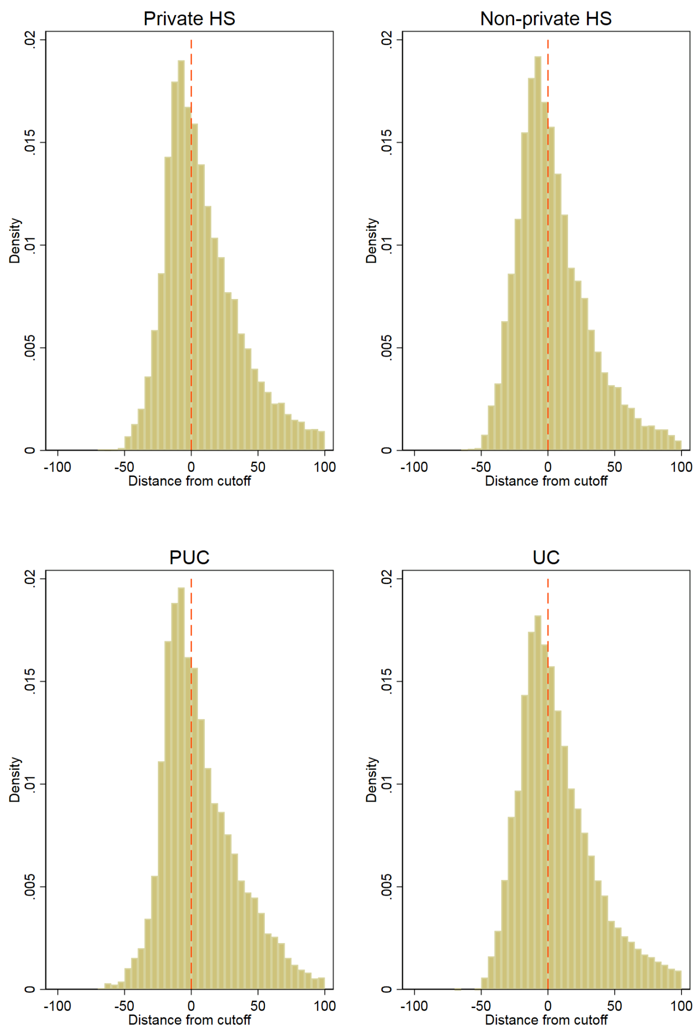
Online Appendix to 'Elite Colleges and Upward Mobility into Top Jobs and Top Incomes'
Seth Zimmerman

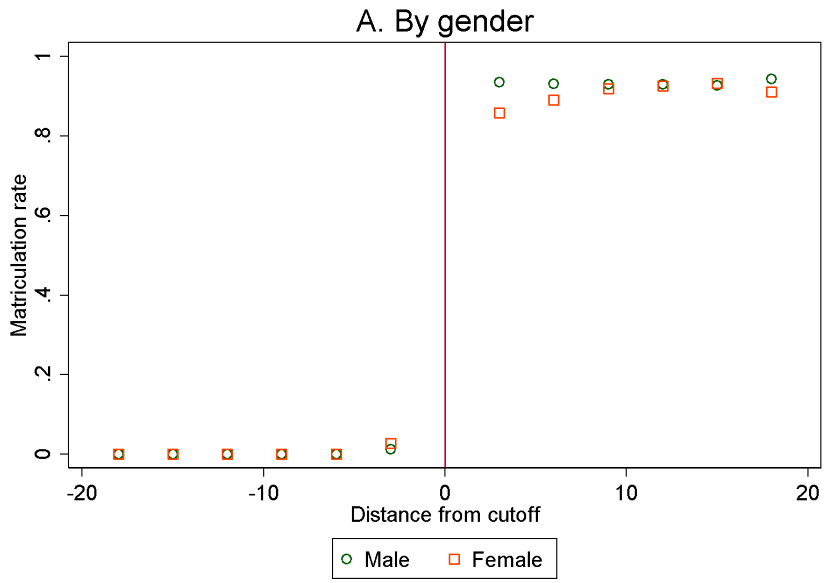
A Additional tables and figures

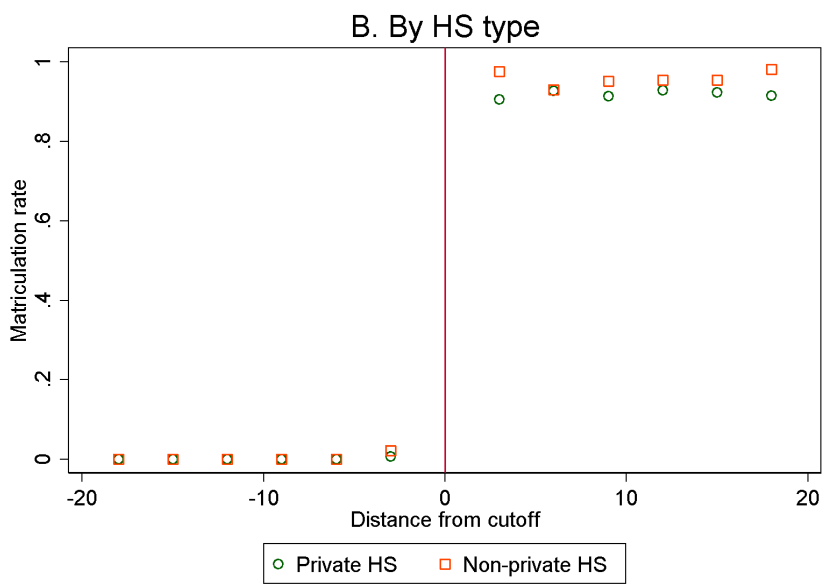
Figure A-1: Histogram of running variable by HS type and institution



Histograms of the running variable by high school type and institution. Upper panel: private vs. non-private high schools, pooled over institutions. Lower panel: PUC vs. UC applicants, pooled over high school types. Densities computed within 5 point bins.

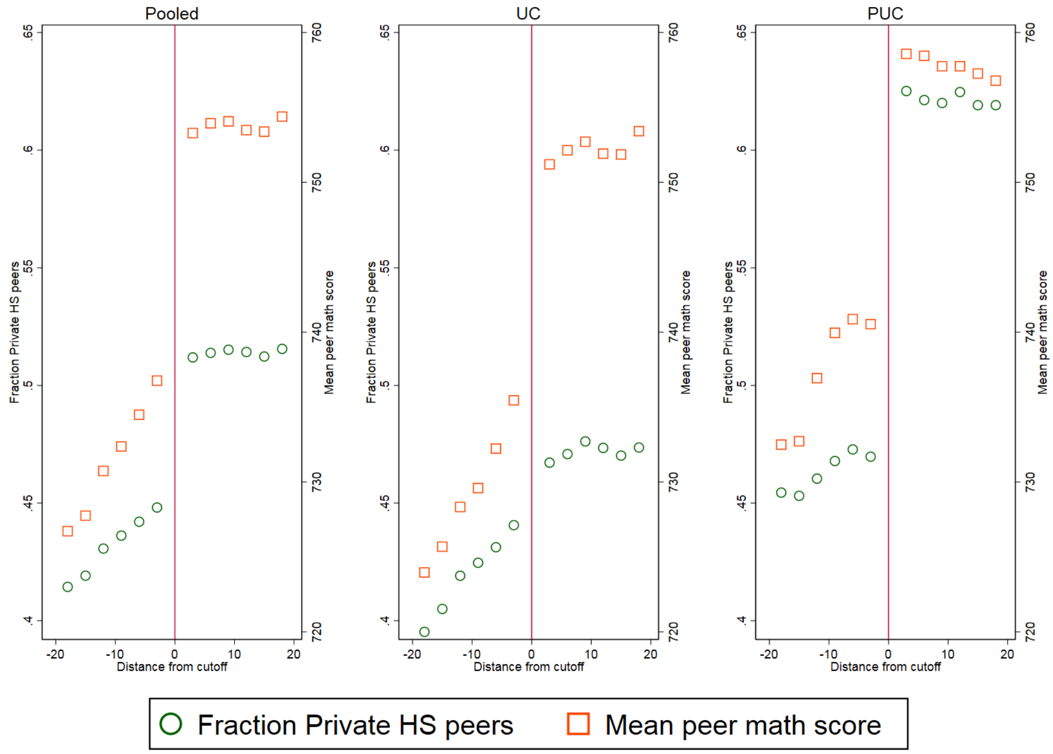
Figure A-2: Matriculation by gender and high school type





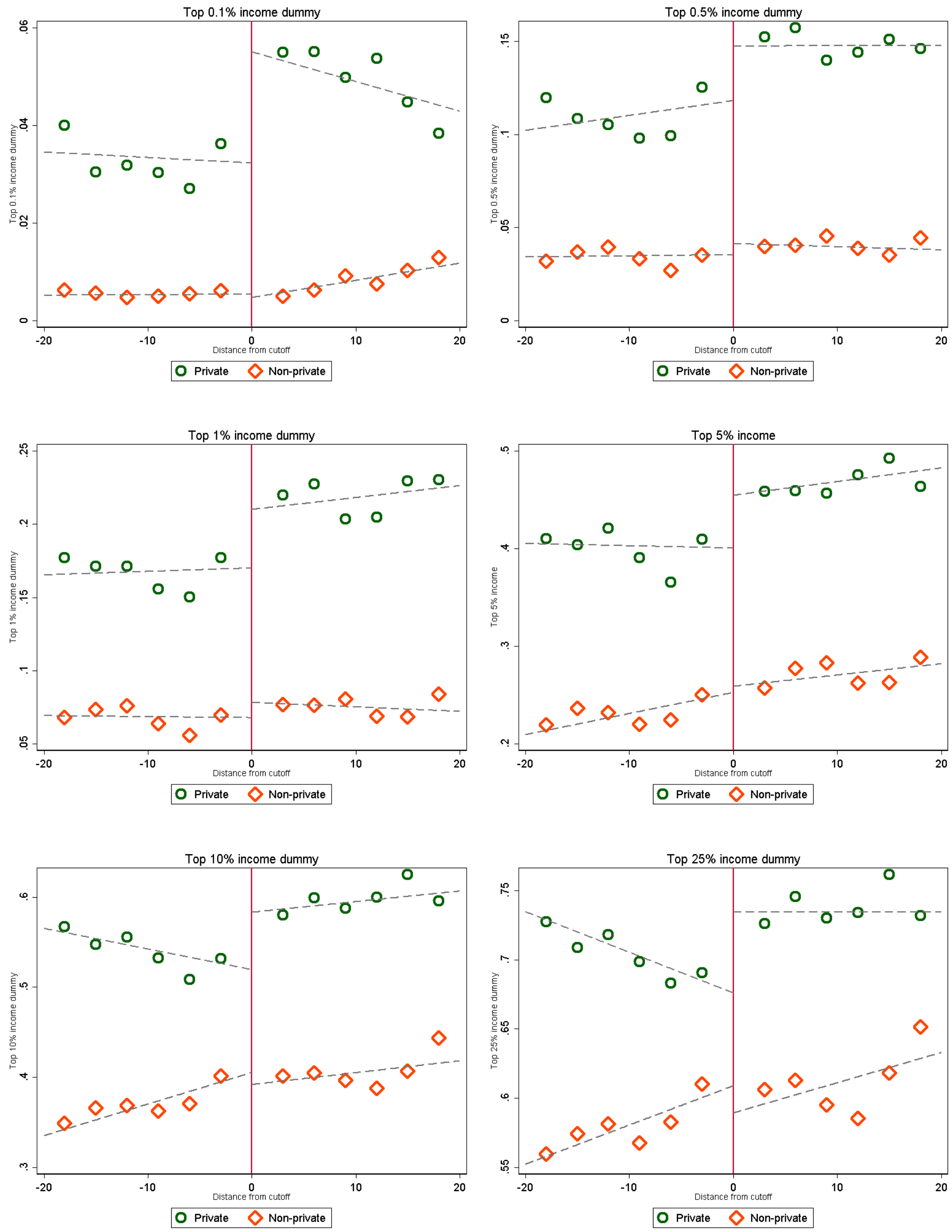
Matriculation to target degree program by gender and position relative to admissions threshold. Sample: 2004 applicants to elite degree programs. Points are mean values within bins of width three on either side.

Figure A-3: Changes in peer characteristics across the admissions threshold



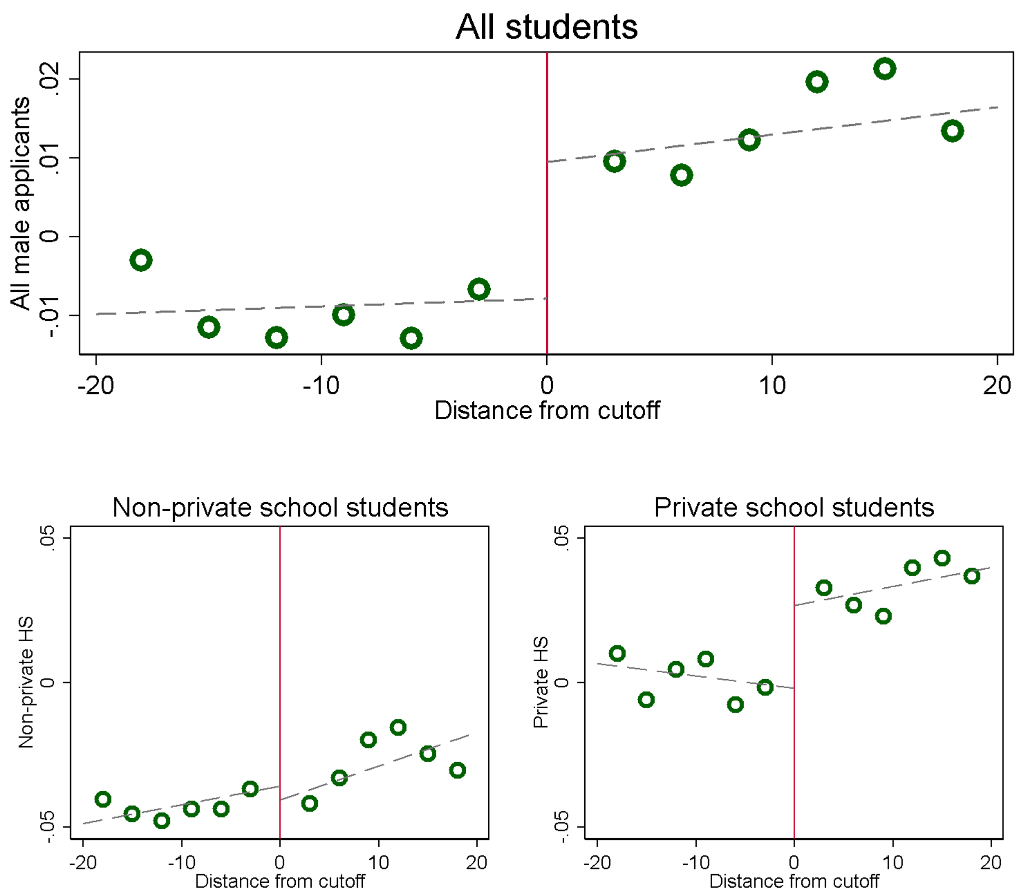
Changes in the fraction of students from private high schools and the mean peer math score at the degree programs to which students are admitted by position relative to threshold. Admissions outcomes include those realized in any application year. Results are similar to Figure 5, which includes only the same-year admissions outcomes. Points are binned means within centered three point windows. Left panel pools across UC and PUC programs. Center and right panel split applications to UC and PUC programs. Fraction HS peers match to left axis; mean test scores to right axis.

