Wp

WORKING PAPERS

BENEFITS TO ELITE SCHOOLS AND THE EXPECTED RETURNS TO EDUCATION: EVIDENCE FROM MEXICO CITY
10/03/2017
N° 2017/02

BENEFITS TO ELITE SCHOOLS AND THE EXPECTED RETURNS TO EDUCATION: EVIDENCE FROM MEXICO CITY

Estrada, Ricardo
Gignoux, Jérémie

DEVELOPMENT BANK OF LATIN AMERICA

WORKING PAPERS

DEVELOPMENT BANK OF LATIN AMERICA

BENEFITS TO ELITE SCHOOLS AND THE EXPECTED RETURNS TO EDUCATION: EVIDENCE FROM MEXICO CITY

Estrada, Ricardo
Gignoux, Jérémie

CAF – Working paper N° 2017/02
10/03/2017

ABSTRACT

We exploit data on the future earnings students at high school completion expect to receive with and without a college education, together with information on learning achievement and college outcomes, to study the benefits from admission into a system of elite public high schools in Mexico City. Using data for the centralized allocation of students into schools and an adapted regression discontinuity design strategy, we estimate that elite school admission increases the future earnings and returns students expect from a college education. These gains in earnings expectations seem to reflect improvement in actual earnings opportunities, as admission to this elite school system also enhances learning achievement and college graduation outcomes. This provides evidence of the earnings benefits from attending elite schools.

LOS BENEFICIOS DE LAS ESCUELAS DE ÉLITE Y LOS RETORNOS ESPERADOS DE LA EDUCACIÓN: EVIDENCIA DE LA CIUDAD DE MÉXICO

Estrada, Ricardo
Gignoux, Jérémie

CAF - Documento de trabajo N° 2017/02
10/03/2017

RESUMEN

En este artículo, estudiamos los beneficios de ser admitido a un sistema público de escuelas de élite del nivel medio superior (grados 10 a 12) en la Ciudad de México. Para ello, combinamos datos sobre las expectativas salariales a futuro que tienen los estudiantes cuando terminan la educación secundaria con información sobre el logro académico en este nivel y la, eventual, obtención de un título universitario. Para identificar un efecto causal, utilizamos un diseño adaptado de regresiones discontinuas que explota el sistema centralizado de asignación de estudiantes a escuelas. Nuestros resultados muestran que ser admitido a este sistema de escuelas de élite aumenta las expectativas que los estudiantes tienen sobre sus salarios futuros y los retornos que esperan obtener de una educación universitaria. Este incremento en los salarios esperados parece ser indicativo de un verdadero aumento en las oportunidades salariales, en tanto la admisión a este sistema de escuelas tiene un efecto positivo también sobre el logro académico y la trayectoria de estudios universitarios. En conjunto, estos resultados proveen evidencia sobre los beneficios que tiene la asistencia a un sistema de escuelas de élite.

Benefits to Elite Schools and the Expected Returns to Education: Evidence from Mexico City

March 10, 2017

Ricardo Estrada[1] and Jérémie Gignoux[2]

Abstract

We exploit data on the future earnings students at high school completion expect to receive with and without a college education, together with information on learning achievement and college outcomes, to study the benefits from admission into a system of elite public high schools in Mexico City. Using data for the centralized allocation of students into schools and an adapted regression discontinuity design strategy, we estimate that elite school admission increases the future earnings and returns students expect from a college education. These gains in earnings expectations seem to reflect improvement in actual earnings opportunities, as admission to this elite school system also enhances learning achievement and college graduation outcomes. This provides evidence of the earnings benefits from attending elite schools.

JEL: D83, D84, I21.

Keywords: elite high schools, earnings expectations, returns to education, college graduation.

[1]Estrada: CAF-Development Bank of Latin America, Avenida Madero 900, Piso 15, Torre Catalinas Plaza, Ciudad de Buenos Aires, Argentina, email: restrada@caf.com.

[2]Gignoux (corresponding author): Paris School of Economics and French National Institute for Agricultural Research, 48 boulevard Jourdan, 75014 Paris, France, email: jeremie.gignoux@psemail.eu.

1 Introduction

The benefits of attending a selective high school are under debate. At least, no clear pattern emerges from the literature on the effect of selective secondary schools on academic achievement. While selective schools seem to produce learning gains in some low- and middle-income countries (see Pop-Eleches and Urquiola, 2013 using data from Romania, Jackson, 2010 for Trinidad and Tobago, and Ding and Lehrer, 2007 for China), this is not the case in others (Lucas and Mbiti, 2014 for Kenya). And most studies in developed countries find no benefits in terms of learning achievements (for the United States, see Cullen et al., 2006 Abdulkadiroglu et al., 2014, Dobbie and Fryer, 2014 and Bui et al. (2011); and Clark, 2010 for the United Kingdom). [3]

These mixed results are at odds with the sustained demand, from parents and students, for selective high schools - which itself generates school selectivity. Such demand has been linked to parents' overestimation of peer effects (Abdulkadiroglu et al., 2014) or valuation of different outcomes, such as crime (Cullen et al., 2006). However, Clark and Bono (2016) assert that, despite having modest effects on learning achievements, elite schools might produce gains in the long run in terms of college graduation and labor income (and maybe also in marriage). This motivates a more comprehensive investigation of the benefits associated to elite high schools. In this line, Abdulkadiroglu et al., 2014 and Dobbie and Fryer, 2014 find little effects on college enrollment of New York and Boston exam schools, while Clark and Bono (2016) find large positive effects on the years of education completed and on the employment and income of women who attended elite secondary schools in the UK in the 1960s. That results vary across contexts suggests that access to different elite school systems have different effects on access to school inputs. For instance, US exam schools may only provide more selected peers, while elite secondary schools in the UK of 1960s or contemporaneous schools in developing countries

[3] Attention has also been devoted to the effects on school completion, and here again the evidence is mixed. There is no clear-cut prediction of the relationship between elite school attendance and high school completion, e.g. more selective schools might increase the returns from attendance but also require higher immediate efforts and investments from students. Using the same dataset that we use but a different sample, including students living further away from selective schools (we return to this in section 3.4), Dustan et al. (2015) find that admission to elite schools increases drop out - however, we find that these effects do not extend to students from Mexico City. Similarly, other studies do not find evidence of effects on school completion (e.g. Pop-Eleches and Urquiola (2013) using Romanian data).

may provide, compared to alternative options, more school inputs or better teaching.

However, the evidence on the long-run benefits of selective schools remains scarce, including on college and labor market outcomes. This steems from the lack of data that follows individuals from high school through college and the labor market, while allowing for credible identification. Most recent studies of elite high schools have exploited data collected in settings with a centralized allocation of students to schools - through a common exam - to warrant the rigorous identification of elite school admission effects. However, such data can rarely be combined with information on college and labor market outcomes.

In this paper, we obtain new evidence on the longer-run effects of selective high schools using data on earnings expectations of students at high school completion, together with information on learning achievements and college outcomes. We argue that data on earnings expectations can inform on actual earnings and provide a broader measure of the benefits from an elite high school education than learning achievement or even years of education completed, one that includes potential gains from the accumulation of non-cognitive skills, access to specific social networks or school reputation. This will hold under some conditions, as (average) expected returns to college will match the (average ex post) returns if students form rational expectations of their future earnings. We obtain and discuss below some evidence suggesting that imperfect information students may have on their future earnings does not play a major role in our setting. In addition, and independently from their rationality, earnings expectations also matter per se as a key determinant of college enrollment decisions and major choices, and of later occupation choices (expectations determine decisions). [4]

To illustrate how earnings expectations can provide new insight and study the benefits from admission to elite high schools in a developing country, we examine the effects of admission into the National Polytechnic Institute ("Instituto Politécnico Nacional" or IPN), a system of 16 elite high schools in the Mexico City Metropolitan Area. IPN schools are the most selective public schools in the Mexican capital city (aside from another system of elite schools, called UNAM, that our data does not allow us to consider). Compared to other public high schools,

[4]Wiswall and Zafar (2015) have empirically verified that measured earnings expectations predict future schooling decisions, and Arcidiacono et al. (2014) have documented that college graduates sort across occupations based on their expected ex-ante monetary returns to occupational (and their major) choices (in addition to non-monetary preferences).

IPN schools give students access to a set of enhanced school inputs, including better qualified teachers, smaller class sizes, more IT equipment, and the presence of higher achieving peers. They also provide education with an emphasis on science and technology. We then ask two questions. First, what are the effects of admission to a high school of this system on expected future earnings and, to inform on actual future earnings, on college graduation outcomes? Second, to what extent the longer-run gains in college and expected earnings we observe are associated with access to improved education inputs, more selected peers, and gains in skills?

We identify the causal effect of admission into the IPN system on the earnings and returns to college that students expect at graduation, after they have spent three years in high school. Graduation is the moment when those expectations will matter, as it is the point when such expectations determine the decision to enroll or not in college. We use a quasi-experimental setting with a panel database that tracks students from high school application to graduation. This data elicits original self-reported information on the wages students expect to obtain with or without a college education. We also estimate the effects on a set of school inputs provided by IPN schools and learning achievement at high school completion, and college outcomes for students who graduate from college and register their diploma in a national registry in the following ten years after they enter high school –and so, that can be interpreted as a measure of on-time graduation. For causal inference, we exploit the allocation of students into high schools based on a common exam and a regression discontinuity (RD) design analysis. Methodologically, our RD design estimates depart from the approach followed in previous studies, notably Abdulkadiroglu et al. (2014) and Pop-Eleches and Urquiola (2013), by identifying the effect of marginal admission to the IPN system rather than to a specific IPN school.

Our results first characterize differences in school environments associated with IPN admission (i.e. the treatment under study) and the effects on learning achievement. They confirm that students marginally admitted to an IPN high school are exposed to better school inputs than students who are marginally rejected. IPN students interact with peers of higher academic ability on average, tend to be taught by more qualified and experienced teachers, in smaller class sizes, and have access to more computers. Furthermore, the IPN system translates these superior inputs into better student outcomes at high school completion. Admission to the IPN system substantially increases learning achievement, with a jump of about .3 standard deviations in mathematics.

Second, we consider measures of the future gains on labor markets to IPN admission - our main outcomes - and find that admission to the IPN elite system substantially increases the wage returns students expect from a college education, with point estimates of about 15 percentage points on the expected college premium in relative terms (while non-IPN students in our sample expect earnings with college 2.0 times as high as earnings with high school, IPN students expect earnings about 2.3 as high with college). This occurs as admission into the IPN system increases earnings expected conditional on having a college education, but has no effect on the earnings expected with only a high school education. In addition, we find that IPN students more often intend (at high school completion) to obtain a graduate diploma and actually graduate more often on time from elite colleges and in engineering, which is is correlated with higher wages. Hence, the increase in earnings expectations we observe is consistently associated with higher actual earnings prospects: IPN students will likely get better paid jobs by studying more often an engineering major and consistently expect higher earnings when graduating. This suggests the earnings expectations of students in this context are to a large extent rational. In addition, we find suggestive evidence that imperfect information on future earnings plays a limited role in this context; specifically students from disadvantaged family backgrounds, who are more likely to have incorrect information and likely underestimate earnings returns to further education, do not adjust their expectations when admitted to elite high schools any more than other students.

Third, we analyze further the importance of peers and school inputs, both of which have been debated in the literature on learning achievement. We estimate the gains from admission into other schools that are as selective as the less selective IPN schools and also bring better peers but fewer school inputs than IPN schools. Attendance at them does not translate into higher earnings expectations, nor does it translate into higher test-scores, so that peers alone do not seem to explain the gains from IPN admission (consistently with the US findings). This set of results suggests that IPN schools provide specific inputs – beyond elite peers – that are not offered in other selective schools and that do affect both skills and future earnings. The focus of IPN schools on science and technology might also play a role by preparing them (and shifting their preferences) for better-paid college majors and jobs. We also examine and find little support for a direct effect of IPN admission, through access to information, on earnings expectations, and discuss the potential benefits from school reputation.

Our methodological contribution is to use information on expectations of future earnings to provide evidence on the economic gains from specific elite high schools (or other sets of schools). We indeed find that the attendance of a system of elite high schools that effectively help students acquire more skills (or ones that are more valuable on the labor market) is valued in their subjective beliefs of future earnings. And the effects we find on college attendance intentions and graduation suggest that these higher expected earnings are matched with gains in actual future earnings. This in turn provides further evidence of the consistency of subjective measures of earnings expectations, and how they are updated systematically so as to internalize gains in returns.

We then contribute to the literature on elite high schools by providing evidence that when they bring better inputs or learning processes that help students accumulate skills, as is the case of IPN schools in Mexico city, such selective schools can affect college attendance decisions and outcomes and increase future earnings opportunities. While our results rely on reduced-form estimates and do not allow decomposing a causal chain from specific features of the IPN schools environment to college and labor market outcomes, the gains from attendance of IPN schools steem at least partly from specific inputs and processes those schools provide and additional skills that students acquire. But other aspects of these schools' environment, such as the focus on science and technology and their reputation, might also matter. These results seem in line with those of Clark and Bono (2016), and can be reconciled with those of studies in the US findings, finding no effects of schools that provide only interactions with more selected peers.

The paper is organized in the following way. Section 2 documents the setting and provides evidence that IPN schools form a well-identified system of elite high schools. Section 3 presents the data we use and describes our main outcome, expectations of future earnings, but also college and learning achievement outcomes. Section 4 presents our empirical strategy.

