ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Anelli, Massimo

Working Paper

The Returns to Elite College Education: A Quasi-Experimental Analysis

IZA Discussion Papers, No. 10192

Provided in Cooperation with:
IZA Network @ LISER, Luxembourg Institute of Socio-Economic Research (LISER)

Suggested Citation: Anelli, Massimo (2016) : The Returns to Elite College Education: A Quasi-Experimental Analysis, IZA Discussion Papers, No. 10192, Institute for the Study of Labor (IZA), Bonn

This Version is available at:
https://hdl.handle.net/10419/147878

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

DISCUSSION PAPER SERIES

IZA DP No. 10192

The Returns to Elite College Education: A Quasi-Experimental Analysis

Massimo Anelli

September 2016

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

The Returns to Elite College Education: A Quasi-Experimental Analysis

Massimo Anelli
Bocconi University
and IZA

Discussion Paper No. 10192
September 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA Discussion Paper No. 10192
September 2016

ABSTRACT

The Returns to Elite College Education: A Quasi-Experimental Analysis*

I take advantage of a sharp discontinuity in the probability of admission to an elite university at the admission score threshold, to estimate causal returns to college education quality. I use a newly constructed dataset, which combines individual administrative records about high school, college admission, college attendance and tax returns. Students with score just above the admission threshold have 52% higher yearly income with respect to just-below-threshold students. This premium is equivalent to a jump from the 44th percentile to the 74th percentile of the income distribution. The richness of the data allows me to explore the counterfactual college career of not admitted students and the potential mechanisms underlying the estimated income premium. I find that students with a just-above-threshold score are less likely to be college dropouts, take six fewer months to graduate, choose different majors and are more likely to have income in the top quartile of the distribution. Cumulated over fifteen years, the time span of income data for my sample, the net premium of attending the elite university amounts to around $120,000.

JEL Classification: I21, I22, I23, I28, J24, J31

Keywords: education quality, returns to education, human capital

Corresponding author:

Massimo Anelli
Department of Policy Analysis and Public Management
Bocconi University
Via Roentgen 1
20136, Milano
Italy
E-mail: massimo.anelli@unibocconi.it

* I am grateful to Paola Borsetto, Enrica Greggio, Erika Zancan for providing data that are crucial for the success of this project. I thank Tito Boeri, Michele Pellizzari and Giovanni Peri for their support with data collection and their comments. I thank also David Autor, Gaetano Basso, Colin Cameron, Scott Carrell, Caroline Hoxby, Hilary Hoynes, Michal Kuerlander, Claudio Lucifora, Paco Martorell, Douglas Miller, Hessel Oosterbeek, Marianne Page, Na'ama Shenhav, Seth Zimmerman, the participants of the CESifo Area Conference on Employment and Social Protection, of the seminars at UC Davis, Stanford, Sonoma State University, LMU, Bocconi University, Bank of Italy, Copenhagen University, LISER, fondazione Rodolfo DeBenedetti and the NBER Education Program Meeting for important comments. I am grateful to CESifo for awarding the "Distinguished Affiliate Award" to this paper.

1 Introduction

Every year, students pay higher tuition to attend elite universities because they observe higher returns for their graduates. But to what extent differences in earnings reflect better students selecting into elite institutions? Estimating causal returns to education quality is a complex endeavor because students select into better universities based on unobservable characteristics that may also be correlated with higher potential earnings. While this selection issue has been extensively analyzed by the literature on the returns to education attainment (Angrist and Krueger 1991, Ashenfelter and Krueger 1994, Duflo 2002, Ichino and Winter-Eber 2004), few studies have tackled it credibly in the context of returns to education quality. Estimates of the returns to college quality obtained without accounting for selection tend to overestimate the returns to attendance of a highly selective college.[1] In the recent years, there has been intense debate about tuition fee hikes and excessive students' borrowing. The Obama administration has long worked to construct a ranking (a "college scorecard"[2]) aimed at helping students choose colleges that *"offer good value for their tuition dollars"*[3] and linking federal funding to it. While the scorecard provides substantial information on the cost of attendance and median earnings of graduates for each university separately, the government has abandoned the goal of publishing an actual ranking of universities. Rankings based on graduates' average or median earnings may indeed give a misleading representation of the value added of different universities and thus cause intense discontent among university administrators.[4]

To isolate the causal returns to educational quality, previous studies have used several econometric techniques. Dale and Krueger (2002,2011) compare earnings of graduates of selective college to those of students admitted at equally selective universities, but chose to

[1] As shown in Hoxby (2009), differences in terms of selectivity among U.S. universities are massive.

[2] The college scorecard can be found at the government website http://collegecost.ed.gov/scorecard/

[3] From the New York Times of August 20^{th}2013

[4] Alternative rankings, not based on earnings, have been proposed by other scholars. Avery et al. (2013) for example make use of college choices of high-achieving students to create a revealed preference ranking of American colleges and universities).

attend less selective institutions. They find positive returns only for low-income students. Brewer et al. (1999) model high school students' choice of university and estimate a 40% earning premium for attending an elite private institution relative to attending public universities. Other studies rely on the comparison of twins: Behrman et al. (1996) focus on female twins and find evidence of positive returns to attending private universities with well-paid senior faculty, while Lindahl and Regner (2005) use Swedish sibling data and show that within-family estimates of the wage premium to college quality are half of the cross-sectional estimates. Andrews et al. (2012) focus instead on quantile treatment effects and compare the wage distribution of graduates from a flagship state university to a counterfactual wage distribution of students from non-selective institutions that assumes same observable characteristics as for the selective university. They find an 11.5% premium for attending the flagship state university. Several other works use a regression discontinuity approach to deal with selection in various educational contexts (Abdulkadiroglu et al. 2014, Hastings et al. 2013). Hoekstra (2009) exploits a discontinuity in the admission rule of a U.S. flagship state university and estimates a 24% premium in earnings for its graduates. Saavedra (2008) examines the extent to which college quality affects students learning, employment and earnings in Colombia. More recent works use also a regression discontinuity design to estimate returns to elite college education on the marriage market (Kaufmann et al. 2013), returns across degree programs (Hastings et al. 2013, Kirkeboen et al. 2014) and the returns to college admission for students that are at the margin of attending university (Zimmermann 2014).

In this paper, I expand and improve on the approach used by Hoekstra (2009). I control for selection into elite college education by exploiting a sharp discontinuity in probability of admission to an elite Italian university at the admission score threshold. I use self-collected administrative data from the city of Milan that combine individual high school data, college admission data, college performance and tax returns to estimate labor market returns to attending the private elite college for graduates between 24 and 39 years old. I am also able

to link family and spatial background information. My quasi-experimental approach relies on the assumption that applicants obtaining a score close to the admission threshold are locally randomly assigned to elite university attendance. Idiosyncratic shocks in admission test performance and relative academic quality make admission close to the unknown threshold as good as random. Several tests consistently validate this fact.

This paper offers several important contributions to the literature. First, I am able to track university choices, academic performance and labor market outcomes of the counterfactual college careers of individuals not admitted.[5] To my knowledge, this is the first paper to use administrative data available for both admitted and not-admitted students to shed light on the mechanisms through which college quality affects individual labor market outcomes. Second, the "sharp" nature of the discontinuity in the admission probability offers a methodological improvement. Admissions at the elite university that I consider are based on a uni-dimensional numerical score. No subjective evaluations such as recommendation letters or essays are considered for admission. This allows me to predict perfectly the admission outcome as a function of the admission score. Third, the nature of the higher education market in Milan provides a common support of students attending universities with substantially different degrees of selectivity and quality. I compare the position in the Italian income distribution of graduates of universities in Milan vis-a-vis the position in the US income distribution of graduates of US universities. I find that the median graduate at the elite university in my setting earns as much as a graduate of Duke University, while the counterfactual student graduating in other non-selective universities in Milan earns approximately like a graduate of Durham Technical Community College (the local college for Duke University). While in the U.S. students who just missed admission at Duke University are most likely attending a similar-quality college elsewhere in the country, in my setting the jump in university quality and selectivity for just-below-threshold students is substantial. Fourth, the fact that I focus on a geographically delimited labor market with low mobility

[5]Hoekstra (2009) did not have the counterfactual college information available

for graduates allows for a proper comparison of income across individuals.[6] This allows me to overcome concerns about regional earnings heterogeneity that affects college quality estimates in markets with high mobility of college graduates such as the U.S. one.

My estimates show that attending the elite university yields an average 52% premium in yearly income. In distributional terms, this premium is equivalent to a shift from the 44^{th} percentile of Italy's national income distribution to the 74^{th} percentile. The U.S.-distribution-equivalent yearly income premium would be $19,952, which corresponds to roughly 65% of the U.S. College/high school premium. I take into account the higher tuition fees for attending the elite university and I estimate that the cumulative discounted net returns to enrolling at the elite university 15 years after graduation are $120,000 (EUR 108,535). I calculate that 40% of the observed difference in average income between graduates of the elite university and other college graduates can be attributed to selection. I also explore the heterogeneity of the R.D. estimates along the income distribution and find that attendance at the elite university is linked to a "lottery effect" at the threshold. Students with an admission score just above threshold are 66% more likely to be in the top quartile of the income distribution.

The effects I find are sizeable. Thanks to the richness of my data, I can exploit counterfactual information of students that were not admitted at the elite university and explore the potential underlying mechanisms. I find that students that are just-above threshold for admission to the elite university are more likely to complete a college degree, take 6 fewer months to graduate and are exposed to a far better selected set of peers. When interpreting the magnitude of the estimated premium it is important to take into account that the elite university focuses on degrees in economics and business. I check for the presence of a discontinuity in the choice of major at the admission score threshold and find that just-above-threshold students are 30 percentage points more likely to choose economics or business as their major. After taking into account the effect of the different major choices on income, I find that the "net" institutional quality effect of attending the elite university

[6]Most of the students graduating from universities in Milan end up working in the Milan metropolitan area, as it is one of the most thriving and productive regions of Europe

is approximately 30%.

These results provide new evidence of the large returns to attending a top quality college compared to not-selective institutions, even after controlling for selection in a quasi-experimental setting. The analysis of the mechanisms and the methodology used in this paper represent important contributions for both the literature and education policy.

2 Data and setting

The master sample in the analysis consists of a group of 30,000 individuals who graduated from college preparatory high schools (in Italian, "Licei") between 1985 and 2005 in the city of Milan (Italy), a large service-oriented metropolitan area in the wealthiest part of Italy. College educated individuals from this city often become part of the Italian elite in business, finance and academia. High school data for these individuals have been collected from 13 different high schools and contain information on high school exit test, class peers, teachers and the location of their home. The high school master data have been linked to multiple sources of administrative data as illustrated in Figure 1. They have been matched (using names and date of birth) to student records at the 5 main universities in Milan. These universities offer a broad set of majors and are considered in the country. Every year, there are around 40,000 freshman students and a total of 160,000 students attending college in Milan (Italian Ministry of Education, 2014). The majority of the students that graduate from college preparatory high schools in Milan applies to one of these 5 universities. In this specific sample, 83% of all high school graduates obtain a college degree and among all college graduates almost 90% of them are matched to the data of at least one of the 5 universities in Milan.[7] For the matched students I am able to reconstruct their university career: whether they graduated, in what year, in what field of study, in what university and with what final exit score. In a further step, the initial 30,000 records of the college preparatory high school

[7]Of the 10% unmatched students graduating from college, 68% attended another Italian university outside Milan, while the rest went abroad for their college studies.

graduates have been linked to internal revenue service data on personal gross income for year 2005, the only year for which I was able to obtain individual administrative records.[8] The home addresses where students lived during high school have been geocoded and matched to house market values.[9] Moreover, for a stratified 10% random sub-sample of all high school graduates, more detailed information has been collected from telephone interviews conducted in June 2011.[10] The additional information include many variables such as family background, parental income, job and education, current employment and current family situation of the individual.

Finally, I have merged these data with records from the admission office of the most selective university among the five in Milan. The admission office performed the match using the name, birth date, high school and year of high school graduation of each individual and returned the matched data after full anonymization. The admission office matched information for all high school students in the master data (both admitted and not admitted). However, the university has kept digital databases for the admission sessions from 1995 onwards, half of the cohorts available in my data. Figure B1 graphically shows the "time-line" of admission information availability. Every year on the time axis represents a cohort of students graduating from Milan college-preparatory high schools and applying to college in that year. As summarized in Figure B1, data on admission for the elite university for the 1995 to 2005 high school cohorts and not for the 1985 to 1994 ones. Background information, high school and college performance data, income for tax year 2005 and phone survey data obtained in 2011 are available for all cohorts.

[8]The administrative file of reported income includes all individuals in the country and it is mandatory to report any income. If a person does not appear is because he/she has no income or he/she lives outside the country. Self-employed are included in the sample.

[9]House value data have been provided by the governmental agency "Agenzia del Territorio" for 55 neighborhoods of the city

[10]The survey was conducted by "Carlo Erminero & Co.", an Italian company specialized in surveys.

2.1 The elite university under analysis

Table 1 summarizes relevant statistics for the elite university vis-a-vis the average characteristics of the other four universities of Milan under analysis. The elite university is characterized by substantially different institutional characteristics, degrees of selectivity and outcomes. It is smaller in size (has around 11,000 students enrolled every year with respect to an average of 39,000 for the other universities), has 93% higher expenditures per student, 13% higher salaries per instructor, but does not have a better student-to-instructors ratio. Net of scholarships and tuition fee discounts, attending the elite university costs 3.7 times more on average than the other universities[11].

Students attending the elite university are highly selected on their pre-college characteristics. Their high school exit scores[12] are substantially (44%) higher. While the average student attending the elite university scores in the 71^{th} percentile of the Milan high school exit score distribution, the average student in the other 4 universities scores in the 48^{th} percentile. When compared to the U.S. distribution in the SAT composite score, this difference is equivalent to a 210 points gap (a score of 1680 vs 1470). Students at the elite university also come from wealthier families, have more educated mothers and their fathers are 42% more likely to work in leadership occupations.[13]

Average differences in students' outcomes are also substantial. Students graduating from the elite university have average annual income (between 1 and 15 years after graduation) that is 87% higher than graduates of the other four universities. In distributional terms, while the average income of an elite university graduate lies at the 85^{th} percentile of the 2005 Italian income distribution, the one of a graduate of the other Milan universities is at the 44^{th} percentile. The U.S.-distribution-equivalent yearly income gap is higher than $40,000 ($64,369 vs $23,886 in the 2005 U.S. income distribution). More than 60% of them (between

[11] Of the other four universities, one is private with nominal tuition fees that are about half those of the elite university under analysis, while the other state universities require low token fees

[12] At the end of high school, all Italian students take an exit exam prepared by the Ministry of Education and graded by external committees. Scores are thus comparable across schools.

[13] Defined as managers, professionals, directors or business-owners.

25 and 45 years old) work in leadership occupations (managers, professionals, directors or business-owners) with respect to 39% of other graduates. They are also more likely to work abroad. However, there is no significant difference in the employment rate of the graduates of the elite university, with very high employment rates (above 90%) of all college graduates in Milan.

In Table 2 I compare labor market outcomes and selectivity of the universities under analysis with the U.S. university system. I use the data provided by the U.S. Department of Education College Scorecard[14] to compare Milan college students with students of US universities. In the top panel, I use the U.S.-income-distribution-equivalent income of Milan college students (as calculated in Table 1) and compare it with the income of U.S. college graduates. While the median graduate of the elite university under analysis is placed in the income distribution similarly to graduates of Duke University and Carnagie Mellon, the median graduate of all other Milan universities is placed much further down in the income distribution at levels comparable to substantially lower ranked colleges in the same states of Duke University and Carnagie Mellon (respectively, Durham Technical Community College and the Community College of Philadelphia). In the bottom panel of Table 2, I exploit the SAT-distribution equivalent score (as calculated in Table 1) to compare the selectivity of the universities under analysis to the selectivity of U.S. colleges. While in terms of graduates' income the elite university is comparable to top-ranked U.S. universities, it is less selective than top-ranked U.S. distribution (comparable to good state universities such as UC-Irvine and UC-Santa Cruz). On the other end, the non-elite universities in Milan have an extremely low level of selectivity compared to the U.S. system. Table 2 shows that the elite vs. non-elite universities selectivity gap in my data is comparable to the difference between UC-Santa Cruz or UC-Irvine and schools such as the California Baptist University of Riverside or Fresno Pacific University. This gap is not surprising since, contrary to the elite university, the other universities in Milan did not have admission tests for most of the

[14]https://collegescorecard.ed.gov/data/

period under consideration.[15]

Two important features of Italian university education (features that are in common with all other European countries) are relevant for this study. First, admission to Italian universities is degree-specific: students apply for a specific field of study at one or more universities. In order to change degree after enrollment, a student must wait for the following round of admissions (once a year) and make a new application. Where an admission test is required, it is administered by the single university. There is no centralized admission system. Second, not every university offers degrees in all fields of study. In Figure B2 I show the breakdown of enrollment at each university and in each field of study in my sample. The elite university offers degrees in Business, Economics and Law[16] with a wide range of specializations, from public policy to "management of arts, culture and fashion". Its mission emphasizes the preparation of managers and leaders with a strong economic background in a very wide range of fields. For this reason the elite university attracts a large part of students with preferences for degrees different from Economics and Business as showed later on in this paper.[17]

Overall, the descriptive statistics show that students who attend the elite university are highly selected on pre-college characteristics and have substantially better labor market outcomes. This suggests that selection bias might play an important role in determining the large gap in student outcomes. The quasi-experimental analysis presented in the next sections will shed light on the relevance of this selection bias.

