | 3454 |

Shares of students with non-missing first- and last-name cultural indices (Jewish, Catholic, Colonial). Panel A is the main 1920s/1930s entering cohort sample. Panel B is the full long-run sample. Name indices are non-missing whenever we observe a name in the 1920s and 1930s population Census records. Public feeder category reflects the extended public feeder definition in the long-run sample. See section B.5 for details.

Table B.10: Harvard sample compared to similarly aged men

| | Men 27-37 | Men 27-37 w col. 1+ | Men 27-37 w col. 4+ | Harvard | Private feeder |
|---|---|---|---|---|---|
| *A. Census: individual* | | | | | |
| Yrs. of col. 4+ | 0.072 | 0.532 | 1.000 | 0.851 | 0.834 |
| Yrs. of col. 5+ | 0.024 | 0.177 | 0.332 | 0.340 | 0.280 |
| In school | 0.012 | 0.037 | 0.040 | 0.030 | 0.025 |
| Cen. Occ.: Doc. | 0.007 | 0.048 | 0.086 | 0.074 | 0.052 |
| Cen. Occ.: Law. | 0.006 | 0.041 | 0.070 | 0.125 | 0.095 |
| In labor force | 0.949 | 0.962 | 0.963 | 0.954 | 0.951 |
| Non-farm self emp. | 0.084 | 0.160 | 0.191 | 0.207 | 0.193 |
| Farm | 0.191 | 0.064 | 0.042 | 0.037 | 0.074 |
| Non-wage inc.50+ | 0.253 | 0.301 | 0.331 | 0.503 | 0.646 |
| Has wage income | 0.758 | 0.774 | 0.757 | 0.726 | 0.712 |
| Wage income | 1167 | 1841 | 2044 | 2747 | 2961 |
| Wage inc. 5000+ | 0.008 | 0.034 | 0.047 | 0.174 | 0.247 |
| Household head | 0.692 | 0.704 | 0.699 | 0.718 | 0.776 |
| Own home | 0.188 | 0.209 | 0.201 | 0.210 | 0.271 |
| Home value | 2720 | 5297 | 6284 | 13574 | 19091 |
| Monthly rent | 60.1 | 89.6 | 99.2 | 103.1 | 109.4 |
| *B. Census: enumeration district* | | | | | |
| Central city | 0.193 | 0.226 | 0.245 | 0.512 | 0.543 |
| Dist. share farm | 0.200 | 0.078 | 0.058 | 0.038 | 0.062 |
| Dist. share non-farm self emp. | 0.106 | 0.146 | 0.151 | 0.163 | 0.159 |
| Dist. share non-wage inc. 50+ | 0.304 | 0.290 | 0.287 | 0.325 | 0.357 |
| Dist. 50p wage income | 1050 | 1408 | 1492 | 1800 | 1732 |
| Dist. share wage 5000+ | 0.014 | 0.036 | 0.045 | 0.109 | 0.122 |
| Dist. 50p home value | 2946 | 4376 | 4865 | 8541 | 9979 |
| Dist. share college 4+ | 0.060 | 0.128 | 0.151 | 0.234 | 0.259 |
| Dist. share Harvard | 0.001 | 0.002 | 0.003 | 0.045 | 0.062 |
| Dist. N men 27-37 | 129.3 | 135.1 | 135.8 | 115.1 | 110.7 |
| N | 11160198 | 1479570 | 787530 | 5723 | 1389 |

Demographic and labor market comparison of Harvard students to the broader Census population of similarly-aged men. Panel A: individual Census records. Panel B: attributes of Census enumeration districts. Columns are samples. Men 27-37: all Census men of listed age. Men 27-37 w col. 1+: men with at least one year of college. Men 27-37 w col. 4+: men with at least four years of college. Harvard: all Harvard students. Private feeder: Harvard students from private feeder schools. Degree attainment is not available in the 1940, so years of schooling serves as a proxy. Mean wage income and share with top-coded wages are conditional on reporting positive wage income. Home ownership, home value, and monthly rent are conditional on the individual being the head of their household.

Table B.11: Selection into upper-year social clubs

| | Hasty | Sel. fin. club | Sel. fin. club \| Hasty |
|---|---|---|---|
| Private feeder | 0.237 | 0.118 | 0.164 |
| | (0.014) | (0.011) | (0.036) |
| Harvard legacy | 0.090 | 0.045 | 0.037 |
| | (0.008) | (0.006) | (0.023) |
| Jewish name | -0.039 | -0.012 | -0.281 |
| | (0.005) | (0.003) | (0.133) |
| Catholic name | -0.020 | -0.023 | -0.232 |
| | (0.010) | (0.006) | (0.053) |
| Colonial name | 0.053 | 0.030 | 0.046 |
| | (0.007) | (0.006) | (0.023) |
| Low acad. rank | 0.046 | 0.030 | 0.133 |
| | (0.011) | (0.008) | (0.049) |
| High acad. rank | -0.033 | -0.012 | -0.063 |
| | (0.006) | (0.003) | (0.054) |
| Not ranked | -0.015 | -0.005 | -0.046 |
| | (0.007) | (0.004) | (0.049) |
| Social Leader | 0.325 | 0.144 | 0.170 |
| | (0.019) | (0.015) | (0.037) |
| Private feeder=1 $\times$ Low acad. rank | -0.037 | -0.001 | -0.068 |
| | (0.027) | (0.024) | (0.062) |
| Private feeder=1 $\times$ High acad. rank | -0.042 | -0.060 | -0.044 |
| | (0.027) | (0.020) | (0.071) |
| Private feeder=1 $\times$ Not ranked | -0.123 | -0.054 | 0.074 |
| | (0.022) | (0.018) | (0.067) |
| Private feeder=1 $\times$ Social Leader=1 | 0.053 | 0.175 | 0.071 |
| | (0.027) | (0.024) | (0.047) |
| Observations | 11494 | 11494 | 1769 |

Coefficient estimates from linear probability regressions with for selection into upper-year social clubs. Outcomes, listed in the columns, are Hasty Pudding sophomore society membership (left column) and selective final club membership (right two columns). Rightmost column restricts sample to Hasty Pudding members, from which nearly all final club members are selected. All columns control for cohort fixed effects in addition to listed coefficients. "Harvard legacy" is indicator for having a father or brother who attended Harvard. Jewish, Catholic, and Colonial name indicators are based on Census name frequencies. "Low acad. rank" is an indicator for first-year academic rank in bottom group. "High acad. rank" is an indicator for having first-year academic rank in the top three groups. "Not ranked" is a dummy for not receiving a first-year academic rank."Middle rank" category (groups four and five) is omitted. "Social leader" is indicator for participating in a social organizing committee or being the president of a club or captain of a team in freshman year. See main text for details. Robust standard errors in parentheses.