Section 5 documents the effects of IPN admission on school inputs received by students and learning achievements. Section 6 reports our main results for the effects on expected earnings and college outcomes. Section 7 gives further evidence on the inputs that drive our results, notably by comparing the effects of admission to IPN and other public high schools. Section 8 concludes.

2 Setting

2.1 IPN elite schools

We focus on the effect of admission into any of the 16 schools of the National Polytechnic Institute ("Instituto Politécnico Nacional" or IPN), a system of elite public high schools (grades 10-12). There are ten institutional systems of senior public high schools in the Mexico City Metropolitan Area. IPN together with the "Universidad Nacional Autónoma de México" (UNAM) system constitute the main system of elite schools as reflected by reputation, higher admission demand, and better school inputs. We do not study admission to UNAM schools because students there do not take the standardised test at high school completion that we use to measure our main outcomes and hence are not followed in our data. Therefore, we investigate the effects of admission to an IPN versus a non-elite (and non-UNAM) school.

IPN schools provide general high school education with a scientific and technical emphasis. IPN is also the main public technological higher education institution in Mexico and, mainly for that reason, the IPN high schools enjoy a wide prestige. But attendance at an IPN high school does not grant access to the IPN higher education institute. Excluding students admitted to UNAM schools, IPN schools admitted about 19,000 students, i.e., less than 10 percent of the roughly 195,000 students admitted to a public high school in Mexico City in 2005, the cohort from which our data was collected.[5]

Table 1 gives descriptive statistics for students' characteristics and school inputs in the 16 IPN and 593 non-IPN schools (based on information from the Mexican census of schools).

Being selected among the best applicants (see Section 2.2), compared to students of other

[5] About 35,000 students were admitted to UNAM schools. We exclude those from our analysis from here on.

schools, IPN students tend to have school peers who are higher-achieving and from a more privileged background. Students' average score at the common entry exam is almost 1.7 standard deviations higher in IPN schools than others, and the dispersion of achievement is lower, with an average school-level test score spread of .52 (unconditional) standard deviations versus .66 in other high schools.[6] Students in IPN schools more often come from a private junior high school (a 9 percentage points difference), and have a higher junior-high GPA (by .77 standard deviation). Two third of them are male. They also more often have parents with a high school education (almost 20 percentage points more than students in other high schools) and/or a white-collar occupation (a 17 percentage point difference).

Admission to an IPN school in addition provides access to a set of enhanced school inputs compared to other public high schools. In particular, IPN schools tend to have smaller class sizes (with an average of 39.6 students per class versus 41.6 in non-IPN schools) and much better access to computers (with 3.5 students per computer versus 9.8 in non-IPN schools). Furthermore, more IPN teachers have a college degree (86 percent) than teachers in other schools (81 percent), and more hold full-time positions (29 percent) than teachers in other schools (13 percent).[7]

2.2 The Comipems student allocation system

Admissions to IPN and other public schools are managed by the Metropolitan Commission of Public Senior-Secondary Education Institutions (Comipems), which oversees nine of the ten systems of public high schools in the Mexico City Metropolitan Area.[8] Since 1996, Comipems has allocated students to public high schools through a single centralized process based on students' schooling choices and scores on a common exam, and the numbers of slots available in each school.

The matching of students to schools follows a procedure which can be modeled by the serial

[6]Also, as discussed below, IPN students have higher expectations than youths in other schools of the wages they would achieve with a college education.

[7]However, IPN teaching might not necessarily benefit all students, as teachers may target specific (likely majoritarian) groups of students (Duflo et al. (2011)).

[8]A recent system of high schools administered by the Mexico City government and targeted to low-achieving student*s* does not belong to Comipems. The Metropolitan Area includes Mexico City and 22 municipalities from the neighboring state of Mexico.

dictatorship algorithm (Pathak, 2011) and operates in the following way. First, before taking the exam, applicants submit their ranking of preferred school choices (for simplicity, we use the term choice set) using a registration form they receive in January and return in February or March of the same year. The numbers of available seats are submitted by schools before students formulate these choices, and schools do not submit any priority criteria over students. Students can select up to 20 school choices (from 634 options in 2005).[9] Second, applicants take a common standardized exam in the last weekend of June. All applicants with at least 31 correct answers out of 128 are allowed to register in a Comipems school.[10] In the third step, students are allocated to schools. The Mexican Center of Evaluation of Education (Ceneval) ranks students according to their exam score and proceeds to allocate students, taken in the order of their ranking in the exam, to the school with available seats that students themselves ranked the highest in their choice sets. As some schools are oversubscribed, not all applicants are admitted into their preferred option(s). In 2005, only one third of applicants gained admission to their first-choice school. Finally, students who chose only schools that happened to be too selective with respect to their test scores –i.e. who miss the admission cutoffs for all their listed choices– can register in the schools with remaining slots in a second-stage allocation process. In 2005, this was the case for 19 percent of the applicants.[11]

This process reveals that IPN schools are in high demand: 40 percent of students admitted to Mexico City public high schools applied to an IPN school (i.e. about 4.1 applicants per seat). Moreover, for each seat offered in the IPN system, there were 1.7 applicants who chose an IPN school as their most preferred option.

Such demand translates into the high level of selectivity of IPN schools. The combination of

[9] They actually submit a list of preferred tracks, as some schools offer more than one track. We use the term school as synonymous with track, though, because most schools, notably IPN schools, have only one at the time of admission.

[10] Applicants who list a school from the UNAM system must take an exam version designed by this institution, while all the other students take an exam design designed by Ceneval (the institution in charge of the assignment process). Both exams are designed to be equivalent in level of difficulty. We do not have information to suggest that some students might prefer taking one version of the test to strategically increase the probability of gaining admission into one of their first choices. We show in the empirical analysis that (marginal) admission into the IPN system is not correlated with the probability of taking the UNAM version of the Comipems exam.

[11] Under this allocation mechanism, students' dominant strategy is to list their chosen schools in a way that is consistent with their true preferences, if there is no limit in the number of schools they can list (*Pathak, 2011*). This later assumption seems satisfied here as, in 2005, the applicants from Mexico City junior high schools submitted 9.2 schooling options on average and only 2 percent listed the maximum number of options allowed (which is 20).

the institutional setting and students' preferences produces a set of admission thresholds, where an oversubscribed school's admission cutoff is the exam score of the last student admitted to the school. Table 2 shows the distribution of schools and students admissions by school selectivity, measured by the admission cutoff scores, distinguishing schools of the IPN and other systems (again UNAM schools are excluded here). The admission cutoffs to the 16 IPN schools range from 66 to 99 with an average of 75.5 exam points (and a standard deviation of 7.37). IPN schools account for the major part of admissions to selective schools, with respectively 70 and 92 percent of admissions into schools with a cutoff of 66 or more and 76 or more (the highest admission cutoff for a non-IPN school is 81). Besides the non-IPN schools that are also very selective provide fewer inputs than the IPN schools (as we study in detail in Section 6).

Aditionally, the admission thresholds are determined ex-post and cannot be predicted beforehand. They hence generate variation in the allocation of students into schools that can be considered as locally (for scores close to the cutoffs) random. This is the variation we exploit to identify the causal effects of admission to selective schools, primarily IPN schools.

Hence, as a whole, the system of IPN schools is more selective, provides access to better inputs than other public schools (except UNAM schools which we do not consider here – we provide causal evidence of this below). We now examine the extent to which admission to the IPN system enhances a youth's opportunities to acquire further education and achieve better economic outcomes on the labor market.

3 Data

3.1 Wage expectations

Our main outcome of interest, expected wages and returns to college, is measured using data from a survey taken by a representative sample of students in the last year of high school, when students take the national Enlace achievement test. [12]

Two of this survey's questions elicit information about the earnings expected with given

[12] In Appendix A, we present in more details the subjective expectations data we use, the focus and limitations of that data, as well as a number of consistency checks. For more information about the measurement of subjective expectations in developing countries, see surveys by Attanasio (2009) and Delavande et al. (2011).

educational attainments by asking about youths' expected future earnings under two scenarios: that they terminate their schooling after completing high school and that they continue their studies and obtain a university degree:

1. *"If you do not obtain a university degree, what monthly income do you expect to have on average five years from now?"*
2. *"If you obtain a university degree, what monthly income do you expect to have on average five years from now?"*

The answers to both questions are given using a pre-codified set of brackets: i) 4,000 pesos (Mex$) or less; ii) 4,001 to 7,000 pesos; iii) 7,001 to 10,000 pesos; iv) 10,001 to 15,000 pesos; v) 15,001 to 20,000 pesos; and vi) more than 20,000 pesos.[13] Those are measures of expectations five years from the moment of the survey, at which point individuals would be about 23 years old.

The survey asks about the "average" (or mean) expected earnings, and provides for each individual a point estimate of his mean expected earnings with high school and college education. Our main measure of earnings returns to college, also defined at the level of individuals, is the ratio of the mean of earnings expected (Y_{coll}) with college and the mean of earnings expected with senior high school (Y_{high}):

$$R = \frac{Y_{coll}}{Y_{high}}$$

which we will refer to, with some abuse of terminology, as the expected return to college (or implied college premium R).[14] To obtain a measure of the implied expected college premium from the data on mean expected earnings in brackets, we assume that each discrete earnings category corresponds to the mean of the two values that define each bracket. For the first and last bracket, which do not have an obvious interval, we assume that the brackets are, respectively, [Mex$3,000–4,000] and [Mex$20,000–27,000]. R is a measure in relative terms of the

[13]In 2008, 7.5 Mexican pesos were equivalent to 1 US dollar in terms of purchasing power parity (OECD).

[14]In principle, expected relative returns should be obtained, at the individual level, as the (subjective) expectation of the ratio of potential earnings with college and high school, or $E(y_{coll}/y_{high})$, which may differ from the ratio of expected earnings with college (y_{coll}) and high school (y_{high}), $R = Y_{coll}/Y_{high} = E(y_{coll})/E(y_{high})$, that we use. We have no way to compute the expected ratio $E(y_{coll}/y_{high})$ using our data, but R should provide a reasonably good proxy. Note that this concern is absent if we use absolute returns (the expectation of which is the difference of expected earnings with college and high school), which we consider in robustness tests (as of course $E(y_{coll} - y_{high}) = E(y_{coll}) - E(y_{high}) = Y_{coll} - Y_{high}$)).

gap between expected earnings with college and senior high school, and we report the results using its log (i.e. $log\ \ Y_{coll} - log\ \ Y_{high}$). We test the robustness of our results to using a measure in absolute terms of that gap by using the the difference in expected earnings with college and senior high school in pesos $R' = Y_{coll} - Y_{high}$ and its log.[15]

For documenting the differentials in wage expectations between students of the elite and non-elite school systems, Figure 1 Panel A plots the conditional means of expected earnings, given learning achievement, with respectively a high school and a college education for students of IPN and the other (non-elite) schools, while Panel B plots the corresponding conditional means of the expected returns to college. Several patterns emerge.

First, even at similar levels of achievement at high school completion, IPN students tend to have higher expected earnings with a college education than students from non-elite schools. The conditional means are taken for each decile of the distribution of achievement and the graphs include lower and upper bounds of 95 percent confidence intervals. Earnings expectations of IPN students are 10 to 20 percent higher than those of non-IPN students across deciles. We observe also that the expected earnings of both groups tend to increase with scholastic achievement.

Second, there are no clear differences between IPN and non-IPN students in the earnings expected with only a high school education. The IPN students have slightly higher expectations, but the 95 percent confidence intervals do not reject the hypothesis that those expectations are equal. Importantly, the factors that make IPN students expect higher wages with college seem not to operate in the labor market for high school graduates. Similarly, while the slopes of the conditional expectation functions for earnings with college follow an upward trend, those for earnings with high school are flat, indicating that students believe that the return to scholastic achievement in the labor market for high school graduates is small or insignificant.

In consequence, as shown in Panel B, IPN students tend to expect higher returns of attending college than non-IPN students, again after conditioning on achievement, with a difference

[15]The Appendix A shows that the measures of the expected returns to college are comparable to the returns that can be estimated with a OLS wage equation using survey data on actual wages of young individuals in Mexico City and State of Mexico. Furthermore, the differentials associated with individual and family characteristics are rather expected, with higher returns to college expected by students with higher learning achievements at 15 years old and by children of more educated, white-collar and wealthier parents.

of 5 to 15 percent in relative returns (for students in the median of the achievement distribution, relative expected returns are approximately $exp(.75) = 2.1$ for non-IPN students and $exp(.85) = 2.3$ for IPN students) that is marginally significant at the right end of the achievement distribution. IPN students in the left tail also have higher expectations, but there are few low-achievers among IPN students and the corresponding conditional means are imprecise.

Similar results are obtained from an OLS regression to test if the observed differentials in expectations between IPN and non-IPN students persist after controlling for family and individual characteristics (see Table **??** in the Appendix). The gaps between IPN and non-elite school students in expected earnings with college and returns to college remain with the controls, while the difference in earnings with only high school is small and not statistically significant.

Overall, these descriptive results indicate that IPN students expect higher earnings if attending college than students of non-elite schools at similar levels of scholastic achievement.

3.2 College attendance intentions and graduation outcomes

To study the labor market benefits from elite school attendance, we use a set of college outcomes obtained from two sources. A first source on college outcomes is the Enlace survey, which contains information on intentions to attend college in general and a graduate school. The Enlace survey is given to a sample of students about to complete high school. 84 percent of students in Mexico City High Schools declare their intention to attend college, while 53 percent do so to attend graduate school – see Table **??** in Appendix A for details on how these intentions correlate with student characteristics.