[15]Exceptions are Engineering, Architecture and Medicine programs in the recent years.

[16]The faculty of Law opened during the years under analysis, but it was very small. Thus, in my analysis I focus on individuals applying for Business and Economics only.

[17]Degrees in Economics and Business are also offered by the other private university in the sample and one of the public universities ("Public university 3" in the diagram). The three public universities represent a large share of all college students in Milan and their degree offerings respond to both specialization and scale criteria: "Public university 1" in fig. B2) focuses on engineering, architecture and design while the other two have a large set of fields. "Private university 2" focuses on social sciences, including Humanities, Economics, Business and Education.

2.2 The admission rule

The admission procedure of the elite university is based on an objective uni-dimensional composite score that equally weights the high school exit exam score and the admission test score.[18] Applicants are ranked according to this score. Every year, the number of admitted students is fixed at N before the admission test is taken and every year the N students with the highest composite score are offered admission.[19] It is very difficult for applicants to predict the admission score cutoff because of idiosyncratic shocks to applicants' performance in the admission test, the relative difficulty of the admission test, cohort size and the quality of students attempting the test each year. The admissions office of the elite university provided the exact score and ranking position of each applicant.

Figure 2 shows the discontinuity in the probability of admission and enrollment at the admission threshold for applicants. The composite admission score has been standardized for each admission session and then re-scaled to be 0 at the admission threshold. The standardized score can thus be interpreted as the number of standard deviations from the admission threshold, allowing me to pool together individuals applying to different admission sessions over the years. While until year 1999 there was a single admission threshold per year for the entire university, starting from the year 2000 admissions were determined separately for each specific degree (e.g. business degree, economics degree, marketing degree, etc.) and thus several rankings and admission thresholds were determined each year. For admissions after 2000, I have thus re-scaled each degree-specific admission threshold to be 0. Given the objective nature of admission scores, I am able to identify a sharp discontinuity in probability of admission. Since not every admitted student ends up enrolling at the elite university, the discontinuity in the probability of enrolling at the elite university is not sharp, although it is very large. About 70% of admitted students in my sample end up enrolling in the elite

[18]The admission test is administered by the elite university simultaneously for all students, is multiple choice and is identical for all students.

[19]In the years under analysis there were on average 3200 students attempting the test with 2200 spots available (68% admission rate).

university. There are two possible reasons for which students admitted to the elite university may decide not to enroll. Either the students prefer to major in a field not offered by the elite university (e.g. engineering) or they may be credit constrained (tuition at the elite university is almost four times higher than tuition at the other universities).

In my analysis, I exploit the sharp discontinuity in the probability of admission to estimate the effect of admission on income. This can be interpreted as an "intent to treat" effect, since not all admitted students actually attend the elite university. I also use a fuzzy discontinuity approach to estimate the effect of enrolling at the elite university on income.

3 Identification strategy

Given the sharp nature of the discontinuity in the probability of admission, I first estimate the effect of admission on income with a parametric RD model with flexible polynomials of different orders (Lee and Lemieux, 2010):

$$log(y_i) = \delta_0 + \rho\, \mathbb{I}(x_{i,t} \geq {x_t}^*) + \delta_1 f(x_{i,t} - {x_t}^*) + \delta_2 f(x_{i,t} - {x_t}^*) * \mathbb{I}(x_{i,t} \geq {x_t}^*) + \epsilon_i \qquad (1)$$

where $log(y_i)$ is the natural log of the adjusted annual income[20] of individuals that have 1 to 15 years of potential experience on the labor market. $\mathbb{I}(x_{i,t} \geq {x_t}^*)$ is a dummy equal to one if the running variable $x_{i,t}$ (the composite admission score) is above the admission threshold ${x_t}^*$ for each year of admission t, $f(x_{i,t} - {x_t}^*)$ is a polynomial (linear or quadratic) in the admission score re-scaled to be zero at the threshold and $f(x_{i,t} - {x_t}^*) * \mathbb{I}(x_{i,t} \geq {x_t}^*)$ is the interaction of the polynomial of the composite admission score with the discontinuity dummy, included to allow the slopes of the polynomial functions to be flexible below and above the threshold. In this sharp discontinuity context $\hat{\rho}$ is the estimate of the admission effect on income. This estimate can be seen as an "intent to treat effect" since not every

[20]Since income is observed for the year 2005 for cohorts of students graduating from high school between 1985 and 2005, I regress individual log income on cohort fixed effects and retrieve the residuals. This allows me to make a more fair comparison of income across cohorts.

admitted student actually enrolls in the elite university.

When interpreting the running variable $x_{i,t}$ in this setting, it is important to keep in mind that the measurement unit of the standardized admission score (standard deviations) refer to the admission composite score which is calculated only for the sub-sample of individuals applying to the elite university. Figure 3 shows the distribution of high school exit scores for students applying to the elite university and those that did not apply and do not take the admission test. The distribution for students applying to the elite university is substantially shifted to the right (mean is 0.4 high school exit score standard deviations higher) and has lower variance. The sub-sample of students on which my RD analysis is run is thus already very homogenous and positively selected with one standard deviation in the admission composite score representing only a relatively small difference in the quality of the students with respect to a standard deviation of the high school exit score.

In order to estimate the enrollment effect on income, I use a fuzzy regression discontinuity approach: I estimate the discontinuity in the probability of enrolling at the elite university and use it to re-scale the discontinuity in income estimated in Specification (1).[21] I implement a two-stage least square approach in which I estimate the predicted probability of enrolling at the elite university $p(\widehat{enrolled_i} = 1)$ with the following first stage regression:

$$p(Enrolled_i = 1) = \theta_0 + \pi \mathbb{I}(x_{i,t} \geq {x_t}^*) + \theta_1 f(x_{i,t} - {x_t}^*) + \theta_2 f(x_{i,t} - {x_t}^*) * \mathbb{I}(x_{i,t} \geq {x_t}^*) + \eta_i \quad (2)$$

in which $p(Enrolled_i = 1)$ is a dummy equal to 1 if the student is observed enrolling at the elite university and $\hat{\pi}$ is the estimate of the discontinuity in the probability of enrollment at the threshold. All other right hand side variables are as in Specification (1).

We can think of the effect of enrolling at the elite university as the estimate $\hat{\tau}$ of the following second stage equation in which $\mathbb{I}(x_{i,t} \geq {x_t}^*)$ serves as an instrument for the probability

[21] In a Wald estimate fashion.

$p(Enrolled_i = 1)$:

$$log(y_i) = \gamma_0 + \hat{\tau}\, p(\widehat{Enr_i = 1}) + \gamma_1 f(x_{i,t} - {x_t}^*) + \gamma_2 f(x_{i,t} - {x_t}^*) * \mathbb{I}(x_{i,t} \geq {x_t}^*) + \epsilon_i \qquad (3)$$

One can easily show that the enrollment effect $\hat{\tau}$ is the result of re-scaling the discontinuity in income $\hat{\rho}$ by the discontinuity in the enrollment probability $\hat{\pi}$: replacing (2) into (3) indeed leads to $\tau\pi = \rho$ and thus $\tau = \frac{\rho}{\pi}$. In Section 5 I present estimates for $\hat{\pi}$, $\hat{\rho}$ and $\hat{\tau}$.

Due to the absence of precise admission data for the cohorts applying to the elite university between 1985 and 1994 and the fact that I only have tax returns data for the year 2005, I can perform regression discontinuity estimation with complete admission information and income for the sample of cohorts applying to the elite university between 1995 and 2001. Figure B1 illustrates the time line of the data. The main set of results presents estimates of the returns to admission, enrollment at the elite university using this sub-sample of younger cohorts.

A second identification strategy allows me to exploit also the information for the full sample. For cohorts applying to the elite university between 1985 and 1994 (as well as for 1995-2005), the high school exit score (which determines the composite admission score for 50%) is available from the high school dataset. Moreover, I have data on enrollment and graduation at the elite university for all cohorts. I exploit a numerical simulation method and the richness of my data to estimate the missing pieces of admission information. I present a formal model for this numerical simulation in the Technical Appendix A, but I summarize the intuition here. I pool all cohorts that applied to college between 1985 and 2001 together and use pre-college characteristics and the subsample of younger cohorts with complete information to predict the missing pieces of information for the older cohorts. These raises two issues. One derives from the fact that I observe students that applied to the elite university only if they enrolled after being admitted. I thus need to identify the applicants in my sample who did not enroll into the elite university (i.e. applicants who were not admitted

and applicants who were admitted but did not enroll). To do so I run a probit model with the probability of applying to the elite university conditional on not enrolling on the left hand side and a wide range of pre-college high school and background variables on the right hand side for the ten post-1994 applying cohorts for which I have complete admission information. I then perform an out-of-sample prediction (conditional on pre-college characteristics) and assign predicted probabilities of applying at the elite university conditional on not enrolling to cohorts graduating from high school before 1994. I also estimate the proportion α of students in my post-1994 sample who apply to the elite university and do not enroll in each year. The second issue is the absence of the exact admission composite score for the elder cohorts. As mentioned earlier, the missing admission composite score is determined by an equally weighted average of the university-specific admission test and the high school exit score, which is available for all cohorts. I thus use only the high school exit score as a running variable for the elder cohorts.[22] Moreover, I do not know the minimum admission score for these cohorts. For each admission session for which I do not observe the admission outcome, I run the following OLS regression for each of k possible thresholds $x_{t,k}{}^*$ in the high school exit score and choose the threshold for which I obtain the highest R^2:[23]

$$P(Enrolled_i) = \beta_0 + \beta_1 \mathbb{I}(x_{i,t} \geq x_{t,k}{}^*) + \epsilon_i \tag{4}$$

with $P(Enrolled_i)$ being a dummy equal to one if student i enrolled at the elite university and $\mathbb{I}(x_{i,t} \geq x_{t,k}{}^*)$ being a dummy equal to one if the high school exit score is higher than the threshold $x_{t,k}{}^*$. All these estimation steps are then repeated a hundred times in a numerical bootstrapping simulation to retrieve estimates for the return to elite education using all

[22] I have run also a more complex simulation in which I predict the score for the university-specific admission test that is missing for elder cohorts using the estimates of the predicted score for younger cohorts for which I have complete admission information. I then estimate the predicted admission test score out of sample for elder cohorts based on a large set of background characteristics. Finally, I use the imputed admission test together with the observed high school exit score to recreate the missing composite score for admission. Results obtained with this second simulation are available upon request.

[23] Hoekstra (2009) uses the same method to estimate the threshold in SAT score that maximizes the discontinuity in probability of admission at a state flagship university.

cohorts in my sample. In this numerical simulation, I first sample $\hat{\alpha} * Cohort_size$ students using the out of sample predicted probabilities of applying to the elite university (conditional on not enrolling) as weights. This forms my control group. Next, I combine the predicted control group with the observed treatment group (students observed enrolling at the elite university) and estimate the threshold using equation (4) separately for each admission session before 1995. Finally, I pool the cohorts together and run the regression discontinuity estimation of equation (3) for the probability of enrollment. I repeat this procedure 100 times and for each repetition b I retrieve the parameter of interest $\hat{\tau}_b$. I take the average simulated coefficient $\hat{\tau_{sim}} = \frac{\sum_{b=1}^{100} \hat{\tau}_b}{100}$ as my estimated effect of enrolling at the elite university and I use the simulated standard deviation of $\hat{\tau}_b$ to perform inference on the estimate of interest $\hat{\tau_{sim}}$.

4 Validity of the design

The main identification assumption for my empirical strategy is that assignment around the threshold that determines treatment is locally random (Lee and Lemieux, 2010). Before presenting my results, I test the validity of this assumption in several ways. The local randomness condition may be violated if the probability of treatment (i.e. the probability of admission at the elite university) around the threshold depends on observable or unobservable characteristics that are correlated with the outcomes under study. In this context, this would manifest in the manipulation of admission outcomes around the threshold by individuals with specific characteristics (e.g. if students with better potential outcomes can actively target a score right above the admission threshold). In the institutional setting under analysis, given the fixed number of admitted students, the purely objective admission procedure (i.e. no recommendation letters and no soft skill evaluations) and the idiosyncratic variation in ability of each year's applying cohort, it is extremely hard for students as well as for the admission office to predict the cutoff score for admission. However, if there were any manipulation,

we would observe a discontinuity in the density of applicants at the admission threshold (McCrary, 2008). Figure 4 shows that the density of applicants is continuous through the threshold. This is reassuring evidence that scores were not manipulated. Panel A includes data for all years of admission in the data, including the post-year-1999 admissions that were determined at a degree-specific level. As explained in section 2.2, I have re-scaled each degree-specific threshold to be 0, so that for admission years after 1999 there are several cutoff scores per year and the running variable as to be interpreted as standard deviations from each degree-specific threshold. As a result of this, for years 2000 and over the number of individuals in Panel A the density appears a little higher around the cutoffs (on both sides) then away from it. However, the smoothness of the density across the threshold (what matters for the test) is preserved. In panel B of Figure 4, I restrict the density test to admission years before 2000 for which there was a single admission cutoff score for the entire university and show that the density behaves regularly also away from the threshold.

A second test of local random assignment to the treatment verifies that pre-treatment characteristics are smooth across the threshold. This is analogous to a test for balance of background characteristics in an experimental study. Table 3 and the corresponding Figure 5 show that this "smoothness" condition is verified for several relevant background characteristics. Estimates are obtained with a parametric sharp discontinuity specification with flexible linear polynomial as presented in Section 3, Specification (1). Standard errors are clustered at cohort-high school level.[24] The first outcome is a pre-college proxy for family wealth, average house value in the city block where students lived at time of high school attendance. The estimate of the discontinuity coefficient is not significant and small in magnitude (0.045 with respect to an average logarithmic house value of 8 and a standard deviation of 0.33). In Column 2, I focus on high school background characteristics. I ranked the high school institutions present in my sample according to an index of the average performance of students on the high school exit exam. The index varies between zero and

[24] A cluster corresponds to a group of same-cohort students graduating in a given year from one of the high schools in the sample.

one. The results show that the quality of the high school attended by the students applying to the elite university is smooth across the admission threshold. In Column 3, I repeat the same exercise using a proxy for the average wealth of parents by school (i.e. the average house value of students at the school). This is an important check for my design, because it shows that students coming from the most selective high schools in Milan have the same probability of being admitted close to the threshold as other students. It also indicates that the admission procedure is based on objective scores and that there is no informal admission channel, nor extra weight given to students coming from elite high schools. In Column 4, I use a different background measure that takes into account high school peers' characteristics and social economic status, the geographical concentration of high school classmates based on an Herfindhal Index. This index is calculated on the neighborhood where the high school students used to live at the time of high school attendance. The more students in the high school class live in same neighborhoods, the higher the class geographical concentration index. Data show that geographically concentrated high school classes are substantially more common in high schools situated in less affluent neighborhoods of the city. While the best high schools attract the highest performing kids from all neighborhoods in the city, periphery schools only attract a local pool of students.[25] This type of geographical selection into high schools is the reason why students in these geographically concentrated classes also have substantially lower scores on the high school exit exam relative to students of other schools. On the one hand, this index can be seen as a proxy for low socio-economic status and academic performance. On the other hand, students in classes that are more geographically concentrated are more likely to interact also after school (since all students live in the same neighborhood) and to have tighter network. If we are concerned that some students have information that allows them to target the admission threshold, it is more likely that this information is shared in these geographically concentrated classes. If this hypothesis were true, we should observe a larger mass of students with a higher Herfindhal Index just right

[25]There is no geographical restriction to high school attendance in Italy.

of the admission threshold. However, also this background characteristic is continuous at the admission threshold.