Table B.12: Brothers sample description: background and college outcomes

| | All | Wages & rank | Brothers | Mixed membership |
|---|---|---|---|---|
| *A. Demographics* | | | | |
| Have high school data | 0.984 | 0.972 | 0.966 | 0.913 |
| Any private high school | 0.463 | 0.456 | 0.630 | 0.952 |
| Private feeder | 0.240 | 0.242 | 0.375 | 0.783 |
| Any public feeder | 0.226 | 0.234 | 0.175 | 0.043 |
| From MA | 0.509 | 0.515 | 0.486 | 0.522 |
| From NY | 0.172 | 0.164 | 0.221 | 0.217 |
| Have Harvard father | 0.070 | 0.104 | 0.227 | 0.217 |
| Have Harvard brother | 0.204 | 0.227 | 1.000 | 1.000 |
| Jewish name | 0.066 | 0.057 | 0.066 | 0.000 |
| Catholic name | 0.055 | 0.049 | 0.024 | 0.000 |
| Colonial name | 0.293 | 0.315 | 0.398 | 0.739 |
| *B. Census childhood household demographics* | | | | |
| Have Census pre-Harvard | 0.619 | 0.702 | 0.842 | 0.826 |
| S or E Eur. immg. gen. 1-2 | 0.098 | 0.078 | 0.021 | 0.000 |
| Have father's occupation | 0.690 | 0.679 | 0.679 | 0.571 |
| Father's occupation: Doctor | 0.103 | 0.087 | 0.073 | 0.000 |
| Father's Occupation: Lawyer | 0.115 | 0.126 | 0.164 | 0.500 |
| *C. First-year campus location* | | | | |
| Have address data | 0.950 | 0.930 | 0.966 | 1.000 |
| Live on campus | 0.800 | 0.824 | 0.891 | 1.000 |
| Have room attributes | 0.712 | 0.709 | 0.792 | 0.955 |
| Room price per occupant | 209.2 | 198.6 | 213.8 | 224.4 |
| Peer neighborhood price | 213.9 | 208.2 | 216.9 | 213.4 |
| 25th pctile neighborhood rank | 0.256 | 0.256 | 0.304 | 0.424 |
| 75th pctile neighborhood rank | 0.711 | 0.685 | 0.734 | 0.852 |
| *D. Academic class rank groups* | | | | |
| Rank group 1 | 0.018 | 0.018 | 0.017 | 0.000 |
| Rank group 2 | 0.069 | 0.059 | 0.038 | 0.000 |
| Rank group 3 | 0.155 | 0.141 | 0.165 | 0.130 |
| Rank group 4 | 0.242 | 0.237 | 0.268 | 0.217 |
| Rank group 5 | 0.375 | 0.381 | 0.347 | 0.304 |
| Rank group 6 | 0.141 | 0.164 | 0.165 | 0.348 |
| Not ranked year 1 | 0.208 | 0.000 | 0.000 | 0.000 |
| *E. First-year activities* | | | | |
| Have any activity | 0.526 | 0.567 | 0.694 | 0.913 |
| N activities | 1.046 | 1.249 | 1.722 | 3.043 |
| Activity leadership position | 0.064 | 0.080 | 0.103 | 0.304 |
| Sports | 0.367 | 0.388 | 0.491 | 0.652 |
| Social | 0.083 | 0.104 | 0.172 | 0.435 |
| Music | 0.133 | 0.177 | 0.216 | 0.261 |
| First-year activity index | 0.000 | 0.051 | 0.377 | 1.232 |
| *F. Upper-year social clubs* | | | | |
| Hasty Pudding | 0.153 | 0.166 | 0.292 | 0.652 |
| Sel. Fin. Club | 0.070 | 0.076 | 0.151 | 0.522 |
| Any final club | 0.136 | 0.155 | 0.268 | 0.652 |
| Upper-year club index | 0.000 | 0.050 | 0.403 | 1.555 |
| N | 14383 | 3417 | 291 | 23 |

All: full Red Book sample universe. Wages and rank: All students in cohorts 1920-1930 who are matched to year 1 class rank and to the 1940 census with wage income reported. These are the students in the earnings specification of models A-C reported in 4. Brothers: Students in cohorts 1920-1930 matched to year 1 class rank and 1940 census wage income who have at least one brother meeting the same criteria. These are the students in the earnings specification of model E reported in 4. Mixed membership: Students with wages and rank group who have at least one brother who also has waged and rank group and has the opposite membership status in selective final clubs. These are the students whose within family variation identifies the final club membership returns parameter in the earnings specification of model E reported in 4. See text for detailed variable definitions.

Table B.13: Brothers sample description: adult outcomes

| | All | Wages & rank | Brothers | Mixed membership |
|---|---|---|---|---|
| *A. Adult associations* | | | | |
| Any social club | 0.343 | 0.370 | 0.423 | 0.565 |
| Country club | 0.242 | 0.249 | 0.330 | 0.391 |
| Gentleman's club | 0.112 | 0.126 | 0.168 | 0.304 |
| Fraternal order | 0.103 | 0.113 | 0.076 | 0.043 |
| Any honor/prof group | 0.369 | 0.440 | 0.419 | 0.391 |
| Adult association index | 0.000 | 0.009 | 0.214 | 0.570 |
| *B. Occupations* | | | | |
| Have occupation | 0.871 | 0.939 | 0.976 | 0.957 |
| Finance | 0.101 | 0.120 | 0.116 | 0.091 |
| Accounting | 0.108 | 0.125 | 0.116 | 0.318 |
| Medicine | 0.091 | 0.054 | 0.046 | 0.000 |
| Law | 0.123 | 0.100 | 0.085 | 0.091 |
| Higher ed. | 0.070 | 0.075 | 0.077 | 0.000 |
| Teach | 0.080 | 0.108 | 0.113 | 0.000 |
| Government | 0.034 | 0.037 | 0.039 | 0.000 |
| Art/pub | 0.074 | 0.080 | 0.095 | 0.045 |
| Senior management | 0.213 | 0.229 | 0.215 | 0.318 |
| Low management | 0.119 | 0.125 | 0.130 | 0.091 |
| Retail | 0.137 | 0.137 | 0.151 | 0.273 |
| Occupation index | 0.000 | 0.056 | 0.047 | 0.073 |
| *C. Adult census* | | | | |
| In school | 0.030 | 0.026 | 0.017 | 0.000 |
| In labor force | 0.954 | 0.982 | 0.976 | 1.000 |
| Wage income | 2747 | 2771 | 2970 | 3051 |
| Has wage income | 0.726 | 1.000 | 1.000 | 1.000 |
| Non-wage inc. 50+ | 0.503 | 0.398 | 0.509 | 0.619 |
| Wage inc. 5000+ | 0.174 | 0.181 | 0.237 | 0.348 |
| N | 14383 | 3417 | 291 | 23 |

All: full Red Book sample universe. Wages and rank: All students in cohorts 1920-1930 who are matched to year 1 class rank and to the 1940 census with wage income reported. These are the students in the earnings specification of models A-C reported in 4. Brothers: Students in cohorts 1920-1930 matched to year 1 class rank and 1940 census wage income who have at least one brother meeting the same criteria. These are the students in the earnings specification of model E reported in 4. Mixed membership: Students with wages and rank group who have at least one brother who also has waged and rank group and has the opposite membership status in selective final clubs. These are the students whose within family variation identifies the final club membership returns parameter in the earnings specification of model E reported in 4. See text for detailed variable definitions.

Table B.14: Labor market outcomes by academic performance and social success; missing rank data included

| | Has earnings | Earnings | Topcoded | Non-wage |
|---|---|---|---|---|
| *Including unranked students* | | | | |
| Private feeder | 0.003 | 130 | 0.048 | 0.157 |
| | (0.013) | ( 53) | (0.015) | (0.016) |
| Class rank year 1 | -0.001 | 50 | -0.004 | 0.008 |
| | (0.005) | ( 19) | (0.005) | (0.006) |
| No academic rank | -0.058 | 30 | -0.025 | -0.026 |
| | (0.016) | ( 62) | (0.017) | (0.020) |
| Selective final club | 0.005 | 728 | 0.212 | 0.156 |
| | (0.022) | ( 88) | (0.029) | (0.025) |
| Sample mean | 0.470 | 2747 | 0.174 | 0.503 |
| N | 8851 | 4156 | 4156 | 5552 |

OLS regressions of outcomes listed in the column on variables listed in the rows. Specifications parallel Table 4 Panel B, but include students with missing data on first-year grades. 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). All specifications include cohort fixed effects (not reported). "No academic rank" coefficient is relative to lowest rank group. Robust standard errors in parentheses.

Table B.15: Labor market outcomes by academic performance in third year and social success