A second source is the National Registry of Professionals from the Federal Ministry of Education, which provides information on graduation from academic degrees in Mexico, and details the degree level, major and education institution. Using this source, we can investigate actual on-time graduation outcomes from college in general, from elite IPN or UNAM colleges, and from specific majors. Table 3 presents partial correlations between college graduation outcomes and student characteristics. Conditional on achievement and family characteristics, IPN high school students are less likely to (overall) graduate from college and from an elite UNAM college, but more likely (and with larger coefficients) to graduate from an elite IPN college and from an engineering major.

The information on majors of graduation can then be combined with information on mean earnings by graduation major of college graduates from the Employment survey to obtain a measure of earnings returns to IPN attendance - while indirect, this measure is based on actual college graduation outcomes and complements our data on earnings expectations and attendance intentions. We consider in particular graduation from a engineering major, which concerns a great number of IPN graduates and has higher average earnings returns than the other majors combined. We examine this in details in section 6.3.

Hence, besides having higher earnings expectations, IPN students also intend more often to attend graduate colleges and do obtain more often a selective college diploma. As those patterns of higher earnings expectations and college outcomes can be driven by the selection of students into school systems, we seek to identify the causal effects of admission to IPN elite schools on wage expectations. Datasets

Our dataset matches information from the 2005 Comipems admission process to the results from both the 2008 and 2009 national achievement test of 12th graders and from the same year versions of a questionnaire survey of a random sample of students who took that test, and to data on higher education outcomes from the 2012-15 files of a register of college graduates. It thus forms a panel dataset in which students are followed from application (in 2005) to graduation from high school (in 2008 or 2009) (unless they drop out before – we document this in depth below) and, for a share of them, to graduation from college (in 2012-15).

At the end of high school, students in Mexico, including those in private schools, take a national standardized test of achievement called the Mexican Evaluation of Scholastic Achievement of Educational Institutions (or Enlace). The only exception is students enrolled in UNAM schools, which is why they are not considered in our analysis. The purpose of the examination is to evaluate schools, and the educational system as a whole. It hence has no bearing on whether students graduate or are admitted to university. However, Enlace results are widely reported by the media and nongovernmental organizations in Mexico, and are used as the principal input for the creation of school league tables. This publicity provides school personnel with incentives to perform better and makes Enlace a medium-stakes test. De Hoyos and Estrada (2016) show that Enlace test scores predict college enrollment and labor market outcomes using matched data from the ENILEMS 2010 module of the labor force survey – see more in section 5.2.

The Comipems data includes the submitted listing of choices (tracks and schools), the score at the first exam conducted at entry, which will serve as the running variable in the RD analysis, the assignment outcome for all applicants, and also some family background information from a questionnaire attached to the registration form.

The Enlace data contains the score at the second exam conducted among 12th graders, which we use to measure the achievements of our sample when completing high school. We normalize the scores for mathematics and Spanish language sections by exam cohort with mean 0 and standard deviation 1 for the sample of matched 2005 Comipems applicants.

In parallel to the Enlace exam and on the same day, the Federal Ministry of Education conducted a complementary survey among a random sample of students taking the test, gathering information on their characteristics, schooling experience, wage expectations and higher education aspirations.

We also have information about the Comipems schools from the 2005 version of the annual census of high schools carried out by the Secretary of Education (called "Formato 911"). The school census provides information on the characteristics of schools, in particular, class size, teachers' profiles, and information technology equipment.

All four datasets (Comipems 2005, Enlace 2008 and 2009 tests and surveys, and 2005 census of schools) can be matched at the individual level, and this was done by the Federal Ministry of Education (Secretaría de Educación Pública) data administration teams. [16]

We also match the Comipems data to 2012-2015 individual administrative records from the National Registry of Professionals of the Federal Ministry of Education (NRP, Registro Nacional de Profesionistas in Spanish). The NRP was created in 1945 to record and certify the holding of academic degrees, and help potential employers to identify the use of false academic credentials. The NRP issues a professional id (*cedula profesional* in Spanish) that certifies that

[16]The Comipems and Enlace datasets are matched using, in this order, the national population identification code (CURP), combinations of CURP and name when the former is incomplete, and the name and birth date when there are missing values for the CURP. The matching between the Enlace exam and the survey results is straightforward using the exam identification code (available for all observations in both datasets). The school census information is recovered for the specific schools' students attend using school identifiers. We return to sample sizes below.

the holder has received an academic degree from an authorized educational institution with the required coursework for the degree level. The RNP certifies degrees in all fields of study, starting from technical education up to doctoral studies. Individuals wishing to state the holding of an academic degree in a professional activity are obliged to comply with registration at the NRP. Goverment agencies and large private employers typically demand the NRP id when they hire for positions advertised for college graduates. Although obtaining a precise figure is not straightforward given the available data, we estimate that more than two-thirds of people who receive a college degree comply with registration in the NRP. It is worth to notice that tertiary institutions in Mexico typically demand students to finish succesfully all coursework, have approved a undergraduate dissertation and fulfill a required ammount of hours of social and/or professional work to be entitled to receive a college degree. Hence, there is typically a non-trivial delay from the time a person finishes undergraduate coursework and the time when she receives her degree. [17][18]

3.3 Estimation samples

For our main analysis, we focus on the sample of 2005 Comipems applicants who a) applied to at least one IPN school and are eligible for admission to such a school (hence the term of IPN applicants we use below) and b) graduated from a Mexico City junior high school.[19] We restrict the sample to applicants of IPN schools because our identification strategy compares the outcomes of admitted and rejected IPN applicants. [20]

We exclude applicants who graduated from a school located in the State of Mexico (38.4 percent of applicants) or another Mexican state (2 percent of IPN applicants). We discuss this

[17]We can put a rough estimate at 70%, given 327,000 people (for all mexico) received a college degree in the academic year 2011-2012 (irrespectively of when they finish their college studies), while 227,000 people registered a college degree in the NPR in the calendar year 2012 (irrespectively of when they received their college degree).

[18]The 2005 Comipems data and the 2012-2015 NRP data on undergraduates degrees are matched at the individual level using full names. Full names in Mexico are composed by given names (typically two), a paternal surname and a maternal surname.

[19]The eligibility restriction discards both applicants with a junior high school GPA lower than 7 out of 10 (as they are not eligible for IPN admission) and applicants with a Comipems score lower than 31 (as they are not eligible for admission to any Comipems school). In practice, the restriction to entry scores lower than 31 is redundant in our main estimations, as those use only observations “close” to the IPN admission cutoffs and the later are well above 31.

[20]In Section 6, for investigating mechanisms, we provide complementary results that use different samples of applicants to other (non-IPN) selective schools. We focus here on our main sample.

restriction in details below. We also omit students admitted to UNAM schools, as they do not take the Enlace exam and survey at high school completion, and, to avoid any bias in our selection of treatment and comparison groups, students admitted to IPN (and respectively other schools) who would be allocated to UNAM schools if rejected (with a higher score).

These restrictions lead to a sample of 15,660 IPN applicants at the Comipems exam (baseline) – 9,010 admitted IPN applicants and 6,650 rejected. We explain formally how we identify rejected IPN applicants in Section 4.

Only a share of the IPN applicants selected above can be matched to the Enlace test in 2008 or 2009. Restricting our study to them leads to samples of 7,796 students taking the Enlace test (49.8 percent of IPN applicants), and 5,284 students (33.7 percent) answering the Enlace complementary survey. Back-of-the-envelope calculations, presented in Appendix B, suggest that the attrition rate of 50.2 percent between the beginning and end of senior secondary education in our sample is driven mainly by students dropping out of high school before completion. More specifically, it can be roughly decomposed in the following way: at least 38 points are due to school drop-outs, up to 6 points to matching errors and up to 5 points to incomplete Enlace turnout. The lower number of survey respondents compared to test-takers is a result of the use of a random sample for the survey.

We are then able to match 916 (5.8 percent) of the 2005 IPN applicants to the 2012-2015 NPR individual records of college degree holders. This seemingly low rate can be explained by high school and college dropout, entry into the labor market after high school graduation, repetition and out-of-school periods between beginning of senior secondary education and the end of college, plus the delay between the end of college and obtaining the college degree and the RNP professional id.[21] In this sense, our NPR results can be interpreted as a proxy for on-time college degree attainment.

The attrition in the Enlace matching could affect the internal validity of our local experiment estimates if it differs between students admitted to and rejected from IPN schools. We hence have to verify that there is no relationship between the treatment of interest, i.e. admission

[21] Roughly, from a cohort of 100 students entering senior high school in 2005, one can expect 50 to graduate from high school, 30 to enter college, 15 to obtain a college degree and 10 the NRP id. Because college in Mexico lasts generally 4.5 years (9 semesters), an "on time" student would be obtaining her RNP id in early 2014 –assuming no repetition, out-of-school periods and no delay between finishing college and obtaining the degree.

to an IPN school, and attrition, measured by Enlace-exam taking and survey answering. We conduct several checks in Section 5.4 which reject the presence of differential attrition in our sample. De Janvry et al. (2014) use the same data, but consider a sample that differs from ours by including graduates from junior high schools in the State of Mexico. In addition, their sample excludes graduates from private junior high schools and returning-to-school students. In this sample, they do find higher rates of drop-out among IPN admitted students than rejected applicants. They also find that dropping out is strongly associated with distance to school.

In section 5.4.2 and Appendix B, we examine in detail the relationship between IPN admission and dropping out. We find that differential attrition does not occur among applicants to IPN school that graduated from Mexico City junior high schools. Only students from the State of Mexico tend to drop out more often when they are admitted to the IPN system. This is consistent with the heterogeneity in drop-outs by distance to schools found by De Janvry et al. (2014). Because 15 of the 16 IPN high schools are located within Mexico City, students from the State of Mexico have to travel to IPN schools which are further away. They hence have to support higher commuting costs of high school attendance (both direct and opportunity costs). In this context, restricting the study to graduates from Mexico City junior high schools allows us to focus on a sample of students for whom internal validity is warranted. Note that we still consider admission to all 16 IPN schools including the one in the State of Mexico.

Now, attrition does affect the external validity of our analysis, as the estimated parameters will capture the effects of admission to selective high schools conditional on completion. In addition, while private schools do not belong to Comipems and manage admission decisions independently, Comipems applicants can opt for a private school after the admission results or transfer to this system during their senior-secondary education. These transfers remain limited as only 2.7 percent of the applicants in our sample take the Enlace test in a private high school, and, conditional on completing high school, all students are observed at follow-up whatever the school they are attending.

4 Empirical strategy

4.1 Discontinuities in admission to the IPN school-system

Our empirical strategy is based on a regression discontinuity (RD) design and provides estimates of the effects of admission into the elite IPN system.[22] To illustrate the strategy, consider a student i who submits a set of ordered school choices $C_i = \{S_1, S_2, ..., S_{20}\}$ – call it the student's "choice set" –, with school S_k, for $k = 1, ..., 20$, ranked as kth choice. As most students rank fewer than 20 schools, we can allow for a number of choices at the end of the list of submitted preferences to be blank without any effect on the definition of admission cutoffs and empirical strategy. The entire admission process (including number of seats in schools, students' applications and exam scores) generates cutoff scores for admission to all schools as described in Section 3.1. Let $\{c_1, c_2, ..., c_{20}\}$ denote the set admission cutoffs corresponding to the choice set C_i, where c_k, for $k = 1, ..., 20$, is the admission cutoff of school S_k.

The ranking of schools in students' choice sets can differ from a ranking of schools by selectivity, as less selective schools (with a lower admission cutoff) can be preferred to more selective ones. Under the allocation mechanism considered, students have the incentive to rank schools in order of expected utility (Pathak (2011)). So, preferences for geographical location or other features of schools could generate a ranking of schools that is not strictly decreasing in admission cutoffs.[23]When examining a student's allocation to schools, we can disregard schools which are both more selective and less preferred than any other schools in student i's choice set, because those choices will never be relevant for her admission: even if she passes the cutoff for admission into those schools, she would be allocated to the less selective and preferred school.[24]We can thus simplify choice sets by deleting those choices. We perform this simplification in practice in the data. For student i, we can thus assume, with no loss of generality, that C_i is simplified so that $c_1 > c_2 > \ldots > c_{20}$.

[22]In Section 6, we use a similar strategy to analyze the effects of admission to (non-IPN) selective schools.

[23]In addition, even if it was their only choice parameter, students would need to have perfect foresight to produce a ranking that is consistent with school selectivity since admission cutoffs are determined endogenously.

[24]This simplification excludes all choices of schools which are more selective and less preferred than some other choice, and not only than the school the student was allocated to. It only helps identifying binding admission cutoffs (either to schools or systems) by removing choices that would never be relevant. It hence has no bearings on the definition of the treatment and control groups, and in particular does not introduce any selection above or below the c_K cutoff.

Now assume that student i lists at rank K of her (simplified) choice set, a given IPN school S^*, for which the admission cutoff is $c_K = c(S^*)$, so that: $C_i = \{S_1, ..., S_K = S^*, ..., S_{20}\}$, with $c_1 > \ldots > c_{K-1} > c_K = c(S^*) > \ldots > c_{20}$. Assume in addition that she obtained a score s_i at the Comipems entry exam. Then student i will be admitted to the IPN school S^* if her score is equal to or above the $c_K = c(S^*)$ cutoff score and below the c_{K-1} one, i.e. if $s_i \in [c_K = c(S^*); c_{K-1}[]?$. This is a school-specific admission cutoff.