In Appendix Table B1, I show that results of Table 3 are robust to smaller bandwidths and hold for other background characteristics. In Table B2, I repeat the same test of smoothness using the high school curriculum[26] of the students and their gender. The coefficient of interest remains insignificant. A remaining threat to my regression discontinuity is selection of the tax records. If students right below admission threshold are less likely to be matched to tax returns records, my estimates may be biased. As Column 3 of Table B2 shows, this is not the case. I also show that the probability of observing income exactly equal to zero in the tax returns is continuous across the threshold.

5 Results

Table 4 and Figure 6 present my main estimates for the returns to elite college attendance for the sample of cohorts applying to the elite university between 1995 and 2001 (the cohorts for which I have complete admission data). Standard errors are clustered at cohort/high-school level. Columns 1 and 2 present estimates of the parameters as described in Section 3 with a linear and a quadratic polynomial respectively, restricting estimation to a bandwidth of ± 3 standard deviations of the composite admission score around the threshold.[27] The sample of students on which my RD analysis is run, the sub-sample who take the test for admission at the elite university, is very homogenous and positively selected (as shown in Figure 3) with one standard deviation in the admission composite score representing only a relatively small difference in the ability of the students with respect to a standard deviation of the high school exit score. The first row of the table shows the sharp discontinuity estimate in income at

[26] College preparatory high schools in my sample offer either a curriculum in classical studies or one in scientific studies.

[27] All specifications in Table 4 include gender and the family wealth proxy (parents' house value) as controls. As standard robustness check for regression discontinuity estimates I replicate results of Table 4 excluding control variables in Appendix Table B3. Results remain unchanged. As in a fully experimental estimations including background controls should increase the total variance explained by the model, but should not change the estimates of the treatment.

the admission threshold, which represents the admission effect. Given that not all admitted students decide to enroll in the elite university, we can consider this estimate as an "intent to treat" effect. The effect is sizeable, around 41 logarithmic points.[28] The second and the third row present the fuzzy discontinuity estimates for the effect of enrolling at the elite university. Enrollment effects (second row of Table 4) are larger then the admission effect, around 52 logarithmic points. The bottom panel of Table 4 shows the estimated discontinuities in the probability of enrolling. These are the first stage estimates used to re-scale the admission effect to obtain enrollment effect in the "Wald-estimate" setting described in Section 3. Figure 6A shows graphical representation of the estimated discontinuity in income of Table 4, Column 1. While the downward sloping linear fit to the left of the threshold might seem counter-intuitive, a deeper look at the data clarifies why this happens. Although the share of "high-S.E.S." students is quite flat along the distribution of the admission score to the right of the threshold, it is increasing to the left of the threshold and spikes up on the left tail of the distribution. Among the lowest performing students, there is a wide majority of students from very wealthy families. The income for these students might thus be correlated with their social background relatively more than to school performance.[29] Moreover, the negative slope coefficient left of the threshold is not statistically significant.

When estimating the discontinuity using the quadratic polynomial specification in Column 2 of Table 4, the magnitude of the coefficients tend to increase. Although the estimates are not statistically different across the two specifications, Figure 6B gives some guidance on why the estimates for the quadratic polynomial tend to be larger. By examining the figure it appears that the quadratic parametric interpolation for the control group (i.e. the non-admitted student) is driven by the outliers at the very bottom of the admission score

[28] Under the semi-logarithmic specification the effects measured are a good approximation of the effect in percentage change as long as the effects are close to 0. To be more precise, the estimate $\hat{\tau}$ in percentage change must be transformed as follow: $e^{\hat{\tau}} - 1$. In this case the percentage change correspondent to the 41 logarithm point change is 50%

[29] In other words, while low S.E.S. students do not apply and do not even take the elite university admission test unless they know they have a substantial chance to be admitted, high S.E.S. students apply and take the test even if they have very small chances of admission.

distribution furthest from the threshold.[30] In Columns 3 and 4, I test the robustness of my results by restricting the estimation to smaller bandwidths (± 2 and ± 1 standard deviations around the admission threshold). In Column 3, the results are similar in magnitude to those of Column 1, while for an even smaller bandwidth the estimates tend to be larger. Although I am careful in interpreting the coefficients of estimates obtained with smaller bandwidths given the relatively small sample size (especially the specification with the ± 1 bandwidth is very demanding), it is clear that estimates for returns to elite college attendance remain sizeable even in the most demanding specifications.

As a robustness check, I have replicated the estimation using data-driven bandwidth selection methods and non-parametric, bias-corrected RD estimators as proposed by Calonico et al. (2014a, 2014b). All estimates are consistent in sign and significance with the ones found in Table 4 using parametric estimation, but are in all specification slightly larger in magnitude. This suggests that the enrollment effect (around 52 logarithmic points as in second row of Table 4) might represent a lower bound for the income premium generated by attending the elite university. Results for the estimates based on Calonico et al. (2014a, 2014b) are available upon request.

In Figure 7 I test whether my results are truly capturing a causal discontinuity in income at the admission threshold or they are instead result of an idiosyncratic jump in income that happens close to the admission cutoff score. I plot the estimated discontinuity $\hat{\rho}$ at the true cutoff score and at other placebo cutoff scores 0.1 or more standard deviations away from the true cutoff. Whiskers represent precision of each estimate. Panels A-D reflect the same 4 specifications as in Table 4 and Figure 6. These figures confirm that the only significant income discontinuity present in the data is at the true threshold reassuring on the causal interpretation of the estimated returns to admission at the elite university.

[30] These outliers are the consequence of the "thin" left tail in the admission score distribution.

6 Discussion

In this section, I discuss my results in the context of the literature, present a robustness check, study whether the effects are heterogenous across the income distribution and explore potential mechanisms. I also use a simple cost-benefit analysis to provide an estimate for the cumulative discounted net returns to elite college education.

As shown in Table 4, I estimate sizeable returns to admission and enrollment at the elite university (respectively 40 and 52%). These returns are marginally smaller in magnitude than the average returns to college graduation with respect to high school, which are between 45% and 75% according to different causal estimates from the literature (see Oreopoulos and Petronijevic, 2013 for a review of these estimates). In distributional terms the estimated 52% premium is equivalent to a jump from the 44^{th} percentile of Italy's national income distribution to the 74^{th} percentile. The U.S.-distribution-equivalent yearly income premium is \$19,952 (\$43,838 for graduates of the elite university vs. \$23,886 for the other graduates). For comparison, the estimated \$19,952 premium to attending elite education amounts to roughly 65% of the U.S. College/high school premium (\$29,125), as calculated by Autor (2014).

When compared to other studies that estimate returns to college education quality, my estimates are among the largest. Hoekstra (2009) finds that U.S. students attending a flagship state university earn 24% more than those who attend non-flagship state universities, while Brewer et al. (1999) estimate a 40% earning premium to attending an elite private institution relative to attending a low quality public college. When interpreting my results, it is important to take into account the institutional context. In terms of selectiveness and outcomes, the difference between the elite university and the other universities under analysis is comparable to the case of U.S. private elite college versus public or community colleges (see Section 2.1). In addition, the education "quality treatment" is more intense in my setting than in the context of a flagship institution in a homogenous college public system.[31] The

[31]While the median graduate of the elite university under analysis is placed in the income distribution

other institutional feature that may influence the size of the estimated returns is the choice of major. As I will discuss in Section 6.3 students' preferences for institutions and majors[32] are not lexicographic and just-below-threshold students are substantially less likely to major in Economics and Business with respect to just-above-threshold ones.

From a policy perspective, it is interesting to compare my quasi-experimental estimates of the returns to elite college education with the unconditional difference in mean income across universities. The average income of students attending the elite university are 87% higher than those of students attending less selective universities in Milan. Comparing this to the 52% income premium estimated with the regression discontinuity implies that about 40% of the unconditional mean premium is due to selection. This suggests that producing rankings based on graduates' average income would substantially overestimate the value added of attending elite universities.

6.1 Extending estimation to older cohorts

Results are significant and robust to different specifications. However, the small sample size may raise some concerns about the credibility of my estimates. I thus extend my analysis to those cohorts applying to the elite university between 1985 and 1994 for which I observe only a partial component of the admission score, by exploiting an alternative identification strategy as presented in Section 3 and described more in detail in the Technical Appendix A. Including older cohorts in the analysis has two main advantages. First, it substantially increases the number of observations. Second, it allows me to extend estimation of returns to individuals who are 6 to 15 years passed graduation instead of focusing on the first 5 years on the labor market only.[33] In Table 5 I report the estimate obtained by extending the

similarly to graduates of Duke University and Carnagie Mellon, the median graduate of all other Milan universities is placed much further down in the income distribution at levels comparable to lower-ranked colleges in the same states of Duke University and Carnagie Mellon (respectively, Durham Technical Community College and the Community College of Philadelphia)

[32]In the European context students apply for specific majors in specific universities and each university offers a limited bundle of majors.

[33]When exploring returns over the experience profile, it is important to keep in mind that I am not able to properly disentangle the effect of labor market experience from cohort effects since I observe income for

estimation to the older cohorts. Column 1 of Table 5 include the preferred estimate obtained with the sample of younger cohorts for which I have complete admission information (as from Table 4). Column 2 estimates the effect of enrolling in the elite university using all cohorts available and the numerical bootstrapping procedure explained in the Technical Appendix A. The estimates are extremely similar in magnitude, while the standard errors drop by 50%.[34] This exercise reassures me on the validity of the estimates obtained with the smaller sample and provides an estimate of the returns to enrollment in the elite university also for the elder cohorts.

6.2 "Lottery-ticket Effect"

In this section, I explore the possibility that elite college attendance affects income heterogeneously over the income distribution. In particular, I investigate whether attending the elite university increases the probability of reaching the top of the income distribution.[35] Figure 8 sheds light on this question. I estimate the effects of enrolling at the elite university on the probability of being in the q^{th} quantile of the income distribution. Each dot in the graph represents a different specification in which the left-hand-side is a dummy variable that takes on value one if the income observed from tax returns is above or equal to the q^{th} quantile of the income distribution. Whiskers represent precision of estimates. The figure shows that the discontinuity is stronger (around 11 percentage points) for the probability of having income between the 70^{th} and the 80^{th} quantile of the income distribution. We can interpret this as evidence of a lottery-ticket effect that allows students right above the admission threshold to reach the top quartile of the income distribution. On the other hand, this exercise shows that no "insurance effect" is present (there is no discontinuity in the probability of being in

year 2005 only. Therefore, I can analyze returns by years since graduation only under the assumption that returns to experience are constant across cohorts

[34] By increasing the sample size three times, precision doubles. This is the result of a trade-off for which I improve precision by increasing the sample size, while reducing it because of measurement error while estimating the unknown threshold for elder cohorts as described in Section 3.

[35] A recent working paper by Zimmermann (2015) explores a similar channel, estimating the effect of elite college admission on students chances of reaching top management positions

the bottom quartile of the income distribution at the admission threshold)

6.3 Mechanisms

The estimates in Table 4 are large and represent the reduced form estimates of the labor markets returns to attendance at the elite university under study. However, what are the potential mechanisms through which admission to the elite university yields such large returns? In this section, I take advantage of the richness of my dataset to explore how admission to the elite university affects academic decisions and performance. As summarized by Figure 9, the mechanisms I explore in my analysis are college completion, time to graduation, choice of major and peer composition. However, other relevant mechanisms such as signaling and human capital are not easy to analyze empirically with my data and are thus not discussed in this work.

In Table 6 and the relative Figure 10, I explore the presence of discontinuities at the admission threshold for four academic outcomes of interest.[36] I show estimates for the baseline parametric specification with linear flexible polynomials and with a bandwidth of ± 3 standard deviations around the admission threshold (as for the main results Table 4 Column 1). In Column 1 I focus on the probability of graduating from college. Students that were just above the admission threshold were 13.6 percentage points more likely to graduate than just-below threshold ones, with respect to a 73% baseline probability of graduation. Admission to the elite university may have an impact also on the time students take to complete a degree. Tuition fees at the elite university are three times higher than for the other universities. This might constitute a strong incentive to avoid delaying graduation or an incentive to perform better and obtain higher returns.[37] In Table 6, Column 2, I thus check the presence of a discontinuity in "time to graduation".[38] The estimate shows that

[36] In Appendix Table B4, I show that results of Table 6 are robust to smaller bandwidths.

[37] Garibaldi, Giavazzi, Ichino, Rettore (2012), using a Regression Discontinuity Design, find that an increase of 1,000 euro in tuition fees reduces the probability of late graduation by at least 6.1 percentage points with respect to a benchmark average probability of 80%

[38] Since I cannot observe time to graduation for college dropout students, I have imputed as time to

just-above threshold students take 6 fewer months to graduate than just-below threshold ones. College completion and time to graduation appear to be an important channel in explaining the returns to attending the elite university. Although I cannot make causal statements on why the elite university substantially increases probability of graduation and reduce time to degree, there must be some characteristics of the elite university educational environment that determine these outcomes. The elite university may have better incentives in place to encourage students to complete their college education and to complete it earlier. A more intense instructor-student interaction or a more rewarding learning experience might play a role. Looking back at Table 1 might shed some light on the institutional features that might affect these academic outcomes. The level of expenditures per student are clearly higher (93% higher) in the elite university, however only a small fraction of these higher expenditures appear to be invested in better-paid and thus better-quality faculty (the elite university pays only 13% more per faculty member than other universities). The other dimensions on which the elite university invests substantially more than the other universities in Milan are infrastructures (e.g. buildings, computers, labs, student spaces, etc.), services to students (e.g. career service, international exchange programs, etc.) and services to faculty (e.g. administration, teaching and research assistants, etc.). These dimensions might all contribute to increase the probability of completing college and reducing time to degree. Another important factor that goes beyond teaching quality and infrastructures is the quality of peers. In Column 3 I estimate a discontinuity in the average quality of peers (defined as students in one's cohort who enroll in same university and program). Estimates show that just-above threshold students are exposed to peers that had substantially higher (more than a standard deviation higher) score in the high school exit exam. Exposure to higher-skilled, more motivated student peers might thus play a crucial role in explaining the improved academic performance for just-above-threshold students at the elite university.

Although it remains challenging to identify the specific institutional features that posi-

graduation the maximum time to graduation in my data after excluding outliers (more than three standard deviations away from the mean).

tively affect these academic outcomes, higher and faster college graduation, and exposure to higher skilled peers are clearly relevant features to factor in the bundled "value added" of attending this elite university.

In Column 4 of Table 6, I test weather the choice of major is one of the relevant mechanisms in explaining the estimated returns to elite college attendance. More specifically, I focus on the probability of choosing an Economics or Business major as outcome in my specification. The elite university offers a broad range of courses, but all formally listed under the faculty of Business and Economics, which may lead admitted students to major in this field more than others. Column 4 of Table 6 shows that students with a score just below the admission threshold are 33 percentage points less likely to major in Business or Economics. Figure 10 shows this graphically. In Figure B3 I show which majors students not admitted at the elite university chose. Only 58% of them major Business and Economics, while the remaining 42% majored in Law (16%), Humanities (5%), Engineering (4%), Social Sciences (4%) or other majors. This reveals the interesting fact that preferences of students for university and majors are not lexicographic: students do not rank institutions by specific chosen major or choose major conditional on university as it naturally happens in the U.S. context. On the contrary, they rank a bundle of institution and major choices.

Baseline differences in income by major are large and the discontinuity in Column 4 of Table 6 shows that admission to the elite university substantially increases the probability to choose a degree in economics and business. The choice of major is clearly an important mechanism in determining the returns to attending the elite university under study. One might be tempted to run the baseline regression discontinuity specification on income by restricting the analysis to students (both admitted and not admitted) who major in economics and business to get an estimate of the returns to elite college attendance for students graduating in the same major. However, in this context the choice of major is an outcome of the admission process (as shown clearly by the discontinuity in the choice of major at the cutoff) and the very heterogenous selection into different majors shown in Figure B3 for

students that were just below the admission threshold is clearly endogenous. Restricting the analysis to students who choose Economics and Business or including fixed effects for major choice in this context would lead to the same selection issues arising when controlling for an outcome in a standard regression setting.[39] Therefore, I follow an alternative "back of the envelope" method to disentangle the effect of attending the elite university from the effect of choice of major. I exploit the sample of younger cohorts not included in the main regression discontinuity analysis (and thus "exogenous" to the estimated quasi-experimental estimates of returns to elite college attendance) to estimate returns to major choice, net of the effect of university. I do so by using a mincerian specification in which I include dummies for choice of major (omitting Economics and Business) and university dummies. Column 1 of Table 7 shows returns to choice of majors relative to returns to choosing Economics and Business. In column 2 I instead report the share of students in the regression discontinuity sample (i.e. later cohorts for which I have estimated the quasi-experimental estimates of returns to elite college education) that I observe choosing each major (other than Economics and Business) after missing admission to the elite university. In column 3 I calculate the average expected returns to choosing majors other than Economics and Business (weighted for the observe shares of students choosing each alternative major) for the sample of students that were just below threshold for admission to the elite university. This exercise relies on two important caveats: average returns to major are the same for earlier and older cohorts and baseline returns to choosing Economics and Business (net of the university specific effect) are the same for students admitted to the elite university and those that were not admitted.[40] I estimate an average weighted return to choosing majors other than Business and Economics of -0.19% for my specific sample of not-admitted students. Keeping in mind the important caveats mentioned above, we can disentangle the 52% estimated effect for enrolling in the elite university, by attributing a 19% effect to the different choices of major made by not-

[39] See "bad controls", Angrist and Pischke 2009

[40] This latter caveat conceptually reflects the baseline regression discontinuity assumption for students who are in a close interval around the admission threshold.

admitted students and by interpreting the remaining 33% as the residual institutional effect of enrolling at the elite university, net of the major choice. Such an estimate is close to (still slightly larger) those estimated by previous literature in similar contexts (i.e. Hoekstra, 2009).