Figure A-4: Regression discontinuity plots by top income category



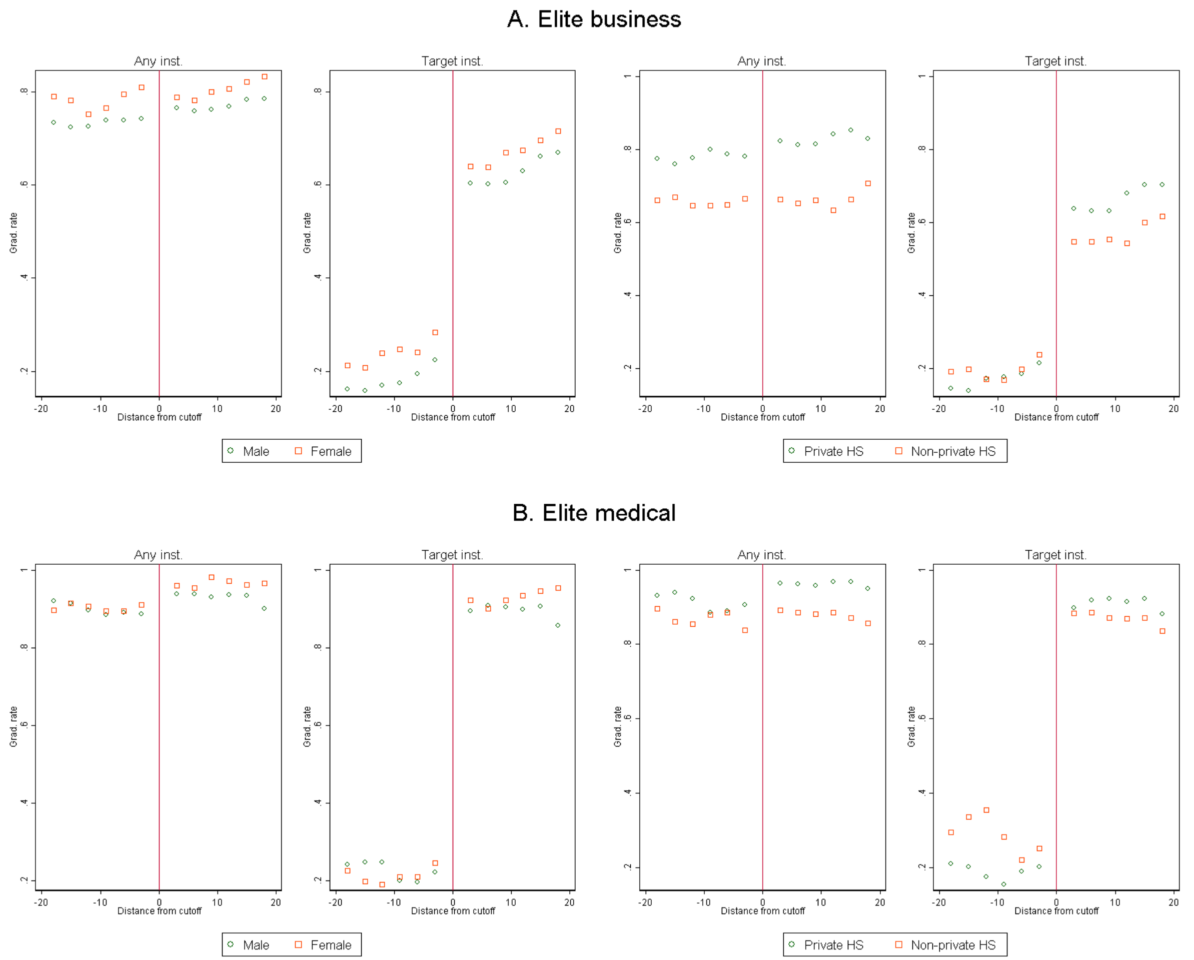
Fraction of male students with incomes falling into different top income groups by position relative to the threshold and high school type. Graphs pool applications across elite degree programs. Points reflect average outcomes for applicants within three points on either side of the horizontal axis value. Fitted values from BW=20 specification.

Figure A-5: Leadership RD graphs with residualized outcomes



Residualized count of leadership positions by position relative to the threshold for all men (upper panel) and by high school type (lower panel). Residuals computed by regressing leadership outcomes on indicators for target program (i.e., the school to which students are applying, whether or not they are admitted) and application year. Graphs pool applications across elite degree programs. Points reflect average outcomes for applicants within three points on either side of the horizontal axis value. Fitted values from BW=20 specification. Compare to Figure 6.A, which produces the same figures but without residualizing first.

Figure A-6: Completion rates by gender, high school type, and degree type



10 year graduation rates at any institution ('any inst' panels) and target institution ('target inst' panels), position relative to threhsold, and degree type. Left two graphs in each panel split by gender, right two graphs split by high school type among male students. Sample: 2000-2005 application cohorts, with graduation outcomes observed between 2000 and 2015. Points are mean values within bins of width three on either side. See section 5.5 for details

Table A-1: Chile in international context

Country	GDP	Top share	Bus. Rank	Informal payments	Tertiary completion	OECD
Panel A: 2014						
Chile	11.8	41.5	48	0.7	38	Yes
United States	55.3	30.2	7		42	Yes
Mexico	10.2	38.9	38	11.6	22	Yes
Argentina	9.1	30.6	121	18.1		No
Brazil	7.1	41.8	116	11.9	12	No
Italy	34.2	26.3	45		21	Yes
Poland	13.4	25.6	25	14.7	37	Yes
Panel B: 1980						
Chile	4				19	
United States	31.1				41	
Mexico	8.1				12	
Argentina	6.3					
Brazil	5.1				9	
Italy	25				11	
Poland					13	

Cross-national comparison of educational and economic aggregates. All data from World Bank (2016) except tertiary completion, which is from OECD (2012; Table A1.3a). GDP is per capita and is from 2014 (upper panel) and 1980 (lower panel) in 1000s of 2014 US dollars. 'Top share' is the percentage of income earned by the top 10 percent of the income distribution. Data from 2012 and 2013. 'Bus. Rank' is the World Bank Ease of Doing Business Index; countries are ordered from 1 to 189 with a top rank of 1. Data from 2015. 'Informal payments' reports the percentage of businesses reporting informal payments to government officials. Data from 2009-2010. Tertiary completion for 2012 panel is given by completion rates for 25-34 y.o. in 2010. Tertiary completion for 1980 panel is given by completion rates for 55-64 y.o. in 2010. 'OECD' column reports OECD membership as of 2016.

Table A-2: Effects of admission on matriculation

	Matriculate at target	Matriculate anywhere
All	0.912	0.219
	(0.011)	(0.021)
	1317	1317
Male	0.928	0.223
	(0.013)	(0.025)
	878	878
Female	0.877	0.206
	(0.024)	(0.038)
	439	439
Male v. female	0.055	0.702
Private HS	0.914	0.252
	(0.018)	(0.034)
	522	522
Non-private HS	0.949	0.163
	(0.018)	(0.038)
	316	316
Private v. non-private	0.156	0.080

Effects of admission on matriculation. Upper row in each cell: point estimate. Middle row: standard errors in parentheses. Lower row: N. Effects of elite admission on matriculation and graduation outcomes. All estimates from BW=10 specification. 'Matriculate at target' column is the effect of admission to an elite degree program on matriculation to that degree program. 'Matriculate any' column is effect of admission to an elite degree program on matriculation to any degree program in the centralized system. Rows denote samples. 'Male vs. female' and 'Private vs. non-private' rows display p-values from tests of null hypothesis of equal effects. Matriculation sample is from 2004 admissions cohort.

Table A-3: Effect of threshold-crossing on co-leadership rates

	Peers	Same program	Same cohort
A. All private high schools			
Point estimate	4.02	-0.19	0.09
Standard error	1.95	0.62	1.14
p-value	0.039	0.761	0.939
Intercept	2.12	2.37	1.75
N	1411903	17865662	1163411
B. Elite private high schools			
Point estimate	22.51	-2.24	1.16
Standard error	11.35	3.85	4.15
p-value	0.047	0.561	0.779
Intercept	3.05	10.86	2.10
N	167118	1986229	157736

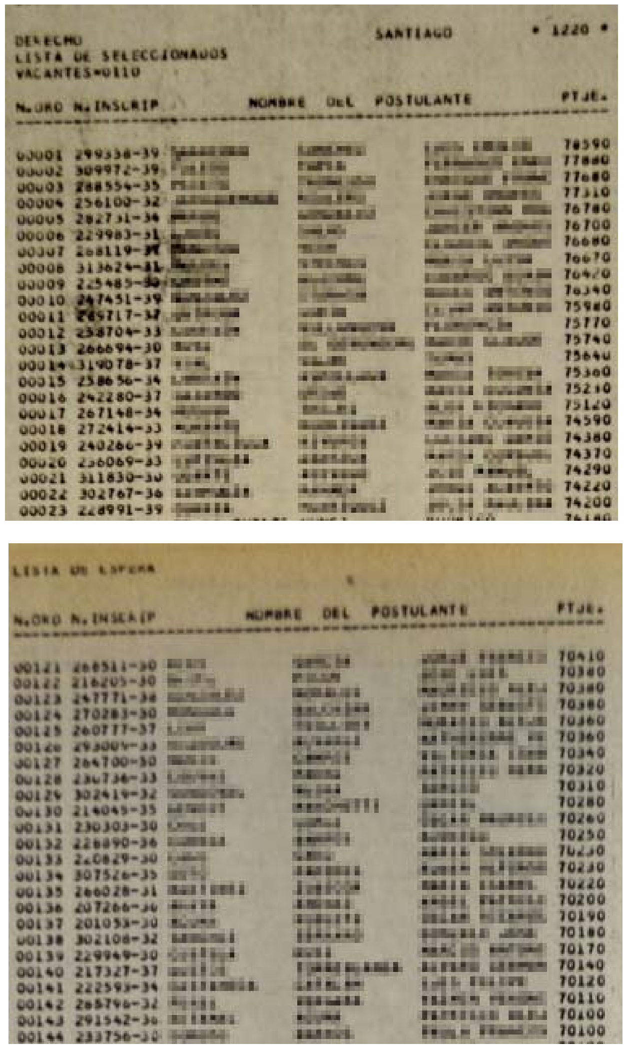
Estimates of effect of admission on co-leadership rates per 100,000 pairs in sample of male students. See Section 6 for details. Left column is for peers in same cohort and same program. Middle column for non-peers in same year but different program. Left column is for non-peers in same program but different year. Sample includes only pairs of students in the same field. Panel A consists of pairs of students where both are from any private high school, Panel B where both are from an elite private high school. Specifications include separate slope terms above and below the threshold. p-values report tests of null that threshold-crossing effect is equal to zero. N refers to pairs. Standard errors clustered at person-person level. 'Intercept' row reports below-threshold mean co-leadership rate.

B Data construction

B.1 Elite application records

Application records were digitized from newspaper records of application outcomes. The Chilean newspaper *El Mercurio* prints admission and waitlist outcomes for each CRUCH degree program in each year. Figure B-1 provides an example of admissions and waitlist records for the PUC law program in 1984. Worth noting is that although there were only 110 spots open in the program this year ('vacantes=110' in the upper panel), 120 students were admitted (the first rejected student in the lower panel has a rank of 121). This allows for some students to turn down admissions offers without waitlist movement. This feature of the process a mitigates concerns about the endogeneity of offer counts that arise in some school choice settings (de Chaisemartin and Behaghel 2015).

Figure B-1: PUC Law admissions and waitlist announcements, 1984



DERECHO
LISTA DE SELECCIONADOS
VACANTES=0110

SANTIAGO • 1220 •

N.ORD	N.INSCRIP	NOMBRE DEL POSTULANTE	PTJE.
00001	299338-39	[illegible]	78590
00002	309972-39	[illegible]	77880
00003	288554-35	[illegible]	77680
00004	256100-32	[illegible]	77310
00005	282731-34	[illegible]	76780
00006	229983-31	[illegible]	76700
00007	268119-[illegible]	[illegible]	76680
00008	313624-31	[illegible]	76670
00009	225485-30	[illegible]	76420
00010	247451-39	[illegible]	76340
00011	245717-37	[illegible]	75980
00012	238704-33	[illegible]	75770
00013	266694-30	[illegible]	75740
00014	319078-37	[illegible]	75640
00015	238656-34	[illegible]	75300
00016	242280-37	[illegible]	75210
00017	267148-34	[illegible]	75120
00018	272414-33	[illegible]	74590
00019	240266-34	[illegible]	74380
00020	236069-33	[illegible]	74370
00021	311830-30	[illegible]	74290
00022	302767-36	[illegible]	74220
00023	228991-39	[illegible]	74200

LISTA DE ESPERA

N.ORD	N.INSCRIP	NOMBRE DEL POSTULANTE	PTJE.
00121	268511-30	[illegible]	70410
00122	216205-30	[illegible]	70380
00123	247771-34	[illegible]	70380
00124	270283-30	[illegible]	70380
00125	260777-37	[illegible]	70360
00126	293009-33	[illegible]	70360
00127	264700-30	[illegible]	70340
00128	236736-33	[illegible]	70320
00129	302419-32	[illegible]	70310
00130	214045-35	[illegible]	70280
00131	230303-30	[illegible]	70260
00132	226890-36	[illegible]	70250
00133	240829-30	[illegible]	70230
00134	307526-35	[illegible]	70230
00135	266028-31	[illegible]	70220
00136	207266-30	[illegible]	70200
00137	201053-30	[illegible]	70190
00138	302108-32	[illegible]	70180
00139	229949-30	[illegible]	70170
00140	217327-37	[illegible]	70140
00141	222593-34	[illegible]	70120
00142	268796-32	[illegible]	70110
00143	291542-36	[illegible]	70100
00144	233756-30	[illegible]	70100

Upper panel is the beginning of the list of admitted students, ordered by score. Following this list the newspaper prints the list of waitlisted students, also ordered by score. Though these data are publicly available, names are blurred to preserve confidentiality. Source: *El Mercurio*, March 1984.

I use data only in years for which newspaper records could be located in the Biblioteca Nacional de Chile. Application records are missing for some programs in some years. Table B-1 describes data availability for each of the six elite degree programs by year, from 1974 through 2001. Out of 168 total program-year cells between 1974 and 2001, I have data on 153. The regression discontinuity analysis uses data on students who applied in 1991 or earlier. Over this period I have data on 97 out of a possible 108 program-year combinations.

Table B-1: Data availability by year

	PUC			UChile		
	Law	Eng.	Bus.	Law	Eng.	Bus.
1974	0	0	0	1	1	1
1975	0	0	0	0	0	1
1976	1	1	1	1	1	1
1977	1	1	1	1	1	1
1978	1	1	1	1	1	1
1979	1	1	1	1	1	1
1980	1	1	1	1	1	1
1981	1	1	1	1	1	1
1982	1	1	1	1	1	1
1983	1	1	1	1	1	1
1984	1	1	1	1	1	1
1985	1	1	1	1	1	1
1986	1	1	1	1	1	1
1987	0	0	0	1	1	1
1988	1	1	1	1	1	1
1989	1	1	1	1	1	1
1990	1	1	1	1	1	1
1991	1	1	1	1	1	1
1992	1	1	1	1	1	1
1993	1	1	1	1	1	1
1994	0	0	1	1	1	1
1995	1	1	1	0	0	1
1996	1	1	1	1	1	1
1997	1	1	1	1	1	1
1998	1	1	1	1	1	1
1999	1	1	1	1	1	1
2000	1	1	1	1	1	1
2001	1	1	1	1	1	1

Data availability by degree program and application year. 1 indicates data is available for a program-year cell. 0 indicates it is not.

Table 2 in the main text describes the set of applicants who are matched to national identification numbers. This is the sample used in the discontinuity analyisis. The upper panel of Table B-2 describes the match from the raw application records to the identification numbers. Over the full sample of years 1974-2001, I observe 100,060 applications and match 94,004, or 94%, to identifiers. Beginning in 1989, these identifiers were printed in the newspaper alongside application results. For the years 1974 through 1988, I obtain the identifiers by matching printed results to administrative application records based on a unique application identifier. Match rates over the 1974-1991 period used for the discontinuity analysis are 90.6%. They are close to one in the post-1991 period.

Table B-2: Merge rates to identifiers

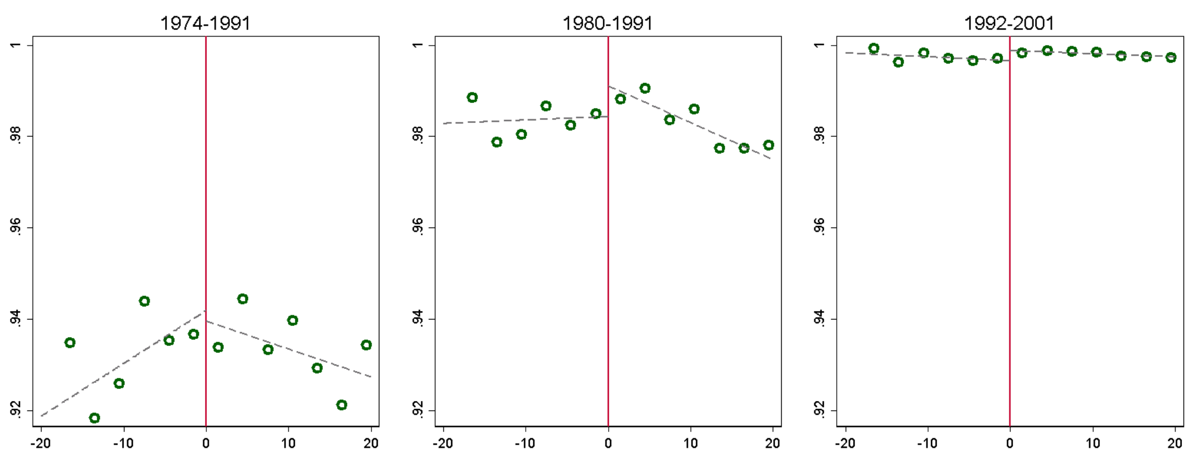
A. Applications	Original data	ID match	Match share
All years	100060	94004	0.939
1974-1991	63922	57940	0.906
1992-2001	36138	36064	0.998

B. Application-outcome year data	Original data	Tax ID match	Income report	Income report share
Age $>=$ 30	541040	533685	444684	0.833
Age $>=$ 40	241668	235809	189672	0.804
$30 \leq$ Age < 40	299372	297876	255012	0.856

Cells are counts except for the 'share' column. Panel A: Observations at application level. 'Original data' counts all applications observed in newspaper records. 'Matched to ID' counts applications where I obtain an national ID record. Panel B: observations at application-outcome year record. There are up to 9 outcome years per application, corresponding to the set of years 2004-2013 for which an applicants was over age 30. See text for description of tax form variable.