We can also define some school system-specific admission cutoffs, in particular for admission to the IPN system. Indeed, for any given simplified choice set C that includes at least one IPN school, there is a cutoff score above which a student with choice C would be allocated to an IPN school and below which the student would be allocated to a non-IPN school (except in cases where no other schools are listed after IPN schools in the choice set[25]).

Assume that the choice set C_i of student i is such that the first L schools are all IPN schools and the next are all non-IPN schools. In this case, the student would be allocated to one of the IPN schools so long as his score is above the admission cutoff c_L of the Lth ranked least selective IPN school in the choice set, i.e. if $s_i \geq c_L$. Formally, for such a choice set C_i, denoting Θ^{IPN} the set of IPN schools, we can define a cutoff score for admission to the IPN system as the lowest cutoff of the chosen IPN schools:

$$c_{IPN} = min \quad \left\{c(S_k), \;\; S_k \in C_i \cap \Theta^{IPN}\right\} \tag{1}$$

In practice, for a student admitted to a school in the IPN system, the IPN admission cutoff is obtained by going down in the choice set (reducing his score in a thought experiment), starting from the school he was admitted to until reaching the last IPN school before the student would be allocated to a non-IPN school. The IPN admission cutoff is the cutoff of the last (i.e. less selective) IPN school. Similarly, for a student who was rejected from all chosen schools of the IPN system, the IPN admission cutoff is obtained by going up in the choice set (counterfactually increasing his score) starting from the school he was admitted to until reaching the first IPN school. The IPN admission cutoff is that of that first preferred IPN school.[26]

[25] In this case, the student would be counterfactually allocated to the second-stage admission process.

[26] We need to deal with the existence of another system of elite schools, the UNAM system. Choice sets include such UNAM schools and some students who are not admitted to IPN schools are allocated to an UNAM

To simplify exposition, we have considered above a choice set in which IPN schools are always preferred to non-IPN schools, so that the two types of schools do not alternate in the ordered choice set. Two variations can occur, but neither of these affect the way IPN admission cutoffs are obtained in practice. First, students can in principle list one or several more selective non-IPN schools before any IPN school in their choice sets. Because the most selective high schools are IPN schools, these cases are very rare: only 0.1 percent of students who applied to IPN schools have such choice sets in our data. We can still take into account these students by ignoring the cutoff between the more selective non-IPN and and IPN schools and focusing only on the cutoff between IPN schools and the less-selective non-IPN schools and observations around that cutoff. Second, a slightly more tricky situation occurs when the set of IPN schools is disjoint, with non-IPN schools inserted between IPN schools. Such choice sets are slightly more frequent in our data, representing about 7 percent of students admitted to IPN schools. For these students, we can still identify some intervals of scores over which they are admitted to the IPN system, and, while thereare now several IPN admission cutoffs, the one that is relevant depends on the actual score achieved at the exam. This does not affect the practical identification of the relevant IPN admission cutoffs, though.

The admission process then generates discontinuities in the relation between students' scores and their admission to an IPN school: students' choices (and the admission process) determine their IPN admission cutoffs, but then the probability that a student is admitted to an IPN school will depend only on her score, and jump from 0 to 1 at her IPN admission cutoff.

4.2 Local experiments

Importantly, applicants are not able to manipulate their IPN admission status because, although they can certainly influence their entry exam score through effort, they cannot know and/or precisely determine ex-ante the cutoffs and the relative position of their scores with respect to

school. Similarly, some students are admitted to IPN schools but would enter an UNAM school given a slightly lower score. For those students, our strategy would identify the effects of admission to one elite system (IPN) versus another (UNAM). In order to focus on the comparison of admission to an elite school system (IPN) versus admission to a school of a non-elite system, we exclude these observations (which correspond to 662 students rejected from the IPN system who were allocated to an UNAM school and 1,582 students admitted to the IPN system who would have been allocated to an UNAM school given a lower score). In addition, although it would be interesting to test whether the admission to UNAM schools has the same effects as the admission to IPN schools, as noted above, our data does not track UNAM students.

those. This would require both total control over their own score and perfect knowledge and anticipation of other applicants' scores and school choices.

To verify this empirically, we follow the procedure proposed by McCrary (2008). Given that our running variable is already discrete (its support being the set of integer values), we directly work with a finely gridded histogram for the densities of students by distance to the IPN admission cutoff (this is similar to obtaining an histogram from continuous data using a bin size of 1). We then smooth this histogram using local linear regressions, with a triangle kernel, separately on both sides of the cutoff. We finally test the statistical significance of the log difference in height of the density functions on both sides of the cutoff. Figure 2 shows the histogram and local linear regressions for the densities for a bandwidth of 4 (we obtain similar results with larger bandwidths – we discuss the bandwidth selection in the next subsection). No discontinuities are apparent at the cutoff (zero), and the log discontinuity estimate is .6 with a standard error of .7 so it is not statistically significant. This confirms local randomness: students who score close to the cutoffs are unable to precisely predict those and manipulate their scores and position to the cutoffs.

For another empirical test, Figure **??** in Appendix B plots the means of a series of observable characteristics of IPN applicants by distance to the IPN admission cutoff. None of those graphs exhibits any clear discontinuity: students scoring just below and above the IPN admission cutoffs have similar characteristics (we check this using estimation techniques below). This provides complementary evidence that admission is locally random and that IPN applicants are unable to influence the allocation process and self-select into treatment.

The discontinuities in allocations to school systems make it possible to identify the causal effects of admission to a school in the IPN system on given outcomes using a RD design strategy (Imbens and Lemieux, 2008 and Lee and Lemieux, 2010). We can then identify the effect of IPN admission, for every group of students with choices that lead to the same IPN admission cutoffs, by comparing the outcomes of students who were admitted to those of students who were rejected because they achieved slightly lower scores. The variations in treatment near the cutoffs are as good as random. As applicants cannot locally self-select into the IPN system, their expected potential outcomes whether they attended an IPN school or not (i.e. with and without treatment) as a function of exam scores will be continuous at the cutoff, and a discontinuity in the outcomes can be attributed to the admission to different school systems. [27]

While their admission to the IPN elite schools system is completely determined by the relative position of their score compared to the admission cutoffs, the probability that students actually attend this elite school system will present a discrete jump at the cutoff (from zero below) but may not be complete above it for two reasons. Students can decide not to attend the public school they were allocated to and opt instead for a private school, or change school after the first or second year of high school (to undersubscribed schools), so that our parameters of interest have an intent-to-treat interpretation. However compliance is almost complete. In Figure **??** in Appendix C, we plot the probability of taking Enlace in an IPN school, conditional on taking the exam in 2008 or 2009, by distance to the IPN admission cutoff. The probability jumps from zero to about .95 at the cutoff. Thus, although there is some drop out (see Section 5.4.2 below), the great majority of IPN admitted students who complete high school do attend an IPN school until completion.[28]

4.3 RD-design estimates

As students with scores above the cutoffs are admitted with certainty to an IPN school, the setting is one of sharp regression discontinuity. We can thus estimate the effects of IPN admission at any cutoff using local linear regressions. In practice, for a subsample of students with the same IPN admission cutoff c (as defined above), we restrict the sample to observations with scores in an interval $[c-h;c+h]$ around the cutoff and estimate the average effect of IPN admission on an outcome Y_i using a single regression. Denoting $d_i = s_i - c$ the distance between a student score and the IPN admission cutoff (the forcing variable), and $W_i = 1\{s_i \geq c\} = 1\{d_i \geq 0\}$ an indicator for his admission to the IPN system, the model to

[27]Note that we do not need all treatment and comparison students to have stated the exact same choice sets for preferred schools. Whether students' scores fall above or below the cutoffs is locally random. Hence, if one considers a sufficiently narrow interval around the admission cutoff, the structure of choices (among other characteristics) of the students above and below the cutoffs will be identical (or balanced) in average, as in a randomized controlled trial, so long as we remain in the neighborhood of the cutoff.

[28]Because compliance is almost complete we do not use a two-stages IV strategy to distinguish the effects of actual attendance.

be estimated is:

$$Y_i = \alpha + \tau W_i + \beta d_i + \gamma d_i.W_i + \varepsilon_i \tag{2}$$

where βd_i and $\beta d_i + \gamma d_i.W_i$ are distinct linear control functions for the slopes of the relationship between scores and outcome Y_i on the left- and right-hand sides of the admission cutoff c . The discontinuity in the outcome is obtained by extrapolating those slopes at the cutoff and taking the difference between the left- and right-hand extrapolations. The parameter τ captures this discontinuity, and OLS estimates of Equation (2) are consistent for this parameter.

Now, the sample consists of students who applied and were admitted or rejected to different IPN schools, so that there are multiple IPN admission cutoffs c, and our primary parameter of interest is the average effect of IPN admission over the different cutoffs. To estimate this parameter, we aggregate all students with different IPN admission cutoffs into a single sample, incorporate cutoff fixed-effects in Equation (2), and estimate the model:

$$Y_{ic} = \alpha + \tau W_i + \delta_c + \sum_c 1\{c\}.\left[\beta_c d_i + \gamma_c d_i.W_i\right] + \varepsilon_i \tag{3}$$

where δ_c are fixed effects for the cutoffs relevant for each student, $1\{c\}$ is an indicator that cutoff c is relevant for student i, and $\beta_c d_i$ and $\beta_c d_i + \gamma_c d_i.W_i$ are now linear control functions which are specific to each different cutoff. Identification remains within groups of students with the same cutoffs, and the parameter τ now captures the average discontinuity in the outcomes at the cutoffs and is again consistently estimated by OLS.

The choice of the bandwidth derives from a tradeoff between bias and precision. Imbens and Kalyanaraman (IK) have proposed an optimal bandwidth (Imbens and Kalyanaraman, 2011), a particular solution of this tradeoff obtained by minimizing the mean squared error. In our data, when using an edge-weighting kernel for estimating the local linear regressions, the IK bandwidth is about 2 and we consider this as our benchmark. However, although its support includes a large number of values (all integers between 0 and 128), our running variable is discrete. The IK framework is based on a continuous running variable, and hence does not apply perfectly here. We thus consider two larger bandwidths of 4 and 8 (obtained by multiplying tthe IK bandwidth by 2 and 4), that allow for higher precision. In addition, we check the robustness of our main results to the use of any bandwidth between 1 and 10.

Working with a discrete running variable adds another complication. As shown by Lee and Card (2008), the reliance on parametric extrapolations (using the local linear regressions) for estimating the counterfactual outcomes of marginally admitted students (had they not been admitted) using observations just below the cutoffs then introduces a specification error, and this source of variability can be accounted for by clustering the standard errors by the values of the running variable, the distances to the cutoffs. We also need to account for common unobserved shocks on the outcomes of students admitted to the same schools. Hence, when estimating Equation (3), we cluster the standard errors at the level of the interactions of distances to cutoffs and senior-secondary schools to which students are admitted.

For those estimates, the sample of 15,560 students who applied and were either admitted or rejected from an IPN school is reduced by the bandwidths around the IPN admission cutoffs. For the three bandwidths of 2, 4 and 8 points, we have samples of respectively 1,636, 2,992 and 5,506 observations of applicants, and of 517, 946 and 1,753 observations when restricting to students taking the Enlace survey.

Two points should be emphasized about our strategy. First, the estimated parameter τ captures the local effects of IPN admission for marginally admitted students. It is the relevant estimator for the effect of a policy change that would consider a marginal increase in the number of available slots in the IPN elite school system, with those slots distributed across IPN schools as the existing ones. Second and importantly, we identify the effects of admission to the IPN system rather than to a specific IPN school. We do this by considering the cutoff for admission to the IPN system, which can differ from the cutoff for admission to the specific school to which admitted students were allocated, and a control group of students who were not admitted to an IPN school. This makes the interpretation of the estimated effects more straightforward: while other studies estimate the effects of admission to more selective elite schools, we estimate the effects of admission to any school of the elite system. In addition, in the context of this study, the differences in schooling environments and received inputs are more pronounced between school systems than between schools. This approach thus provides more variation with which to identify the effects of school inputs.

5 School inputs and learning

Before studying the economic benefits from IPN admission, we document the changes in the school environment associated with such admission and the effects on learning achievements.

5.1 Higher school inputs

In Section 2, Table 1, we showed that IPN high schools have on average richer school environments than other public high schools in the city. We document now that – it is also true that – students marginally admitted to the IPN system are exposed to different school inputs and peers than students who are marginally rejected. For a visual representation, Figure 3 plots the averages of a series of school characteristics by distance to the IPN admission cutoff. The graphs suggest that IPN students interact with peers of a higher academic ability on average, tend to be taught by more qualified and experienced teachers, in smaller class sizes, and have access to more computers.

Table 4 gives RD design estimates of the effects of admission to an IPN school on the same measures of educational inputs received in senior-secondary schools. The estimates confirm that IPN students have on average peers with test scores about 1.1-1.2 standard deviations higher at the entry Comipems exam, but not more or less variable. These peers have more priviledged family backgrounds: they are respectively 11 and 5 percent points more likely to have parents with a high school degree and to have attended a private junior high school (for concision, we did not include these variables in Figure 3). They benefit from smaller classes by 1.6–1.9 students according to the estimates with larger bandwidth (the estimate with the bandwidth of 2 is smaller and not statistically significant; non-admitted students have on average class sizes of 42 students), and from much better access to computers with about 7-8 less students per computer (3-4 in average against 11 for non-admitted students). They are taught by teachers who are 8-9 percent more likely to have graduated from college (20 percent of teachers of non-admitted students are not college graduates) and 11–17 percent more likely to work full time in the school (21 percent of teachers of non-admitted students are in a full-time position).