Besides the more academic mechanisms analysed in this section, an array of other quality components affecting income remain unexplored in this work. Institutional reputation and signaling may influence individual income, especially in the first years on the labor market. Peer effects may also be relevant for the labor market outcomes. As shown in the descriptive statistics, just-above-threshold students are exposed to peers that have substantially higher academic performance and come from relatively wealthier and more educated families. This may affect the type of networks available to the just-above-threshold students before and after graduation.

6.4 Net discounted returns

While the annual income premium of attending the elite university is large and significant, it is interesting to evaluate the net returns after taking into account the substantially higher tuition fees. In Table 8 I present a "cost-benefit" calculation that takes into account expected net yearly tuition, expected time to graduation and a baseline 3% discount rate to translate the estimated annual income premium to the net returns to investment. The calculation is based on the income premium estimated on all cohorts available for the first 15 years after graduation (Table 5) and on the discontinuity in time to graduation as estimated in Table 6. Table 8 shows that the cumulative discounted net returns to enrolling at the elite university measured 15 years after graduation are equal to $119,389 (EUR 108,535). Alternatively, this means that the net returns to enrolling in the elite university are positive already in the 3^{rd} year after graduation. By repeating the same exercise using the "pure" institutional estimate (after taking into account the effect of the choice of major) presented in Section 6.3, the net returns to attending the elite university 15 years after graduation are $71,729

and turn positive in the 4^{th} year after graduation.

7 Conclusions

I exploit a newly collected administrative Italian dataset following students from high school through college and into labor market, to estimate returns to enrolling at an elite university in the city of Milan. The administrative dataset has been matched to individual records from the admission office of the elite university under analysis. I take advantage of a sharp discontinuity in the probability of admission at the admission score threshold to obtain quasi-experimental estimates of returns to enrollment. This regression discontinuity approach allows me to overcome concerns of selection bias due to students selecting into elite universities based on unobservable characteristics correlated with potential income.

I estimate a 52% annual income premium for students enrolling at the elite university under analysis. In distributional terms, this premium is equivalent to a shift from the 44^{th} percentile of Italy's national income distribution to the 74^{th} percentile. The U.S.-distribution-equivalent yearly income premium would be \$19,952 (\$43,838 for graduates of the elite university vs. \$23,886 for the other graduates). This premium amounts to roughly 65% of the \$29,125 U.S. College/high school premium, as estimated by Autor (2014).

I explore the mechanisms underlying such a large effect. I find that students obtaining an admission score just above the admission threshold are more likely to complete a college degree and take six fewer months to graduate. I also estimate a large discontinuity in the choice of major. Just-above threshold students are 30 percentage points more likely to choose economics or business as their major. This suggests that preferences of students for institutions and majors are not lexicographic. Students do not rank institutions after choosing a specific major, but make bundled "institution/major" enrollment decisions. I show that, net of the major choice, the residual institutional quality effect of attending the elite university is 33%. This estimate is in line with those previously estimated by other

papers in similar contexts.

I explore how attending the elite university affects income across the income distribution and find evidence of a significant "lottery-ticket effect". Just-above threshold students have higher probability to have income in the top quartile of the national income distribution. On the other hand, they are not significantly less likely to end up in the bottom quartile of the income distribution.

Breaking down the effect of college quality on income into single quality components is a complex task. The estimates I obtain must be interpreted as the effect of the "bundled" quality of the elite university. This includes multiple educational inputs such as faculty, research and organization, and also the effect of signaling, reputation and peer effects. Interestingly, the discontinuity in enrollment determines a radical change in the average characteristics of college peers. The evidence I find of a discontinuous increase in the probability of completing college for just-above-threshold students suggests that the educational environment of the elite university may motivate students more, ceteris paribus. Competitive and motivated peers or more effective teachers might play a role in this.

I also present a cost-benefit analysis that takes into account yearly tuition, average time to graduation and a discount rate to translate the estimated annual income premium to discounted net returns to investment. The estimate shows that cumulative net returns to attending the elite university measured 15 years after graduation (the span of income data for my sample) are equal to $120,000 (EUR 108,535).

Finally, approximately 40% of the observed unconditional difference in mean income across universities is due to selection. In a policy perspective, this suggests that ranking institutions according to graduates' average income would substantially overestimate the value added of attending elite universities.

References

[1] Abdulkadiroglu, A., Angrist, J. & Pathak, P. 2014, "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools", Econometrica, vol. 82, no. 1, pp. 137-196.

[2] Andrews, R.J., Li, J. & Lovenheim, M.F. 2012, Quantile Treatment Effects of College Quality on Earnings: Evidence from Administrative Data in Texas, National Bureau of Economic Research, Inc, NBER Working Papers: 18068.

[3] Avery, Christopher, Mark Glickman, Caroline M. Hoxby, and Andrew Metrick (2013), A revealed preference ranking of us colleges and universities. Quarterly Journal of Economics, 128(1), 145.

[4] Angrist, J.D. & Krueger, A.B. 1991, "Does Compulsory School Attendance Affect Schooling and Earnings?", Quarterly Journal of Economics, vol. 106, no. 4, pp. 979-1014.

[5] Angrist, J.D. & Pischke, J. 2009, Mostly Harmless Econometrics: An Empiricist's Companion, Princeton and Oxford: Princeton University Press.

[6] Ashenfelter, O. & Krueger, A.B. 1994, "Estimates of the Economic Returns to Schooling from a New Sample of Twins", American Economic Review, vol. 84, no. 5, pp. 1157-1173.

[7] Avery, C, Glickman, M, Hoxby, C, & Metrick, A 2013, 'A Revealed Preference Ranking of U.S. Colleges and Universities', Quarterly Journal Of Economics, 128, 1, pp. 425-467

[8] Autor, David H. 2014, "Skills, education, and the rise ofearnings inequality among the other 99 percent", SCIENCE sciencemag.org 23 MAY 2014 - VOL 344 ISSUE 6186

[9] Behrman, J.R., Rosenzweig, M.R. & Taubman, P. 1996, "College Choice and Wages: Estimates Using Data on Female Twins", Review of Economics and Statistics, vol. 78, no. 4, pp. 672-685.

[10] Brewer, D.J., Eide, E.R. & Ehrenberg, R.G. 1999, "Does It Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings", Journal of Human Resources, vol. 34, no. 1, pp. 104-123.

[11] Calonico, S., M. D. Cattaneo, and R. Titiunik. 2014a. Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. Econometrica 82(6): 2295-2326.

[12] Calonico, S., M. D. Cattaneo, and R. Titiunik. 2014b. Robust Data-Driven Inference in the Regression-Discontinuity Design. Stata Journal 14(4): 909-946.

[13] Dale, S.B. & Krueger, A.B. 2002, "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables", Quarterly Journal of Economics, vol. 117, no. 4, pp. 1491-1527.

[14] Dale, S. & Krueger, A.B. 2011, Estimating the Return to College Selectivity over the Career Using Administrative Earnings Data, National Bureau of Economic Research, Inc, NBER Working Papers: 17159.

[15] Duflo, E. 2001, "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment", American Economic Review, vol. 91, no. 4, pp. 795-813.

[16] Garibaldi, P, Giavazzi, F, Ichino, A, & Rettore, E 2012, 'College Cost and Time to Complete a Degree: Evidence from Tuition Discontinuities', Review Of Economics And Statistics, 94, 3, pp. 699-711

[17] Hastings, J.S., Neilson, C.A. & Zimmerman, S.D. 2013, Are Some Degrees Worth More than Others? Evidence from college admission cutoffs in Chile, National Bureau of Economic Research, Inc, NBER Working Papers: 19241.

[18] Hoekstra, M. 2009, "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach", Review of Economics and Statistics, vol. 91, no. 4, pp. 717-724.

[19] Hoxby, Caroline M. (2009), "The changing selectivity of american colleges." Journal of Economic Perspectives, 23(4), 95118.

[20] Ichino, A. & Winter-Ebmer, R. 2004, "The Long-Run Educational Cost of World War II", Journal of Labor Economics, vol. 22, no. 1, pp. 57-86.

[21] Kaufmann, K., Messner, M. & Solis, A., 2013, "Returns to Elite Higher Education in the Marriage Market: Evidence from Chile", IGIER working paper series.

[22] Lee, D.S. & Lemieux, T. 2010, "Regression Discontinuity Designs in Economics", Journal of Economic Literature, vol. 48, no. 2, pp. 281-355.

[23] Kirkeboen, L, Leuven, E, & Mogstad, M 2014, 'Field of Study, Earnings, and Self-Selection'. NBER working paper series, wp. 20816.

[24] McCrary, J. 2008, "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test", Journal of Econometrics, vol. 142, no. 2, pp. 698-714.

[25] Oreopoulos, P. & Petronijevic, U. 2013, Making College Worth It: A Review of Research on the Returns to Higher Education, National Bureau of Economic Research, Inc, NBER Working Papers: 19053.

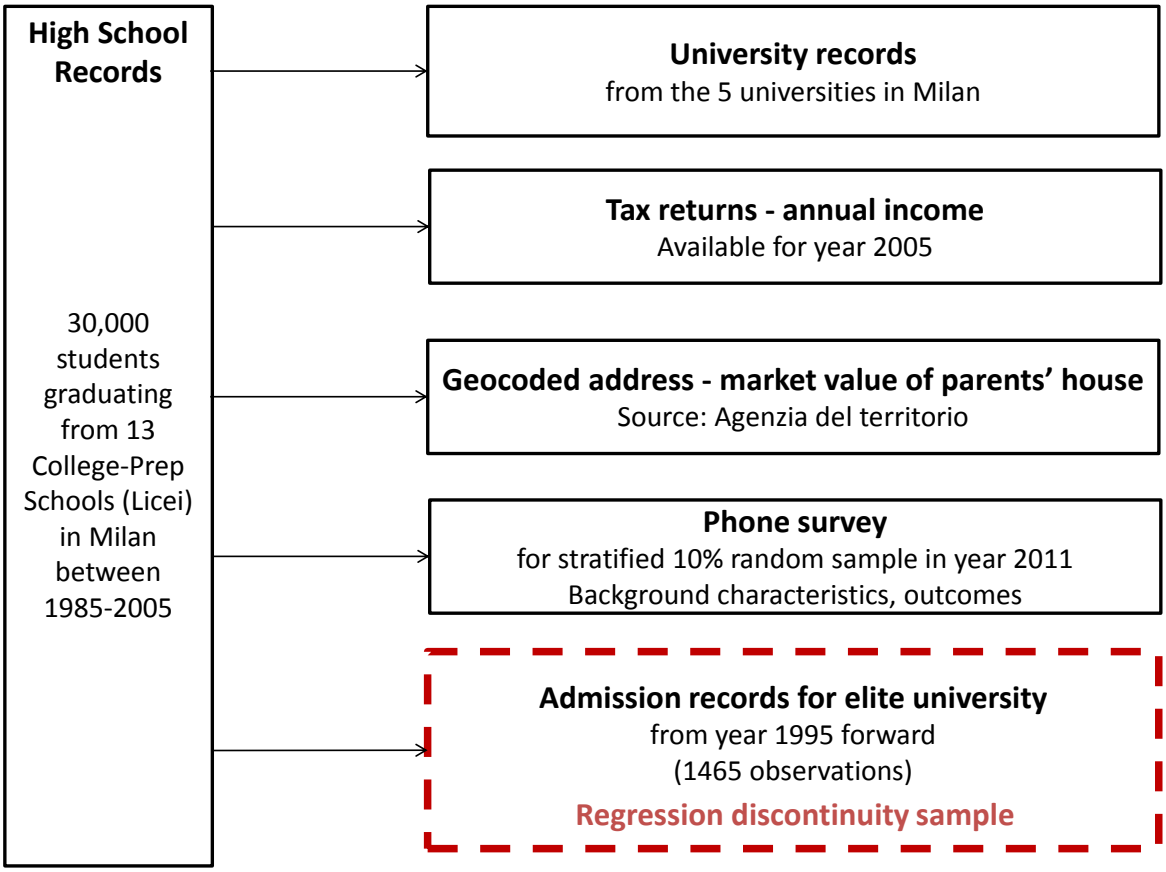
[26] Saavedra, Juan E. 2008, "The Returns to College Quality: A Regression Discontinuity Analysis"

[27] Zimmerman, S.D. 2014, "The Returns to College Admission for Academically Marginal Students", Journal of Labor Economics, vol. 32, no. 4, pp. 711-754.

[28] Zimmerman, S.D. 2015, "Making Top Managers: The Role of Elite Universities and Elite Peers", Mimeo