Discontinuities in match rates at the admissions threshold would raise concerns for the regression discontinity analysis. Figure B-2 presents plots of match rates by position relative to the threshold for the 1974-1991, 1980-1991, and 1992-2001 periods. The first set of applications is used in the leadership analysis and the second for the income analysis. The third set is included for completeness. There is no evidence of discontinuous changes in merge rates to person identifiers across the admissions threshold.

Figure B-2: Balance on merge rates to identifiers



Rates of merge to matched national identifiers by position relative to admissions cutoff. Points are means within centered bins of width three. Fitted lines are from BW=20 specification.

The lower panel of Table B-2 describes match rates to tax data. Observations here are at the application-outcome year level, as in the regression discontinuity analysis of top incomes. Of the 541,040 possible income observations, 533,685 (82.6%) have ID variables that are confirmed to be valid inside the tax authority, and 444,684 (83.3%) are matched to the labor force sample in a given year. For student-year observations corresponding to ages 40 and over (those used in the RD analysis), the match rate to income reports is 80.4% labor force participation rate, as reported reported in the upper panel of Table 2. Recall from above that the labor force participation rate was smooth across the threshold, and also that labor force non-participants are included in the top income analysis (non-participants do not have top incomes).

B.2 Categorizing high schools

This subsection discusses the classification of students into groups by high school type. The majority of this paper distinguishes between two types of high schools: private and non-private high schools. I use the private high school dummy as a proxy for coming from a high SES background. I obtain high school data using administrative application records that contain numeric high school codes. A challenge in mapping from codes to high school types is that the numeric codes change each year and mappings from codes to high school names or type indicators are not available prior to 2000. I address this challenge using the following procedure. First, I use students who apply to college in multiple years to create a set of codes that are consistent across years. I then use data on school type from 2000 to classify schools from earlier cohorts. This procedure will work well to the extent that a) there are at least some multi-year applicants in each high school in each year, b) high school types are stable over time, and c) the set of high schools itself is stable over time. To the extent that the procedure falsely categorizes either private or public high schools, this will bias estimates of differences between the two groups downward, away from my findings of cross-type heterogeneity. Recall from Table 3 that both the probability of non-missing high school type data and the probability of being classified as having a private school background are continuous through the admissions threshold.

In Sections 4.8 and 6 I also consider finer high school classifications. I divide private high schools into an 'elite' category and a non-elite category. The elite category consists of seven schools: St. George's College, Colegio del Verbo Divino, the Grange School, Colegio Sagrados Corazones Manquehue, Colegio Tabancura, Colegio San Ignacio, and the Craighouse School. Each elite private school is located in or near Santiago and charges very high tuition.[27] Several are male only (recall that my analysis focuses only on male students). Admissions can be exclusive. For instance, applications for admission to the pre-kindergarten program at the Grange School require

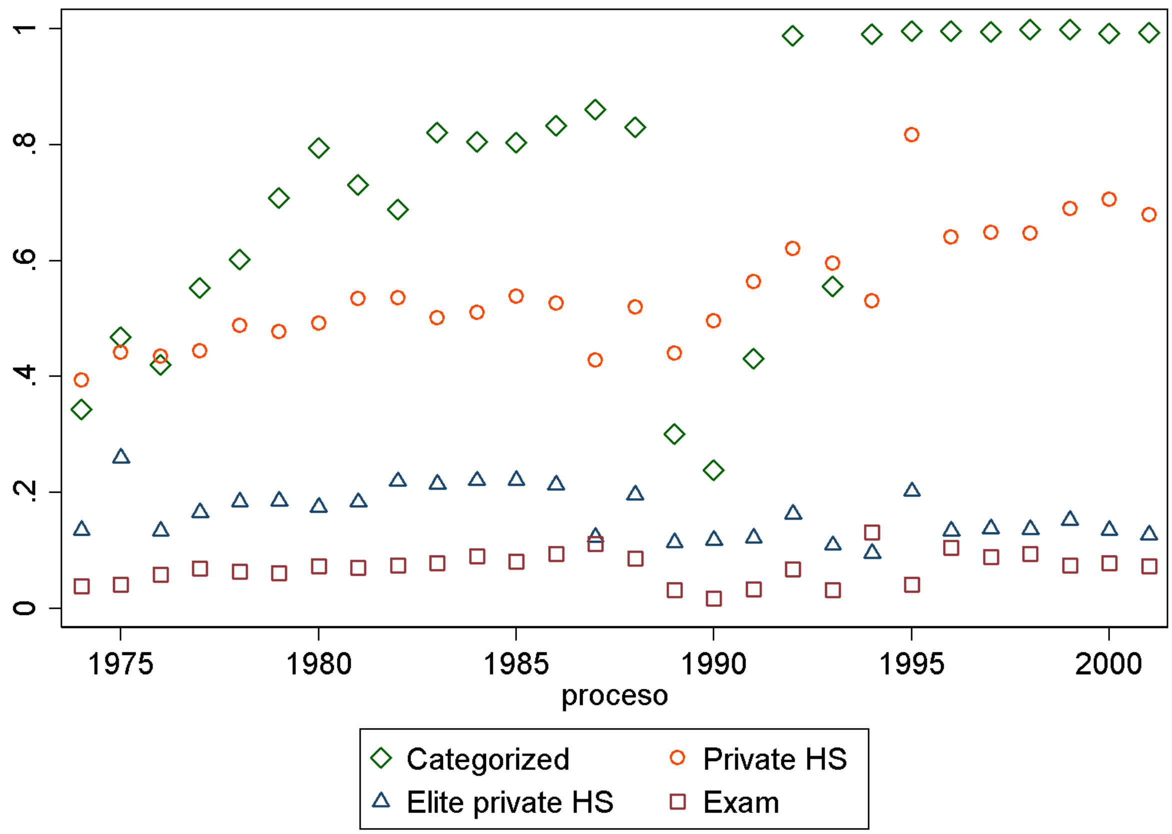
[27] As fraction of per capita GDP, tuition at these schools is similar to tuition at elite U.S. high schools like Deerfield or Phillips-Andover; see Neilson (2013).

a letter of reference from a member of the school community.[28] These schools appear frequently in press accounts and studies of the business elite.[29] Non-elite private high schools consist of all private schools not on this list. I also consider an elite public category that includes only the Instituto Nacional General José Miguel Carrera (henceforth the Instituto Nacional), an exam school located in Santiago. There is no tuition fee at the Instituto Nacional. However, as is the case with exam schools in the US, such as Stuyvesant or Bronx Science, admission depends on students' scores on an entrance exam. The Instituto Nacional is typically the only public school mentioned in studies of the Chilean business elite (SPI).

I evaluate the high school categorization in two ways. First, I describe patterns in successful categorization rates and the fraction of students from private high school backgrounds over time. Figure B-3 plots the fraction of male applicants to elite degree programs who are successfully assigned a high school type by application year, as well as the fraction of students assigned to different types of high schools. As expected given the matching procedure, the fraction of categorized students is higher in more recent years. It is close to one from 1994 through 2001. It is very low in the years 1989, 1990, 1991, and 1993. These are years for which even the year-specific high school codes are not available for all students due to data limitations. The match rate hovers around 80% from 1980 through 1988, then declines steadily to just under 40% from 1974 to 1979. Match rates decline in the 1970s because I can only consider cross-applicants applying to elite programs in multiple years, as opposed to students applying to all programs. Conditional on successful classification, the fraction of students from private high schools is more stable, declining steadily from just under 70% in 2001 to 40% in 1974. The fraction of students from elite private high schools hovers between roughly 10% and 20% for most of the period, and is generally slightly higher in the 1970s and early 1980s than in the 1990s. The fraction of students from the elite public exam school fluctuates between roughly 5% and 10% over the period.

[28] See `http://www.grange.cl/admissions`. Accessed 11/6/2013.
[29] See, e.g., Engel (2013), SPI.

Figure B-3: HS match rates and types over time

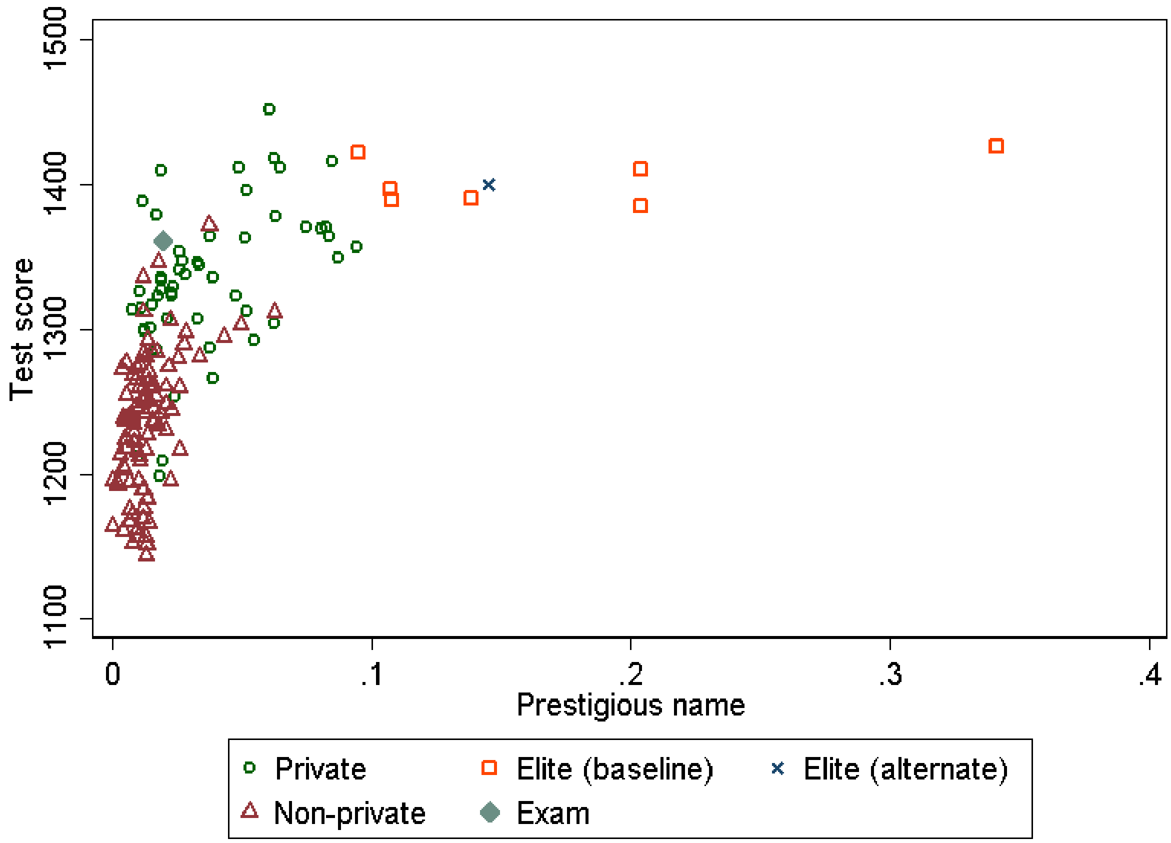


Rates of successful high school categorization and, conditional on categorization, high school type by application year. Sample is admitted male applicants at elite degree programs. 'Categorization' is dummy equal to one if a student can be assigned to a private/non-private high school category. 'Private HS', 'Elite private HS,' and 'Exam' are dummy variables equal to one if a student is assigned a high school of the listed type conditional on assignment. Types described in this section.

I next examine how students from different high school types differ in terms of other observable markers of socioeconomic status. I consider two alternate measures of socioeconomic status. The first is the sum of students' math and reading admissions scores. The second is a measure of prestige based on an applicant's two last names. Núñez and Miranda (2010) shows how European last names are a strong predictor of income in Chile. Here, I use a data driven measure that takes the ratio of the frequency with which a last name appears in the Chile section of 'Who's Who in Latin America' (Hilton, 1971) to the frequency with which it appears in the Chilean voter record file in 2012. The goal is to create a measure of 'name prestige' based on information predating the careers of applicants in my data. I create the index by taking the average index value over an applicant's last names, of which there are typically two.

Figure B-4 plots high schools by the fraction of individuals with name prestige indices within the top 1% of the voting population (horizontal axis) and their mean combined math and reading scores (vertical axis). I conduct this exercise within the sample of applicants to elite degree programs, so test scores are truncated from the bottom. Nevertheless, several patterns emerge. First, almost all private schools have higher average test scores than almost all public schools. Second, all of the schools with the highest fractions of students with prestigious names are private high schools. Third, within the set of private high schools, students from the high schools in the 'elite' category have among the highest test scores and all of the highest fractions of prestigious names. There is one high school not in the predetermined elite category that has comparably high test scores and prestige rankings to the elite high schools; it is marked with an 'x'. This is the Redland School, a private religious school in the wealthy Las Condes neighborhood of Santiago. Fourth, the elite public exam school has the highest average scores of any non-private high school.

Figure B-4: Test scores and name prestige by high school type



Each point represents a high school of listed type. Sample is applicants to elite degree programs 1974-2001. Horizontal axis is fraction of students from the listed high school with names in the top 1% of the name prestige index. See description in this section. The vertical axis is the mean of the sum of math and reading admissions test scores. See above text for description of high school types.

The conclusion from this analysis is that private high school attendance provides an imperfect but useful measure of student socioeconomic background. Within the set of private high schools, the 'elite private' designation is strongly correlated with alternate measures of student SES background. Similarly, the 'elite public exam' designation isolates the highest performing public school in the applicant dataset. Though the accuracy of these measures is likely lower for students in earlier application cohorts, categorization errors will create bias toward a finding of no heterogeneous effects.

B.3 Leadership data

Publicly traded companies in Chile are required to disclose the identities of top executives and board members to the Superintendicia de Valores y Seguros (SVS), the Chilean analogue to the Securities and Exchange Commission in the US. I obtain leadership data using a web scrape of the SVS website.[30] I conducted this scrape in March of 2013. The SVS website allows users to search historical filing records by date for each firm. I searched for all executive managers and directors who served between January 1st, 1975 and January 1st, 2013. Figure B-5 provides an example of the data recovered using this scrape. Records include national identifiers and names for directors and C-suite executives. They also include data on position title and date of appointment.

The search primarily yields records for current leadership teams, and most firms do not provide leadership records for the earlier part of this period. 10 percent of all corporate leaders in my sample began their leadership roles in 2012 or 2013, and the median leader was hired in 2009. 92 percent of leaders were hired in 1998 or later. A departure date is recorded for 47% of leaders in the full sample with a median departure year of 2011. Within the sample of leadership positions matched to elite applicants, the share of leaders with a departure date is 19%, with a median departure year of 2012. The large majority of leadership positions in the data reflect leadership teams at the time of collection or else leaders who departed within the previous two years.

Figure B-5: Example of data from web scrape

GERENTES, EJECUTIVOS PRINCIPALES

(Fecha Informe: 29/09/2013)

RazÃ³n Social: AES GENER S.A.
RUT: 94272000

Listado histÃ³rico Gerentes, Ejecutivos Principales de la Sociedad entre el 01/01/1975 y el 01/01/2013

Rut	Nombre	Cargo	Cargo Ejecutivo Principal	Fecha Nombramiento	Fecha termino
6.921.313-8	DANIEL STADELMANN ROJAS	Gerente General Subrogante		07/04/2011	
0-E (Extranjero)	MICHAEL WHITTLE -	Ejecutivo Principal	GERENTE DESAROLLO	26/01/2011	
12.458.775-1	IVAN JARA CARRASCO	Ejecutivo Principal	GERENTE DE INGENIERIA Y CONSTRUCCION	24/09/2010	
12.240.551-6	MARIANA PAZ SOTO ESPINOSA	Ejecutivo Principal	GERENTE ASUNTOS CORPORATIVOS	01/09/2010	
7.054.225-0	ALBERTO ZAVALA CAVADA	Ejecutivo Principal	FISCAL Y MANDATARIO JUDICIAL	24/05/2010	
23.202.311-2	VICENTE JAVIER GIORGIO	Ejecutivo Principal	GERENTE DE EXPLOTACION Y GERENTE GENERAL SUBROGANTE	26/05/2009	
6.921.313-8	DANIEL STADELMANN ROJAS	Ejecutivo Principal	GERENTE DE FINANZAS Y GERENTE GENERAL SUBROGANTE	25/02/2009	
6.375.799-3	LUIS FELIPE CERON CERON	Gerente General		29/08/2001	

Source: SVS filings. `http://www.svs.cl/sitio/mercados/consulta.php?mercado=V&entidad=RVEMI`. Accessed 9/29/2013.

[30] `http://www.svs.cl/sitio/mercados/consulta.php?mercado=V&entidad=RVEMI`. Accessed 9/23/2013.

I merge these data to records of participation in the Santiago Stock Exchange (SSE) in January 2013 and to publicly disclosed asset data from December 2012. Table B-3 describes the population of C-suite executives and directors in the leadership sample and the subset matched to 1974-1991 applicants to elite degree programs (the sample used in the discontinuity design). Panel A presents counts of executives and directors. There are 10,220 total leadership positions in the dataset, of which 7,008 are directorships and 3,212 are C-suite roles. 1974-1991 applicants account for just over a quarter (2,522) of all leadership roles. Roughly one third of firms in both the full sample and the applicant sample are listed on the SSE. Applicants hold positions in 619 different firms, so there are many firms where more than one applicant holds a top position.