While other specificities of the IPN school environment are probably missed by those few indicators, these findings confirm that IPN students are exposed to a set of enhanced school inputs and more selected peers.

5.2 Higher levels of skills

We turn now to investigate whether IPN schools translate their higher school inputs (and more selected peers) into more learning. Panel A of Figure 4 provides visual evidence of a discontinuity at the cutoff for achievement in mathematics. The estimates in Table 5 (columns 1-3) confirm that admission to the IPN system increases achievement in mathematics at high school completion by .27 to .30 standard deviations, a large effect, statistically significant at the 5 percent level for all bandwidths. On the other hand, as seen in Panel B of Figure 4 and columns 4-6 of Table 5, we do not find a significant effect on achievement in Spanish, which may reflect the emphasis of IPN schools on scientific and technical fields. Panel A of Figure **??** in Appendix D shows that the estimates of the effect on achievement in mathematics are robust to using bandwidths of 1 to 10.

Studies of charter school attendance in the U.S. find effects on achievement in mathematics of a similar magnitude to those we observe for IPN admission (see for example Dobbie and Fryer, 2013). In contrast, studies of elite schools in the US as well as other developed countries (e.g. Cullen et al., 2006, Bui et al., 2011, Abdulkadiroglu et al., 2014, and Dobbie and Fryer, 2014) find no effects or a positive effect restricted to minority students. For instance, Abdulkadiroglu et al. (2014) find gains in Boston exam schools (of .17 standard deviations) restricted to Blacks and Hispanics in English. The gains in achievement in mathematics associated with IPN admission thus lie in the upper part of the range identified in the literature, which confirms the capacity of this elite school system to combine the inputs it provides in a way to increase learning.

Elite schools can also enhance non-cognitive skills such as future-orientedness, self- confidence, leadership, etc. We do not have a comprehensive set of measures for these at the end of high school, but we can construct two indexes – using principal component analysis – for students' self-organization and attitude to school using related questions from the Enlace survey. Table **??** in Appendix D reports the corresponding RD design estimates for the effect of IPN system admission in these two proxies for non-cognitive skills. The point estimates are all positive but their statistical significance is not robust to the use of different bandwidths. The evidence on these effects is thus insufficient to form a conclusion on any effect on non-cognitive skills.

Summing up, IPN schools provide a higher quality education, with enhanced schools inputs which produce, at least, more cognitive skills. In contemporaneous work, De Hoyos and Estrada (2016) show that higher Enlace test scores map into better college and labor market outcomes, using merged data from the Enilems 2010 module of the labor force survey, given to a national representative sample of youth aged 18-20 years old. As Figure **??** in Appendix E shows, a 1 standard-deviation increase in the Enlace mathematics test score is correlated with a 10 a percentage points increase in the probability of being enrolled in college and a decrease of 3 percentage points in the probability of being neither in school nor at the work force.

6 Expected earnings and college outcomes

6.1 Higher earnings expectations

We now examine the economic benefits of IPN admission and consider first our main outcomes, expectations of future earnings. To begin with earnings expectations, Panels A to C of Figure 5 plot the local averages of the expected earnings with a high school and college education and the associated expected college premium. Students admitted to IPN schools have consistently higher expectations of earnings with college and returns, with marked discontinuities at the admission cutoff. Table 6 reports the RD design estimates of the effect of IPN admission on earnings expectations. Admission into the IPN system increases the earnings expected with college (columns 4-6). The estimates of the three local linear regressions show large increases with point estimates from 10 to 21 percent (depending on the bandwidth), all statistically significant at the 5 percent level. In comparison, average expected earnings below the cutoff are about 14,000 pesos (exponential of 9.573, around 900 USD). In contrast, we do not find any effect on the wages expected conditional on staying with a high school education (columns

1-3). The magnitude of the point estimates is small and its sign changes when using different bandwidths, between -4 and 6 percent, and the effects are never statistically significant at the 10 percent level.

The RD design results also confirm that IPN admission increases the expected returns to college derived from the expected earnings with a college and a high school education (columns 7-9). Compared to average relative expected returns of about 190 percent below the cutoff - i.e. non-IPN students expect earnings with college 2.0 times (log of 0.70) as high as earnings with high school, the point estimates indicate a substantial effect of 13 to 15 percent - i.e. IPN students expect earnings with college 2.3 times (log of 0.85) as high as earnings with high school. The statistical significance diminishes as we restrict the number of observations in the sample to the bandwidth of 2, but the estimates with the two larger bandwidths are statistically significant at the 5 and 1 percent levels respectively. Figure **??** in Appendix D plots the estimated parameter of interest with bandwidths of 1 to 10 exam points above/below the IPN system admission cutoff (see Panels B–D). Results are robust to the bandwidth selection.[29]

In a human capital model, an increase in skills (productivity) would induce an increase in the expectation of future earnings. Accordingly, students seem to update their beliefs in a way that is consistent with the internalization of the improved skills they obtain in IPN schools.

The observation of simultaneous gains in cognitive skills and expected wages with college suggests that students believe that the labor markets for college and high school graduates reward cognitive skills differently. This is in line with the descriptive findings presented in Section 3.1, Figure 1, which uses data from all Enlace survey respondants in Mexico City. Figure 1 shows that the slopes of the conditional expectation functions for earnings with college on Enlace scores follow an upward trend, while those for earnings with high school on Enlace scores are flat. This contrasting pattern indicates that students believe that the return to scholastic achievement in the labor market for high school graduates is small or insignificant, and not only that the specific skills that IPN schools provide have some value only if a college education follows.[30]IPN graduates could also anticipate that the additional skills that they ac-

[29]Table **??** tests the robustness of the effects on expected returns by estimating the effects on absolute returns and finds that IPN admission increases the expected returns to college by about 2700 pesos compared to expected returns of about 7400 pesos in monthly earnings for non-IPN students.

[30]Measurement error in earnings expectations could bias downward the correlation between learning achieve-

quired in IPN schools may help them enter a more selective college (by performing better at the college entry exams) and/or perform better in college, which may generate further wage gains. We provide evidence on this below.

6.2 Higher college attendance intentions

To get a sense of the actual long-term benefits associated with admission to an IPN school, we would ideally want to observe the outcomes of individuals in our sample on the labor market. Our data does not allow tracking students after they started working, but it contains two pieces of information on their college outcomes.

First, the Enlace survey contains information on the intentions of students at high-school completion to attend college. The visual evidence in Panel D of Figure 5 is again suggestive of positive effects. Table 7 gives the results of RD design estimates of the effects of IPN admission on intentions to obtain undergraduate (columns 1-3) and graduate (columns 4-6) degrees. The estimate for the effect on the probability of intending to obtain an undergraduate education has a positive sign, but its magnitude is relatively small (4 to 9 percentage point) and the effect is statistically significant at 10 percent only with the larger bandwidth. The already high share of students who intend to obtain such education and are below the cutoff, about 90 percent, likely explains this result. However, there is evidence for a positive effect on the probability of intending to obtain a graduate education, with an increase of 13 to 17 percentage points significant at the 1 percent level for the two larger bandwidths. This effect is to be compared with a share of about 58 percent intending to obtain a graduate education.

Thus, the effects of admission to an IPN school on the financial gains from a college education seem to translate into intentions to obtain a higher education. This is a first piece of evidence suggesting that elite schools can affect decisions to enroll at college and have long-lasting consequences on individual outcomes.

ment and expectations, but this does not explain our observation of a positive correlation for expected earnings with college as there is no obvious reason why measurement error would be higher for expected earnings with high school only.

6.3 College graduation outcomes

Second, we have information on actual college outcomes with the list of students of our sample who registered their college degree (on-time) in 2012-2015 .

The evidence in Figure 6 suggests that IPN students both more often graduate from an IPN elite college and obtain an engineering degree. Table 8 reports RD Design estimates of the effects of IPN admission on obtaining a college degree (column 1), graduation from an elite IPN or UNAM college (column 2 and 3), and graduation from an engineering major (column 4). These estimates confirm that IPN high school students are more likely to graduate from the IPN college (about 4 times more than non-IPN students) and to obtain an engineering diploma (about 3 times more often). The magnitude of these effects must be interpreted with care since these are unconditional means with zeros for all students who, for different reasons including not having a college degree, do not register their college diploma between 2012 and 2015. The coefficient on graduation from UNAM (other elite) college is negative but marginally significant and smaller in magnitude. Hence these findings provide direct evidence that admission to an IPN schools increases students' chances to graduate from an elite IPN college and with an engineering graduate degree. One must remind here that attendance of an IPN high school does not grant access to an IPN diploma. But IPN high school students, as we observed, are learning more Mathematics, which should prepare them better for engineering studies, and their high school education migh shift their preferences towards engineering.

Table 9 then reports regression estimates of the wage premium to an engineering diploma using data from the ENOE employment survey. These observational estimates suggest that that college graduates with a major in engineering earn 15 percent higher wages than college graduates with another major. [31]

Hence IPN admittees are more likely to graduate from an IPN elite college and with an engineering diploma - which is in turn associated with higher earnings even compared to other college graduates. Hence, while our data is insufficient to estimate precisely the gains in actual earnings and compare them to gains in earnings expectations, these results on college outcomes

[31] We use quarterly rounds from third quarter of 2006 until second quarter of 2010 of the ENOE and restrict to a sample of individuals aged 23-35 years old who graduated from college and reside in the Federal District with positive earnings. We control for gender, age, age squared, and state of birth.

suggest that the economic benefits that accrue to IPN students are real and consistent with their higher earnings expectations at high school completion.

Consistent patterns are obtained using observational data for a sample in which we can identify IPN high school graduates. Table 10 gives regression estimates of correlations between attendance of an IPN high school and higher education and job market outcomes among individuals aged 23 to 35 years who graduated from a public senior high school in Mexico City. It uses data from the ENTELEMS labor force survey. IPN graduates are more likely to attend college han students in non-elite public high schools, with a differential of 34 percentage points compared to a baseline college attendance of 32 percent, and to achieve better outcomes on the labor market. IPN graduation is correlated with higher labor market participation (by 13.5 percentage points), lower unemployment (by 6 percentage points), and higher hourly wages (by 51 percent). These results also confirm that IPN graduates are more likely (by 16 percentage points) to have been obtained an engineering degree. Table **??** in the Appendix provides more detailed information about what college majors are attended by IPN and non-IPN high school graduates. This descriptive evidence must be at least partly driven by individual and family characteristics associated with the selection of students into high schools, but it tends to confirm the causal estimates we obtained.

Wiswall and Zafar (2015) and Arcidiacono et al. (2014) have used data on expected returns to different college majors to investigate determinants of college majors and occupation choices respectively. In this spirit, we can provide observational evidence of the extent to which earnings expectations are associated with college attendance intentions and graduation. Appendix Table **??** reports a multivariate correlation analysis showing that higher expected college premium are positively and statistically signicantly associated with higher intentions to attend college and a graduate diploma and also with the actual changes to register as a college graduate with a diploma in engineering and from an IPN college. This suggests that earnings expectations do translate into intentions to attend college and graduate studies and actual college graduation. [32]

[32]We thank an anonymous referee for this suggestion.

7 Robustness checks

7.1 Validity of the RD design at assignment

We analyze the validity of the regression discontinuity design following Lee and Lemieux (2010). We have already verified in Figure 2 that there is no sign of discontinuities in the number of students with scores just above the IPN admission cutoffs, which confirms that individuals have no precise control over the assignment process. In addition, we verify that the characteristics of the students marginally admitted and rejected to the IPN system are balanced.

We formally test the local balance of baseline covariates across both sides of the IPN admission cutoffs by using a vector of baseline covariates as dependent variables in our main econometric specification. We expect the coefficients for treatment not to be different from zero if students are unable to sort above the IPN cutoffs. Columns 1-3 of Table 11 give the results of three sets of RD estimates, using the bandwidths of respectively 2, 4 and 8, for a set of baseline covariates (similar to the one used before) as dependent variables (the respective graphical evidence can be seen in Figure **??** in Appendix B). We also test the joint significance of discontinuities in the full set of covariates using seemingly unrelated regressions (SUR). We do not find evidence that any groups of students sort above the IPN admission threshold, and the results are consistent across specifications. With only one exception across the nine regressions reported in each column, the admission coefficient is not statistically significant at the 10 percent level. The Chi-square test for the discontinuity indicator being zero in all equations takes p-values of .16, .85, and .83 using the three bandwidths. This is strong evidence in support of the locally random-like variation of assignment to treatment.

7.2 Differential attrition

We discussed in Section 3 the presence of attrition in our sample, due mainly to dropping out of high school. Attrition correlated with IPN admission would be particularly problematic for identification, as it would impair the comparability at follow-up of students below and above the cutoffs.

Differential attrition would reflect in a discontinuity, at the admission cutoff, in the distribution of the Comipems score among students completing senior-secondary education. To investigate this, we again follow the McCrary (2008) procedure. Similarly to Figure 2 and using the same specification, Figure 7 plots local linear regressions on the two sides of the cutoffs respectively for the density distributions of students answering the survey. While the densities are lower on the right of the cutoffs, the graphs show no evidence of discontinuities at the cutoffs. The test of the log discontinuities are not statistically significant, with point estimates (and standard errors) respectively of -9.4 (11.2) percentage points for the density of test takers and 9.8 (13.1) for the density of survey takers.