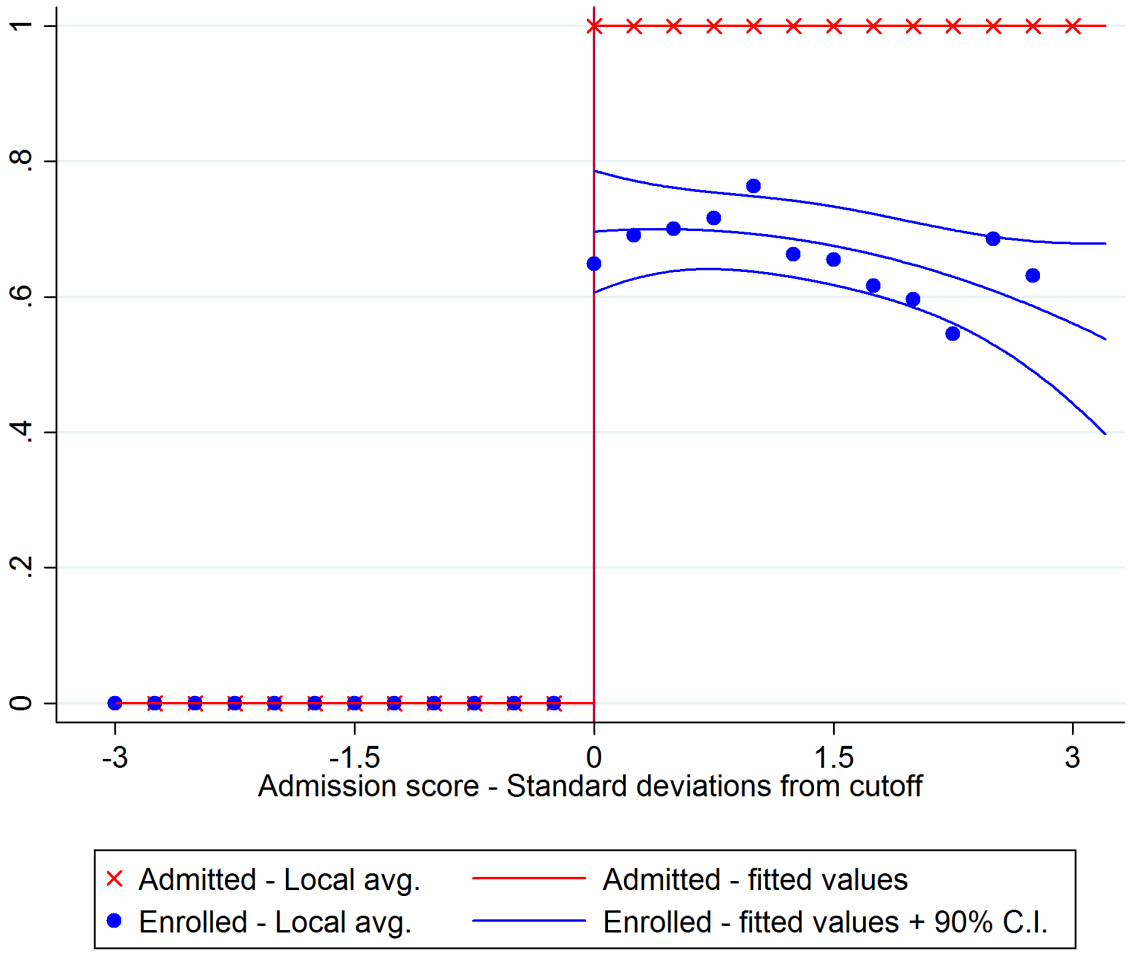
Figures

Figure 1: Dataset structure



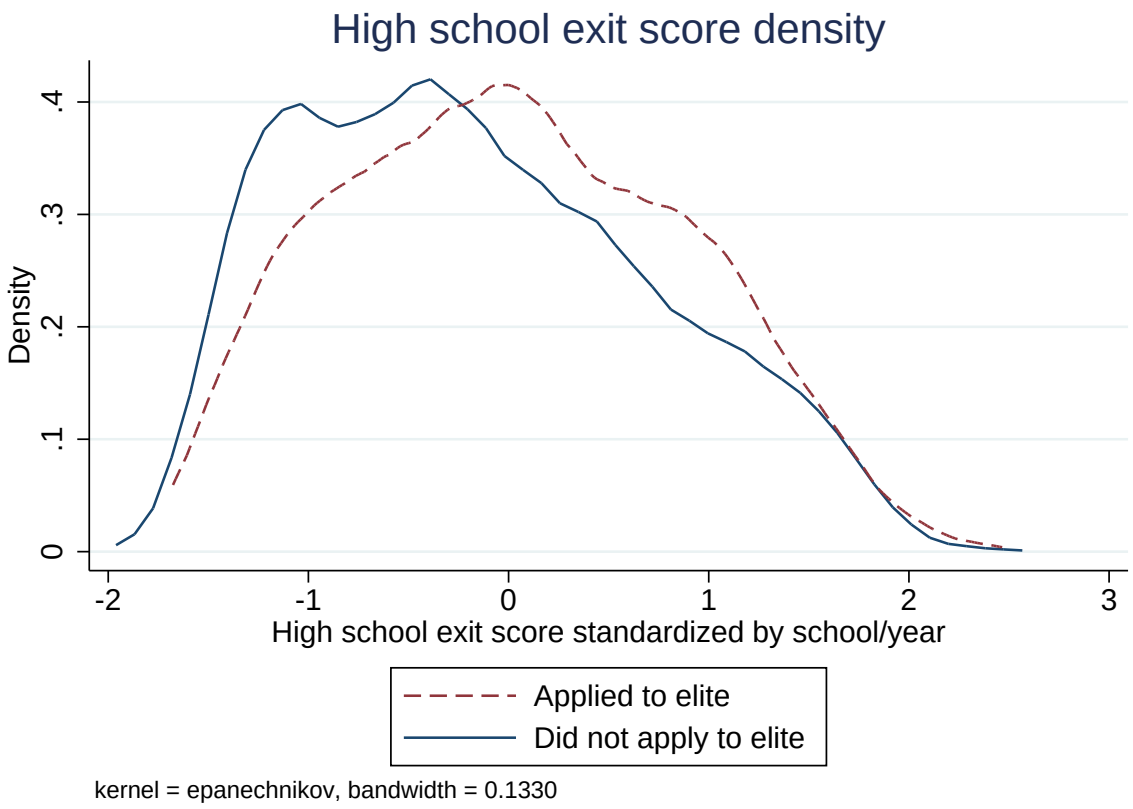
Note: This diagram presents the different sources of administrative data I have assembled. High school administrative records represent the master dataset to I which I have linked the other datasets. The dash-line box highlights the sub-sample of individuals (those who applied to the elite university) on which the regression discontinuity analysis is run.

Figure 2: Discontinuity in the probability of admission and of enrollment at the elite university



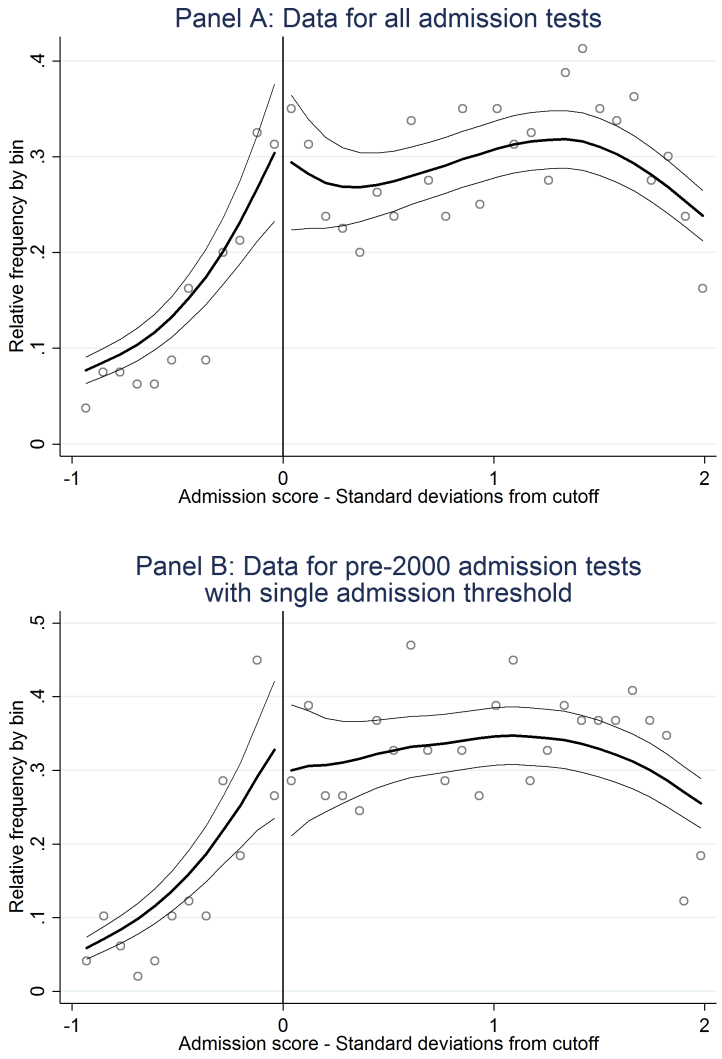
Note: The discontinuity in the probability of being admitted at the elite university is sharp at the admission threshold given the objective and one-dimensional nature of the admission score. However, not every admitted student chooses to enroll. The discontinuity in the probability of enrolling is around 70%. **Running variable:** Composite admission score (0.5*high school exit score + 0.5*admission test) for the elite university in standard deviations from admission threshold. The composite admission score has been standardized by year of admission and re-scaled to be zero at each year's admission threshold
Sample: Students graduating from college-preparatory high schools in Milan applying to elite university

Figure 3: The regression discontinuity sample



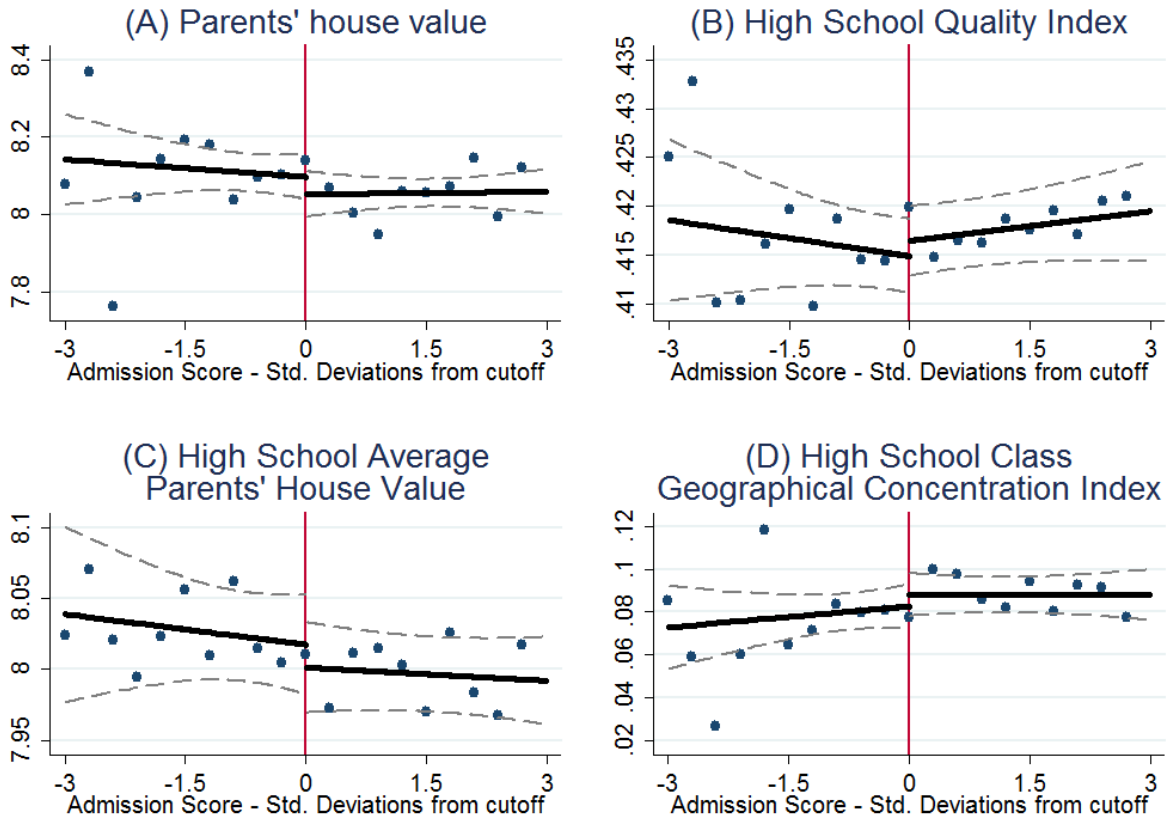
Note: This figure shows the distribution of high school exit scores for students who applied to the elite university (the sub-sample of individuals on which the regression discontinuity analysis is run) and those that did not apply. The dotted line distribution represents students who applied to the elite university, while the solid line represents those who did not apply. Since there are multiple cohorts of students attending different high schools in the data, the high school exit score was standardized by year of high school graduation and school. The distribution for students applying to the elite university is substantially shifted to the right (mean is 0.4 high school exit score standard deviations higher) and has lower variance.

Figure 4: Density test around the admission threshold



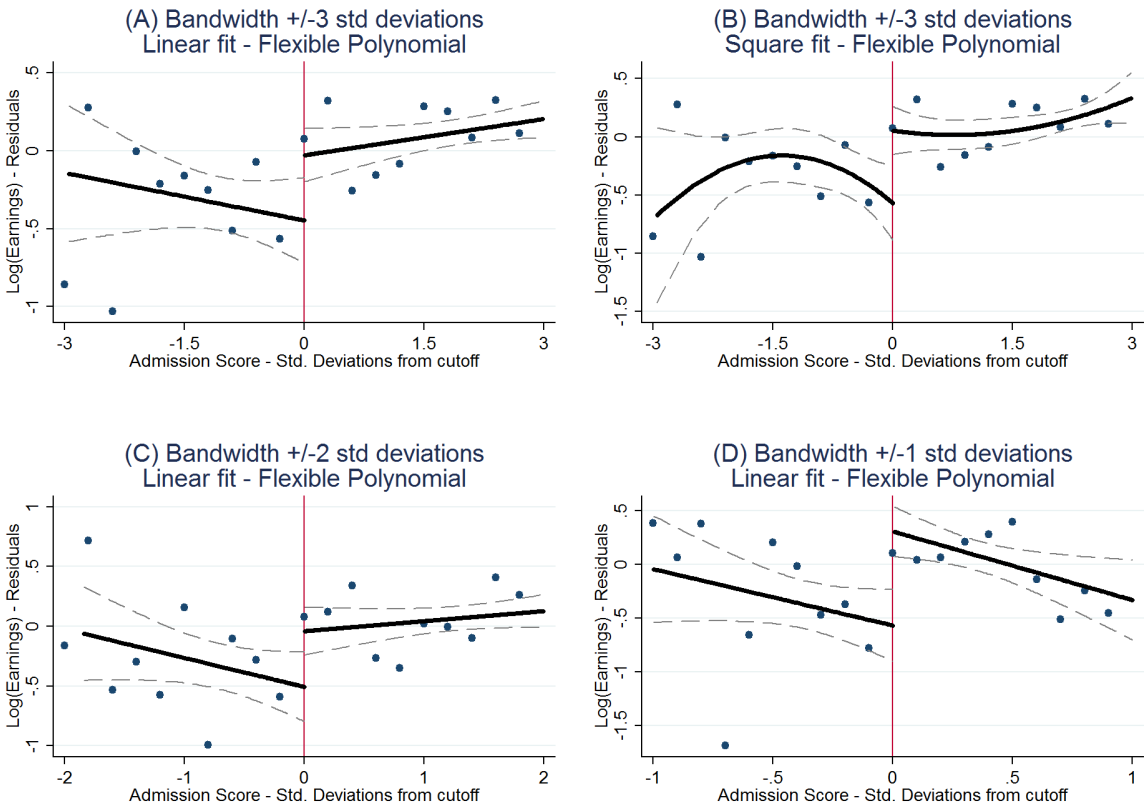
Note: Test for the presence of manipulation of admission score close to the threshold. Density of observations is continuous through the admission threshold. No bunching above the cutoff score.
Sample: In Panel A Students graduating from college-preparatory high schools in Milan applying to elite university between 1995 and 2001. This sample includes all years of admission available, including 2000 and 2001. In these two years, admissions were determined at a degree-specific level. As explained in section 2.2, I have re-scaled each degree-specific threshold to be 0, so that for admission years after 1999 there are several cutoff scores per year and the running variable as to be interpreted as standard deviations from each degree-specific threshold. In Panel B Students graduating from college-preparatory high schools in Milan applying to elite university between 1995 and 1999 when there was a single admission cutoff score for the entire university
Running variable: Composite admission score (0.5*high school exit score + 0.5*admission test) for the elite university in standard deviations from admission threshold. The composite admission score has been standardized by year of admission and re-scaled to be zero at each year's admission threshold

Figure 5: Validity test - smoothness of background characteristics at discontinuity



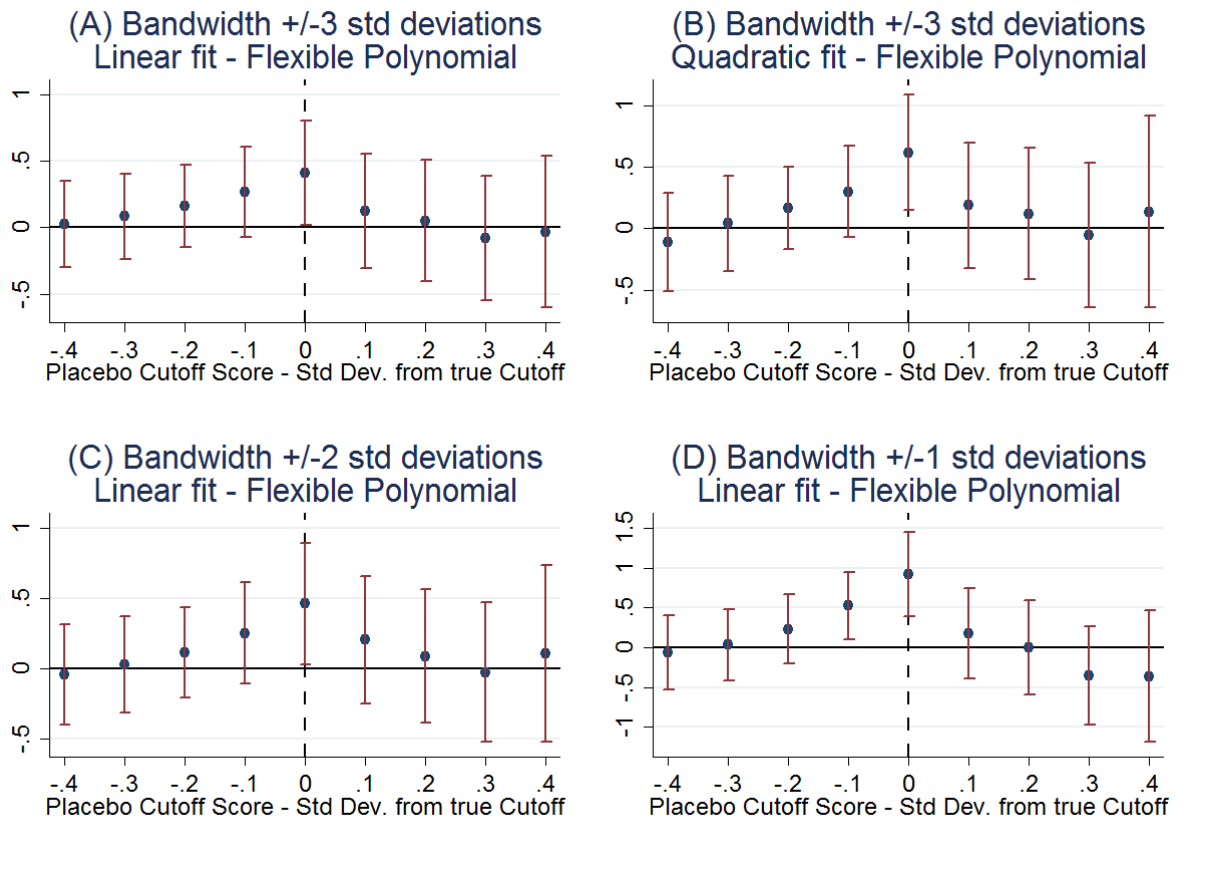
Note: Figures refer to baseline parametric regression discontinuity specification with flexible linear polynomials and bandwidth ±3 standard deviations from admission threshold. Standard errors clustered at high school/cohort level.
Dependent Variables: in Fig. A market value of average house in the neighborhood where the student used to live at time of high school, in Fig. B high school quality index based on average academic performance of students in the standardized national exit score, in Fig. C high school social economic status index based on average family wealth of students attending each school and in Fig. 4 an herfindhal index based on the neighborhoods where classmates used to live by high school class.
Running variable: Composite admission score (0.5*high school exit score + 0.5*admission test) for the elite university in standard deviations from admission threshold. The composite admission score has been standardized by year of admission and re-scaled to be zero at each year's admission threshold
Sample: Students graduating from college-preparatory high schools in Milan applying to elite university