Asset data is available for roughly half of represented firms. Conditional on data availability, applicants lead firms with somewhat higher asset values than the full leadership sample. As shown in Panel B of Table B-3, the 25th percentile firm has assets valued at USD $170 million, with the largest firms having assests valued at nearly USD $50 billion (Values in 2012 USD). Panel C describes the titles leaders hold. These statistics cover the full leader sample.

The companies in these data span a wide variety of industries, and the largest firms in the dataset, such as Quiñenco, Antarchile, and Falabella, routinely appear in the Forbes Global 2000 list of the world's largest companies. Quiñenco, the largest of these as measured by assets, had assets of more the $40 billion in December 2011. Quiñenco held controlling stakes in a) Banco de Chile, which merged with the Chilean subsidiary of Citibank in 2008; b) CCU, a joint beverage venture with Heineken, and c) Madeco, an international manufacturer of flexible packaging. The state-owned copper company Codelco is included in this dataset, as is Enersis (one of the largest private energy providers in South America), and Falabella (a South American department store chain headquartered in Chile).

Table B-3: Leadership descriptive statistics

	All	Applicants
A. Description of positions		
All positions	10220	2522
Directorships	7008	1543
C-suite	3212	979
SSE positions	3087	865
Have hire date	0.978	1.00
Median hire year	2009	2009
Have departure date	0.469	0.188
Median departure year	2011	2012
Unique firms	819	619
B. Description of firm assets		
Have asset data	0.489	0.533
25th percentile	0.1	0.17
50th percentile	0.49	0.64
75th percentile	1.61	1.98
95th percentile	10.8	12.46
Max	46.68	46.68
C. Titles for directors and C-suite positions		
Directorship titles		
Director	0.63	
Substitute director	0.14	
President	0.11	
General manger	0.06	
Vice President	0.06	
C-suite titles		
Principal Executive	0.79	
General Manager	0.2	
Other	0.01	

Descriptive statistics on leadership position data. Panel A: count of positions by listed type for full leadership dataset ('All' column) and subset of leadership data matched to 1974-1991 applicants ('Applicants' column). Panel B: position-weighted descriptive statistics on firms in leadership dataset. Percentiles reported in billions of 2012 USD. Panel C: listed titles for full leadership data sample.

Almost all board positions are classified as either 'directors' or 'substitute directors,' with a minority designated as 'president,' 'general manager,' or 'vice-president.' C-suite positions are divided into two broad categories: 'Principal executives' (79%) and 'General managers' (20%). Within these broad designations are a wide variety of more specific titles. Of the 79% of positions for which the more specific title is available, the most common positions are financial officers (14.0% of positions), 'gerente comercial' or business manager positions (8.6%), operations officers (8.4%), planning and development officers (5.3%), general managers or general directors

(4.2%), and technology officers (3.8%). Legal officer positions (3.2%) and accounting or auditing positions (3.6%) are also fairly common.

B.4 Income data descriptive statistics

I use percentile ranks to document differences in earnings outcomes by major and high school background in section 3, and to describe threshold-crossing effects in section 4.5. In this subsection I provide additional detail on the income levels associated with different percentile rankings, and on how income sources vary by ranking.

Table B-4 presents percentiles of the income distribution in 2014 USD by outcome year. These statistics are for the full income sample: all students I observe taking the admissions exam in 1980 or later who have incomes equal to one half of the 12-month minimum wage, or roughly USD $2,300. Median income in the test taker population rises from from $12,400 in 2005 to $15,700 in 2013. Income at the 99.9th percentile rises from $258,300 in 2005 to $333,600 in 2013. I use year-specific values as the income cutoffs for top income shares in the main analysis.

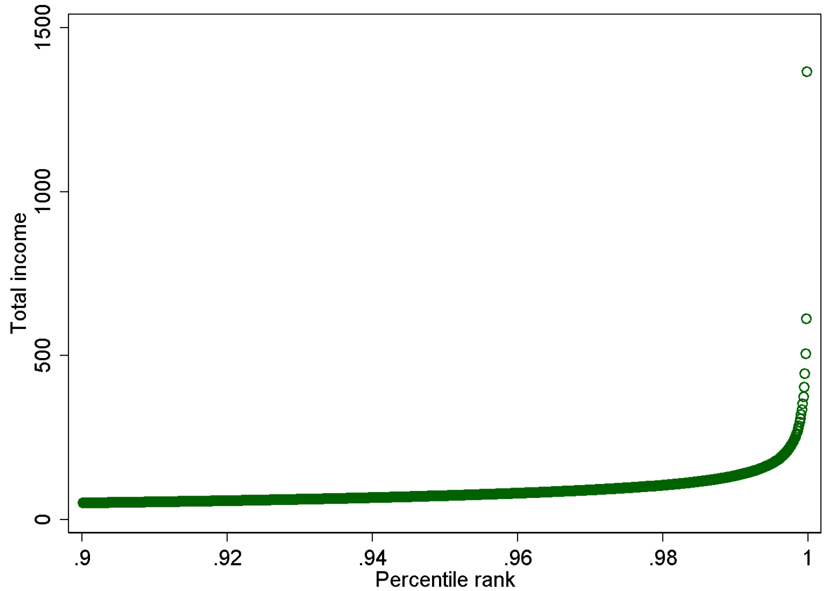
Table B-4: Percentiles of the income distribution by year

Year	Minimum	10	25	50	75	90	95	99	99.5	99.9
2005	2.3	4.3	7.0	12.4	23.1	44.0	63.8	117.6	149.3	258.3
2006	2.3	4.3	7.1	12.5	23.6	45.2	66.3	121.0	153.8	266.3
2007	2.3	4.5	7.4	13.2	25.0	48.2	69.4	128.4	163.6	295.5
2008	2.3	4.5	7.4	13.3	25.3	48.7	70.2	130.5	167.4	307.4
2009	2.3	4.5	7.5	13.5	25.5	49.0	71.3	131.7	169.3	305.9
2010	2.3	4.6	7.8	14.2	26.8	51.1	73.9	135.8	174.3	317.4
2011	2.3	4.8	8.1	14.7	27.9	53.4	76.5	142.0	182.4	337.0
2012	2.3	4.9	8.3	15.1	28.7	54.9	80.0	145.7	186.9	342.6
2013	2.3	5.1	8.7	15.7	29.6	55.6	79.3	144.3	184.2	333.6

Percentiles of income distribution by listed income year, in 1000s of 2014 USD. 'Minimum' value is cutoff for sample inclusion, other columns are percentiles. Sample is students taking the college admissions exam after 1980 who are at least 30 years of age at the time of observation and have incomes below the minimum cutoff, which is set to 50% of the value of the (annualized) monthly minimum wage.

Section 3 also shows how income densities vary by high school and college background even within the top 1% or top 0.1%. Figure B-6 explores this variation in more detail by presenting mean income values within each 0.01% percentile range. Reported values are evenly-weighted averages over the 2005-2013 period. Income rises rapidly within the top 1%, from a lower bound of $133,000 at the 99.00th percentile to a value of $307,000 at the 99.9th percentile to a value of $1,365,000 at the 99.99th percentile.

Figure B-6: Mean income by percentile



Mean total income within 0.01% bins of the income distribution in 1000s of 2014 USD. Unweighted average over 2005-2013 earnings years. Sample is students taking the college admissions exam after 1980 who are at least 30 years of age at the time of observation and have incomes below the minimum cutoff, which is set to 50% of the value of the (annualized) monthly minimum wage.

Following procedures developed by the Chilean tax authority and the World Bank (Cossio and Andres 2016), I construct individual income measures that include labor earnings (reported to the tax authority by employers) as well as income from pensions, rents, taxable capital gains, dividends, and distributed profits. Income data are not topcoded. My income measure omits business profits that are reinvested in firms, which may lead to underestimates of top income shares (Fairfield and Jorratt de Luis 2015). For workers employed in long-term contracts, records also contain basic employer characteristics such as sector. Data are available on an annual basis for the years 2005 through 2013.

C Model selection

This appendix considers the selection of optimal polynomial controls and bandwidths for the regression discontinuity analysis. I also consider the addition of controls for other covariates. There are four main findings. First, an analysis of out-of-sample fit based on leave-one-out cross validation suggests a preferred specification that does not include slope terms. Second, specifications based on optimal bandwidth calculations and bias corrections from Calonico et al. (2014, henceforth CCT) and Calonico et al. (2016, henceforth CCFT) yield similar results to those presented in the main text. Third, allowing for Lee-Card clustered standard errors does not affect inference. Fourth, adding controls for a full set of degree program by cohort interaction terms does not materially affect point estimates or inference.

I first consider the selection of optimal polynomial controls. Tables C-1 and C-2 present estimates of average mean squared errors (MSEs) from a 500-fold cross-validation procedure conducted at bandwidths of 10, 20, and 30 points with polynomials of degree zero, one, two, and three. This procedure evaluates candidate models based on out-of-sample fit by repeatedly estimating RD specifications of the form given in equations 1 and 2 , omitting one five hundredth of applicant pool at a time. I compute predicted values for the omitted applicants, compare them to the observed values, and present average MSEs over all applicants in the listed RD sample. Table C-1 reports results where top 0.1% income is the dependent variable, while Table C-2 presents results for leadership. The minimum AMSE within each bandwidth/sample combination is in bold.

For the top income dependent variable, polynomials of degree zero offer best out of sample fit at each bandwidth in the pooled sample, the sample of male students, and the sample of female students. Within the splits by high school type the degree zero polynomial gives best out of sample fit at three of six possible bandwidths. In all the zero-degree polynomial is the best fit in 12 of 15 tested bandwidth/sample combinations. When the outcome is leadership, the degree zero polynomial offers the best fit for each bandwidth in the pooled sample, the male private high school sample, and eight of 15 tested bandwidth/sample combinations overall. Best-fit specifications in the remaining bandwidth/sample combinations are split between degree one and degree two polynomials. These findings motivate the choice of a mean comparison that excludes slope terms as my preferred specification. They are consistent with the observation from Figure 6 that top outcomes are relatively flat in admissions scores. Concerns that degree zero polynomials may not offer best out of sample fit in leadership specifications for students not from private high schools are mitigated by the observation that effects are close to zero for these students in all specifications.

Table C-1: Polynomial choice by AMSE for top 0.1% income

	Degree 0	Degree 1	Degree 2	Degree 3
A. All				
BW=10	**0.0166716**	0.0166762	0.0166817	0.0166839
BW=20	**0.0157177**	0.0157203	0.0157224	0.0157247
BW=30	**0.0155156**	0.0155172	0.0155182	0.0155201
B. Male				
BW=10	**0.0208004**	0.0208047	0.0208107	0.0208168
BW=20	**0.0197112**	0.0197153	0.0197166	0.0197196
BW=30	**0.0210835**	0.0210852	0.0210881	0.0210911
C. Female				
BW=10	**0.0023341**	0.0023360	0.0023360	0.0023381
BW=20	**0.0021725**	0.0021727	0.0021730	0.0021755
BW=30	**0.0025685**	0.0025694	0.0025689	0.0025695
D. Private HS				
BW=10	0.0419108	**0.0419069**	0.0419309	0.0419614
BW=20	**0.0366472**	0.0366688	0.0366574	0.0366669
BW=30	**0.0366701**	0.0366800	0.0366962	0.0367168
E. Non-private HS				
BW=10	**0.0042356**	0.0042385	0.0042390	0.0042391
BW=20	0.0083771	**0.0083767**	0.0083796	0.0083823
BW=30	0.0062804	**0.0062790**	0.0062806	0.0062825

Sample average mean squared errors from 500-fold cross-validation estimates of equation 2 , by bandwidth, student background, and polynomial degree. Top 0.1% income share is dependent variable.

Table C-2: Polynomial choice by AMSE for leadership

	Degree 0	Degree 1	Degree 2	Degree 3
A. All				
BW=10	**0.0662082**	0.0662151	0.0662190	0.0662322
BW=20	**0.0793931**	0.0794044	0.0794064	0.0794053
BW=30	**0.0780767**	0.0780789	0.0780839	0.0780822
B. Male				
BW=10	0.0892685	0.0892724	**0.0892478**	0.0892620
BW=20	**0.0938492**	0.0938696	0.0938713	0.0938794
BW=30	0.0969588	**0.0969468**	0.0969559	0.0969563
C. Female				
BW=10	**0.0194693**	0.0194883	0.0194924	0.0194851
BW=20	0.0217335	**0.0217109**	0.0217228	0.0217349
BW=30	0.0207777	**0.0207776**	0.0207824	0.0207888
D. Private HS				
BW=10	**0.1738483**	0.1739648	0.1739495	0.1740420
BW=20	**0.1559175**	0.1560274	0.1560638	0.1561568
BW=30	**0.1771359**	0.1772127	0.1772568	0.1773235
E. Non-private HS				
BW=10	0.0315208	0.0315230	**0.0315180**	0.0315708
BW=20	0.0340981	0.0340996	**0.0340803**	0.0340825
BW=30	0.0345780	0.0345874	**0.0345770**	0.0345868

Sample average mean squared errors from 500-fold cross-validation estimates of equation 1 , by bandwidth, student background, and polynomial degree. Count of leadership positions is the dependent variable.

I next consider optimal bandwidth selection conditional on polynomial degree. I use the mean squared error minimizing selection procedure described in CCT and CCFT. I allow for separate bandwidths to the left and to the right of the threshold because, as shown in Figure 3, the range of support for the running variable is larger for admitted than for rejected students in the available data. Panel A of Table C-3 presents estimates of equations 1 and 2 for the preferred degree zero specifications using the CCT optimal bandwidths. These estimates also incorporate the bias corrections to point estimates and standard errors described in CCT. These corrections account for bias in confidence interval coverage caused by the selection of 'large' bandwidths. Standard errors cluster at the student level, as described in CCFT. Columns denote subsamples of the applicant population. The 'test' column reports p-values of tests of the null hypothesis that the coefficients in the private high school and non-private high school specifications are equal.

Optimal bandwidth choices generally range from 6 points to 11 points across samples. The exceptions are for female and non-private high school students, for whom CCT bandwidths are sometimes very narrow. These groups have very low rates of leadership and top income attainment. As expected, optimal bandwidth sizes are larger to the right of the threshold (i.e. for admitted students) than to the left (rejected students). Focusing first on leadership outcomes, effect estimates are similar to those reported in the BW=10 specification from Table 5. Inference is also similar: I reject the null of no effect in the full sample, male sample, and male private high-school sample at (at least) the ten percent level. The test of equality between private and non-private high-school coefficients has a p-value of 0.011 here, compared to 0.039 in the main analysis. The main difference between these results and those in the main text is that there is some evidence here of negative leadership effects for women and non-private school men.

Patterns in point estimates and inference are similar in the analysis of top incomes. I reject the null hypothesis that admissions effects are equal for private and non-private high school students at the 5% level for both outcomes. The one notable difference between CCFT specifications and those presented in the main text is that the CCFT approach chooses a very narrow left-hand-side bandwidth for female applicants (1.4 test score points) in the income specification and and finds evidence of a small positive effect.

Panel B of Table C-3 examines the effect of clustering on running variable value on inference as in Card and Lee (2010). The goal of this procedure is to account for correlations in error terms driven by misspecification of the polynomial control function. The running variable in this application is finely measured: it is based on an average of several component scores and recorded to the hundredth of a decimal point. We would therefore not expect clustering on this dimension to significantly alter estimated standard errors. Point estimates are identical to those in Table 5. Inference is unaffected.

Panel C of Table C-3 reports estimates of the BW=10 specification, adding a set of control variables that consists of all interactions between program identifiers and application years. As expected in the context of a valid RD, adding these controls does not substantively affect my findings.

Table C-3: Alternate regression discontinuity models

	All	Male	Female	Private HS	Non-private HS	Test
A. Optimal bandwidth selection						
Leadership positions						
Effect	0.010	0.016	-0.011	0.036	-0.015	0.011
SE	0.006	0.008	0.007	0.019	0.007	
p-value	0.081	0.038	0.103	0.055	0.042	
BW (left)	7.30	7.16	6.46	6.85	4.28	
BW (right)	8.95	9.27	10.34	11.26	4.00	
N	14657	10351	3704	3381	1853	
Top 0.1% income						
Effect	0.006	0.009	0.003	0.022	-0.001	0.027
SE	0.003	0.004	0.002	0.010	0.003	
p-value	0.094	0.032	0.076	0.027	0.750	
BW (left)	6.72	7.67	1.41	6.98	8.88	
BW (right)	8.74	10.08	8.24	10.52	5.07	
N	58255	46686	9291	16493	15552	
B. Card-Lee standard errors						
Leadership positions						
Effect	0.013	0.019	-0.005	0.032	0.002	0.047
SE	0.004	0.006	0.005	0.014	0.005	
p-value	0.002	0.001	0.305	0.023	0.654	
N	18266	12933	4548	3853	4462	
Top 0.1% incomes						
Effect	0.007	0.009	0.000	0.022	0.001	0.004
SE	0.002	0.003	0.001	0.007	0.002	
p-value	0.002	0.002	0.966	0.002	0.811	
N	77045	54134	19581	19507	21412	
C. BW=10 Spec. with controls						
Leadership						
Effect	0.012	0.017	-0.007	0.035	0.003	0.027
SE	0.006	0.007	0.007	0.014	0.005	
p-value	0.027	0.011	0.314	0.011	0.576	
N	10637	8315	2322	3853	4462	
Top 0.1% income						
Effect	0.008	0.010	0.000	0.022	-0.000	0.003
SE	0.003	0.004	0.001	0.007	0.003	
p-value	0.002	0.004	0.618	0.002	0.909	
N	52936	40919	12017	19507	21412	

Panel A: Point estimates, robust standard errors, p-values, and optimal bandwidths from CCT optimal bandwidth calculations. Panel B: estimates of equations 2 and 1 using Card-Lee standard errors. Panel C: estimates of equations 2 and 1 that add controls for predetermined student characteristics. Columns split by gender student high school background. Dependent variables as noted. 'Test' column reports results from two-sided test of equality between private HS and non-private HS admissions effects. Sample sizes in Panel A vary with bandwidth choice. Sample sizes in Panel C vary across columns due to missing data on gender and high school type.