We also formally test the presence of differential attrition using RD estimates of the effect of IPN admission on the probabilities of taking the Enlace exam and survey. Results are reported in Table 12. The point estimate for the effect of IPN admission on Enlace exam-taking (columns 1-3), while larger at about 5 negative points for the two larger bandwidths, is close to zero for the smaller bandwidth of 2 and is never statistically significant at the 5 percent level. The estimates for the effect of IPN admission on the probability of answering the Enlace survey (in column 4-6), which determines the observation of our main outcomes – wage expectations –, have a small magnitude and are not different from zero at conventional levels of significance. This confirms the absence of differential attrition associated with IPN admission in our data. The students in our sample admitted to IPN schools are not more or less likely than the non-admitted students to drop out before completing high school. Again, this is in contrast with the results recently obtained by Dustan et al. (2015) who – including students from outside Mexico City – find that admission to IPN schools increases the probability to drop out. The difference can be explained by the samples used, as we show in Appendix B, and notably in Table **??**.

Now, even with no effect of IPN admission on the probability of taking the Enlace exam on average, some heterogeneous effects across specific groups remain possible. These would be present if, for example, IPN schools were relatively better at keeping in school some types of students, but are relatively worse at keeping other types. We thus test for the balance of baseline covariates at graduation. Table 11 gives the results of RD estimates, together with the SUR joint-significance test, of differences at the cutoff in the observable characteristics for the samples of Enlace test-takers (columns 4-6) and survey-respondants (columns 7-9). For the sample of test-takers, some statistically significant differences appear only for junior high school GPA (again for the bandwidth of 8) and graduate school intentions at the end of junior high school (significant for the bandwidth of 2). For survey respondants, the sample for our main outcomes, the only variables for which there are statistically significant differences are junior high school GPA (significant only for the bandwidth of 8) and attendance of a private junior high school (significant for the bandwidths of 2 and 4). The SUR tests reject the joint significance for the full set of characteristics for the three bandwidths for the sample of test-takers and for the two smaller bandwidths for the sample of survey respondants, so that overall students below and above the admission thresholds are not systematically different at Enlace taking.

We examine further the robustness of our results by adding a vector of the same set of observable characteristics used previously as explanatory variables in our econometric model. In particular, we control for junior high school GPA (using a vector of junior high school GPA deciles fixed effects), past attendance at a private junior high school and baseline intentions to attend a graduate school, the only variables for which there were some statistically significant differences between IPN admitted and rejected candidates at follow-up. Results, presented in Table **??** in Appendix D, are very similar to the ones reported in Tables 5 to 6.

Finally, we estimate our main model restricting to applicants from public junior high schools (results are given in Table **??** in Appendix D). This constitutes another check that the differences in the junior school system of origin in the composition of the population of students at follow-up does not drive our results. Again the estimates of IPN admission on our main outcomes remain unchanged.

Overall, these checks confirm that the discontinuities in school quality indicators, learning achievement, earnings expectations at completion of high school and college outcomes we have documented can be causally attributed to the locally exogenous allocation of students, generated by the exam-based allocation process, to schools in the IPN system.

8 Further evidence

We now provide further evidence to investigate which inputs IPN schools provide generate the economic benefits expected by IPN students and their better college outcomes. We also test for potential direct effects of IPN attendance on earnings expectations through access to information.

8.1 Peers versus school inputs: comparison with non-IPN elite schools

To examine which inputs explain the gains we observe of IPN admission, we investigate the benefits from admission into non-IPN selective schools. As documented in Table 2, IPN schools account for the largest share of enrollment into elite high schools in Mexico City – and all enrollment in the most selective non-UNAM schools. However, some high schools are equally selective as the less-selective IPN schools. We can use a similar empirical strategy to estimate the effects of admission into these other selective schools.

We consider the set of non-IPN high schools with an admission cutoff of 66 entry exam points – the minimum for an IPN school – and define an admission cutoff for each of these schools. Table **??** (columns 3 and 4) in Appendix C gives the distribution of applicants to these non-IPN schools by school selectivity. We then obtain RD estimates of the effects of admission to any of those schools.[33]

Table 13 reports the RD estimates for the effect of admission into a non-IPN elite school on the school environment experienced by the student. Admission into non-IPN elite high schools also provides access to peers with higher average achievement (by .69–.72 standard deviations) and from more priviledged family backgrounds (the gaps in peers' parental education and attendance of a private junior high school are very close to the ones associated with IPN admission), but not to systematically better other school inputs: although class sizes tend to be slightly smaller (by 2.8–2.9 students per class), those schools do not provide better access to computers (there are actually 2.6-4.7 more students per computer), and fewer teachers work full time than

[33] An alternative – and complementary – experiment would be to look at the effect of admission into a non-IPN school more preferred and selective than some IPN school. However, we cannot estimate such effects because there are very few observations (only 11) in which students assigned to a non-IPN school would have been allocated to an IPN school given a marginally lower score.

in other schools. So, these other selective schools seem to provide fewer educational inputs than IPN schools do.

Table 14 gives the RD estimates for the effect of admission to these schools on student outcomes for the bandwidth of 4. Admission into such schools has no effect on high school completion, expected earnings, learning achievement or intentions regarding further education. The point estimates for the effects on expected returns (7.6 percentage points), wages expected with college (7.0 percent) are positive, but these estimates are not statistically significant. The effect on achievement in mathematics is close to zero. Thus, non-IPN elite schools also provide access to more selected and academically better peers, but do not generate the same gains as IPN schools.

IPN schools hence seem to provide some specific inputs – beyond more selected peers – that are not offered in other selective schools and that do affect both the skills and the returns students expect from college. This is in contrast to the findings of studies of elite high schools in the United States (Abdulkadiroglu et al. 2014; Dobbie and Fryer 2014), which observed that these schools mainly provide interactions with more selected peers. This suggests that, in our setting, it is the additional school inputs that IPN students receive which help them build skills to which they attach some economic value. IPN schools might also offer improved learning processes that enhance students' skills. The focus on science and technology might also help them acquire skills that are better valued on the labor market.

8.2 Direct effects of IPN schools on students' information and/or beliefs

The higher expected returns to college of IPN students seem to derive from the better school inputs and coincident acquisition of skills, in particular in mathematics, in those schools and their future value. However, the attendance of an IPN elite school could also directly shape students' expected earnings through access to better information on earning opportunities provided by a college education, which may in turn affect college application decisions and long-run outcomes.

In particular, if students from disadvantaged family backgrounds underestimate their real returns to education (see for example Jensen, 2010), admission to IPN schools could correct this downward bias through interactions with more informed peers or direct informational activities. If this is the case, the perceived returns to college of youths with less educated parents would increase more than those of their more privileged peers, given exogenous IPN admission.

We do observe lower average earnings expectations among youth from more disadvantaged backgrounds. For instance, even after controling for academic achievement, we observe that youth with more educated parents expect higher earnings with college, while youth of indigenous origin expect lower earnings with college (see multivariate regression analysis in Table **??** in Appendix A). However, we do not find that students from more disadvantaged backgrounds increase their earnings expectations any more than other students. For instance, the estimates in Table 15, with interaction terms, show that the effects do not differ by parental education in any of the considered outcomes. Youths with more educated parents benefit in terms of learning achievement, expected earnings, and expected returns to college as much as those with less educated parents. Similar results (available from the authors) are obtained using other background characteristics, such as parental occupation and graduation from a private school at the junior-high level.

Disadvantaged youths do not seem to correct their potentially downward-biased perceived returns to education when exposed to better information in IPN schools. Information could still be an underlying channel if it affects all IPN students in same way. But this would require that all students receive and treat information similarly, which is unlikely given that we observe gaps in expectations by family background.

Besides, in line with a signaling model, the higher reputation of elite schools, and the diploma they award, may give their graduates a signal of higher productivity. In general, this signal could increase expected earnings directly in the labor market, with access to better paid jobs at least early in graduates' careers. While we cannot test it directly using our data, such a mechanism is difficult to reconcile with our finding that attendance of an elite school increases expectations of wages if obtaining a college degree but not if entering the labor market after high school.

While we cannot exclude that they play a role, given that IPN schools generate gains in skills, these mechanisms seem at best to be complementary to the internalized gains from higher skills.

9 Conclusions

We find that admission to a system of elite public high schools in Mexico City (IPN) substantially increases the future earnings and returns students expect to receive with a college education. Consistently with the assumption that students form rational expectations of their future earnings, these higher earnings expectations seem to follow actual earnings gains, as we also find that IPN admission increases the chances of intending, at high school completion, to obtain an elite college degree, and of graduating from college with a major in engineering. At the same time, we find that IPN admission has no effect on the earnings expected with high school education alone, which seems to reflect that students do not expect significant returns to skills if working with a high school degree only. In addition, we find that IPN admission gives access to better school inputs and that increases learning achievement at high school completion.

Our results suggest that, in such a setting, information on expectations of future earnings can be used to document the economic gains from specific elite high schools (or other sets of schools) when information on labor market outcomes is lacking. Our results indeed indicate that students internalize the future benefits on the labor market from a privileged learning environment, the additional or specific skills they acquired, and their better college education opportunities, into their subjective beliefs. In addition, earnings expectations might provide a broader measure of the benefits on labor markets from an elite high school education than learning achievement or even completed education - one that includes potential gains from the accumulation of non-cognitive skills, access to specific social networks or school reputation. In particular, while we observe clear gains in scholastic skills, IPN schools could also increase non-cognitive skills ; the emphasis of their education on science and technology might also have value ; and we cannot exclude that in addition they also generate economic gains through the signaling effect their reputation can generate.

Our findings also contribute to the debate on elite schools: we provide evidence that attending a school in a system that provides better inputs, and not only interactions with more selected peers, increases at the same time students' skills, their college outcomes and the earnings they expect to obtain with a college education. In particular, when comparing the effects of IPN and other selective schools in the city which also provide more selected peers but not better school inputs, we find that the later do not generate the same gains in expectations. This finding echoes recent studies (for instance Abdulkadiroglu et al., 2014) that observe that schools only providing interactions with more selected peers have almost no effect on students' learning.

Acknowledgements

We are grateful to François Bourguignon, Denis Cogneau, Julie Cullen, Adeline Delavande, Francisco Ferreira, Julien Grenet, Marc Gurgand, Rafael de Hoyos, Katja Kaufmann, Karen Macours and Sylvie Lambert for helpful comments on earlier drafts. Insightful comments were also received at seminars at the Paris School of Economics, Gothenburg University, University of Essex, ZEW Mannheim University and the European University Institute, and at the ECINEQ, LACEA, ESPE, EALE, ESSLE, EUDN and EWMES conferences. Three anonymous reviewers provided comments and suggestions that helped improve and clarify the manuscript. We thank the staff at the Mexican Ministry of Education for kindly providing us access to data. We gratefully acknowledge financial support from Cepremap. Estrada also wishes to acknowledge funding from CONACYT and the Max Weber Program at EUI.

References

Abdulkadiroglu, Atila, Joshua Angrist, and Parag Pathak, "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools," *Econometrica*, 01 2014, *82* (1), 137–196.

Arcidiacono, Peter, V. Joseph Hotz, Arnaud Maurel, and Teresa Romano, "Recovering Ex Ante Returns and Preferences for Occupations using Subjective Expectations Data," NBER Working Papers 20626, National Bureau of Economic Research, Inc October 2014.

Attanasio, Orazio P., "Expectations and Perceptions in Developing Countries: Their Measurement and Their Use," *American Economic Review*, May 2009, *99* (2), 87–92.

Bui, Sa A., Steven G. Craig, and Scott A. Imberman, "Is Gifted Education a Bright Idea? Assessing the Impact of Gifted and Talented Programs on Achievement," Working Paper 17089, NBER May 2011.

Clark, Damon, "Selective Schools and Academic Achievement," *The B.E. Journal of Economic Analysis and Policy*, 2010, *10* (1), 1–38.

_ and Emilia Del Bono, "The Long-Run Effects of Attending an Elite School: Evidence from the United Kingdom," *American Economic Journal: Applied Economics*, January 2016, *8* (1), 150–76.

Cullen, Julie B., Brian A. Jacob, and Steven Levitt, "The Effect of School Choice on Participants: Evidence from Randomized Lotteries," *Econometrica*, 09 2006, *74* (5), 1191–1230.

Delavande, Adeline, Xavier Giné, and David McKenzie, "Measuring subjective expectations in developing countries: A critical review and new evidence," *Journal of Development Economics*, March 2011, *94* (2), 151–163.

Ding, Weili and Steven F. Lehrer, "Do Peers Affect Student Achievement in China's Secondary Schools?," *The Review of Economics and Statistics*, May 2007, *89* (2), 300–312.

Dobbie, Will and Jr. Fryer Roland G., "Getting beneath the Veil of Effective Schools: Evidence from New York City," *American Economic Journal: Applied Economics*, 2013, *5* (4), 28–60.

_ and Roland G. Fryer, "The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools," *American Economic Journal: Applied Economics*, July 2014, *6* (3), 58–75.

Duflo, Esther, Pascaline Dupas, and Michael Kremer, "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya," *American Economic Review*, August 2011, *101* (5), 1739–74.

Dustan, Andrew, Alain De Janvry, and Elisabeth Sadoulet, "Flourish or Fail? The Risky Reward of Elite High School Admission in Mexico City," Working Papers, UC Berkeley 2015.

Hoyos, Rafael De and Ricardo Estrada, "Do test scores predict labor market outcomes? Evidence from a national examination in Mexico," mimeo 2016.

Imbens, Guido and Karthik Kalyanaraman, "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *Review of Economic Studies*, 2011, *79* (3), 933–959.