| | Has earnings | Earnings | Topcoded | Non-wage | Imputed wages | Topcodes x 1.4 |
|---|---|---|---|---|---|---|
| *A. Baseline* | | | | | | |
| Private feeder | -0.013 | 403 | 0.114 | 0.187 | 194 | 287 |
| | (0.016) | ( 63) | (0.019) | (0.019) | ( 64) | ( 102) |
| Class rank year 3 | 0.004 | 83 | 0.009 | 0.005 | 150 | 237 |
| | (0.005) | ( 20) | (0.006) | (0.007) | ( 21) | ( 33) |
| Sample mean | 0.487 | 2797 | 0.177 | 0.507 | 3397 | 4193 |
| N | 5595 | 2722 | 2722 | 3685 | 3203 | 3203 |
| *B. Add most elite final clubs* | | | | | | |
| Private feeder | -0.016 | 226 | 0.065 | 0.149 | 58 | 88 |
| | (0.017) | ( 66) | (0.020) | (0.021) | ( 68) | ( 109) |
| Class rank year 3 | 0.004 | 94 | 0.012 | 0.007 | 159 | 250 |
| | (0.005) | ( 20) | (0.006) | (0.007) | ( 21) | ( 33) |
| Selective final club | 0.015 | 744 | 0.203 | 0.164 | 588 | 865 |
| | (0.027) | ( 100) | (0.035) | (0.030) | ( 97) | ( 160) |
| Sample mean | 0.487 | 2797 | 0.177 | 0.507 | 3397 | 4193 |
| N | 5595 | 2722 | 2722 | 3685 | 3203 | 3203 |
| *C. Add Census family background controls* | | | | | | |
| Private feeder | -0.026 | 106 | 0.033 | 0.100 | 28 | 39 |
| | (0.028) | ( 105) | (0.030) | (0.033) | ( 107) | ( 170) |
| Class rank year 3 | 0.024 | 92 | 0.014 | -0.000 | 90 | 143 |
| | (0.008) | ( 32) | (0.009) | (0.010) | ( 33) | ( 51) |
| Selective final club | 0.023 | 733 | 0.211 | 0.129 | 708 | 1070 |
| | (0.042) | ( 163) | (0.055) | (0.051) | ( 158) | ( 257) |
| Sample mean | 0.526 | 2802 | 0.183 | 0.495 | 3366 | 4150 |
| N | 2668 | 1364 | 1364 | 1854 | 1596 | 1596 |
| *D. Private feeders with HS FEs, legacy indicators* | | | | | | |
| Class rank year 3 | 0.007 | 86 | 0.017 | 0.004 | 109 | 181 |
| | (0.013) | ( 50) | (0.015) | (0.015) | ( 49) | ( 78) |
| Selective final club | 0.002 | 614 | 0.161 | 0.052 | 518 | 742 |
| | (0.036) | ( 141) | (0.045) | (0.040) | ( 138) | ( 222) |
| Have Harvard father | -0.034 | 31 | 0.038 | 0.085 | 236 | 421 |
| | (0.035) | ( 136) | (0.043) | (0.038) | ( 136) | ( 220) |
| Sample mean | 0.484 | 3062 | 0.257 | 0.650 | 3480 | 4310 |
| N | 1271 | 615 | 615 | 836 | 691 | 691 |
| *E. Within family* | | | | | | |
| Class rank year 3 | -0.018 | 53 | 0.013 | 0.028 | 141 | 215 |
| | (0.024) | ( 91) | (0.031) | (0.029) | ( 109) | ( 169) |
| Selective final club | -0.042 | 954 | 0.328 | 0.094 | 1017 | 1740 |
| | (0.123) | ( 590) | (0.146) | (0.183) | ( 525) | ( 783) |
| Sample mean | 0.508 | 2938 | 0.222 | 0.618 | 3480 | 4348 |
| N | 748 | 216 | 216 | 393 | 295 | 295 |
| *F. Within Hasty Pudding (approximate applicant pool)* | | | | | | |
| Private feeder | -0.044 | 133 | 0.023 | 0.115 | 73 | 123 |
| | (0.036) | ( 137) | (0.044) | (0.042) | ( 138) | ( 220) |
| Class rank year 3 | 0.016 | 162 | 0.067 | 0.022 | 160 | 305 |
| | (0.016) | ( 61) | (0.020) | (0.017) | ( 60) | ( 96) |
| Selective final club | -0.030 | 454 | 0.104 | 0.124 | 375 | 515 |
| | (0.034) | ( 124) | (0.042) | (0.039) | ( 125) | ( 205) |
| Sample mean | 0.501 | 3348 | 0.353 | 0.661 | 3694 | 4653 |
| N | 932 | 467 | 467 | 619 | 517 | 517 |

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). All specifications include cohort fixed effects (not reported). Panel C adds controls for father's occupation, father's and mother's state or non-US country of birth, family size, parental presence, home ownership, presence of domestic employees, and farm status; sample is restricted to students for whom these records are non-missing. Panel D restricts the sample to students who attended private feeder high schools and includes fixed effects for each high school. Panel E restricts the sample to students from families where multiple brothers attended Harvard during our sample period, and includes family fixed effects. Panel F includes only students who are members of the Hasty Pudding club. Robust standard errors in parentheses. Standard errors of model D are clustered at the family level.

Table B.16: Entryway rank effects on short-run outcomes

| | All | Private | Non-private | Test |
|---|---|---|---|---|
| *A. First-year activities* | | | | |
| Have any activity | 0.052 | 0.194 | -0.007 | 0.003 |
| | (0.035) | (0.059) | (0.042) | |
| N activities | 0.182 | 0.525 | 0.040 | 0.078 |
| | (0.120) | (0.252) | (0.124) | |
| Activity leadership position | 0.017 | 0.055 | 0.003 | 0.238 |
| | (0.019) | (0.040) | (0.022) | |
| Social | 0.054 | 0.199 | -0.015 | 0.001 |
| | (0.026) | (0.057) | (0.025) | |
| Sports | -0.012 | 0.098 | -0.065 | 0.024 |
| | (0.036) | (0.067) | (0.039) | |
| Music | 0.015 | 0.059 | -0.004 | 0.291 |
| | (0.027) | (0.054) | (0.032) | |
| Other activities | 0.027 | 0.025 | 0.041 | 0.815 |
| | (0.032) | (0.059) | (0.036) | |
| First-year activity index | 0.152 | 0.535 | -0.030 | 0.004 |
| | (0.086) | (0.179) | (0.092) | |
| N | 9640 | 2850 | 6633 | |
| *B. Upper-year social clubs* | | | | |
| Selective final club | 0.035 | 0.127 | -0.009 | 0.022 |
| | (0.024) | (0.057) | (0.019) | |
| Less selective final club | -0.020 | -0.066 | 0.004 | 0.186 |
| | (0.021) | (0.050) | (0.021) | |
| Hasty Pudding Inst. 1770 | 0.001 | 0.083 | -0.023 | 0.147 |
| | (0.033) | (0.066) | (0.035) | |
| Upper-year club index | 0.056 | 0.253 | -0.019 | 0.156 |
| | (0.089) | (0.181) | (0.088) | |
| N | 8886 | 2629 | 6110 | |
| *C. First-year academic rank* | | | | |
| Rank group 1 | 0.005 | -0.006 | 0.010 | 0.394 |
| | (0.011) | (0.009) | (0.016) | |
| Rank group 2 | 0.011 | 0.018 | 0.007 | 0.722 |
| | (0.017) | (0.022) | (0.022) | |
| Rank group 3 | -0.006 | -0.033 | 0.008 | 0.365 |
| | (0.023) | (0.035) | (0.029) | |
| Rank group 4 | -0.008 | 0.069 | -0.053 | 0.028 |
| | (0.026) | (0.043) | (0.034) | |
| Rank group 5 | 0.073 | 0.071 | 0.076 | 0.942 |
| | (0.034) | (0.055) | (0.043) | |
| Rank group 6 | -0.023 | -0.033 | -0.025 | 0.879 |
| | (0.023) | (0.043) | (0.028) | |
| Rank listed year 1 | 0.053 | 0.086 | 0.023 | 0.289 |
| | (0.027) | (0.050) | (0.033) | |
| Class rank year 1 | 0.036 | 0.032 | 0.069 | 0.863 |
| | (0.106) | (0.152) | (0.141) | |
| N | 7268 | 2103 | 5012 | |

Coefficients on entryway price rank from regressions of form given in equation 2. Specifications are identical to those reported in Table 6 but with ranks based on much larger entryway groups, not peer neighborhood. See section B.7 for details. Standard errors clustered at entryway level.