D Heterogeneous effects analysis

D.1 Heterogeneity by institution and major

The results presented so far pool students applying to business, law, and engineering fields at PUC and UC. It is possible that effects on top labor market achievement differ by field of study and institution. Effects by institution are of particular interest because they correspond to the 'any elite' and the 'top program' treatment concepts described in Section 4.1. For brevity I focus on the sample of male students and differences by high school type within this group. Panels A and B of Table D-1 present estimated effects on top income, leadership, and log income outcomes for students admitted to the elite degree programs at PUC and UC, respectively. Program-specific effects are more noisily estimated than pooled effects due to the reduction in sample size. Patterns of leadership and log income effects are broadly similar across the two institutions, with positive and statistically significant effects in the male sample and larger point estimates for students from private high school backgrounds. The leadership (log income) effect for the pooled high school sample is 0.024 (0.135) for PUC and 0.018 (0.113) for UC. Patterns for top income effects are somewhat different. While PUC admission raises the chance a student will have income in the top 0.1% by 2 percentage points, admission to UC raises the top income probability by only 0.5 percentage points. Similarly, PUC raises the chance of attaining a top 1% income by 5.3 percentage points, compared to 1.5 percentage points for UC admission.

Table D-1: Effect of elite admission by high school by institution and major

	Top 0.1%	Leadership	Top 10%	Top 1%	Log inc.
A. PUC					
Male	0.020	0.024	0.034	0.053	0.135
	(0.008)	(0.012)	(0.019)	(0.014)	(0.045)
	16387	3630	16387	16387	13628
Male private	0.037	0.040	0.051	0.084	0.211
	(0.015)	(0.027)	(0.028)	(0.025)	(0.067)
	6843	1300	6843	6843	6021
Male non-private	0.002	0.007	0.005	0.024	0.047
	(0.008)	(0.014)	(0.034)	(0.021)	(0.076)
	5346	1016	5346	5346	4625
Test	0.042	0.269	0.296	0.066	0.106
B. UC					
Male	0.005	0.018	0.037	0.015	0.113
	(0.003)	(0.006)	(0.012)	(0.007)	(0.028)
	37747	9303	37747	37747	31018
Male private	0.014	0.028	0.072	0.037	0.197
	(0.007)	(0.014)	(0.021)	(0.015)	(0.047)
	12664	2553	12664	12664	10999
Male non-private	0.000	0.001	0.018	0.002	0.060
	(0.001)	(0.005)	(0.018)	(0.008)	(0.041)
	16066	3446	16066	16066	13393
Test	0.042	0.078	0.055	0.048	0.027
C. Business					
Male	0.022	0.040	0.027	0.038	0.157
	(0.008)	(0.013)	(0.021)	(0.016)	(0.052)
	13793	3422	13793	13793	11057
Male private	0.030	0.071	0.047	0.061	0.192
	(0.016)	(0.028)	(0.031)	(0.028)	(0.080)
	5631	1148	5631	5631	4816
Male non-private	0.000	0.007	0.012	0.004	0.065
	(0.007)	(0.013)	(0.036)	(0.023)	(0.081)
	4572	960	4572	4572	3848
Test	0.080	0.040	0.462	0.116	0.264
D. Engineering					
Male	0.003	0.008	0.042	0.018	0.101
	(0.003)	(0.006)	(0.015)	(0.009)	(0.032)
	27065	6557	27065	27065	22752
Male private	0.012	0.009	0.071	0.036	0.190
	(0.008)	(0.015)	(0.024)	(0.019)	(0.053)
	9231	1857	9231	9231	8130
Male non-private	-0.000	-0.008	0.022	0.005	0.040
	(0.003)	(0.006)	(0.023)	(0.010)	(0.047)
	11220	2405	11220	11220	9665
Test	0.157	0.287	0.143	0.155	0.035
E. Law					
Male	0.009	0.027	0.022	0.029	0.115
	(0.005)	(0.012)	(0.021)	(0.011)	(0.047)
	13276	2954	13276	13276	10837
Male private	0.027	0.039	0.057	0.063	0.197
	(0.012)	(0.033)	(0.035)	(0.023)	(0.079)
	4645	848	4645	4645	4074
Male non-private	0.003	0.021	-0.006	0.019	0.063
	(0.002)	(0.009)	(0.032)	(0.013)	(0.074)
	5620	1097	5620	5620	4505
Test	0.045	0.600	0.187	0.097	0.216

Upper row in each cell: point estimate. Middle row: standard error in parentheses. Lower row: N. Estimates of threshold-crossing effects from equations 1 and 2 by target institution and field of study. Estimates are from BW=10 specification with sample of male students. Columns correspond to dependent variables. Rows define sample populations. 'Test' row reports p-value of test of equality between the private HS and non-private HS coefficients. Standard errors clustered at student level.

Panels C, D, and E of Table D-1 present separate results by field. Aggregating across high school types, effects on top income and leadership outcomes are largest in business programs, and fairly small in engineering programs. For law programs, leadership effects are nearly as large as for business programs, but top income effects are closer to engineering programs. Log earnings effects are positive across the three fields, ranging from 10 percent to 16 percent.

Effects may also vary with finer measures of student socioeconomic background. Private high school students make up nearly half of marginal applicants. It is possible that a narrower sample of students from very rich backgrounds account for a large share of the observed effects. I explore this possiblity by estimating separate effects for students from elite and non-elite private high schools, and for students who attend the selective public exam school. High school type definitions are described in Section 2.3. Results are reported in Table D-2. Though estimates are too imprecise to reject the null of equal effect sizes, point estimates of top 0.1% income, leadership, and top 1% are substantially larger for students from elite private high schools than from other private high schools. Differences are most pronounced for the leadership outcome, where admission for students from elite private high schools raises the mean count of positions held by 0.060 while admission for students from non-elite private high schools does not raise leadership counts at all. For students from elite private high schools (other private high schools), admission raises the probability of a top income by 0.036 (.012) percentage points and log earnings by 0.231 (0.177).

Table D-2: Leadership outcomes by high school type

	Top 0.1%	Leadership	Top 10%	Top 1%	Log inc.
Elite private	0.036	0.060	0.054	0.070	0.231
	(0.014)	(0.031)	(0.026)	(0.023)	(0.061)
	7400	1395	7400	7400	6648
Other private	0.012	0.012	0.066	0.040	0.177
	(0.007)	(0.011)	(0.022)	(0.016)	(0.050)
	12107	2458	12107	12107	10372
Exam	-0.005	0.007	0.036	-0.011	0.022
	(0.005)	(0.017)	(0.042)	(0.022)	(0.079)
	3474	709	3474	3474	3101

Upper row in each cell: point estimate. Middle row: standard error in parentheses. Lower row: N. Estimates of threshold-crossing effects from equations 1 and 2 by disaggregated high school type. Columns correspond to dependent variables. Rows define sample populations. Estimates are from BW=10 specification. 'Elite private' high schools are a set of seven highly selective and expensive private high schools. 'Other private' is the set of other private high schools. 'Exam' is an elite public exam school. See section 4.8 and Appendix B.2 for more on high school classification. Standard errors clustered at student level.

Table D-3: Effect of elite admission on other leadership outcomes

	C-suite	Board	Any lead	No topcode	Winsorized log inc.
All	0.005	0.009	0.008	0.014	0.093
	(0.002)	(0.005)	(0.002)	(0.005)	(0.028)
	18266	18266	18266	18266	77045
Male	0.007	0.013	0.011	0.020	0.108
	(0.003)	(0.006)	(0.003)	(0.007)	(0.034)
	12933	12933	12933	12933	54134
Female	-0.002	-0.002	-0.003	-0.005	0.027
	(0.003)	(0.003)	(0.003)	(0.005)	(0.052)
	4548	4548	4548	4548	19581
Test	0.015	0.036	0.002	0.004	0.190
Private	0.012	0.021	0.016	0.033	0.181
	(0.006)	(0.014)	(0.007)	(0.016)	(0.054)
	3853	3853	3853	3853	19507
Non-private	0.001	-0.000	0.003	0.001	0.070
	(0.004)	(0.003)	(0.003)	(0.006)	(0.049)
	4462	4462	4462	4462	21412
Test	0.103	0.146	0.080	0.058	0.325

Upper row in each cell: point estimate. Middle row: standard error in parentheses. Lower row: N. Estimates of effects of admission on leadership outcomes by high school type. Estimates are from BW=10 specification. See section 4.1 for more details on estimation. Columns denote dependent variables. 'C-suite' is the count of c-suite positions an applicant holds. 'Directorship' is the count of directorship positions an applicant holds. 'Any' is a dummy variable equal to one if an applicant holds any leadership position. 'No topcode' is count of leadership positions with no topcoding. 'Winsorized log inc.' is log income with missing values and values below 50% of the annual minimum wage set to equal 50% of the annual minimum wage, or approximately $2,300 USD. Observations are at the application level. 'Test' row reports p-values from tests that the estimates for private and non-private HS students are equal. Standard errors clustered at person level.

Effects on each of these outcomes are very close to zero for students from the public exam school, mirroring estimates for public school students from lower-scoring schools. The finding that returns are no higher for exam school students than other public high school students complements the literature on the effects of exam school admission on schooling outcomes (Abdulkadiroğlu et al. 2014, Dobbie and Fryer 2014, Zhang 2013). This research shows that exam school admission has little impact on academic outcomes such as high school completion, scores on college entrance exams, or college attendance.

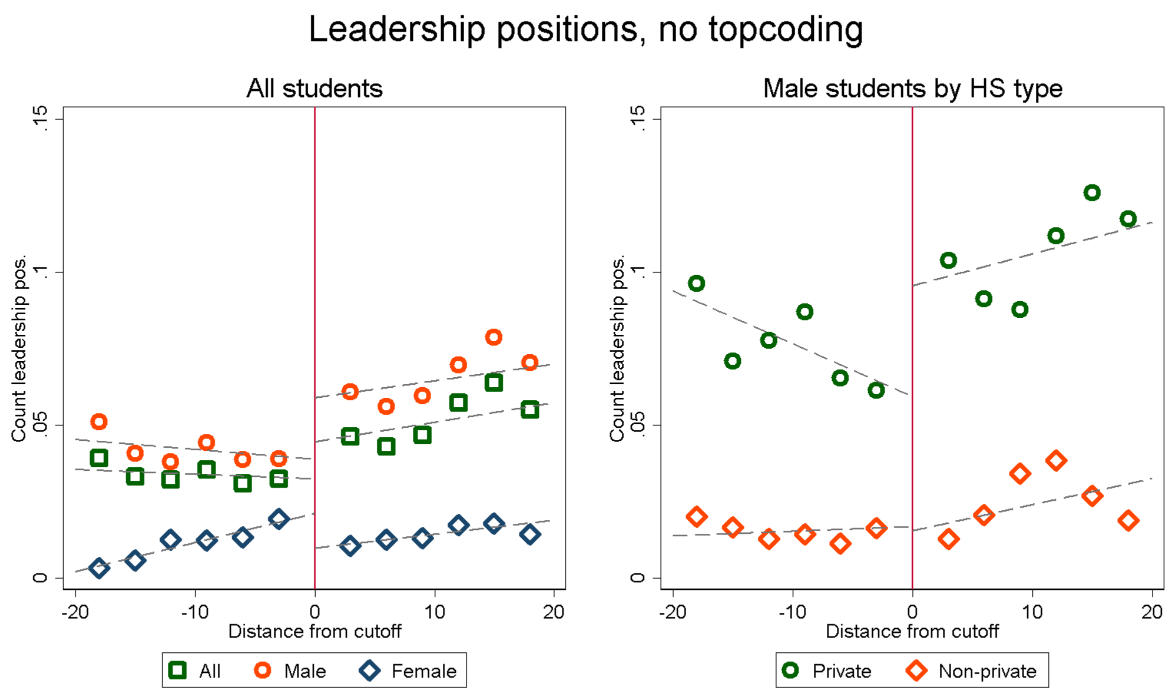
D.2 Alternate leadership and income definitions

The left four columns of Table D-3 report estimates for supplemental measures of leadership outcomes: the count of C-suite positions, the count of directorship positions, a dummy variable equal to one if an applicant holds any leadership position, and the non-topcoded count of leadership positions. Admission to an elite degree program raises counts of both C-suite and directorship positions. Point estimates are roughly twice as large for directorship positions, proportional to their larger share among leadership positions overall. Effects are large and positive for private high school students and almost exactly zero for other students. Elite admission also raises the rate at which students hold any leadership position, again only for students from private high schools. Looking at the male sample, the effect estimate of 0.011 for any leadership is smaller than the estimate of 0.020 for the count of positions. Admission affects leadership outcomes on both the extensive and intensive margins.

Eliminating topcoding slightly raises effect estimates in the pooled and private high school samples relative to those reported in Table 5. Inference is not affected. The only substantive effect of switching from topcoded to non-topcoded values is to generate a negative slope in the conditional means for outcomes below the threshold. Though fairly large in magnitude, the slope is statistically insignificant (p=0.58 in the full sample and 0.26 in the private high school sample) and is driven by a very small number of individuals who hold a large number of board seats. Compare figure D-1 below to figure 6 in the main text, or to Appendix figure A-5.

The rightmost column of Table D-3 shows estimates of log income effects using a winsorized log income variable.The log income specifications in Table 5 drop person-year observations outside of the labor force, defined as those for which income is less than 50% of the annual minimum wage, or roughly $2,300 USD (compare to a sample mean of $78,300). This sample restriction focuses the analysis on labor market participants, while simultaneously limiting the effects of arbitrarily large variation in log income at low income values on results. Winsorizing is an alternate approach that addresses the second issue but not the first. I code all values below the $2,300 threshold to the threshold value, and re-estimate the log income specifications. Point estimates are almost the same as in Table 5 across all samples, while standard errors are somewhat larger. Both of these findings make sense: labor force participation is balanced across the threshold (Table 3) so adding non-participants to the sample should not affect the difference in conditional means. Adding observations with very low log incomes raises the residual variance. Note that observation counts are identical to those for the top income categories in the upper panel of Table 5. In the top income specifications, labor force non-participants are coded as not having top incomes.

Figure D-1: Effect of admission on leadership attainment, no topcoding



Count of leadership positions and fraction of students with incomes in the top 0.1% of the population distribution by position relative to the threshold. Graphs pool applications across elite degree programs. Points reflect average outcomes for applicants within three points on either side of the horizontal axis value. Fitted values from BW=20 specification.

D.3 Leadership results by cohort

Table D-4 displays estimates of the effects of admission on leadership outcomes, splitting the sample into earlier (1974-1982) and later (1983-1991) application cohorts. Each cell of the table displays the estimated effect, the standard error and p-value associated with that effect, and the below-threshold mean value. Patterns by high school type are the same across cohort groups. Admissions effects are larger in earlier cohorts, as are below-threshold means. For private high school students, admission to an elite degree program raises the count of positions held by 0.044 in the earlier cohort group (61% on base of 0.072) and by 0.024 in the later cohort group (50% on a base of 0.048). These findings on the changing rates of leadership attainment over the life course are consistent with the observation from Figure 2 that rates of top income attainment rise later in students' life cycles.

Table D-4: Effect of elite admission on leadership by cohort split

Time period	Male students	Private HS	Other HS
1974-1991	0.019	0.032	0.002
	(0.006)	(0.014)	(0.005)
	0.036	0.059	0.015
	12933	3853	4462
1982 and earlier	0.025	0.044	0.004
	(0.009)	(0.025)	(0.006)
	0.041	0.072	0.018
	6448	1745	2156
1983 and later	0.014	0.024	0.001
	(0.006)	(0.014)	(0.007)
	0.031	0.048	0.013
	6485	2108	2306

Upper row in each cell: point estimate. Second row: (standard error). Third row: below-threshold mean. Estimates of effects of admission on leadership outcomes by high school type and application cohort.

D.4 Top jobs in less remunerative sectors

Table D-5 shows results from estimates of equation 2 in which the dependent variable is equal to one if a student has an income within the top 0.1% of the sector-specific income distributions in teaching and public administration. Top within-sector income here is taken as a proxy for top attainment in these fields. I find no evidence of gains in these alternate measures of career success.