Imbens, Guido W. and Thomas Lemieux, "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, February 2008, *142* (2), 615–635.

Jackson, C. Kirabo, "Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago," *Economic Journal*, 2010, *120* (549), 1399–1429.

Jensen, Robert, "The (Perceived) Returns to Education and the Demand for Schooling," *The Quarterly Journal of Economics*, 2010, *125* (2), 515–548.

Lee, David S. and Thomas Lemieux, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, June 2010, *48* (2), 281–355.

Lucas, Adrienne M. and Isaac M. Mbiti, "Effects of School Quality on Student Achievement: Discontinuity Evidence from Kenya," *American Economic Journal: Applied Economics*, July 2014, *6* (3), 234–63.

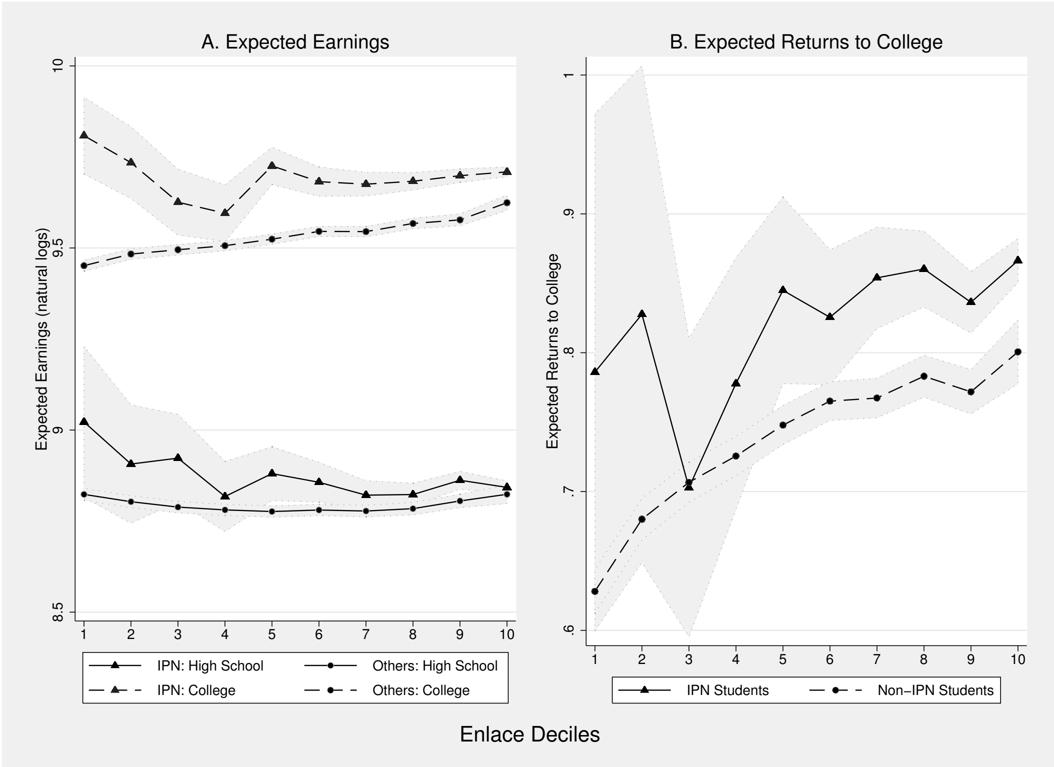
McCrary, Justin, "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, February 2008, *142* (2), 698–714.

Pathak, Parag, "The Mechanism Design Approach to Student Assignment," *Annual Review of Economics*, Sept 2011, *3*, 513–536.

Pop-Eleches, Cristian and Miguel Urquiola, "Going to a Better School: Effects and Behavioral Responses," *American Economic Review*, 2013, *103* (4), 1289–1324.

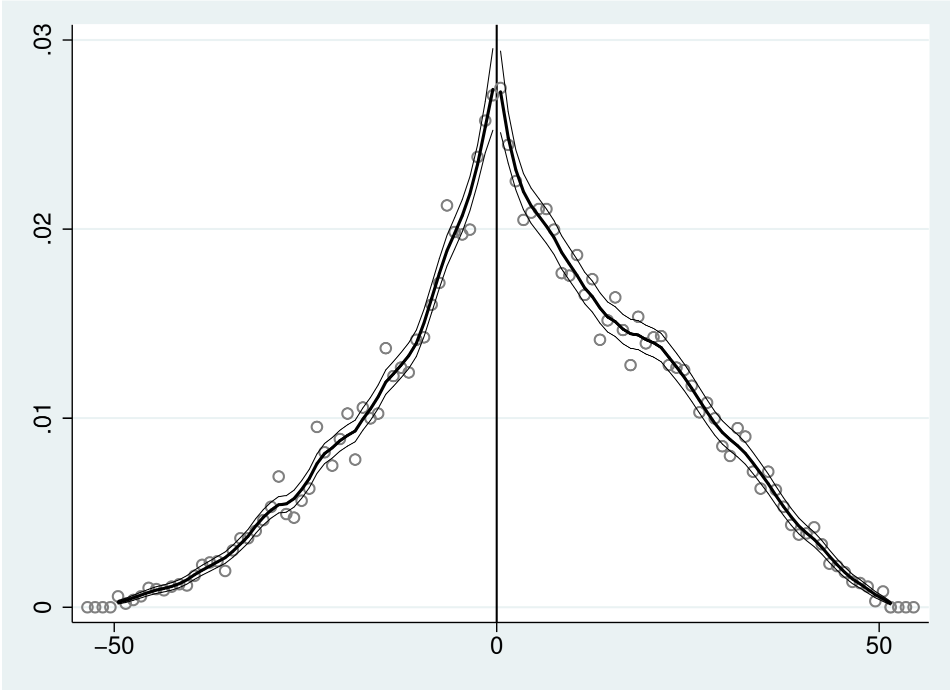
Wiswall, Matthew and Basit Zafar, "How Do College Students Respond to Public Information about Earnings?," *Journal of Human Capital*, 2015, *9* (2), 117 – 169.

Figure 1: Local means of expected returns to college and earnings



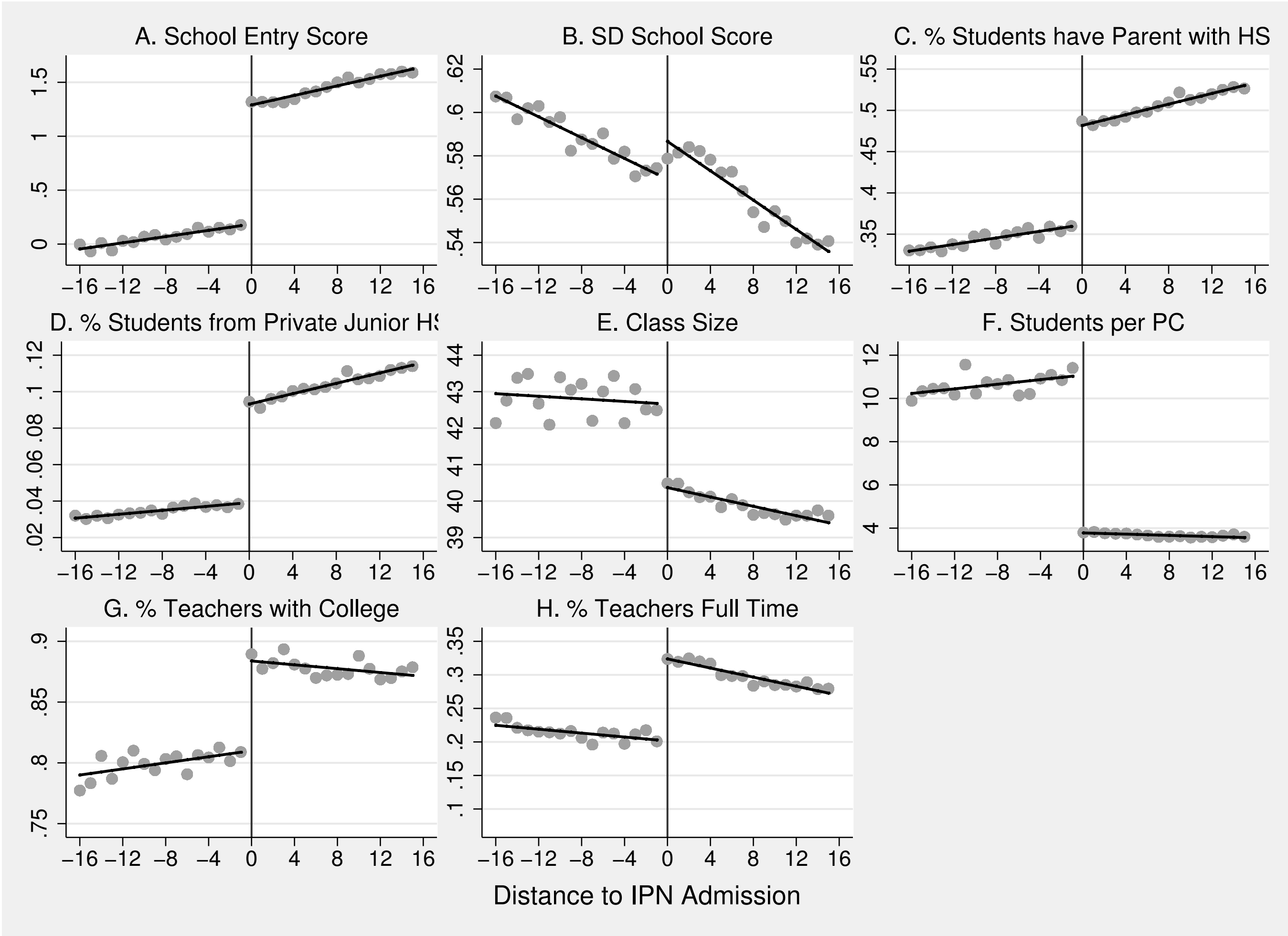
Notes: Conditional means, by decile of high school completion (ENLACE) achievement test, of expectations of log wages expected with high school and college degrees (Panel A) with 95 percent confidence interval and log of relative earnings returns to college (Panel B) of students of IPN and other public high schools. Source: ENLACE 2008 surveys. Sample: 3rd year high school students of Mexico City public high schools.

Figure 2: Distribution of the distance of students' scores from IPN admission cutoffs at assignment to high school



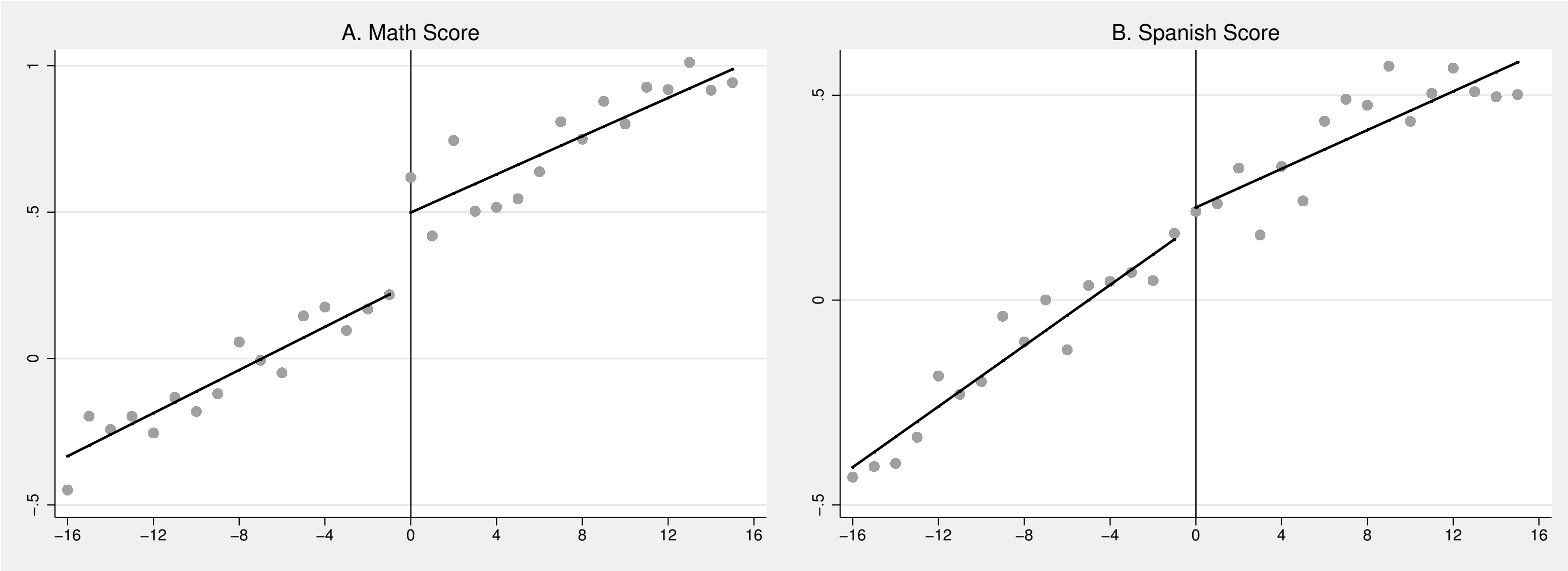
Notes: Distribution of the distance of students' scores at the COMIPEMS entry exam and from their cutoff scores for admission into the IPN system. Source: COMIPEMS 2005 school choices and entry exam scores. Sample: applicants from Mexico City junior high schools to COMIPEMS public high schools.