Table B.17: Entryway rank effects on long-run outcomes

| | All | Private | Non-private | Test |
|---|---|---|---|---|
| *A. Adult associations* | | | | |
| Any social club | 0.043 | 0.134 | -0.001 | 0.105 |
| | (0.036) | (0.069) | (0.046) | |
| Country club | 0.040 | 0.122 | 0.003 | 0.120 |
| | (0.034) | (0.066) | (0.041) | |
| Gentleman's club | 0.060 | 0.091 | 0.037 | 0.404 |
| | (0.025) | (0.058) | (0.026) | |
| Fraternal order | -0.015 | -0.009 | -0.016 | 0.877 |
| | (0.022) | (0.033) | (0.029) | |
| Any honor/prof group | -0.034 | 0.035 | -0.073 | 0.143 |
| | (0.035) | (0.061) | (0.043) | |
| Prof. Association | -0.034 | 0.020 | -0.072 | 0.208 |
| | (0.034) | (0.060) | (0.042) | |
| Honor society | -0.003 | 0.034 | -0.016 | 0.252 |
| | (0.019) | (0.036) | (0.024) | |
| Adult association index | 0.193 | 0.307 | 0.138 | 0.386 |
| | (0.082) | (0.179) | (0.088) | |
| N | 8453 | 2497 | 5797 | |
| *B. Occupation choice* | | | | |
| Finance | 0.030 | 0.173 | -0.048 | 0.001 |
| | (0.024) | (0.058) | (0.026) | |
| Medicine | 0.006 | 0.010 | -0.004 | 0.696 |
| | (0.023) | (0.030) | (0.030) | |
| Higher ed. | -0.013 | -0.031 | -0.005 | 0.541 |
| | (0.019) | (0.032) | (0.027) | |
| Law | -0.030 | -0.056 | -0.020 | 0.514 |
| | (0.027) | (0.047) | (0.034) | |
| Business | -0.008 | 0.026 | -0.042 | 0.447 |
| | (0.040) | (0.075) | (0.049) | |
| Teach | -0.018 | -0.027 | -0.008 | 0.664 |
| | (0.020) | (0.034) | (0.027) | |
| Government | 0.002 | -0.007 | 0.015 | 0.523 |
| | (0.014) | (0.030) | (0.018) | |
| Art/pub | 0.020 | 0.021 | 0.028 | 0.878 |
| | (0.020) | (0.036) | (0.026) | |
| Occupation index | 0.099 | 0.534 | -0.124 | 0.001 |
| | (0.079) | (0.185) | (0.088) | |
| N | 7285 | 2112 | 5017 | |
| *C. Adult income* | | | | |
| Wage income | -102.9 | -113.4 | -112.8 | 0.999 |
| | (160.8) | (349.6) | (201.4) | |
| Wage inc. 5000+ | 0.026 | 0.047 | 0.001 | 0.684 |
| | (0.045) | (0.098) | (0.055) | |
| Non-wage inc. 50+ | 0.045 | 0.093 | -0.012 | 0.364 |
| | (0.056) | (0.091) | (0.070) | |
| Class Report wage index | 2.7 | 52.1 | -26.8 | 0.222 |
| | (30.0) | (57.3) | (34.0) | |
| N | 2552 | 705 | 1754 | |

Coefficients on entryway price rank from regressions of form given in equation 2. Specifications are identical to those reported in Table 7 but with ranks based on much larger entryway groups, not peer neighborhood. See section B.7 for details. Standard errors clustered at entryway level.

Table B.18: Nearest neighbor rank effects on short-run outcomes

| | All | Private | Non-private | Test |
|---|---|---|---|---|
| *A. First-year activities* | | | | |
| Have any activity | 0.014 | -0.039 | 0.032 | 0.334 |
| | (0.036) | (0.060) | (0.045) | |
| N activities | 0.169 | 0.458 | 0.044 | 0.128 |
| | (0.117) | (0.246) | (0.129) | |
| Activity leadership position | 0.032 | 0.082 | 0.005 | 0.142 |
| | (0.021) | (0.050) | (0.021) | |
| Social | 0.025 | 0.154 | -0.045 | 0.003 |
| | (0.027) | (0.064) | (0.025) | |
| Sports | -0.015 | 0.006 | -0.033 | 0.640 |
| | (0.038) | (0.069) | (0.047) | |
| Music | 0.054 | 0.096 | 0.039 | 0.398 |
| | (0.029) | (0.057) | (0.036) | |
| Other activities | -0.002 | -0.081 | 0.049 | 0.066 |
| | (0.032) | (0.063) | (0.037) | |
| First-year activity index | 0.064 | 0.256 | -0.068 | 0.116 |
| | (0.088) | (0.189) | (0.093) | |
| N | 9117 | 2744 | 6225 | |
| *B. Upper-year social clubs* | | | | |
| Selective final club | 0.037 | 0.114 | -0.014 | 0.043 |
| | (0.023) | (0.062) | (0.018) | |
| Less selective final club | -0.025 | -0.046 | -0.015 | 0.582 |
| | (0.021) | (0.052) | (0.023) | |
| Hasty Pudding Inst. 1770 | 0.013 | 0.113 | -0.031 | 0.042 |
| | (0.030) | (0.065) | (0.031) | |
| Upper-year club index | 0.061 | 0.298 | -0.083 | 0.057 |
| | (0.081) | (0.188) | (0.080) | |
| N | 8409 | 2536 | 5735 | |
| *C. First-year academic rank* | | | | |
| Rank group 1 | 0.001 | -0.015 | 0.003 | 0.420 |
| | (0.011) | (0.016) | (0.014) | |
| Rank group 2 | 0.019 | 0.009 | 0.020 | 0.765 |
| | (0.017) | (0.026) | (0.022) | |
| Rank group 3 | 0.009 | -0.006 | 0.009 | 0.775 |
| | (0.025) | (0.044) | (0.031) | |
| Rank group 4 | 0.014 | 0.114 | -0.027 | 0.018 |
| | (0.031) | (0.049) | (0.038) | |
| Rank group 5 | 0.003 | -0.130 | 0.080 | 0.006 |
| | (0.034) | (0.065) | (0.042) | |
| Rank group 6 | 0.027 | 0.060 | -0.006 | 0.266 |
| | (0.025) | (0.050) | (0.030) | |
| Rank listed year 1 | 0.073 | 0.032 | 0.078 | 0.482 |
| | (0.031) | (0.055) | (0.038) | |
| Class rank year 1 | 0.028 | 0.007 | 0.031 | 0.915 |
| | (0.109) | (0.195) | (0.133) | |
| N | 6871 | 2028 | 4698 | |

Coefficients on nearest-neighbor price rank from regressions of form given in equation 2. Specifications are identical to those reported in Table 6 but with ranks based on smaller nearest-neighbor groups, not peer neighborhood. See section B.7 for details. Standard errors clustered at nearest-neighbor level.

Table B.19: Nearest neighbor effects on long-run outcomes

| | All | Private | Non-private | Test |
|---|---|---|---|---|
| *A. Adult associations* | | | | |
| Any social club | 0.053 | 0.224 | -0.020 | 0.003 |
| | (0.037) | (0.069) | (0.048) | |
| Country club | 0.052 | 0.174 | -0.010 | 0.021 |
| | (0.034) | (0.068) | (0.043) | |
| Gentleman's club | 0.009 | 0.067 | -0.017 | 0.204 |
| | (0.028) | (0.060) | (0.030) | |
| Fraternal order | -0.002 | 0.036 | -0.015 | 0.302 |
| | (0.024) | (0.039) | (0.030) | |
| Any honor/prof group | 0.016 | 0.046 | 0.015 | 0.683 |
| | (0.034) | (0.064) | (0.043) | |
| Prof. Association | 0.012 | 0.034 | 0.008 | 0.731 |
| | (0.033) | (0.063) | (0.042) | |
| Honor society | -0.019 | 0.026 | -0.025 | 0.239 |
| | (0.020) | (0.035) | (0.026) | |
| Adult association index | 0.070 | 0.286 | -0.043 | 0.089 |
| | (0.080) | (0.172) | (0.093) | |
| N | 7985 | 2405 | 5430 | |
| *B. Occupation choice* | | | | |
| Finance | 0.000 | 0.113 | -0.063 | 0.011 |
| | (0.027) | (0.063) | (0.031) | |
| Medicine | -0.009 | -0.062 | 0.014 | 0.123 |
| | (0.024) | (0.038) | (0.032) | |
| Higher ed. | -0.017 | -0.022 | -0.029 | 0.880 |
| | (0.022) | (0.037) | (0.029) | |
| Law | -0.031 | 0.028 | -0.056 | 0.137 |
| | (0.028) | (0.045) | (0.038) | |
| Business | 0.085 | 0.119 | 0.053 | 0.487 |
| | (0.042) | (0.081) | (0.051) | |
| Teach | -0.008 | -0.016 | -0.002 | 0.789 |
| | (0.023) | (0.042) | (0.029) | |
| Government | 0.013 | 0.037 | 0.011 | 0.449 |
| | (0.015) | (0.029) | (0.019) | |
| Art/pub | -0.012 | -0.023 | 0.002 | 0.585 |
| | (0.022) | (0.036) | (0.028) | |
| Occupation index | 0.014 | 0.374 | -0.185 | 0.010 |
| | (0.087) | (0.194) | (0.103) | |
| N | 6885 | 2037 | 4700 | |
| *C. Adult income* | | | | |
| Wage income | -26.1 | -440.5 | 191.0 | 0.138 |
| | (191.4) | (367.8) | (236.3) | |
| Wage inc. 5000+ | 0.020 | -0.029 | 0.073 | 0.369 |
| | (0.050) | (0.099) | (0.061) | |
| Non-wage inc. 50+ | 0.053 | 0.081 | 0.044 | 0.753 |
| | (0.052) | (0.099) | (0.064) | |
| Class Report wage index | 33.6 | 128.8 | -17.2 | 0.011 |
| | (26.7) | (48.1) | (33.8) | |
| N | 2396 | 697 | 1611 | |

Coefficients on nearest-neighbor price rank from regressions of form given in equation 2. Specifications are identical to those reported in Table 7 but with ranks based on smaller nearest-neighbor groups, not peer neighborhood. See section B.7 for details. Standard errors clustered at nearest-neighbor level.