Figure 6: Returns to elite college education



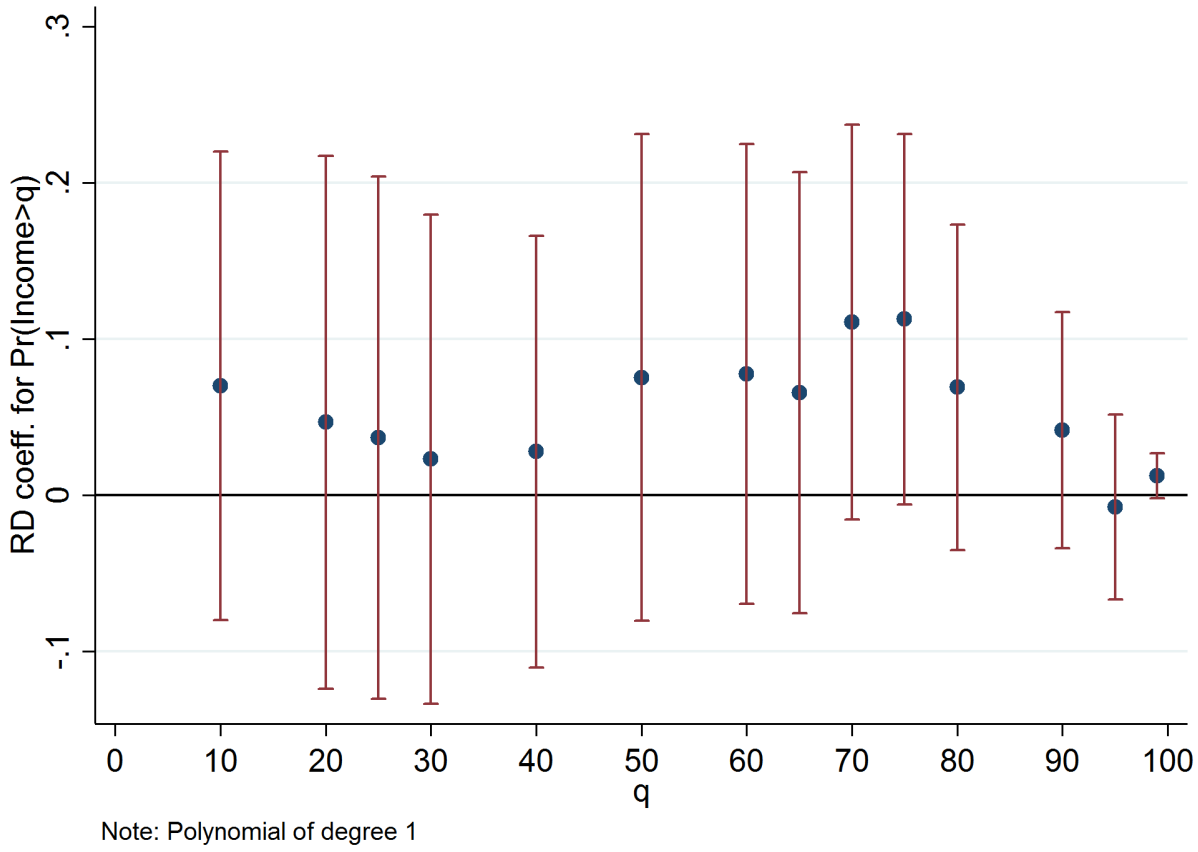
Note: baseline parametric sharp regression discontinuity estimation.
Specifications: linear flexible polynomial with a bandwidth of ±3 standard deviations from threshold for Fig. A, quadratic flexible polynomial with a bandwidth of ±3 for Fig. B, linear flexible polynomial with a bandwidth of ±2 for Fig. C, linear flexible polynomial with a bandwidth of ±1 for Fig. D
Sample: Students graduating from college-preparatory high schools in Milan applying to elite university
Dependent variable: logarithm of personal income (residuals after regression on cohort fixed effects), as revealed to the internal revenue service in year 2006 for year 2005
Running variable: Composite admission score (0.5*high school exit score + 0.5*admission test) for the elite university in standard deviations from admission threshold. The composite admission score has been standardized by year of admission and re-scaled to be zero at each year's admission threshold
Controls: dummy female, a dummy equal to 1 if parents' house value is in the top decile of the house value distribution and another dummy equal to 1 when it is in the bottom decile. A dummy equal to 1 if students used to commute from outside the city of Milan

Figure 7: Estimation with placebo cutoff scores



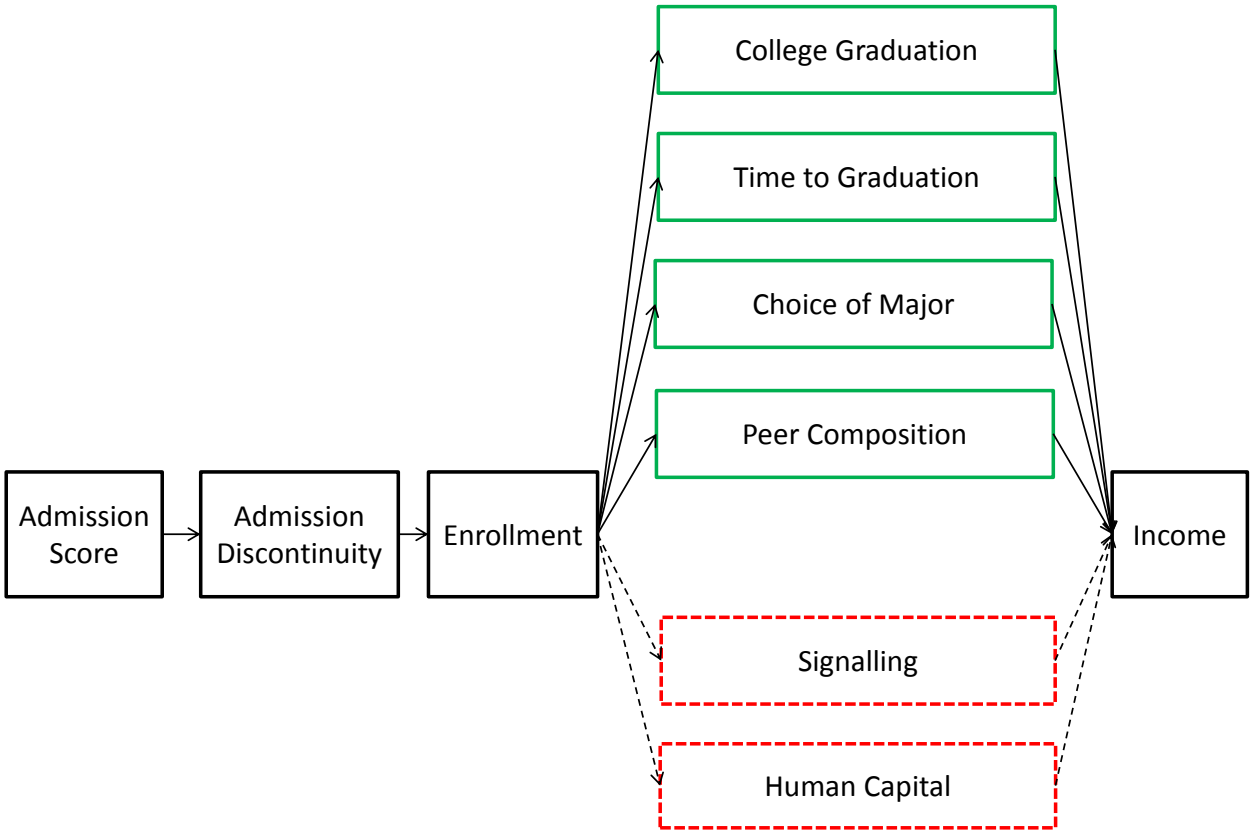
Note: dots represent the estimated discontinuity $\hat{\rho}$ at the true cutoff score and at other placebo cutoff scores 0.1 or more standard deviations away from the true cutoff. Whiskers represent 90% confidence intervals.
Specifications: linear flexible polynomial with a bandwidth of ±3 standard deviations from threshold for Fig. A, quadratic flexible polynomial with a bandwidth of ±3 for Fig. B, linear flexible polynomial with a bandwidth of ±2 for Fig. C, linear flexible polynomial with a bandwidth of ±1 for Fig. D
Sample: Students graduating from college-preparatory high schools in Milan applying to elite university
Dependent variable: logarithm of personal income (residuals after regression on cohort fixed effects), as revealed to the internal revenue service in year 2006 for year 2005
Running variable: Composite admission score (0.5*high school exit score + 0.5*admission test) for the elite university in standard deviations from admission threshold. The composite admission score has been standardized by year of admission and re-scaled to be zero at each year's admission threshold
Controls: dummy female, a dummy equal to 1 if parents' house value is in the top decile of the house value distribution and another dummy equal to 1 when it is in the bottom decile. A dummy equal to 1 if students used to commute from outside the city of Milan

Figure 8: Effect of enrolling at the elite university on the probability of being in the q^{th} quantile of the income distribution



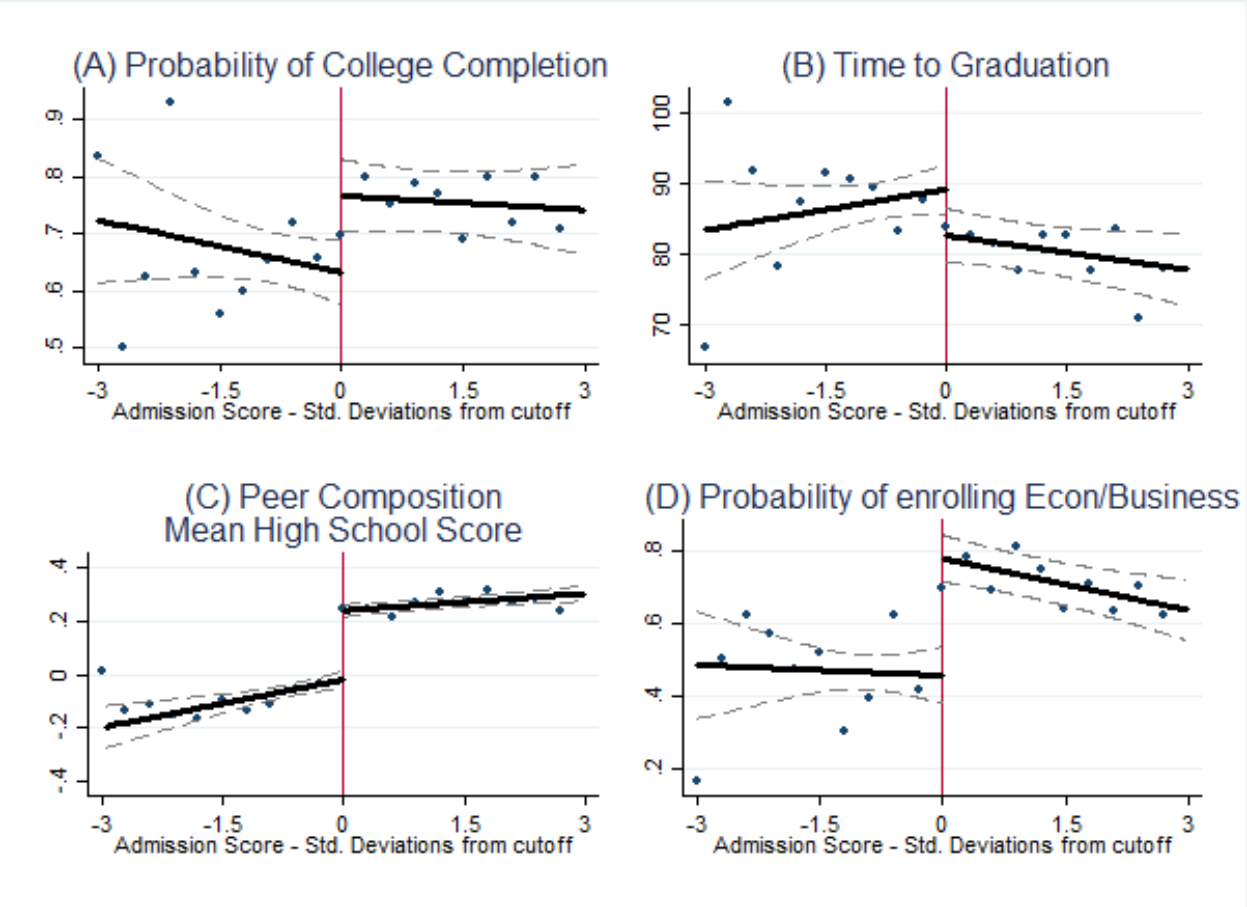
Note: Each dot in the graph represents a different specification in which the left-hand-side is a dummy variable that takes on value one if the income observed from tax returns is above or equal to the q^{th} quantile of the income distribution. Whiskers represent 90% confidence intervals
Sample: Students graduating from college-preparatory high schools in Milan applying to elite university

Figure 9: Mechanisms - identified channels



Note: This diagram is meant to illustrate the mechanisms which I am able to tackle with my data and those (signaling and human capital) that are instead hard to identify with the empirical setting available

Figure 10: Mechanisms



Method: parametric OLS sharp discontinuity with flexible linear polynomials, robust std. errors clustered at high school/cohort level and bandwidth ±3 standard deviations from admission threshold.
Sample: Students graduating from college-preparatory high schools in Milan applying to elite university.
Dependent variables: in Fig. A probability of graduating from college, Fig. B time taken from enrollment to final graduation in months, Fig. C average high school exit score of peers (defined as students in one's cohort who enroll in same university and program), Fig. D probability of enrolling in an economics or business degree after the admission test.
Running variable: Composite admission score (0.5*high school exit score + 0.5*admission test) for the elite university in standard deviations from admission threshold. The composite admission score has been standardized by year of admission and re-scaled to be zero at each year's admission threshold
Controls: dummy female, a dummy equal to 1 if parents' house value is in the top decile of the house value distribution and another dummy equal to 1 when it is in the bottom decile. A dummy equal to 1 if students used to commute from outside the city of Milan.