Table D-5: Effect of elite admission on sector-specific top incomes

	Public administration	Teaching
A. BW=10		
All	0.001	-0.017
	(0.001)	(0.011)
	49353	49353
Male	0.001	-0.024
	(0.001)	(0.015)
	36273	36273
Female		
Male private		-0.027
		(0.027)
		13702
Male non-private	0.002	0.002
	(0.002)	(0.002)
	14723	14723
B. BW=20		
All	0.001	-0.037
	(0.003)	(0.021)
	90197	90197
Male	0.003	-0.050
	(0.003)	(0.029)
	66956	66956
Female	-0.002	
	(0.002)	
	20903	
Male private	-0.001	-0.064
	(0.001)	(0.053)
	25444	25444
Male non-private	0.009	-0.008
	(0.009)	(0.008)
	26688	26688

Upper row in each cell: point estimate. Second row: (standard error). Third row: N. Estimates of effects of admission on leadership outcomes by high school type. Panel A reports estimates from BW=10 specification and Panel B from BW=20 specification. See section 4.1 for more details on estimation. Columns denote dependent variables. Dependent variables are indicators equal to one if a student has a main job in the listed sector and has an income within the top 0.1% of workers in that sector. Note that values are multiplied by 100 so that units are in percentage points. Standard errors clustered at person level.

E Heterogeneous effects by next option

E.1 Data description

This section analyzes heterogeneity in early-career log earnings effects by below threshold next option, supplementing section 5.3.2 in the main text. I cover five topics. First, I describe the procedure used to determine students' next options. Second, I present descriptive statistics for the applicants in the discontinuity analysis from Table 8, with an emphasis on describing below-threshold options. Third, I present balance and first stage estimates for this sample. Fourth, I present regression discontinuity graphs corresponding to Table 8. Fifth, I present alternate estimates for Table 8 using the BW=10 specification.

I construct students' next options using data on preference lists and subtest scores. Students applying to CRUCH prior to 2004 take up to eight subtests, which are combined with students' high school grades using degree-specific weights to obtain the index used for admissions. Students submit these scores along with a list of up to eight preferences. I define a next option as the degree to which an applicant at a particular degree program would have been admitted if they are rejected from the target degree. For example, if an application is a student's second-choice program, his next option is the next-ranked program (beginning with the third rank) to which he would have received admission had he been considered for a spot.

To construct next options, I first back out test subscore weights at each degree program using a linear regression of composite scores on score components. I then use the weights to construct simulated composite scores for each choice on a student's application. Then, starting with next ranked school after the target application, I work my way down the preference list. I assign the 'next option' designation to the most-preferred degree below the target to which a student has an admissions score greater than the lowest-ranked admitted student.

Table E-1 describes the marginal students in the 'next option' sample. These are students who applied to college between 2000 and 2003. Observations are at the application-outcome year level, and include only observations at least 10 years after the application year. Just under 90% of observations are associated with some next option. The remaining ten percent would not be admitted any degree program if they are rejected at their target program. These students may intend to re-apply in the following year if they are not admitted to their top choices, or attend an institution outside of the CRUCH system (such options were more prevalent by the early 2000s than in earlier decades). Roughly 30% of applications to elite programs have another elite program as their next option. As discussed in the main text, these are largely applicants to PUC programs to have the less selective UC program in the same field as their next choice. 43% have a program in a different broad area than their target as the next option, while the remaining 57% have a program in the same area. 70% have a next option in one of the business fields– business, law, or technology. The average gap in mean peer scores between the target and the next option is 26 points, with a standard deviation of 22 points. The average difference in private high school share is 0.13, with a standard deviation of 0.19. 86% of observations match to the labor force sample.

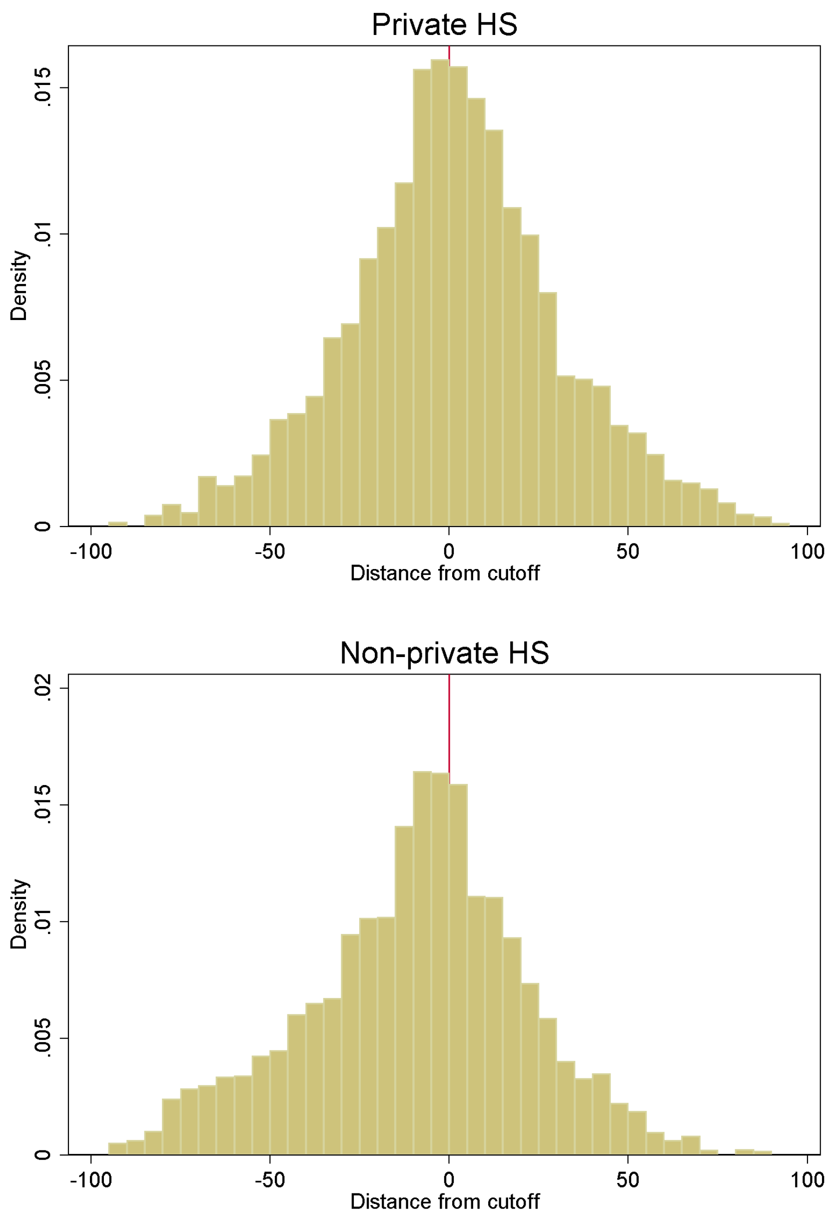
Table E-1: Description of below-threshold sample

	All	Private HS	Non-private HS
Any next option	.896	.871	.947
Next elite	.313	.353	.235
Next different area	.427	.448	.385
Next non-business area	.301	.323	.257
Score gap	26.2	23.8	30.3
SD of score gap	22.3	21.6	23
Peer gap	.128	.117	.149
SD of peer gap	.193	.185	.206
Have earnings	.86	.878	.826
Mean log earnings	10.4	10.5	10.1
N	15412	10201	5211
N w/ inc+next option	11839	7752	4087

Descriptive statistics for marginal students (within 20 points on either side of the admissions cutoff) in the below-threshold analysis sample. All statistics are means unless otherwise noted. Sample includes marginal applicants to elite degree programs in the 2000 through 2003 cohorts in outcome years at least 10 years after the application year. Observations are at the application-outcome year level. 'Any next option' is an indicator equal to one if a student would be admitted any alternate degree if rejected from their target. 'Next elite' is a dummy equal to if a student's next option is another one of the six elite degrees.'next non-business area' is in indicator equal to one if a degree is not in one of the law, business, or technology fields. 'Score gap' is the difference in mean scores for admitted students between the target and next option degree. 'Peer gap' is the difference in share of private high school students between target and next choice option.

Figure E-1 shows the density of admissions scores for students in 2000-2003 application cohorts by high school type. There is no evidence of clumping above the cutoff value. Panel A of Table E-2 shows how predetermined student characteristics change across the admissions cutoff. These estimates use the BW=20 specification of equation 2. The sample includes only students who are assigned to some next option. Sample size differs between this table and Table 8 because these results include students in the full sample applicants with next option, not just those in the labor force sample. There is little evidence of changes in student characteristics across the cutoff. Students own private high school background is continuously distributed across the cutoff. The differences between peer characteristics at the target and next option degree programs are also smooth.

Figure E-1: Density of admissions scores in below-threshold analysis sample



Density of admissions scores for private high school (upper panel) and non-private high school (lower panel) students in 2000-2003 application cohorts. Bins have width 5.

Panel B of Table E-2 shows the effects of admission to an elite program on characteristics of the admitted program. The very small number of students who are not admitted to a next choice despite simulation prediction that they will be admitted are omitted from the sample. These findings are consistent with the below threshold descriptive statistics reported in Table E-1. For example, elite admission raises the chance of being admitted to a business degree program by 28.4 percentage points. Table E-1 reports that 30.1% of students in the sample have a next option in a non-business-related field. These findings indicate that the predicted next options accurately capture true next options.

Table E-2: Balance and first stage

	All	Private	Non-private
A. Balance			
Peer score gap	-0.711	-0.895	0.034
	(1.189)	(1.502)	(1.888)
Peer HS gap	-0.006	0.002	-0.019
	(0.010)	(0.012)	(0.016)
Own private HS	0.025		
	(0.028)		
B. First stage			
Admitted to business	0.284	0.270	0.308
	(0.018)	(0.023)	(0.032)
Admitted to elite	0.708	0.633	0.834
	(0.018)	(0.024)	(0.025)
Peer score	27.569	23.204	34.631
	(1.260)	(1.590)	(1.885)
Peer private HS	0.148	0.127	0.177
	(0.009)	(0.010)	(0.014)
N obs.	13675	8767	4908
N ind.	4548	2815	1733

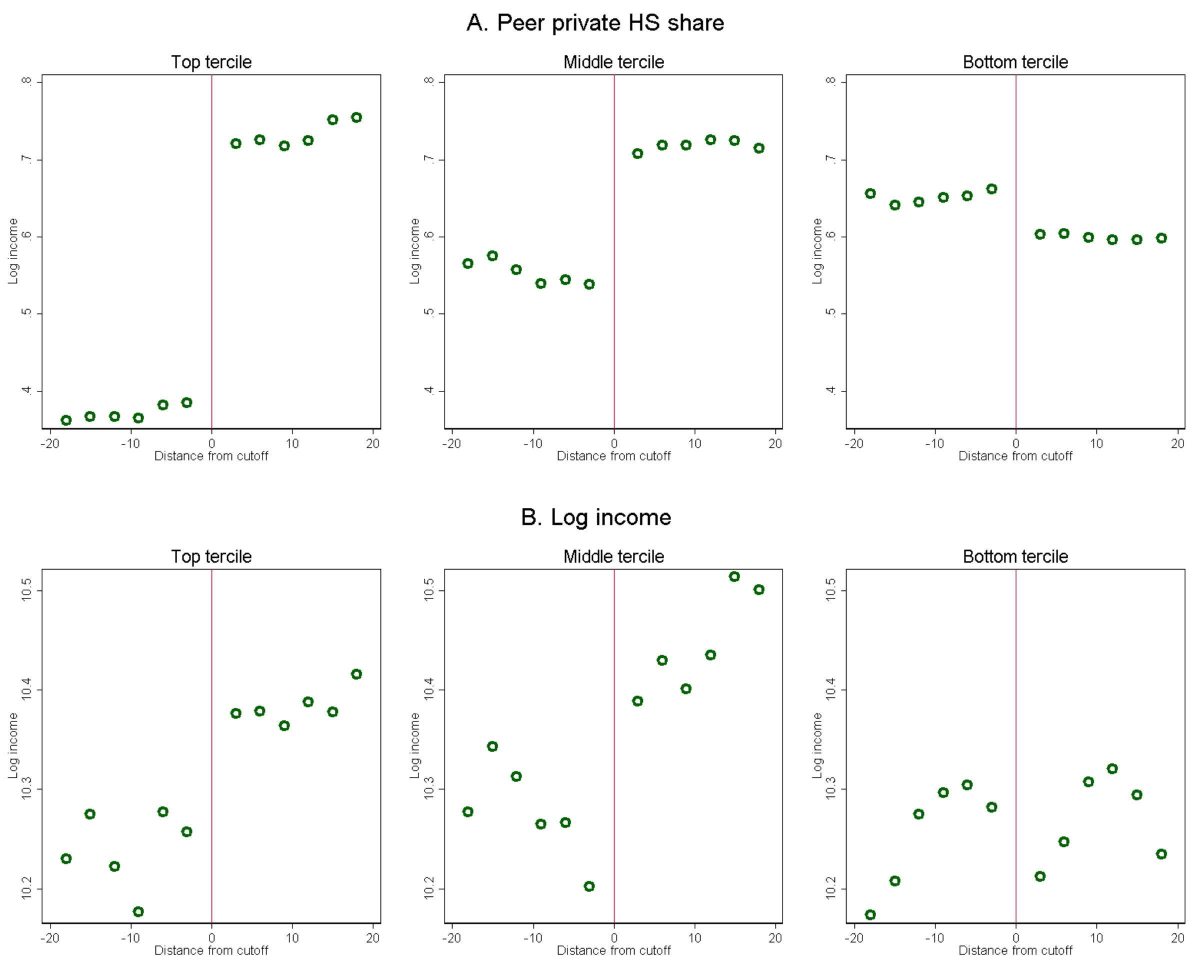
Standard errors in parentheses. Balance tests and first stage estimates from equation 2 in the next option sample using BW=20 specification for male students. Columns define sample populations. Upper panel: 'Score gap' is the difference in mean scores for admitted students between the target and next option degree. 'Peer gap' is the difference in share of private high school students between target and next choice option. 'Own private HS' is a dummy for a student's own HS status. Lower panel: 'admitted to business' and 'admitted to elite' are dummy variables describing degree to which student is admitted. 'Peer score' and 'Peer private HS' describe mean scores and HS backgrounds at admitted option.

E.2 Results

Table 8 in the main text presents regression-discontinuity estimates of earnings effects obtained using the BW=20 specification in these data. See section 5.3.2 for a discussion. Figure E-2 presents graphical representations of the findings for the pooled sample in Panel B. Findings by subsample are excluded for brevity. Panel A of Figure E-2 shows how the share of private high schools peers at the degree program to which a student is admitted changes across the threshold, splitting by tercile of the cross-threshold gap in peer scores. As expected, the largest gains are for students in the top tercile, while students in the bottom tercile attend schools with lower peer private high school shares if they are admitted to their target program.

Panel B of Figure E-2 shows earnings effects. The figure shows discontinuous increases in earnings for admitted students for whom the gap between the private high school share at their target program and next-choice option falls in the upper or middle tercile of the gap distribution, and flat or falling earnings for students in the lower tercile. These findings provide visual support for results reported in Table 8.

Figure E-2: Log income and peer composition effects



Threshold-crossing effects on peer private high school share and log income by tercile of peer private high school share gap. Sample is all male students. Graphs within each panel correspond to terciles of the private high school share gap between the target and next option degrees. Panel A: outcome is private high school share at admitted degree program. Panel B: outcome is log income. Points reflect average outcomes for applicants within three points on either side of the horizontal axis value. See Section 5.3.2 for details.

E.3 Alternate bandwidths

Table E-3 reports the same estimates as in Table 8 but for BW=10 specification. For the full sample and for students in private high schools, the estimates are fairly similar and more precisely estimated. For students in non-private high schools, estimates are larger in the BW=10 specification. The finding that specification choice matters when the outcome variable is log income parallels the finding from section 4 in the sample of older applicants.