Table B.20: Peer neighborhood effects on outcome-by-grade interactions for private feeder students

| | Rank group 1-4 | | Rank group 5-6 or unlisted | |
|---|---|---|---|---|
| *A. Upper-year social clubs* | | | | |
| Hasty Pudding Inst. 1770 | 0.091 | (0.050) | -0.003 | (0.058) |
| Selective final club | 0.129 | (0.039) | 0.054 | (0.051) |
| N | 2457 | | 2457 | |
| *B. Occupation choice* | | | | |
| Finance | 0.113 | (0.034) | 0.030 | (0.049) |
| Medicine | -0.044 | (0.021) | 0.026 | (0.027) |
| Law | -0.004 | (0.032) | -0.006 | (0.032) |
| Higher ed. | -0.049 | (0.026) | 0.023 | (0.020) |
| N | 2094 | | 2094 | |
| *C. Adult associations* | | | | |
| Any social club | 0.159 | (0.048) | 0.082 | (0.061) |
| Prof. Association | 0.048 | (0.043) | -0.016 | (0.054) |
| N | 2678 | | 2678 | |

Coefficients on peer neighborhood price rank from regressions of the form given in equation 2. Outcomes are interactions between the social or career outcome listed in the row and indicators for academic performance levels listed in the column. All specifications include randomization block and large feeder dummies; see section 5.1 for details. See section B.8.1 for description of outcome variables. The sample is private feeder students only. Standard errors clustered at peer neighborhood level are reported in the columns to the right of coefficients.

Table B.21: Final club effects on adult outcomes for pre-Depression and Depression graduates

| | (1) Finance | (2) Medicine | (3) Higher ed. | (4) Law | (5) Occ. index | (6) Any social club | (7) Country club |
|---|---|---|---|---|---|---|---|
| Priv. fdr. | 0.072 | -0.038 | -0.022 | -0.031 | 0.276 | 0.107 | 0.127 |
| | (0.008) | (0.006) | (0.006) | (0.008) | (0.027) | (0.011) | (0.010) |
| Sel. fin. club | 0.100 | -0.023 | -0.007 | -0.044 | 0.323 | 0.160 | 0.154 |
| | (0.026) | (0.013) | (0.012) | (0.017) | (0.077) | (0.029) | (0.029) |
| Sel. fin. club=1 $\times$ Depression graduate=1 | 0.051 | -0.002 | -0.026 | 0.063 | 0.202 | 0.077 | 0.081 |
| | (0.033) | (0.017) | (0.015) | (0.023) | (0.103) | (0.036) | (0.036) |
| Observations | 10247 | 10247 | 10247 | 10247 | 10247 | 11667 | 11667 |

Linear regressions of Class Report occupation adult occupation and social club outcomes on listed covariates and cohort fixed effects. "Depression graduates" are graduating cohorts from 1930-1939. See section B.8.2 for details. Robust standard errors in parentheses.

Table B.22: Residential peer effects for private feeder students before and during the Great Depression

| | Pre-depression cohorts | Depression cohorts | Test |
|---|---|---|---|
| Selective final club | 0.134 | 0.167 | 0.773 |
| | (0.093) | (0.069) | |
| Any social club | 0.214 | 0.258 | 0.751 |
| | (0.114) | (0.082) | |
| Occupation index | 0.694 | 0.340 | 0.348 |
| | (0.305) | (0.224) | |
| Class Report wage index | 141.824 | 104.717 | 0.679 |
| | (75.546) | (49.115) | |
| N | 739 | 1355 | |

Estimates of the effect of residential peer neighborhood price rank from equation 2 in the sample of private feeder students, split by graduating cohort. All specifications include randomization block fixed effects and fixed effects for major feeder high schools. "Pre-Depression cohorts" are those graduating 1929 and earlier; "Depression cohorts" are those graduating 1930 and later. "Test" column reports the p-value from a statistical test that estimated effects for pre-Depression and Depression cohorts are equal. Robust standard errors in parentheses. See section B.8.2 for details.

## Table B.23: Characteristics of students choosing different majors

| | Cohorts w/ maj. | Have Major | Cohorts w/ int. | Have intent data | Economics | STEM/Eng. | Humanities | Soc. Science |
|---|---|---|---|---|---|---|---|---|
| Have major | 0.831 | 1.000 | | | 1.000 | 1.000 | 1.000 | 1.000 |
| Share in major | | | | | 0.156 | 0.261 | 0.458 | 0.125 |
| *A. Demographics* | | | | | | | | |
| Private feeder | 0.238 | 0.246 | 0.236 | 0.229 | 0.211 | 0.165 | 0.311 | 0.222 |
| Public feeder | 0.321 | 0.323 | 0.333 | 0.349 | 0.368 | 0.460 | 0.234 | 0.341 |
| From MA | 0.506 | 0.516 | 0.508 | 0.518 | 0.545 | 0.561 | 0.478 | 0.527 |
| From NY | 0.171 | 0.166 | 0.170 | 0.162 | 0.141 | 0.132 | 0.194 | 0.168 |
| Have Harvard father | 0.057 | 0.058 | 0.064 | 0.066 | 0.040 | 0.055 | 0.073 | 0.033 |
| Have Harvard brother | 0.212 | 0.222 | 0.211 | 0.220 | 0.209 | 0.200 | 0.246 | 0.194 |
| *B. College outcomes* | | | | | | | | |
| Class rank year 1 | 4.221 | 4.197 | 4.238 | 4.186 | 4.490 | 3.913 | 4.219 | 4.274 |
| N activities | 0.911 | 0.966 | 0.952 | 1.049 | 0.922 | 0.776 | 1.071 | 1.036 |
| Social | 0.050 | 0.053 | 0.058 | 0.062 | 0.034 | 0.028 | 0.072 | 0.063 |
| Sports | 0.374 | 0.392 | 0.371 | 0.397 | 0.424 | 0.326 | 0.409 | 0.428 |
| Music | 0.101 | 0.107 | 0.114 | 0.130 | 0.098 | 0.111 | 0.111 | 0.093 |
| Hasty Pudding Inst. 1770 | 0.139 | 0.166 | 0.137 | 0.163 | 0.118 | 0.089 | 0.226 | 0.164 |
| Selective final club | 0.063 | 0.075 | 0.063 | 0.071 | 0.039 | 0.026 | 0.116 | 0.066 |
| *C. Career intent at grad.* | | | | | | | | |
| Have intent data | 0.625 | 0.747 | 0.626 | 1.000 | 0.811 | 0.785 | 0.699 | 0.762 |
| Finance | 0.050 | 0.050 | 0.052 | 0.052 | 0.148 | 0.007 | 0.047 | 0.023 |
| Business | 0.259 | 0.260 | 0.284 | 0.284 | 0.495 | 0.112 | 0.259 | 0.273 |
| Higher education | 0.052 | 0.052 | 0.051 | 0.051 | 0.011 | 0.158 | 0.012 | 0.008 |
| Medicine | 0.142 | 0.143 | 0.137 | 0.137 | 0.016 | 0.381 | 0.058 | 0.081 |
| Law | 0.183 | 0.183 | 0.185 | 0.185 | 0.195 | 0.039 | 0.214 | 0.378 |
| *D. Adult outcomes* | | | | | | | | |
| Have soc. club | 0.466 | 0.499 | 0.470 | 0.527 | 0.536 | 0.463 | 0.517 | 0.463 |
| Prof. Association | 0.374 | 0.405 | 0.375 | 0.440 | 0.349 | 0.543 | 0.357 | 0.364 |
| Honor society | 0.082 | 0.090 | 0.083 | 0.094 | 0.063 | 0.127 | 0.081 | 0.081 |
| Have occupation | 0.796 | 0.831 | 0.808 | 0.857 | 0.861 | 0.834 | 0.835 | 0.774 |
| Finance | 0.090 | 0.089 | 0.092 | 0.092 | 0.137 | 0.029 | 0.103 | 0.102 |
| Medicine | 0.098 | 0.103 | 0.096 | 0.118 | 0.017 | 0.270 | 0.046 | 0.071 |
| Law | 0.117 | 0.123 | 0.116 | 0.135 | 0.153 | 0.043 | 0.127 | 0.246 |
| Higher ed. | 0.076 | 0.080 | 0.073 | 0.074 | 0.039 | 0.089 | 0.097 | 0.046 |
| *E. Earnings* | | | | | | | | |
| Has earnings | 0.475 | 0.501 | 0.479 | 0.508 | 0.547 | 0.468 | 0.502 | 0.490 |
| Wage income | 2207.860 | 2234.189 | 2399.751 | 2494.878 | 2341.664 | 2023.465 | 2273.353 | 2345.492 |
| Non-wage inc. 50+ | 0.449 | 0.450 | 0.468 | 0.469 | 0.366 | 0.466 | 0.495 | 0.357 |
| Wage inc. 5000+ | 0.073 | 0.073 | 0.099 | 0.103 | 0.071 | 0.048 | 0.088 | 0.063 |
| N (full sample) | 7526 | 6265 | 9155 | 5738 | 976 | 1636 | 2870 | 782 |