Tables

Table 1: Summary statistics by university-type

	Elite university	Other universities	Elite/Others
Institutional characteristics[♯]:			
# Students	11841	39092	0.30
# Students / # Instructors	19.25	18.20	1.06
Total expenditures/ # Students ($)	12699	6563	1.93
Avg. yearly expenditure per instructor ($)	83058	73224	1.13
Yearly tuition revenues / # Students ($)	5068	1370	3.70
Background characteristics of students:			
Avg. high school exit score (on a scale of 0-1) (Standardized by year)	0.58	0.40***	1.44
Avg. high school exit score (Percentile of Milan Distribution)	71	48	
Composite SAT Equivalent	1680	1470	
Parents' house value ($\$/m^2$)	3518	3397***	1.04
Parents' monthly income ($)	3256	2859**	1.14
(Mother is college graduate=1)	0.41	0.29*	1.42
Students' outcomes:			
Average yearly income (2005) ($)	45982	24622***	1.87
Median yearly income (2005) ($)	32326	17784***	1.82
Percentile of 2005 National Income Distribution[†]	85	44	
Equivalent income for 2005($) U.S. National Income Dist.[✠]($)	64396	23886	
(Employed=1)	0.948	0.913	1.04
(Top occupation=1)	0.64	0.39***	1.65
(Work abroad=1)	0.16	0.07*	2.47

Note: *** Mean is statistically different from elite university at 99% confidence level, ** at 95% and *at 90%., [♯] Source: Italian Ministry of Education (MIUR), 2001. [†] Source: Ministry of Economics and Finance, Department of Finance (MEF-Dipartimento delle Finanze) 2005. [✠] Source: U.S. Census Bureau, 2005 American Community Survey

Sample: Students graduating from college-preparatory high schools in Milan enrolled in one of the five universities in Milan

Table 2: Comparison with U.S. universities

State	University	Median Income	University	Median Income
NY	Cornell	61900	Hunter Business School	26200
IN	Notre Dame	63100	Indiana Business College	25300
NC	Duke	63500	Durham Technical Comm Coll	29400
Italy	**Elite University**	64396	**Non-Elite Universities**	23886
PA	Carnegie Mellon	64900	Comm Coll of Philadelphia	29000
CA	Caltech	66200	Golden State College	24500
CA	Stanford	68100	Berkeley City College	23700
MA	Harvard	76200	Boston Conservatory	23400
MA	MIT	82100	Salter School	24400

State	University	Avg. SAT	University	Avg. SAT
CA	UC-Santa Cruz	1635	California Baptist Univ., Riverside	1460
Italy	**Elite University**	1680	**Non-Elite Universities**	1470
CA	UC-Irvine	1685	Fresno Pacific Univ.	1479
FL	Florida International Univ.	1710	The Univ. of West Florida, Pensacola	1515
IN	Purdue, Main Campus	1775	Purdue, Fort-Wayne	1465
CA	Stanford	2210	San Diego Christian College	1473
MA	Harvard	2260	Fitchburg State Univ.	1470
MA	MIT	2235	Suffolk University, Boston	1510

Note: this table compares Milan college students with students of US universities. Source of the data is the U.S. Department of Education College Scorecard. In the top panel, I use the U.S.-income-distribution-equivalent income of Milan college students (as calculated in Table 1) and compare it with the income of U.S. college graduates. In the bottom panel I compare the SAT-distribution equivalent score of Milan college students (as calculated in Table 1) with the average SAT score of students attending some comparable U.S. colleges.

Table 3: Discontinuity in predetermined characteristics

	(1) Parents' house value	(2) High School Quality Index	(3) High School Avg. parents' house value	(4) High School Class geographical concentration
Discontinuity at threshold				
$\mathbb{I}(x_{i,t} \geq x_t^*)$	-0.045	0.002	-0.016	0.005
	(0.051)	(0.002)	(0.018)	(0.007)
N. Control	141	220	220	163
N. Treatment	548	724	724	638
Bandwidth	±3 sd	±3 sd	±3 sd	±3 sd
Polynomial	Flexible	Flexible	Flexible	Flexible
y mean	8.066	0.418	8.004	0.086
y std. dev.	0.324	0.020	0.143	0.057

Method: parametric sharp RD estimation with flexible linear polynomials, robust std. errors clustered at high school/cohort level. *** p<0.01, ** p<0.05, * p<0.1.
Sample: high school graduates of Milan applying to Elite University between 1995 and 2001.
Dependent variable: in Column 1 market value of average house in the neighborhood where the student used to live at time of high school, in Column 2 high school quality index based on average academic performance of students in the standardized national exit score, in Column 3 high school social economic status index based on average family wealth of students attending each school and in Column 4 an herfindhal index based on the neighborhoods where classmates used to live by high school class.
Running variable: Composite admission score (0.5*high school exit score + 0.5*admission test) for the elite university in standard deviations from admission threshold. The composite admission score has been standardized by year of admission and re-scaled to be zero at each year's admission threshold
Bandwidth: ±3 standard deviations around admission threshold

Table 4: Returns to elite college education

Dependent Variable: Log of Earnings (Residuals)				
	(1)	(2)	(3)	(4)
	Linear	Quadratic	Linear	Linear
Treatment Effects ↓	+/-3sds	+/-3sds	+/-2sds	+/-1sds
Admission Effect (ITT)	0.411*	0.615**	0.463*	0.971***
	(0.240)	(0.292)	(0.267)	(0.324)
Enrollment Effect	0.515*	0.889**	0.620*	1.493***
	(0.303)	(0.432)	(0.359)	(0.530)
First Stage Estimates:				
Disc. in P(Enrolled)	0.799***	0.691***	0.746***	0.650***
	(0.060)	(0.081)	(0.063)	(0.091)
N. Control	116	116	107	82
N. Treatment	490	490	390	185
Bandwidth	±3 s.d.s	±3 s.d.s	±2 s.d.s	±1 s.d.s
Polynomial	Lin. Flex.	Quad. Flex.	Lin. Flex.	Lin. Flex
Reference Fig.	Fig. 6A	Fig. 6B	Fig. 6C	Fig. 6D

Method: parametric sharp RD estimation for Admission Effect, fuzzy RD estimation (2 stage least squares) for Enrollment Effect, robust std. errors clustered at high school/cohort level. *** p<0.01, ** p<0.05, * p<0.1.
Specifications: linear flexible polynomial with a bandwidth of ±3 standard deviations from threshold for Column 1, quadratic flexible polynomial with a bandwidth of ±3 for Column 2, linear flexible polynomial with a bandwidth of ±2 for Column 3, linear flexible polynomial with a bandwidth of ±1 for Column 4
Dependent variable: logarithm of personal income (residuals after regression on cohort fixed effects), as revealed to the internal revenue service in year 2006 for year 2005.
Running variable: Composite admission score (0.5*high school exit score + 0.5*admission test) for the elite university in standard deviations from admission threshold. The composite admission score has been standardized by year of admission and re-scaled to be zero at each year's admission threshold
Controls: dummy female, a dummy equal to 1 if parents' house value is in the top decile of the house value distribution and another dummy equal to 1 when it is in the bottom decile. A dummy equal to 1 if students used to commute from outside the city of Milan.
Sample: high school graduates of Milan applying to Elite University between 1995 and 2001.

Table 5: Returns to elite college education. Extended analysis with simulated data. All cohorts applying between 1985 and 2000.

	Log of Income	
	(1)	(2)
	Cohorts 1995-2001	All Cohorts 1985-2001
Enrollment Effect	0.515*	0.465***
	(0.303)	(0.159)
N. Control	116	687
N. Treatment	490	1390
Polynomial	Linear	Linear
Bandwidth	±3	±3

Method: in Column 1 fuzzy RD estimation with linear flexible polynomial and a bandwidth of ±3 standard deviations from threshold. In Column 2 bootstrap simulation method with parametric fuzzy RD estimation and flexible polynomials, 100 replications as described in the Technical Appendix A. Robust std. errors clustered at high school/cohort level. *** p<0.01, ** p<0.05, * p<0.1
Sample: In Column 1, high school graduates of Milan applying to Elite University between 1995 and 2001 for which precise admission data are available. In Column 2 high school graduates of Milan applying to Elite University between 1985 and 2001
Dependent variable: logarithm of personal income (residuals after regression on Admission Session Fixed Effects), as revealed to the internal revenue service in year 2006 for year 2005
Running variable: Standardized Composite admission score (0.5*high school exit score + 0.5*admission test) for individuals applying to elite university between 1995 and 2001 and High School Exit Score standardized at high school and cohort level for students applying between 1985 and 1995
Controls: dummy female, a dummy equal to 1 if parents' house value is in the top decile of the house value distribution and another dummy equal to 1 when it is in the bottom decile. A dummy equal to 1 if students used to commute from outside the city of Milan
Bandwidth: ±3 standard deviations from admission threshold

Table 6: Mechanisms

Dependent Variable: Log of Earnings (Residuals)				
	(1)	(2)	(3)	(4)
VARIABLES	P(Graduated=1)	Time to Graduation	Mean H.S.score of university peers	P(Econ/Business=1)
Discontinuity at threshold	0.136**	-6.633*	0.260***	0.326***
$\mathbb{I}(x_{i,t} \geq {x_t}^*)$	(0.061)	(3.609)	(0.032)	(0.073)
N. Control	220	220	188	220
N. Treatment	724	724	618	724
Bandwidth	±3 sd	±3 sd	±3 sd	±3 sd
Polynomial	Flexible	Flexible	Flexible	Flexible
y mean	0.732	82.009	0.190	0.655
y std. dev.	0.443	28.636	0.229	0.475

Method: parametric OLS sharp discontinuity with flexible linear polynomials, robust std. errors clustered at high school/cohort level. *** p<0.01, ** p<0.05, * p<0.1.
Sample: high school graduates of Milan applying to Elite University between 1995 and 2001.
Dependent variables: in Column 1 probability of graduating from college, Column 2 time taken from enrolment to final graduation in months, Column 3 average high school exit score of peers (defined as students in one's cohort who enroll in same university and program), Column 4 probability of enrolling in an economics or business degree after the admission test.
Running variable: Composite admission score (0.5*high school exit score + 0.5*admission test) for the elite university in standard deviations from admission threshold. The composite admission score has been standardized by year of admission and re-scaled to be zero at each year's admission threshold
Controls: dummy female, a dummy equal to 1 if parents' house value is in the top decile of the house value distribution and another dummy equal to 1 when it is in the bottom decile. A dummy equal to 1 if students used to commute from outside the city of Milan.
Bandwidth: ±3 standard deviations from admission threshold.

Table 7: Decomposing the effect of choice of major

Earlier cohorts		R.D. sample			
	Log(Income)	Share of not admitted students by alternative major choice	Average weighted returns to major given alternative major choices of non-admitted students	Estimated elite premium	"Net effect"
Architecture	-0.973***	0.048	-0.19	0.52	**0.33**
	(0.067)				
Law	-0.364***	0.175			
	(0.060)				
Engineering	-0.227***	0.037			
	(0.046)				
Math/Phyisics	-0.356***	0.027			
	(0.086)				
Medicine	-0.077	0.011			
	(0.071)				
Social Sciences	-0.438***	0.037			
	(0.073)				
Humanities	-0.648***	0.0582			
	(0.073)				
Agricultural Sciences	-0.613***	0.016			
	(0.113)				
Natural Sciences	-0.428***	0			
	(0.071)				
Observations	6,781				
R-squared	0.219				

Note: Column 1 shows returns to choice of majors relative to returns to choosing Economics and Business. Column 2 reports the share of students in the regression discontinuity sample (i.e. later cohorts for which I have estimated the quasi-experimental estimates of returns to elite college education) that I observe choosing each major after missing admission to the elite university. **Method**: In Column 1 OLS, robust std. errors clustered at high school/cohort level. Business/Economics is the omitted variable. *** p<0.01, ** p<0.05, * p<0.1.
Sample: for Column 1 high school students in Milan graduating between 1985 and 1994 (Earlier cohorts with respect to those used for the main regression discontinuity analysis). For Column 2 students who applied to the elite university, but did not get admitted.
Dependent variables: Log of Income in 2005 as reported in Tax Returns
Controls: university fixed effects, female, potential experience, high school exit score, high school/cohort fixed effects, teacher fixed effects.

Table 8: Discounted Cumulative Net Returns

		(A) Elite	(B) Non-elite	(A)-(B)	(A)/(B)
	Yearly tuition	\$5068	\$1370		
	Average time to graduation (years)	5.87	6.42		0.9✠
(1)	Total attendance cost	\$29747	\$8795	\$20953	3.4
	Yearly income (Average for the first 15 years after graduation)	\$39198	\$24622	\$16793	$e^{0.47\dagger}$
	Discount rate	0.03	0.03		
(2)	Expected cumulated premium 15 years after graduation			\$140341	
(2)-(1)	Cumulative net return to elite college education 15 years after graduation			\$119389	

Note:† Elite college premium as estimated in RD log(income) specification for all cohorts available. Tuition fees are net of scholarships and tuition fee remissions. ✠ The difference in time to degree is calculated based on the RD estimate obtained in Table 6.

A Technical Appendix - Numerical simulation for extending estimation to older cohorts

As explained in Section 3, data from admission office of the elite university have been matched only for cohorts of students in my sample who graduated from high school between 1995 and 2005. For the cohorts of students graduating from high school between 1985 and 1994 I know only partial admission information: I know which university they attended, their high school exit score (that determines admission score with a weight of 50%), while I miss the score obtained in the elite university admission test. Moreover I do not know which students applied to the elite university but were not admitted.

In this appendix I describe the numerical simulation method used to combine partial information for older cohorts with the complete data available for earlier cohorts to produce estimates of returns to elite education for the full sample of cohorts graduating from high school between 1985 and 2001. Since income is available for year 2005 only, pooling all cohorts together not only substantially increases the number of observations, but also allows to estimates returns for individuals that have spent up to 15 years on the labor market after graduation.

Numerical simulation

Let i be the index for one of the 30,000 individuals in my sample graduating from high school in Milan between 1985 and 2005. t ($1985 \leqslant t \leqslant 2005$) represents the year of high school graduation for individual i and also the year in which i takes university admission tests.

For each individual i I define four different statuses defining her/his relationship with the elite university under study:

$Applied_i$: Indicator=1 if student i is observed applying to the elite university

$Admitted_i$: Indicator=1 if student i is admitted at the elite university

$Enrolled_i$: Indicator=1 if student i is observed enrolling in the elite university

$Graduated_i$: Indicator=1 if student i is observed graduating from elite university

Following the dataset structure I categorize individuals in five types:

1. Type 0000: ($Applied_i$=0; $Admitted_i$=0; $Enrolled_i = 0$; $Graduated_i = 0$) students graduating from one of the high schools sampled in the data who did not apply at the elite university and hence did not take the test.

2. Type 1000: ($Applied_i$=1; $Admitted_i$=0; $Enrolled_i = 0$; $Graduated_i = 0$) students applying at the elite university, but failed the test.

3. Type 1100: ($Applied_i$=1; $Admitted_i$=1; $Enrolled_i = 0$; $Graduated_i = 0$) students applying at the elite university who passed the test, but decided to enroll in a different university.

4. Type 1110: ($Applied_i$=1; $Admitted_i$=1; $Enrolled_i = 1$; $Graduated_i = 0$) students applying at the elite university who passed the test, enrolled at the elite university, but dropped out from elite university.

5. Type 1111: ($Applied_i$=1; $Admitted_i$=1; $Enrolled_i = 1$; $Graduated_i = 1$) students graduating from elite university.

The variables involved in the numerical simulation are defined as follows:

y_i: Logarithm of annual income observed in year 2005 for individual i

$x^h s_{i,t}$: High school exit score for individual i completing high school in year t (observed for all cohorts in the sample)

$x^u{}_{i,t}$: Elite university admission test score for individual i in admission session of year t (observed for individuals applying to elite university after 1994, $Applied_i$=1)

$x_{i,t} = \frac{1}{2}x^h s_{i,t} + \frac{1}{2}x^u{}_{i,t}$: Composite uni-dimensional university admission score determining admission (observed for individuals applying to elite university after 1994, $Applied_i$=1)

${x_t}^*$: Admission threshold for admission session of year t such that individual i is admitted at elite university if $x_{i,t} \geq {x_t}^*$

For individuals i graduating in $1995 \leq t \leq 2005$ all above variables are available.

For individuals i graduating in $1985 \leq t < 1995$ I still observe ${x^h}{s_{i,t}}$, $Enrolled_i = 1$, $Graduated_i = 1$ and y_i, but the following variables are missing: $Applied_i$, $Admitted_i$, ${x^u}_{i,t}$, $x_{i,t}$, ${x_t}^*$

The empirical issues arising from the absence of these variables for the older cohorts can be summarized as follows:

Missing variable problem 1: I observe $Enrolled_i$, $Graduated_i$. I do not observe $Admitted_i = 0$. I can infer $\mathbb{I}(Applied_i = 1|Enrolled_i = 1)$ and $\mathbb{I}(Admitted_i = 1|Enrolled_i = 1)$. I need to estimate $\mathbb{I}(Applied_i = 1|Enrolled_i = 0)$

Estimation step 1: estimate $\mathrm{P}(Applied_i = 1|Enrolled_i = 0,\ {x^h}_{i,t},$ Pre college characteristics Z_i,$t)$ for $1995 \leq t \leq 2005$ and predict them out of sample to older individuals graduating from high school in $1985 \leq t \leq 1994$. Use $P(Applied_i = 1|\widehat{Enrolled}_i = 0, {x^h}_{i,t}, Z_i)$ as individual weights to draw $B = 100$ times a sample of individuals predicted as applying to elite university conditional on observing them not enrolling in the elite university: $\mathbb{I}_{\mathbb{b}}(\widehat{Applied_{i,b}} = 1|Enrolled_{i,b} = 0)$ with b being each of 100 repetitions in the simulation. What follows is also repeated at each repetition, however I drop the subscript b to keep notation simple.