Table E-3: Estimates of heterogeneous effects by next choice, BW=10 specification

	All	Private	Non-private
A. Admissions effects by attribute			
Main effect	0.091	0.111	0.125
	(0.052)	(0.061)	(0.086)
Peer score gap	0.001	-0.000	0.002
	(0.002)	(0.002)	(0.003)
Peer HS gap	0.662	0.614	0.179
	(0.171)	(0.197)	(0.307)
Non-business falback	0.172	0.111	0.066
	(0.089)	(0.104)	(0.147)
Elite fallback	-0.266	-0.245	-0.422
	(0.085)	(0.097)	(0.158)
B. Split by tercile of private HS gap			
Top tercile	0.148	0.114	0.146
	(0.049)	(0.059)	(0.080)
Middle tercile	0.130	0.112	0.153
	(0.053)	(0.061)	(0.093)
Bottom tercile	-0.092	-0.081	-0.059
	(0.052)	(0.056)	(0.101)
N	6758	4409	2349

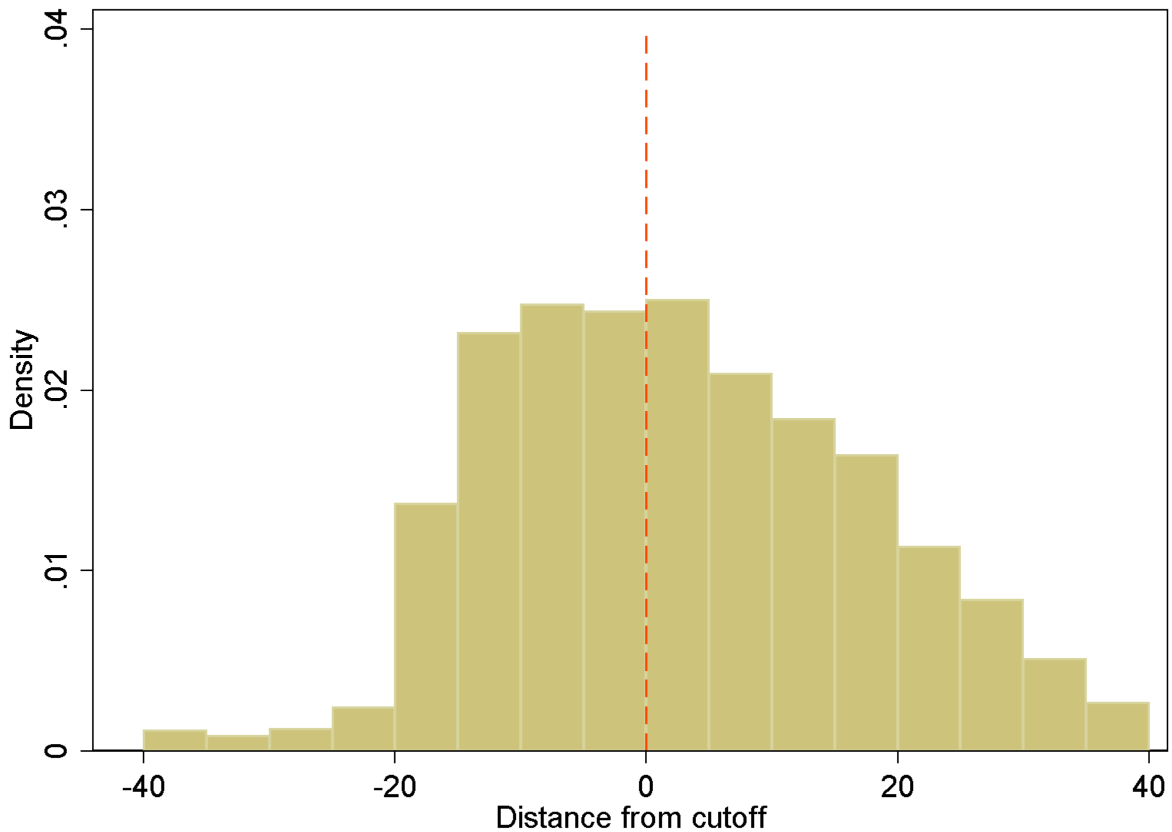
Standard errors in parentheses. Estimates of Equation 3 by HS type using 2000-2003 application data and the BW=10 specification. Dependent variable is log income. Panel A reports estimated main admissions effect and estimates of interactions between admission and the listed variables. 'Non-business fallback' is a dummy equal to one if a students' next-choice degree is not in business, law, or engineering. 'Peer score gap' is the difference between mean math scores at the target degree program and mean math scores at the next option. 'Private HS gap' is the difference between the fraction of students from private high schools at the target program and the fraction at the next option. Score gap and HS gap variables are demeaned (using means within the BW=20 sample). See Online Appendix E for descriptive statistics. 'Elite fallback' is a dummy variable equal to one if a students' next option is another elite degree program. Panel B reports estimates of equation 2 splitting by terciles of peer private HS gap. Sample pools over applications to all elite degree programs, and excludes both admitted and rejected students who would not be admitted to any degree program if they were rejected from the target.

F Comparison to elite medical programs

This subsection presents additional results from the regression discontinuity analysis of admissions to elite medical programs in section 5.4. It parallels the analyses from Tables 3 and 5 for business applicants. See main text for details. Figure F-1 displays a histogram of the score distribution by position relative to the admissions cutoff. There is no evidence of a discontinuous jump above the threshold. The p-value from a McCrary (2008) test of the null of no discontinuity in density is 0.39. Table F-1 shows that predetermined covariates (gender and high school type) are smooth across the admissions threshold. Labor force participation rates are generally smooth as well. There is some evidence of small increases in participation in the male and non-private high school sample in the BW=10 specification, but these effects become quantitatively smaller and statistically insignificant after controlling for linear slope terms in the BW=20 specification. Recall that individuals who do not participate in the labor force are excluded from specifications with log income as the dependent variable but included (as zero values) when the dependent variable is an indicator for surpassing some quantile of the income distribution.

Table F-2 reports regression discontinuity estimates of admissions effects on labor market outcomes. The 'Top 0.1%' and 'Top 10%' columns correspond to Figure 8 in the main text. As with estimates for elite business applicants, effect estimates are consistent across specifications with more precision in the BW=10 specification. For top 10% and log income outcomes, income gains are larger for women than for men, and for non-private high school students than for private high school students. Consider the case of log income. Admission to an elite medical program raises earnings by 25% in the pooled sample, by 28% for women, and by 21% for men. Earnings rise by 14% men from private high school backgrounds, and 28% for men not from private high school backgrounds. For 1% attainment, effects are larger and statistically significan only for men, though I cannot reject the hypothesis of equal effects across genders. Effects have similar size by high school type. In contrast to lower levels of income, admission to an elite medical program does not raise the chances students will have income in the top 0.1% of the distribution at all.

Figure F-1: Histogram of scores relative to cutoff for elite medical applications



Density of scores for 1982-1991 applicants to elite medical degree programs. Densities reported in bins of width 5.

Table F-1: Balance and labor force participation for medical degree applicant sample

	All	Male	Private	Non-private
A. BW=10				
Male	0.035			
	(0.020)			
Have HS	0.004	0.006		
	(0.007)	(0.010)		
Private HS	0.009	0.008		
	(0.020)	(0.026)		
Index	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
N	2425	1439	718	673
In LF sample	0.020	0.032	0.002	0.038
	(0.013)	(0.016)	(0.017)	(0.020)
Intercept	0.882	0.889	0.940	0.904
N	13689	8189	4017	3870
B. BW=20				
Male	0.036			
	(0.029)			
Have HS	0.007	0.009		
	(0.010)	(0.014)		
Private HS	-0.016	-0.024		
	(0.030)	(0.039)		
Index	-0.000	-0.001	-0.001	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
N	4242	2470	1219	1171
In LF sample	0.011	0.024	0.009	0.019
	(0.019)	(0.024)	(0.026)	(0.029)
Intercept	0.891	0.900	0.946	0.920
N	23900	14143	6778	6874

Upper row in each cell: point estimate. Second row: (standard error). Third row: N. 'Male,' 'Have HS,' and 'private HS' rows are dummy variables for indicated covariate. Index is a leadership-weighted linear index of these observable characteristics and program by cohort dummy indicators. Observations for these variables are at the application level. 'In LF sample' is an indicator equal to one for student-year observations in the labor force sample (positive income >$2,300). Columns split by gender and within male students by high school background. Sample: applicants to medical programs at PUC and UC for 1982 and later applications cohorts who are 40 or over in a given outcome year. Standard errors clustered at student level.

Table F-2: Income effects of admission to elite medical degrees

	Top 0.1%	Top 1%	Top 10%	Log inc.
A. BW=10				
All	-0.001	0.017	0.162	0.252
	(0.001)	(0.007)	(0.019)	(0.031)
	13689	13689	13689	12209
Male	-0.001	0.021	0.142	0.209
	(0.002)	(0.010)	(0.024)	(0.039)
	8189	8189	8189	7415
Female	-0.000	0.008	0.170	0.282
	(0.001)	(0.010)	(0.031)	(0.049)
	5434	5434	5434	4754
Test	0.790	0.359	0.482	0.244
Private	-0.002	0.020	0.071	0.139
	(0.003)	(0.015)	(0.030)	(0.047)
	4017	4017	4017	3782
Non-private	-0.000	0.021	0.200	0.283
	(0.001)	(0.015)	(0.036)	(0.059)
	3870	3870	3870	3569
Test	0.658	0.951	0.006	0.058
B. BW=20				
All	0.000	0.027	0.139	0.231
	(0.001)	(0.011)	(0.028)	(0.045)
	23900	23900	23900	21337
Male	0.000	0.031	0.118	0.194
	(0.002)	(0.015)	(0.036)	(0.058)
	14143	14143	14143	12849
Female	-0.000	0.015	0.137	0.237
	(0.002)	(0.013)	(0.045)	(0.069)
	9610	9610	9610	8389
Test	0.862	0.426	0.750	0.637
Private	-0.000	0.024	0.037	0.108
	(0.004)	(0.024)	(0.047)	(0.076)
	6778	6778	6778	6347
Non-private	0.001	0.038	0.184	0.272
	(0.002)	(0.021)	(0.052)	(0.083)
	6874	6874	6874	6401
Test	0.825	0.672	0.035	0.145

Upper row in each cell: point estimate. Second row: (standard error). Third row: N. Estimates from equation 2 of effects of admission to medical programs at UC and PUC on log income and top income attainment, by student high school background. Sample: applicants to medical programs at PUC and UC for 1982 and later applications cohorts who are 40 or over in a given outcome year. Upper panel: BW=10 specifications. Lower panel: BW=20 specifications. Standard errors clustered at student level.

G Supplementary peer effects analysis

G.1 Alternate standard errors

I estimate equations 4 and 5 using paired data that is unique at the application by application level. A priori it is not clear how best to compute standard errors. Leadership error terms are correlated within all observations for the same individual. It is also reasonable to think that errors might be correlated across observations within the same program cohort, since students within program-cohort cells have access to the same or similar peer contacts. Prior studies using similar data have taken a variety of approaches. Bayer et al. (2008) compute standard errors using a pairwise bootstrap. Shue (2013) presents results using both two-way clustering at the person-person level (as in Cameron et al. 2011 or Peterson 2009) and a non-parametric placebo test. The two procedures yield similar results. Fracassi (2012) uses data on pairs of firms, and computes standard errors using two-way clustering at the firm-firm level.

In the main text, I present standard errors that allow for two-way clustering at the person-person level. To check that inference is not compromised by error correlations within peer groups, I present alternate estimates that allow for two-way clustering at the program-year by program-year level in Table G-1 below. Inference does not materially change.

Table G-1: Mean co-directorships for student pairs, by type of pair

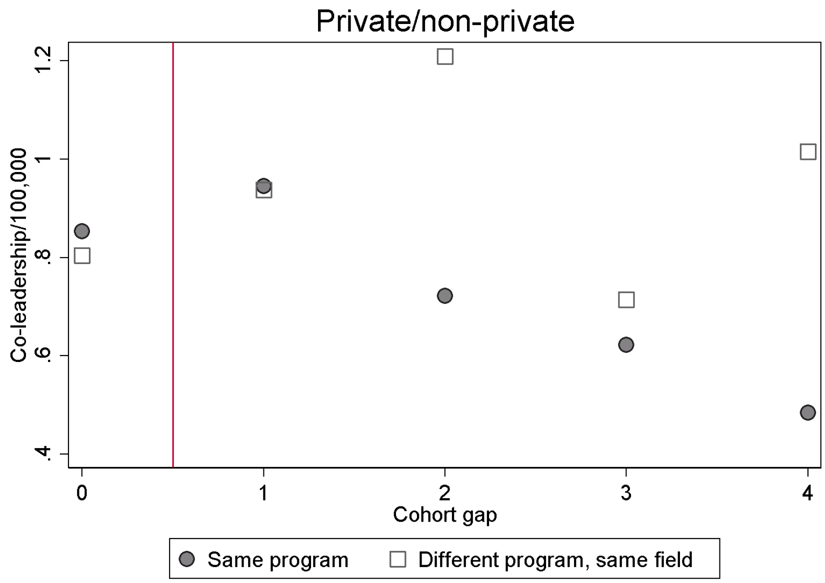
	Private/ Private	Elite/ Elite	Private/ Non-private	Non-private/ Non-private
A.Single difference				
Same cohort	3.42	15.72	0.37	0.29
	(1.34)	(9.66)	(0.29)	(0.32)
One year gap	0.92	3.91	0.46	0.15
	(0.77)	(4.14)	(0.27)	(0.18)
Two year gap	0.94	-0.33	0.24	0.16
	(0.39)	(3.48)	(0.24)	(0.25)
Three year gap	0.24	1.74	0.14	-0.22
	(0.80)	(5.74)	(0.27)	(0.08)
N	5761326	658422	13119911	8207263
B.Difference in differences				
Same cohort	3.64	20.93	0.58	0.56
	(1.69)	(10.61)	(0.59)	(0.35)
One year gap	0.19	5.58	0.54	-0.02
	(1.03)	(5.07)	(0.37)	(0.40)
Two year gap	-0.33	0.14	0.04	-0.49
	(0.92)	(4.45)	(0.41)	(0.56)
Three year gap	1.51	6.43	0.44	-0.19
	(0.81)	(5.44)	(0.39)	(0.22)
N	10609222	1317245	22022462	11568265

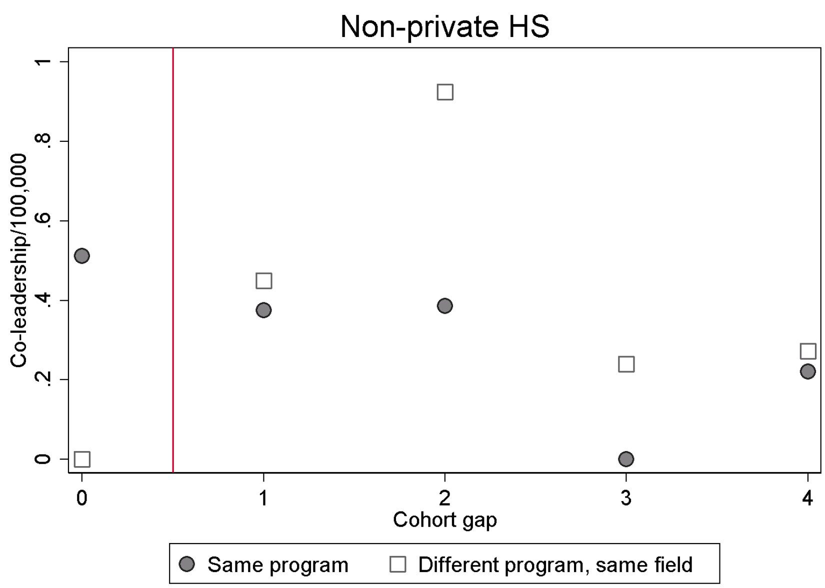
Standard errors in parentheses. Standard errors use two-way clustering at degree-year level. Sample is male students. Units are co-leadership rates per 100,000 pairs. Estimates of equations 4 and 5 by sample listed in column. 'Private/private' column consists of pairs of private high school students. 'Elite/elite' column uses pairs of students where both members are from an elite private high school. 'Private/non-private' considers pairs where one student is from a private HS and the other is not. 'Non-private/non-private' is pairs of students both from non-private schools.

G.2 Additional tables and figures

Figure G-1 shows co-leadership rates by cohort distance for pairs of students in which one member is from a private high school and one is not (upper panel) and from which both members are not from private high schools (lower panel). In contrast to the analysis of private high school pairs, there is no evidence of elevated co-leadership rates for same cohort peers.

Figure G-1: Co-leadership rates by gap





Rates of co-leadership per 100,000 pairs by absolute difference in application cohort for students in the same field at either the same institution or a different institution. Sample is male students. Upper panel: pairs of admitted students where one is from private high school and one not. Lower panel: pairs where both members are from a private high school.

Table G-2 explores rates of co-leadership for pairs of students from private high schools who attended the same and different high schools. Rates of co-leadership are much higher for students who attended the same high school, and peer effects are larger in magnitude as well. Students who attended the same high school and are in the same college cohort are 444% more likely to lead the same firm than students from the same high school and at least four cohorts apart. There is also evidence that students from different high schools experience gains in co-leadership if they are peers. Students from different high schools are 90% more likely to lead the same firms if they are peers in the same program than if they are four or more years apart. As in the main results from Section 6, findings are similar in the single difference and difference-in-differences specifications.

Table G-2: Mean co-directorships for pairs of private high school students

	All	Same HS	Dif HS	C-suite	Director
A.Single difference					
Same cohort	3.42	29.23	2.33	-0.01	3.34
	(1.37)	(14.48)	(1.24)	(0.47)	(1.30)
One year gap	0.92	7.55	0.64	0.66	0.30
	(0.85)	(7.60)	(0.80)	(0.56)	(0.65)
Two year gap	0.94	7.86	0.65	0.39	0.45
	(0.76)	(7.30)	(0.71)	(0.42)	(0.63)
Three year gap	0.24	8.87	-0.09	0.17	-0.02
	(0.77)	(9.87)	(0.69)	(0.46)	(0.64)
N	5761326	205164	5556162	5761326	5761326
B.Difference in differences					
Same cohort	3.64	34.80	2.39	-0.16	3.86
	(1.67)	(17.70)	(1.55)	(0.68)	(1.52)
One year gap	0.19	8.26	-0.11	0.13	0.10
	(1.20)	(11.65)	(1.09)	(0.74)	(0.94)
Two year gap	-0.33	8.31	-0.64	0.16	-0.43
	(1.06)	(6.98)	(1.04)	(0.59)	(0.89)
Three year gap	1.51	22.82	0.74	0.89	0.67
	(0.92)	(10.82)	(0.87)	(0.49)	(0.79)
N	10609222	369671	10239551	10609222	10609222

Standard errors in parentheses. Standard errors use two-way clustering at person-person level. Units are co-leadership rates per 100,000 pairs. Sample includes pairs of male students where both members are from a private high school. 'All' column is all such pairs. 'Same HS' column is pairs where both student are from the same high school. 'Dif HS' is pairs where students are from different high schools.