Characteristics of students in full sample (first column), in the sample of students for whom major is non-missing (second column), and by broad major grouping (remaining columns). Majors are observed for cohorts entering in 1927 or later. "Share in major" row reports shares of students for whom major data is available choosing the listed major. All other cells are sample means. Panel A reports baseline demographics. Panel B reports college outcomes. Panel C reports stated career intent at time of graduation. "Have intent data" is an indicator equal to one for people who report intent in their senior class register. Panel D reports adult outcomes from Class Report records. Panel E reports outcomes observed in Census data. The sample for Census outcomes is restricted to entering cohorts from 1930 and earlier. See section B.8.3 for details.

Table B.24: Major-specific wage and occupational premia

| | (1) | | (2) | | (3) | | (4) | |
|---|---|---|---|---|---|---|---|---|
| | $\beta$ | SE | $\beta$ | SE | $\beta$ | SE | $\beta$ | SE |
| *A. Wage income* | | | | | | | | |
| Humanities | 210.3 | ( 80.4) | 395.5 | ( 92.0) | 366.1 | ( 92.5) | 365.9 | ( 92.6) |
| Soc. Science | 340.3 | (124.8) | 494.0 | (132.9) | 477.2 | (134.0) | 477.4 | (134.0) |
| Economics | 319.3 | ( 94.3) | 503.6 | (106.4) | 506.4 | (106.6) | 510.0 | (114.6) |
| Private feeder | 290.9 | ( 80.0) | 300.4 | ( 84.1) | 174.7 | ( 89.1) | 178.4 | (100.8) |
| Class rank year 1 | | | 55.5 | ( 32.0) | 59.2 | ( 32.0) | 59.1 | ( 32.0) |
| Sel. fin. club | | | | | 531.4 | (163.7) | 529.9 | (164.6) |
| Private Feeder=1 × Economics=1 | | | | | | | -17.1 | (199.4) |
| Observations | 1380 | | 1191 | | 1191 | | 1191 | |
| *B. Topcoded income* | | | | | | | | |
| Humanities | 0.034 | (0.016) | 0.041 | (0.018) | 0.034 | (0.018) | 0.034 | (0.018) |
| Soc. Science | 0.019 | (0.024) | 0.011 | (0.024) | 0.007 | (0.024) | 0.007 | (0.024) |
| Economics | 0.023 | (0.019) | 0.027 | (0.021) | 0.028 | (0.021) | 0.029 | (0.022) |
| Private feeder | 0.062 | (0.019) | 0.058 | (0.020) | 0.029 | (0.020) | 0.030 | (0.022) |
| Class rank year 1 | | | -0.008 | (0.007) | -0.007 | (0.007) | -0.007 | (0.007) |
| Sel. fin. club | | | | | 0.124 | (0.045) | 0.124 | (0.045) |
| Private Feeder=1 × Economics=1 | | | | | | | -0.007 | (0.049) |
| Observations | 1380 | | 1191 | | 1191 | | 1191 | |
| *C. Finance* | | | | | | | | |
| Humanities | 0.060 | (0.008) | 0.056 | (0.008) | 0.047 | (0.008) | 0.047 | (0.008) |
| Soc. Science | 0.068 | (0.013) | 0.068 | (0.014) | 0.064 | (0.013) | 0.064 | (0.013) |
| Economics | 0.101 | (0.013) | 0.099 | (0.014) | 0.098 | (0.014) | 0.096 | (0.015) |
| Private feeder | 0.088 | (0.011) | 0.082 | (0.012) | 0.049 | (0.012) | 0.047 | (0.012) |
| Class rank year 1 | | | -0.011 | (0.003) | -0.009 | (0.003) | -0.009 | (0.003) |
| Sel. fin. club | | | | | 0.164 | (0.025) | 0.164 | (0.025) |
| Private Feeder=1 × Economics=1 | | | | | | | 0.013 | (0.037) |
| Observations | 5171 | | 4436 | | 4436 | | 4436 | |

Regressions of outcomes listed in panel title on major type, social outcomes, and academic outcomes. Sample in all panels is restricted to entering cohorts from 1927 and later, for whom we observe major. Panels A and B use wage outcomes reported in the 1940 Census, and further restrict the sample to individuals entering Harvard in 1930 or earlier (i.e., at least six years past scheduled college completion by 1940) and who merge successfully to Census wage records. "Topcoded income" is an indicator equal to one for individuals who report the Census maximum of wage income, which is $5,000. Panel C uses Class Report data and takes as an outcome and indicator equal to one if the individual reports working in finance. Omitted major category in all specifications is STEM. Class rank year 1 is reverse-signed, so higher ranks correspond to better grades. See section B.8.3 for details. Robust standard errors in parentheses.

Table B.25: Peer neighborhood effects on major choice and stated career intent

| | All | Private | Non-private | Test |
|---|---|---|---|---|
| *A. Major & Intended Occ. Indices* | | | | |
| Major index | 0.045 | 0.118 | -0.005 | 0.566 |
| | (0.110) | (0.172) | (0.141) | |
| Intended Occ. index | 0.150 | 0.377 | 0.020 | 0.201 |
| | (0.111) | (0.259) | (0.123) | |
| Major + Intended Occ. index | 0.276 | 0.464 | 0.145 | 0.259 |
| | (0.121) | (0.256) | (0.143) | |
| Finance index | 0.013 | 0.014 | 0.006 | 0.669 |
| | (0.007) | (0.018) | (0.008) | |
| N | 3333 | 929 | 2330 | |
| *B. Major Choice* | | | | |
| Economics | 0.032 | -0.052 | 0.075 | 0.094 |
| | (0.038) | (0.058) | (0.049) | |
| STEM/Eng. | 0.024 | 0.007 | 0.047 | 0.625 |
| | (0.042) | (0.067) | (0.054) | |
| Humanities | 0.022 | 0.059 | -0.003 | 0.561 |
| | (0.055) | (0.086) | (0.070) | |
| Social Science | -0.030 | 0.070 | -0.079 | 0.025 |
| | (0.032) | (0.055) | (0.041) | |
| Double Major | 0.029 | 0.057 | 0.019 | 0.179 |
| | (0.012) | (0.021) | (0.017) | |
| N | 4481 | 1339 | 3061 | |
| *C. Intended Occupation* | | | | |
| Finance | 0.029 | 0.082 | -0.005 | 0.231 |
| | (0.027) | (0.070) | (0.028) | |
| Business | 0.045 | 0.021 | 0.046 | 0.825 |
| | (0.047) | (0.100) | (0.057) | |
| Higher Ed. | -0.012 | 0.008 | -0.031 | 0.347 |
| | (0.020) | (0.033) | (0.027) | |
| Medicine | -0.006 | -0.084 | 0.037 | 0.079 |
| | (0.036) | (0.054) | (0.047) | |
| Law | -0.034 | 0.001 | -0.026 | 0.768 |
| | (0.044) | (0.075) | (0.057) | |
| N | 3878 | 1087 | 2697 | |

Coefficients on peer neighborhood price rank from regressions of form given in equation 2. The first three columns denote samples. Rows are outcome variables. All specifications include randomization block and dummies for large feeder high schools; see section 5.1 for details."Test" column reports the p-value from a test of the null that the coefficients reported in the private and non-private columns are equal. Panel A reports indices based on combinations of graduating major and stated career intent. "Major index" is a standardized index that uses major choice to predict whether a student came from a private feeder school. "Intended occupation index" is a standardized index that uses stated occupational intent to predict private feeder status. "Major+Intended Occ index" uses both major and stated occupational intent to predict private feeder status. "Finance index" uses both major and stated occupational intent to predict whether a student reports a finance occupation in Class report data; this index is not standardized. Panels B and C report specifications where the outcome is the listed major/stated career intent. Data on major available for 1927 and later entering cohorts only. Data on occupation intent available for 1924 and later entering cohorts. Samples restrict to observations where the relevant combination of major/intent data is available. Sample sizes vary across specifications in Panel A due to differential availability of major and intent data. Reported sample sizes are for the joint major/intent index. See section B.8.3 for details.