Missing variable problem 2: I do not observe university admission test score ${x^u}_{i,t}$ and as a consequence the admission composite score $x_{i,t}$ determined for 50% by the admission test score

Estimation step 2: I use high school exit score ${x^{hs}}_{i,t}$ (which determines 50% of the overall ad-

mission score) as running variable for running my regression discontinuity at each repetition b.

Missing variable problem 3: I do not observe the admission threshold x_t^*

Estimation step 3: I pool the individuals predicted as applied at the elite university conditional on observing them not enrolling $\mathbb{I}(\widehat{Applied_i} = 1|Enrolled_i = 0)$ and those who are observed enrolling $Enrolled_i = 1$ for 1985$\leq t \leq$ 1994. For each t and for each possible threshold $min(x^{hs}{}_{i,t}) < x^{hs*}_{t,k} < max(x^{hs}_{i,t})$ I estimate:
$Enrolled_i = 1 = \lambda + \mathbb{I}_i(x^{hs}_{i,t} \geq x^{hs}{}^*_{t,k}) = 1) + \epsilon_i$ and retrieve $\widehat{R^2_k}$. Then at each repetition b I choose threshold $\widehat{x^{hs*}_{t,b}} = x^{hs}{}^*_{t,k}$ such that $\widehat{R^{2*}} = max(\widehat{R^2_k})$ (i.e. the threshold that best explain the probability of admission)

Once the control group of not admitted applicants to the elite university and admission thresholds have been estimated for the older cohorts, at each repetition b I recombine the data as follows. For cohorts graduating from high school after year t=1994 I keep the original data with complete admission information. For older cohorts of students applying to elite university between 1985 and 1994 I include individuals that have been observed enrolling at the elite university as the treatment group and I take as the control group those individuals predicted as not admitted applicants.

After pooling these three sub-smaples together at each repetition b of the numerical simulation, I estimate my main regression discontinuity specifications:
First Stage

$$p(Enr_i = 1) = \phi + \pi\mathbb{I}(\widetilde{x_{i,t}} \geq \widetilde{x_t{}^*}) + \theta f(\widetilde{x_{i,t}} - \widetilde{x_t{}^*}) + \delta f(\widetilde{x_{i,t}} - \tilde{x}_t{}^*) * \mathbb{I}(\widetilde{x_{i,t}} \geq \widetilde{x}_t{}^*) + \eta_i$$

Second Stage

$$log(y_i) = \alpha + \tau_b\, p(\widehat{Enr_i} = 1) + \beta f(\widetilde{x_{i,t}} - \widetilde{x_t{}^*}) + \gamma f(\widetilde{x_{i,t}} - \widetilde{x_t{}^*}) * \mathbb{I}(\widetilde{x_{i,t}} \geq \widetilde{x_t{}^*}) + \epsilon_i$$

with

$$\widetilde{x_{i,t}} = \begin{cases} {x^{hs}}_{i,t} & \text{(High school exit score) if } 1985 < t < 1994, \\ x_{i,t} & \text{(Composite admission score) if } 1995 < t < 2001. \end{cases}$$

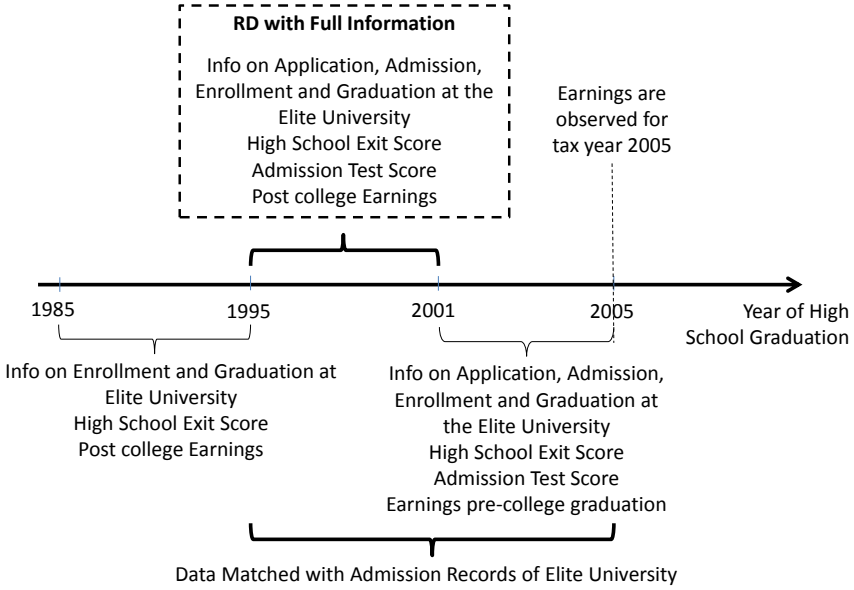
and

$$\widetilde{{x_t}^*} = \begin{cases} \widehat{{x^{hs}}^*_t} & \text{(Estimated admission threshold) if } 1985 < t < 1994, \\ {x_t}^* & \text{(Observed admission threshold) if } 1995 < t < 2001. \end{cases}$$

At each repetition b I then store the estimate of the return to enrolling in the elite university $\hat{\tau}_b$. After repeating estimation steps 1,2 and 3 a hundred times to retrieve $\hat{\tau}_b$ I estimate the expected return to elite education as $\hat{\tau} = \frac{\sum_{b=1}^{100} \hat{\tau}_b}{100}$ and use the simulated standard deviation of $\hat{\tau}_b$ to perform inference on the estimate of interest $\hat{\tau}$.

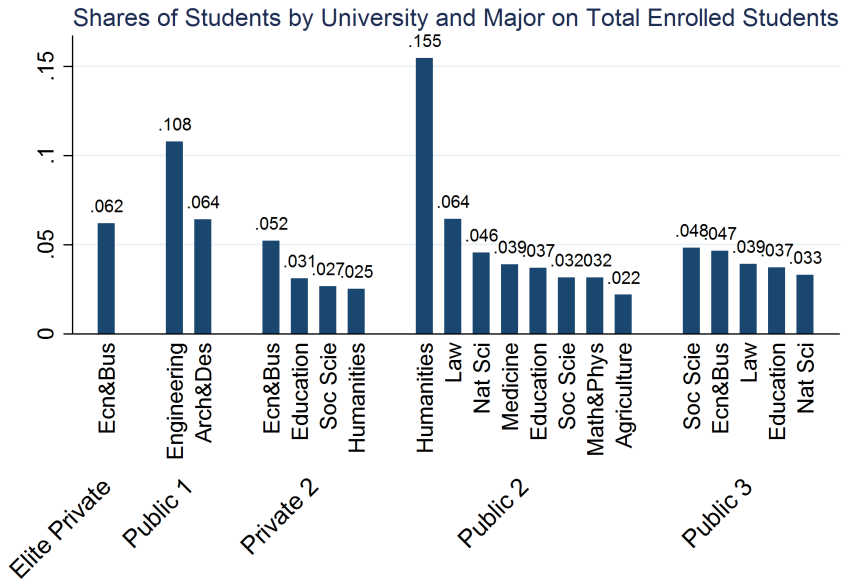
B Appendix Figures and Tables

Figure B1: Admission and income data availability



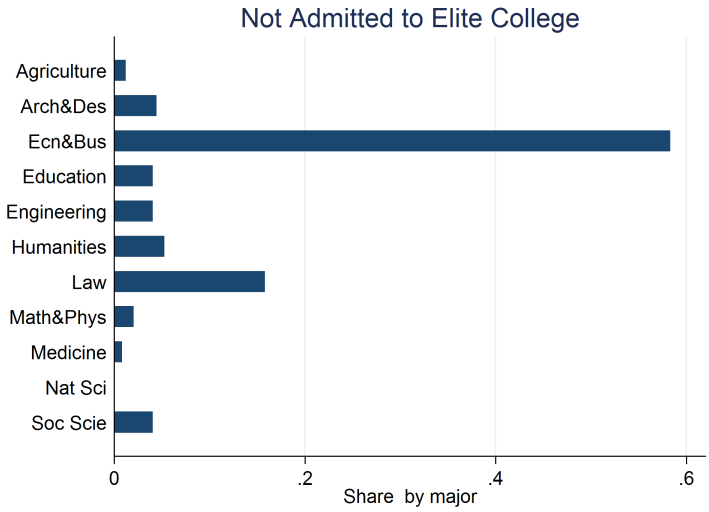
Note: This diagram summarizes availability of income and admission data across cohorts of high school graduates. Administrative data have been matched to individual admission data of the elite university only for cohorts graduating from high school and applying to college between 1995 and 2005. Yearly income is available for all cohorts only for the tax year 2005. For older cohorts only high school performance, college enrollment and graduation are available.

Figure B2: College institutional setting



Note: Breakdown of total students enrolled in college by university and field of study. This diagram shows which degrees are available by university and the relative number of students enrolled in each degree.
Sample: Students graduating from college-preparatory high schools in Milan and enrolling in one of the 5 main universities based in the same city

Figure B3: Choice of major of students not admitted



Note: Share of students who applied to elite university and were not admitted by major chosen at other universities.

Table B1: Discontinuity in predetermined characteristics

Specifications ↓	(1) Parents house value	(2) High School Quality Index	(3) High School Avg. parents house value	(4) High School Class geographical concentration
Bandwidth ±2				
$\mathbb{I}(x_{i,t} \geq x_t^*)$	-0.021	0.001	-0.014	0.005
	(0.061)	(0.003)	(0.021)	(0.008)
N. Control	130	197	197	150
N. Treatment	436	585	585	512
Bandwidth ±1				
$\mathbb{I}(x_{i,t} \geq x_t^*)$	0.052	0.004	-0.004	-0.004
	(0.075)	(0.004)	(0.030)	(0.011)
N. Control	97	137	137	115
N. Treatment	201	273	273	236
Polynomial	Linear Flexible	Linear Flexible	Linear Flexible	Linear Flexible
y mean	8.066	0.418	8.004	0.086
y std. dev.	0.324	0.020	0.143	0.057

Method: parametric sharp discontinuity with flexible linear polynomials, robust std. errors clustered at high school/cohort level. *** p<0.01, ** p<0.05, * p<0.1.
Sample: high school graduates of Milan applying to Elite University between 1995 and 2001. Dependent variable: in Column 1 market value of average house in the neighborhood where the student used to live at time of high school, in Column 2 high school quality index based on average academic performance of students in the standardized national exit score, in Column 3 high school social economic status index based on average family wealth of students attending each school and in Column 4 an herfindhal index based on the neighborhoods where classmates used to live by high school class.
Running variable: Composite admission score (0.5*high school exit score + 0.5*admission test) for the elite university in standard deviations from admission threshold. The composite admission score has been standardized by year of admission and re-scaled to be zero at each year's admission threshold
Bandwidth: top panel ±2 standard deviations around admission threshold, bottom panel ±1 standard deviations around admission threshold

Table B2: Discontinuity in predetermined characteristics and relevant variables

	(1) Scientific High School=1	(2) Female=1	(3) Missing Income=1	(4) Zero Income=1
Discontinuity at threshold				
$\mathbb{I}(x_{i,t} \geq {x_t}^*)$	0.079	-0.026	-0.086	0.043
	(0.063)	(0.056)	(0.058)	(0.038)
N. Control	220	220	220	220
N. Treatment	724	724	724	724
Bandwidth	±3 sd	±3 sd	±3 sd	±3 sd
Polynomial	Linear	Linear	Linear	Linear
	Flexible	Flexible	Flexible	Flexible
y mean	0.6	0.4	0.2	0.1
y std. dev.	0.5	0.5	0.4	0.3

Method: parametric sharp discontinuity with flexible linear polynomials, robust std. errors clustered at high school/cohort level. *** p<0.01, ** p<0.05, * p<0.1.
Sample: high school graduates of Milan applying to Elite University between 1995 and 2001. Dependent variable: in Column 1 dummy=1 if students attended high school with scientific curriculum and zero if classical studies curriculum, in Column 2 dummy female, in Column 3 dummy equal to 1 if tax returns record have not been matched, in Column 4 dummy equal to 1 if income was zero in tax returns.
Running variable: Composite admission score (0.5*high school exit score + 0.5*admission test) for the elite university in standard deviations from admission threshold. The composite admission score has been standardized by year of admission and re-scaled to be zero at each year's admission threshold
Bandwidth: ±3 standard deviations around admission threshold

Table B3: Returns to elite college education - Excluding controls

Dependent Variable: Log of Earnings (Residuals)				
	(1)	(2)	(3)	(4)
	Linear	Quadratic	Linear	Linear
VARIABLES	+/-3sds	+/-3sds	+/-2sds	+/-1sds
Admission Effect (ITT)	0.418*	0.623**	0.466*	0.875***
	(0.240)	(0.283)	(0.263)	(0.318)
Enrollment Effect	0.524*	0.912**	0.624*	1.338***
	(0.306)	(0.425)	(0.357)	(0.503)
First Stage Estimates:				
Disc. in P(Enrolled)	0.797***	0.683***	0.747***	0.653***
	(0.058)	(0.084)	(0.060)	(0.096)
N. Control	116	116	107	82
N. Treatment	490	490	390	185
Bandwidth	±3 s.d.s	±3 s.d.s	±2 s.d.s	±1 s.d.s
Polynomial	Lin. Flex.	Quad. Flex.	Lin. Flex.	Lin. Flex

Method: parametric sharp RD estimation for Admission Effect, fuzzy RD estimation for Enrollment Effect, robust std. errors clustered at high school/cohort level. *** p<0.01, ** p<0.05, * p<0.1.
Specifications: linear flexible polynomial with a bandwidth of ±3 standard deviations from threshold for Column 1, quadratic flexible polynomial with a bandwidth of ±3 for Column 2, linear flexible polynomial with a bandwidth of ±2 for Column 3, linear flexible polynomial with a bandwidth of ±1 for Column 4.
Sample: high school graduates of Milan applying to Elite University between 1995 and 2001. Dependent variable: logarithm of personal income (residuals after regression on Admission Session Fixed Effects), as revealed to the internal revenue service in year 2006 for year 2005.
Running variable: Composite admission score (0.5*high school exit score + 0.5*admission test) for the elite university in standard deviations from admission threshold. The composite admission score has been standardized by year of admission and re-scaled to be zero at each year's admission threshold

Table B4: Mechanisms

	(1)	(2)	(3)	(4)
	P(Graduated=1)	Time to Graduation	Experience	P(Econ/Business=1)
Bandwidth ±2				
$\mathbb{I}(x_{i,t} \geq {x_t}^*)$	0.136**	0.461	-0.768	0.335***
	(0.068)	(2.730)	(6.070)	(0.075)
N. Control	197	130	184	197
N. Treatment	585	438	575	585
Bandwidth ±1				
$\mathbb{I}(x_{i,t} \geq {x_t}^*)$	0.091	-0.665	1.698	0.254**
	(0.091)	(3.348)	(7.005)	(0.106)
N. Control	137	90	127	137
N. Treatment	273	204	269	273
Polynomial	Linear Flexible	Linear Flexible	Linear Flexible	Linear Flexible
y mean	0.732	67.432	32.244	0.655
y std. dev.	0.443	18.111	37.073	0.475

Method: parametric sharp discontinuity with flexible linear polynomials, robust std. errors clustered at high school/cohort level. *** p<0.01, ** p<0.05, * p<0.1.
Sample: high school graduates of Milan applying to Elite University between 1995 and 2001.
Dependent variables: in Column 1 probability of graduating from college, Column 2 time taken from enrolment to final graduation in months, Column 3 average high school exit score of peers (defined as students in one's cohort who enroll in same university and program), Column 4 probability of enrolling in an economics or business degree after the admission test.
Running variable: Composite admission score (0.5*high school exit score + 0.5*admission test) for the elite university in standard deviations from admission threshold. The composite admission score has been standardized by year of admission and re-scaled to be zero at each year's admission threshold
Bandwidth: top panel ±2 standard deviations from admission threshold, bottom panel ±1 standard deviations from admission threshold.