G.3 Model of leadership hiring

I present a simple model hiring in which hiring depends on student skills and referrals from school peers. This exercise has two goals. The first is to show how an analysis of changes in co-leadership rates across the admissions threshold such as that presented above can provide insight into the relative importance of the peer effects channel. The second is to develop a formula relating differences in co-leadership rates for pairs of students who are and are not college peers to the overall gain in leadership positions associated with admission. This will provide the basis for a decomposition of the total effect of admission into a 'skill' component and a 'peers' component. I emphasize that this is one of a number of possible models of peer effects in hiring.

In the model, peer ties help students by providing firms with information about student productivity. Peer ties could also operate through other channels. I design the model to capture three intuitive features of top management hiring decisions. First, admission to an elite degree program may help students develop skills applicable at many firms, and may also connect students to peers who make top management hiring decisions at specific firms. Second, because a very small number of students are in a position to influence the top management hiring process at any particular firm, students from different degree programs may have connections to different firms even if each degree program is large. Third, even within degree programs, access to connections may depend on students' friend groups, which in turn may depend on social class.

To ensure the model is tractable, I consider a very simple framework for hiring decisions and the formation of peer connections.[31] Model setup is as follows. Students may attend either elite or non-elite high schools, indexed by $d \in \{h, l\}$. They may attend either elite or non-elite college degree programs; there are P elite degree programs, each with a measure of students. There are F total firms hiring top managers. Hiring is independent across firms, and students may hold positions at multiple firms. The latter point is consistent with the observed data.

Hiring depends on students' skills and on referrals from college peers. Students take one of two skill types: productive or unproductive. I use the term 'productivity' loosely; it may reflect firm profitability, but could also reflect the incentives of those making the hiring decision. The probability a student is the productive type depends on what kind of high school and college he attended. A student who attended a type d high school and a non-elite college is productive with probability γ_d. If that student attends an elite college, the probability he is productive rises to $\gamma_d + b_d$. The model captures skill complementarities by permitting b_h to be larger than b_l; i.e., skill gains can be larger for students from elite high schools. Skill endowments are independent across students.

Firm incentives are such that a firm hires a worker if and only if the firm knows the worker is productive. Fraction π of firms observe students' skill types and do not face an informational constraint. Fraction $(1 - \pi)$ do not observe skill types. These firms receive information

[31] Strong restrictions on the hiring process and the form of peer connections (or complementarities) are often necessary to obtain tractable solutions in agglomeration models. The closest parallel in the existing literature is Oyer and Schaefer (2012), who adopt the model of Ellison and Glaeser (1997). This analysis differs from Oyer and Schaefer in that my goal is to relate extensive-margin changes in hiring outcomes to gains from peer ties, whereas Oyer and Schaefer conduct their analysis within a sample of individuals already employed at law firms.

on skills from referrals. Referrals, which reveal student productivity without error, are available only from same-type college peers with ex-ante connections to particular firms. For simplicity, I assume that there is precisely one student per firm who can provide referrals, and that referral-providing students always attend elite degree programs. The former assumption can be relaxed; the key feature is that referrals cannot be so common that they are available at all elite degree programs. I discuss the empirical basis for the latter restriction below. One could think of the students providing referrals as having family connections to firms, or as advisers to hiring committees chosen prior to management hiring. The probability that the referral-providing student for firm f is of high school type d and attends elite degree program p is r_d. Such students provide referrals to all of their same-type college peers.

The network setup described here captures in a simple way the idea that connections to firms are more common in elite degree programs and vary across both elite degree programs and social groups. I now consider the effects of elite college admission on hiring outcomes. Let Y_{idpf} be a dummy variable equal to one if student i from high school d attending elite college degree program p is hired for a leadership role at f, and let Y_{id0f} be a dummy variable equal to one if i from high school d attending a non-elite college degree program is hired at f. Then we may write the effect of elite admission on expected total leadership positions for students from high school type d as

$$E\left[\sum_f (Y_{idpf} - Y_{id0f})|d\right] = F \times (b_d\pi + r_d(\gamma_d + b_d)(1-\pi))$$

or

$$\Delta_d = S_d + C_d \tag{G.1}$$

where $\Delta_d = E\left[\sum_f (Y_{idpf} - Y_{id0f})|d\right]$, $S_d = F \times (b_d\pi)$, and $C_d = F \times (r_d(\gamma_d + b_d)(1-\pi))$. The total gain in leadership positions is equal to the sum of skill component S_d, which is equal to zero if skill gain b_d is zero, and connections component C_d, which is equal to zero if connections gain r_d is zero. The evidence in Section 4 suggests that either $S_h > S_l$, $C_h > C_l$, or both.

Co-leadership outcomes can provide insight into the relative importance of S_d and C_d in Δ_d. I focus on two model implications. The first model implication is that, for students gaining admission to elite degree program p who would otherwise have attended a non-elite college, a) any increase in the rate of co-leadership with students at some other elite degree program q is due to skill gains, and b) any additional increase in the rate of co-leadership with college peers at p is attributable to network effects. More formally, let κ_d^{dif} be the change in rates of co-leadership with non-peers associated with elite admission, and let κ_d^{same} be the change in rates of co-leadership with peers. For example, consider an experiment in which a student is randomly assigned admission to PUC Law rather than admission to some non-elite degree program. κ_d^{dif} reflects the expected gain in co-leaders from some other elite degree program (e.g., UC Law), while κ_d^{same} reflects the expected gain in co-leaders also from PUC Law.

Then, for a pair of students $i \neq j$, and two elite degree programs $p \neq q$, we may write

$$\begin{aligned}\kappa_d^{dif} &= E\left[\sum_f \left(Y_{idpf}Y_{jdqf} - Y_{id0f}Y_{jdqf}\right) |d\right] \\ &= F(\gamma_d + b_d)b_d\pi \end{aligned} \tag{G.2}$$

and

$$\begin{aligned}\kappa_d^{same} &= E\left[\sum_f \left(Y_{idpf}Y_{jdpf} - Y_{id0f}Y_{jdpf}\right) |d\right] \\ &= \kappa_d^{dif} + r_d F(\gamma_d + b_d)^2(1-\pi) \end{aligned} \tag{G.3}$$

The admissions gain in co-leadership rates with non-peers, κ_d^{dif}, is equal to zero if skill gains b_d are equal to zero, or if $\pi = 0$ and there are no firms that observe skill gains without referrals. The admissions gain in co-leadership with peers, κ_d^{same}, is equal to κ_d^{dif} plus a term that is greater than zero only if network gains r_d are greater than zero and if some firms do not observe skill perfectly.

The second model implication is that the connection effect term C_d can be expressed as the difference in co-leadership rates for pairs of college peers relative to non-peers multiplied by a scaling term. Specifically,

$$C_d = \tau_d \times \frac{E[\sum_f Y_{idpf}|d]}{E[\sum_f Y_{idpf}Y_{jdpf}|d]} \tag{G.4}$$

where $\tau_d = E\left[\sum_f \left(Y_{idpf}Y_{jdpf} - Y_{idpf}Y_{jdqf}\right) |d\right]$. The scaling term, which is equal to $(\gamma_d + b_d)^{-1}$, accounts for the fact that students can only form co-leadership pairs if both members are the productive type.

This model abstracts from potentially important forms of heterogeneity by assuming that degree programs and firms are homogeneous. It is possible that students admitted to different degree programs or fields of study have skill sets that match particularly well to certain firms. It could also be the case that students from particular cohorts are more valuable to some firms if career paths for students of a certain age coincide with a firm's management hiring schedule. I address these concerns in my empirical work using first difference and difference-in-differences approaches that compare pairs of students in the same degree programs but different cohorts and different degree programs in the same field and the same cohort. Below, I extend the model to the case in which productivity at specific firms varies across degree programs and cohorts. As is standard in difference-in-difference analyses, the key assumption required to obtain estimates of causal effects is that firm-program-type effects and firm-cohort-type effects are additively separable; i.e., that there are not differential changes in the skill match between degree programs and firms over time.

G.4 Decomposing gains from peer connections

The first model implication maps fairly closely to results presented in Panel B of Figure 9. The model indicates that, if elite admission increases rates of co-leadership with students who are not peers at the targeted elite degree program, this can be interpreted as evidence of skill gains. If co-leadership gains with peers at the elite degree program targeted for admission exceed co-leadership gains with non-peers, this is evidence of gains from peer connections. Figure 9 shows that admission to an elite degree program only increases co-leadership rates with college peers at that degree program, not with students who attend that program in different years or who attend different programs in the same field. This suggests a limited role for skill effects in driving overall gains from admission, and a potentially large role for peer effects.[32]

The second model implication is that the connection effect component of total admissions gains, C_d, can be expressed as the difference in co-leadership rates for pairs of college peers relative to non-peers, τ_d, multiplied by a scaling term. I estimate differences in co-leadership rates for elite college peers and non-peers using the difference-in-difference specification given in Equation 5. Results of this estimation procedure are discussed in Section 6. With the added structure of the hiring model, I can use results of the difference-in-differences analysis to estimate the total contribution of peer ties to leadership rates. For private high school students, the scaling term

[32]One complication in mapping the model to the regression discontinuity analysis is that threshold-crossing into admission at one degree program is associated with a reduced probability of attending the same-field program in the other institution. We see some evidence in the slight cross-threshold reduction in rates of co-leadership with same-year, other-institution peers. There is no effect on rates of co-leadership with same-program students from other cohorts, consistent with the idea that students do not substitute across this margin when admitted to the target program.

is equal to the mean leadership rate for admitted applicants (0.128) divided by the mean co-leadership rate for pairs of same-year peers (an application-weighted 1.36 $\times 10^{-4}$). Multiplying the scaling term by the estimated private high school effect from Table 10 of 3.42×10^{-5} yields the estimate $C_{private} = 0.032$. This value is close to the estimated threshold-crossing effects of 0.033 (BW=10 specification) and 0.036 (BW=20 specification) for private high school students reported in Table 5. This analysis is consistent with the hypothesis that peer effects account for a large share of overall leadership gains associated with admission.

G.5 Model extension to match heterogeneity

This section extends the model to allow for heterogeneous skill match between firms and degree programs and between firms and cohorts. The presence of separable degree-firm and cohort-firm match effects motivates the difference-in-differences empirical approach.

Model setup is as above, with three changes. First, students attend elite or non-elite degree programs in T different cohorts, denoted by t. Second, skills are firm-specific, and may depend on which elite degree program a student attends and when. The probability a student from high school type d in cohort t who does not attend an elite degree program is productive at firm f is γ_{dtf}. If that student attends an elite college, the probability he is productive at f rises to $\gamma_{dtf} + b_{dpf}$. The skill gains from admission to a particular elite degree program for each high school type are constant within firms over time. Third, referrals occur only within degree-program-cohort groups; i.e., within students who attend the same degree program at the same time. The probability that there is a type d referral-provider for firm f in elite degree program p and cohort t is again r_d.

Let Y_{idptf} be a dummy variable equal to one if student i from high school type d attending elite degree program p in cohort t is hired at f, and let Y_{id0tf} be a dummy equal to one if i from high school type d in cohort t attending a non-elite degree program is hired at f. Define Δ_d^{pt} as the gain from admission to elite degree program p rather than a non-elite degree program for a cohort-t student. Then we may write

$$\begin{aligned}
\Delta_d^{pt} &= E[\sum_f (Y_{idptf} - Y_{id0tf})|d] \\
&= F \times \left(\bar{b}_{dpf}\pi + r_d(\overline{\gamma}_{dtf} + \bar{b}_{dpf})(1-\pi)\right) \\
&= S_d^{pt} + C_d^{pt} \qquad \text{(G.5)}
\end{aligned}$$

where $\bar{b}_{dpf} = F^{-1}\sum_f b_{dpf}$, $\bar{\gamma}_{dtf} = F^{-1}\sum_f \gamma_{dtf}$, $S_d^{pt} = F\bar{b}_{dpf}\pi$, and $C_d^{pt} = Fr_d(\bar{\gamma}_{dtf} + \bar{b}_{dpf})(1 - \pi)$. As above, we can think of S_d^{pt} and C_d^{pt} as reflecting the skill and connections contributions, respectively, to the total effect.

Now consider the first model implication, which held that the admissions gains in co-leadership rates with non-peers were equal to zero if skill gains were zero, and that admissions gains in co-leadership rates with peers were equal to gains with non-peers plus a term that depended on network gains r_d. Let $\mu_{dpt}^{dp't'} = E\left[\sum_f Y_{idptf} Y_{jdp't'f} | d, i \neq j\right]$ be the expected co-leadership rate for a pair of high school d students at elite degree programs p and p' in cohorts t and t'. Further, let $\mu_{d0t}^{dp't'}$ be expected co-leadership rate for a high school type d student in cohort t who is not admitted to an elite program and some student with high school-program-cohort category $dp't'$. Then for some program $p' \neq p$, we may write

$$
\begin{aligned}
\kappa_{dpt}^{dp't} &= \mu_{dpt}^{dp't} - \mu_{d0t}^{dp't} \\
&= F\pi\overline{b_{dpf}(\gamma_{dtf} + b_{dp'f})}
\end{aligned}
\tag{G.6}
$$

where $\overline{b_{dpf}(\gamma_{dtf} + b_{dp'f})} = F^{-1}\sum_f b_{dpf}(\gamma_{dtf} + b_{dp'f})$. This is the gain in co-leadership rates with students from other degree programs in the same cohort associated with elite admission. Similarly,

$$
\begin{aligned}
\kappa_{dpt}^{dpt'} &= \mu_{dpt}^{dpt'} - \mu_{d0t}^{dpt'} \\
&= F\pi\overline{b_{dpf}(\gamma_{dt'f} + b_{dpf})}
\end{aligned}
\tag{G.7}
$$

while

$$
\begin{aligned}
\kappa_{dpt}^{dpt} &= \mu_{dpt}^{dpt} - \mu_{d0t}^{dpt} \\
&= F(1-\pi)r_d\overline{\left(\gamma_{dtf} + b_{dpf}\right)^2} + F\pi\overline{b_{dpf}(\gamma_{dtf} + b_{dpf})}.
\end{aligned}
\tag{G.8}
$$

Equations G.6 and G.7 show the gains in co-leadership rates with students who are not college peers associated with elite college admission. As in the model without match effects, these gains would be zero if elite colleges do not make students more likely to be productive; i.e., if the

$b_{dpf} = 0$ for each dpf triplet. Equation G.8 shows the gains in co-leadership rates with college peers associated with elite admission. As in the model without match effects, these gains include a peer connections term that is positive only if $r_d > 0$, as well as a skill effect term. However, in contrast to the simpler model, the skill effect term in κ_{dpt}^{dpt} is not the same as the skill effect terms in $\kappa_{dpt}^{dp't}$ and $\kappa_{dpt}^{dpt'}$. In the former case, the skill effect terms will differ if $b_{dpf} \neq b_{dp'f}$; i.e., if different degree programs teach students different skills. In the latter case, the skill effect terms will differ if $\gamma_{dtf} \neq \gamma_{dt'f}$. In my empirical work, I take steps to reduce the impact of these types of heterogeneity by considering only pairs of students within the same fields of study, and by considering pairs of students from nearby cohorts.

I now turn to the second model implication, which provided a formula for the connections effect term C_d. A similar result goes through here based on a difference-in-differences approach. Specifically, we may define

$$\begin{aligned}\tau_d^{pt} &= \left(\left(\mu_{dpt}^{dpt} - \mu_{dpt}^{dpt'}\right) - \left(\mu_{dpt}^{dp't} - \mu_{dpt}^{dp't'}\right)\right) \\ &= F(1-\pi)r_d\overline{\left(\gamma_{dtf} + b_{dpf}\right)^2} \end{aligned} \tag{G.9}$$

This difference-in-differences approach isolates a term that is proportional to total peer effect contribution C_d^{pt}. The single-difference approach outlined in the model without match effects fails because of cohort match effects (for differences within degree programs across cohorts) and skill match effects (for differences across degree programs within cohorts).

As in the model presented in the main text, it is possible to rescale the difference-in-differences estimates using observable quantities to recover overall gains from peer effects. Let $C_d = E_{pt}[C_d^{pt}|d]$ and $\tau_d = E_{pt}[\tau_d^{pt}|d]$ be the average total peer effect and the difference-in-differences estimate of peer co-leadership effects, respectively, across all degree programs and cohorts. Then

$$C_d = \tau_d \times \frac{E_{ipt}[\sum_f Y_{idptf}|d]}{E_{ijpt}[\sum_f Y_{idptf}Y_{jdptf}|d, i \neq j]} \tag{G.10}$$

as before.