Table B.26: Descriptive statistics on marriage rates and spouse characteristics

| | All | Private feeder | All non-private | Hasty Pudding | Final club |
|---|---|---|---|---|---|
| *A. Marriage rates* | | | | | |
| Married | 0.740 | 0.759 | 0.733 | 0.803 | 0.824 |
| *B. Spouse attributes* | | | | | |
| Have spouse name score | 0.974 | 0.975 | 0.974 | 0.978 | 0.980 |
| Jewish spouse | 0.121 | 0.051 | 0.144 | 0.039 | 0.045 |
| Catholic spouse | 0.083 | 0.074 | 0.086 | 0.061 | 0.058 |
| Colonial spouse | 0.248 | 0.325 | 0.222 | 0.372 | 0.410 |
| Colonial marriage | 0.100 | 0.162 | 0.079 | 0.195 | 0.222 |
| N | 12597 | 3061 | 9536 | 1830 | 858 |

Means of row variables in samples defined by the column. Sample is restricted to Harvard students matched to Class Report records. Panel B restricts to students who we identify as married. See B.8.4 for details.

Table B.27: Spouse attributes and adult career and social attainment

| | (1) Finance | (2) Finance | (3) Social club | (4) Social club | (5) Country club | (6) Country club |
|---|---|---|---|---|---|---|
| Colonial | 0.027 | 0.026 | 0.013 | 0.009 | 0.020 | 0.015 |
| | (0.007) | (0.008) | (0.010) | (0.011) | (0.009) | (0.010) |
| Final club | 0.126 | 0.124 | 0.205 | 0.189 | 0.199 | 0.185 |
| | (0.017) | (0.017) | (0.019) | (0.018) | (0.019) | (0.019) |
| Private feeder | 0.066 | 0.065 | 0.107 | 0.104 | 0.125 | 0.122 |
| | (0.009) | (0.009) | (0.011) | (0.011) | (0.011) | (0.010) |
| Married | | 0.006 | | 0.160 | | 0.119 |
| | | (0.007) | | (0.010) | | (0.008) |
| Colonial spouse | | 0.020 | | 0.027 | | 0.031 |
| | | (0.010) | | (0.015) | | (0.014) |
| Colonial marriage | | -0.003 | | 0.002 | | 0.008 |
| | | (0.017) | | (0.024) | | (0.022) |
| Observations | 10319 | 10319 | 11733 | 11733 | 11733 | 11733 |

OLS regressions of outcome listed in column on own and spouse attributes. "Finance" is indicator for listing a finance occupation in the class reports. "Social club" and "country club" are indicators for listing participation in an adult social organization of the given type in Class Report data. "Colonial" is an indicator equal to one if the individual has a high value of the Colonial last name index. "Colonial spouse" is an indicator equal to one if the indiviudal is married and the spouse has a high value of the Colonial last name index. Married is an indicator equal to one if the individual is married. Marriage data come from Class Reports. All regressions include cohort fixed effects. Sample is individuals for whom own Colonial name indices and final club membership data are non-missing. Spouses for whom indices could not be calculated are included in sample and designated as non-Colonial. See B.8.4 for details. Robust standard errors in parentheses.

Table B.28: College peer effects on spouse attributes

| | All | Private | Non-private | Colonial | Non-Colonial |
|---|---|---|---|---|---|
| Final club | 0.065 | 0.167 | 0.001 | 0.123 | 0.034 |
| | (0.021) | (0.056) | (0.017) | (0.044) | (0.021) |
| Married | 0.009 | 0.070 | -0.033 | 0.133 | -0.047 |
| | (0.032) | (0.058) | (0.039) | (0.057) | (0.040) |
| Colonial spouse | 0.055 | 0.092 | 0.048 | 0.163 | 0.008 |
| | (0.030) | (0.060) | (0.035) | (0.059) | (0.034) |
| Colonial marriage | 0.048 | 0.096 | 0.035 | 0.163 | 0.000 |
| | (0.022) | (0.046) | (0.025) | (0.059) | (0.000) |
| N | 8178 | 2476 | 5551 | 2625 | 5294 |

Coefficients on peer neighborhood price rank from regressions of form given in equation 2. Columns are samples. Rows are outcome variables. All specifications include randomization block and dummies for large feeder high schools; see section 5.1 for details. Column 1 is full sample, and columns 2 and 3 are private feeder and non-private feeder samples. Column 4 is the sample of individuals with Colonial last names and column 5 is the sample of individuals with non-Colonial last names. "Final club" row reproduces the selective final club outcome from the main text. "Married" is an indicator equal to one for reporting being married in the Class Report. "Colonial spouse" is an indicator for being married and having a spouse with a Colonial last name. "Colonial marriage" is an indicator equal to one if the individual has Colonial last name and reports a marriage to a spouse with a Colonial last name. See section B.8.4 for details.

Table B.29: Additional codes for occupations used in long-run time series

| Category | Subcategory | Identifiers |
|---|---|---|
| Financial | Investment | Capital Management, Financing, Fund Management, Fund Manager, Investing, Investor, Private Equity, Security Analyst, Trader, Trading, Venture Capital |
| | Banking | Lending |
| | Firms (ext. definition) | Barclays, Bear Stearns, Berkshire, BMO Capital Markets, BNP Paribas, Carl M Loeb, Citigroup, Clark Dodge, Credit Suisse, Deutsche Bank, E.F.Hutton, Evercore Partners, Goldman Sachs, HSBC, J.P. Morgan, Jefferies Group, John P. Chase, Kidder Peabody, Lazard, Lehman Brothers, Loeb Rhoades, Manhattan Chase, Merrill Lynch, Mizuho, Morgan Chase, Morgan Stanley, Nomura, Perice Fenner Beane, RBC Capital, S. F. Moseley, Scudder Stevens Clark, Solomon Brothers, UBS, Upham & Company, Wainwright & Company, Welch Forbes, Wells Fargo, White Weld |
| Law | Law | Justice, Litigation, Litigator, Solicitor, Supreme Court |
| Medical | Doctor | Cardiologist, Dermatologist, Gastroenterologist, Gynecology, M.D., Medical Director, Medical Officer, Neonatologist, Neurology, Nurse, Oncologist, Ophthalmolology, Opodiatrist, Optometrist, Otolaryngologist, Pathologist, Podiatrist, Pulmonologist, Radiologist, Rheumatologist |
| | Psychiatry | Mental Health Center, Psychological, Psychotherapist |
| Academics/Research | Professor | College Teacher, Faculty, Institute Of Technology, Lecturer, Senior Tutor, University Teacher |

Table B.30: Additional codes of common high schools used in the long-run time series

| School Name | Private Feeder | Private | Public feeder |
|---|---|---|---|
| Baltimore City College HS | 0 | 0 | 1 |
| Bellaire High School | 0 | 0 | 1 |
| Belmont High School | 0 | 0 | 1 |
| Berkeley High School | 0 | 0 | 1 |
| Berkeley Prep | 0 | 1 | 0 |
| Bethesday-Chevy Chase | 0 | 0 | 1 |
| Beverly Hills High School | 0 | 0 | 1 |
| Braintree High School | 0 | 0 | 1 |
| Brighton High School | 0 | 0 | 1 |
| Brooklyn Technical High School | 0 | 0 | 1 |
| Cambridge School of Weston | 0 | 1 | 0 |
| Canterbury School | 0 | 1 | 0 |
| Collegiate School | 0 | 1 | 0 |
| Concord Academy | 0 | 1 | 0 |
| Concord High School | 0 | 0 | 1 |
| Concord-Carlisle | 0 | 0 | 1 |
| Cranbrook School | 0 | 1 | 0 |
| Delbarton School | 0 | 1 | 0 |
| Detroit Country Day | 0 | 1 | 0 |
| Edgemont High School | 0 | 0 | 1 |
| Eton College | 0 | 1 | 0 |
| George School, Pa. | 0 | 1 | 0 |
| Germantown Friends School | 0 | 1 | 0 |
| Gilman School | 0 | 1 | 0 |
| Great Neck North | 0 | 0 | 1 |
| Great Neck South | 0 | 0 | 1 |
| Greenwich High School | 0 | 0 | 1 |
| Harvard-Westlake | 0 | 0 | 1 |
| Henry M. Gunn High School | 0 | 0 | 1 |
| Highland Park High School | 0 | 0 | 1 |
| Hopkins School | 0 | 1 | 0 |
| Horace Greeley High School | 0 | 0 | 1 |
| Hunter College High School | 0 | 0 | 1 |
| Iolani School | 0 | 1 | 0 |
| La Jolla High School | 0 | 0 | 1 |
| Lakeside School | 0 | 1 | 0 |
| Lincoln-Sudbury | 0 | 0 | 1 |
| Lynbrook High School | 0 | 0 | 1 |
| Maimonides School | 0 | 1 | 0 |
| Mamaroneck High School | 0 | 0 | 1 |
| Milton High School | 0 | 0 | 1 |
| National Cathedral School | 0 | 1 | 0 |
| Needham High School | 0 | 0 | 1 |
| New Rochelle High School | 0 | 0 | 1 |
| Newton High School | 0 | 0 | 1 |
| Newton North | 0 | 0 | 1 |
| Newton South | 0 | 0 | 1 |
| Nichols School | 0 | 1 | 0 |

Continued on next page

**Table B.30 – continued from previous page**

| School Name | Private Feeder | Private | Public feeder |
|---|---|---|---|
| Palo Alto High School | 0 | 0 | 1 |
| Paul D. Schreiber | 0 | 0 | 1 |
| Pine Crest School | 0 | 1 | 0 |
| Polytechnic School | 0 | 1 | 0 |
| Princeton Day School | 0 | 1 | 0 |
| Princeton High School | 0 | 0 | 1 |
| Punahou School | 0 | 1 | 0 |
| Putney School | 0 | 1 | 0 |
| Regis High School | 0 | 1 | 0 |
| Ridgewood High School | 0 | 0 | 1 |
| Rye Country Day | 0 | 1 | 0 |
| San Francisco University High School | 0 | 1 | 0 |
| San Marino High School | 0 | 0 | 1 |
| Shaker Heights High School | 0 | 0 | 1 |
| Sidwell Friends School | 0 | 1 | 0 |
| St. Albans | 0 | 1 | 0 |
| St. Andrew's School | 0 | 1 | 0 |
| St. John's Prep | 0 | 1 | 0 |
| St. John's School | 0 | 1 | 0 |
| Staples High School | 0 | 0 | 1 |
| The Brearley School | 0 | 1 | 0 |
| The Chapin School | 0 | 1 | 0 |
| The College Preparatory School | 0 | 1 | 0 |
| The Dalton School | 0 | 1 | 0 |
| The Pingry School | 0 | 1 | 0 |
| The Winsor School | 0 | 1 | 0 |
| Thomas Jefferson | 0 | 0 | 1 |
| Torrey Pines High School | 0 | 0 | 1 |
| Trinity School | 0 | 1 | 0 |
| Troy High Schoo | 0 | 0 | 1 |
| Upper Canada College | 0 | 1 | 0 |
| Walt Whitman High School | 0 | 0 | 1 |
| Wayland High School | 0 | 0 | 1 |
| Westfield High School | 0 | 0 | 1 |
| Weston High School | 0 | 0 | 1 |
| Westwood High School | 0 | 0 | 1 |
| Weymouth High School | 0 | 0 | 1 |
| White Plains High School | 0 | 0 | 1 |
| Winston Churchill High School | 0 | 0 | 1 |
| Wyoming Seminary | 0 | 1 | 0 |
| Xaverian Brothers High School | 0 | 1 | 0 |
| Xaverian High School | 0 | 1 | 0 |

Table B.31: Sports participation by high school background

| | All | Private feeder | Public feeder | Other HS |
|---|---|---|---|---|
| *A. Sport by competition level* | | | | |
| Any Sport | 0.367 | 0.552 | 0.230 | 0.340 |
| Schoolwide | 0.327 | 0.506 | 0.200 | 0.297 |
| Intramural/Dormitory | 0.065 | 0.089 | 0.040 | 0.065 |
| *B. Sport by type* | | | | |
| Crew/Rowing | 0.071 | 0.133 | 0.025 | 0.060 |
| Track | 0.059 | 0.091 | 0.048 | 0.049 |
| Football | 0.051 | 0.108 | 0.020 | 0.038 |
| Baseball | 0.034 | 0.059 | 0.025 | 0.025 |
| Basketball | 0.021 | 0.016 | 0.009 | 0.028 |
| Other sport | 0.164 | 0.212 | 0.113 | 0.165 |

Participation in sports by high school type and type of sport. Cells are means of row variables (all indicators) in sample given by column heading. Panel A displays participation in sports by type. "Schoolwide sports" are those where a person competes for Harvard against another institution; i.e. intercollegiate sports. "Intramural/dormitory" sports are within-Harvard competitions. See section B.10.2 for details.

Table B.32: Sports participation premia for academic, social, and labor market outcomes

| | (1) Class rank | (2) Sel. fin. club | (3) Wage inc. | (4) Finance | (5) Wage inc. | (6) Finance | (7) Wage inc. | (8) Finance |
|---|---|---|---|---|---|---|---|---|
| Schoolwide Sport | -0.22 | 0.089 | 248.7 | 0.042 | 129.7 | 0.021 | | |
| | (0.029) | (0.0064) | (45.6) | (0.0083) | (47.1) | (0.0083) | | |
| Private feeder | -0.36 | 0.19 | 203.2 | 0.096 | 19.2 | 0.063 | 15.2 | 0.063 |
| | (0.031) | (0.0091) | (52.4) | (0.010) | (56.0) | (0.011) | (55.9) | (0.011) |
| Sel. fin. club | | | | | 348.4 | 0.042 | 364.0 | 0.044 |
| | | | | | (109.3) | (0.025) | (110.3) | (0.025) |
| Hasty Pudding | | | | | 339.4 | 0.053 | 326.0 | 0.054 |
| | | | | | (83.5) | (0.017) | (83.1) | (0.017) |
| Social | | | | | 189.4 | 0.062 | 145.1 | 0.061 |
| | | | | | (87.5) | (0.019) | (90.5) | (0.019) |
| Crew/Rowing | | | | | | | 56.6 | 0.0029 |
| | | | | | | | (85.8) | (0.016) |
| Track | | | | | | | 237.9 | 0.016 |
| | | | | | | | (82.1) | (0.016) |
| Football | | | | | | | 208.4 | 0.0034 |
| | | | | | | | (100.0) | (0.022) |
| Baseball | | | | | | | 93.3 | 0.048 |
| | | | | | | | (115.1) | (0.025) |
| Basketball | | | | | | | 270.2 | 0.014 |
| | | | | | | | (130.5) | (0.025) |
| Other sport | | | | | | | 156.8 | 0.016 |
| | | | | | | | (59.2) | (0.011) |
| Observations | 6931 | 8628 | 4046 | 7287 | 4046 | 7287 | 4046 | 7287 |

Linear regressions of outcome listed in column header on variables listed in rows. All specifications also include cohort fixed effects. "Class rank": reverse-signed class rank, so that higher values correspond to better academic performance. "Sel. fin. club": membership in selective final club. "Finance": reports finance occupation on Class Report. "Wage inc.": wage income from the 1940 Census. Sample size varies across specifications due to data availability. Census specifications restrict to entering cohorts in 1930 and earlier. See section B.10.2 for details. Robust standard errors in parentheses.