Figure 3: Means of peer characteristics and school inputs by distance from IPN admission cutoffs



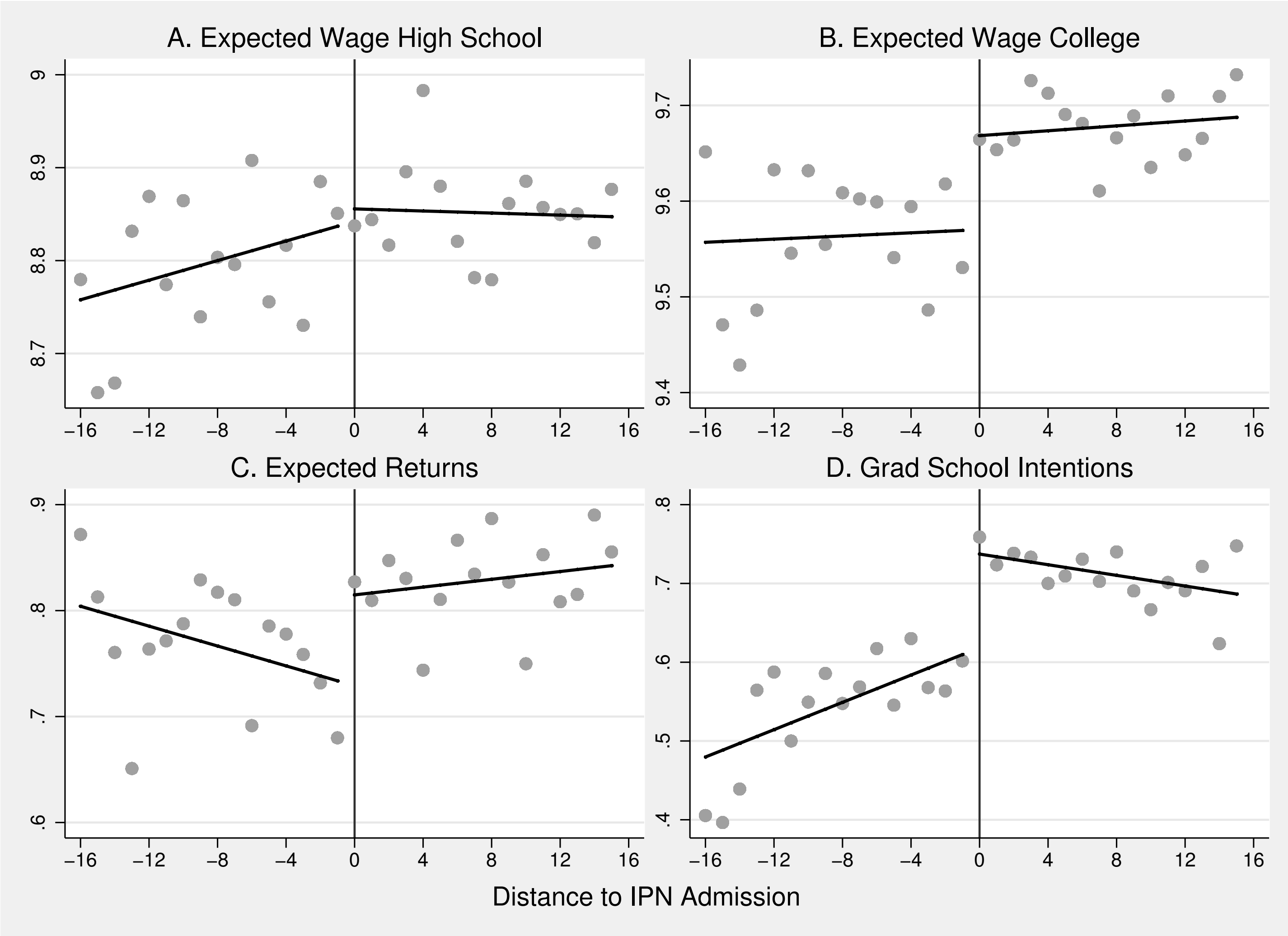
Notes: Conditional means of measures of peer characteristics (school peers' entry scores: mean and standard deviation, and means of students who graduated from a private junior HS and have a parent with HS education) and school inputs (class size, number of students per computer, share of teachers with a college degree, share of teachers employed full time in the school) by distance of IPN entry exam test score from IPN admission cutoff. Source: COMIPEMS 2005, Census of schools. Sample: applicants from Mexico City junior high schools to IPN schools.

Figure 4: Means of learning achievement outcomes by distance from IPN admission cutoffs



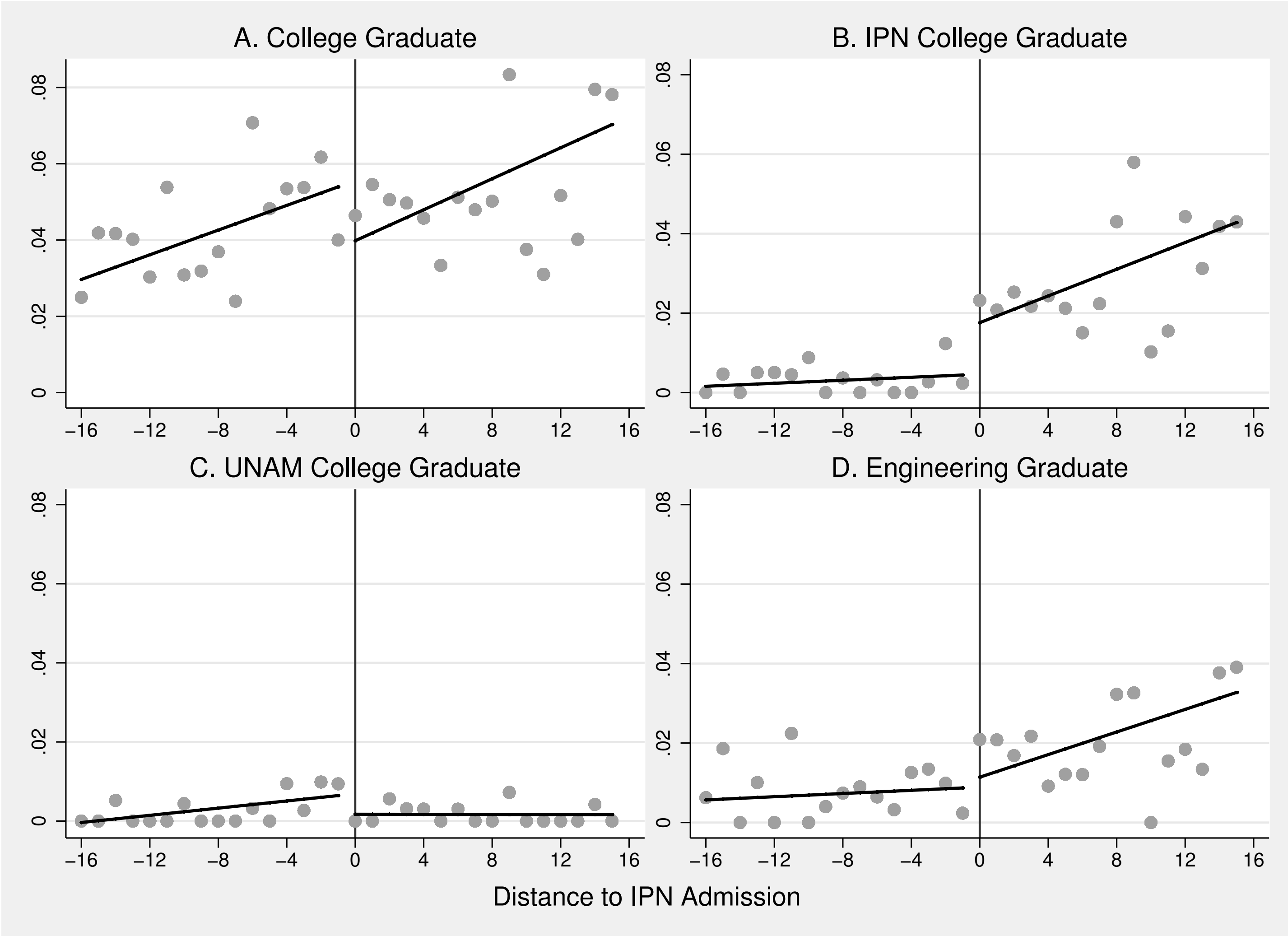
Notes: Conditional means of measures of learning achievement outcomes by distance of IPN entry exam test score from IPN admission cutoff. Source: COMIPEMS 2005, ENLACE 2008 and 2009. Sample: applicants from Mexico City junior high schools to IPN schools observed at ENLACE high school completion achievement test (and, for all outcomes except test scores, responding the ENLACE survey) .

Figure 5: Means of college outcomes by distance from IPN admission cutoffs



Notes: Conditional means of measures of student outcomes by distance of IPN entry exam test score from IPN admission cutoff. Source: COMIPEMS 2005, ENLACE 2008 and 2009. Sample: applicants from Mexico City junior high schools to IPN schools observed at ENLACE high school completion achievement test (and, for all outcomes except test scores, responding the ENLACE survey) .

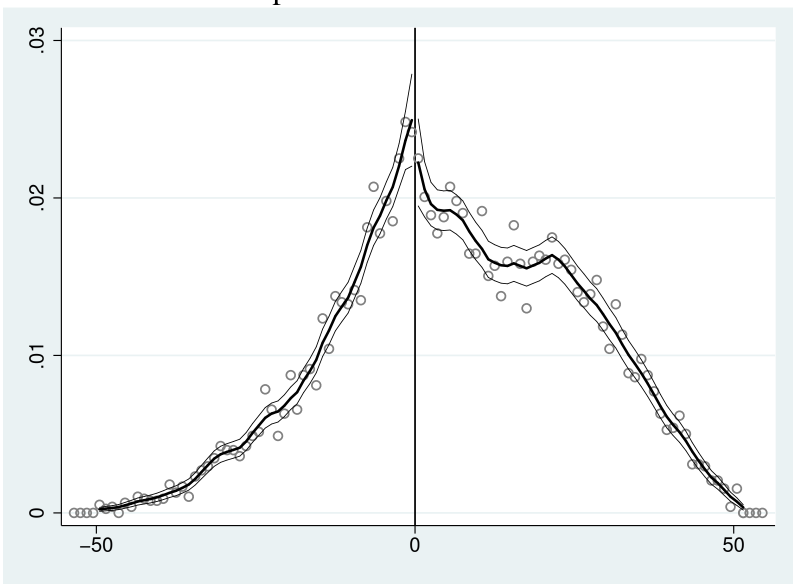
Figure 6: Means of college graduation outcomes by distance from IPN admission cutoffs



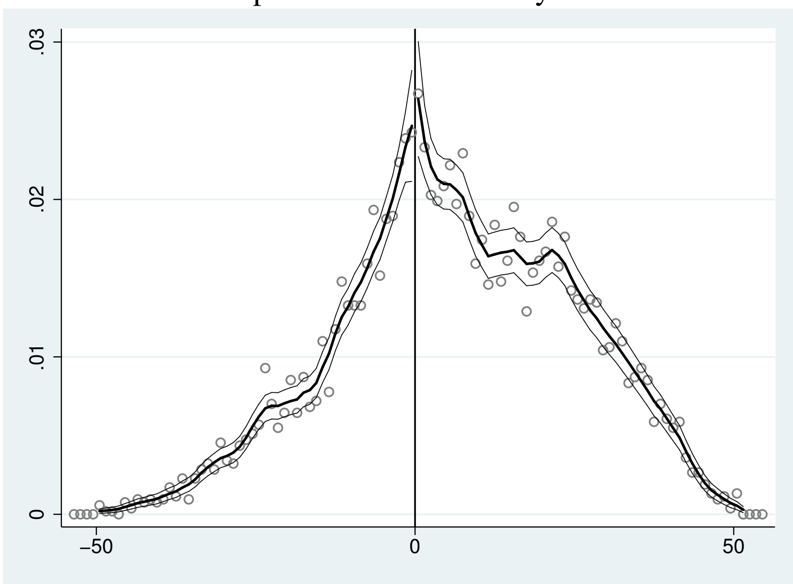
Notes: Conditional means of measures of college graduation outcomes by distance of IPN entry exam test score from IPN admission cutoff. Source: Registro Nacional de Profesionistas 2012-2015. Sample: applicants from Mexico City junior high schools to IPN schools observed at ENLACE high school completion achievement test.

Figure 7: Distribution of the distance of students' scores from IPN admission cutoffs at end to high school

A: sample of ENLACE test-takers



B: sample of ENLACE survey takers



Notes: Distribution of the distance of students' scores at the ENLACE last year test (Panel A) and survey (Panel B) from their cutoff scores for admission into the IPN system. Source: COMIPEMS 2005 school choices and entry exam scores. Sample: applicants from Mexico City junior high schools to COMIPEMS public high schools observed at ENLACE high school completion achievement test (a) and background survey (b).

Tables

Table 1: Student Characteristics and School Inputs

VARIABLES	(1) Other Schools	(2) IPN Schools
Panel A: Student Characteristics		
Junior HS GPA	-0.0636	0.711
	(0.982)	(1.011)
Private JHS	0.0313	0.118
	(0.174)	(0.323)
Female	0.517	0.333
	(0.500)	(0.471)
IPN applicant	0.341	1
	(0.474)	(0)
At least one parent has senior HS	0.269	0.487
	(0.444)	(0.500)
At least one parent is white-collar	0.204	0.368
	(0.403)	(0.482)
Observations	176,760	19,042
Panel B: School Inputs		
School Entry Score	-0.146	1.677
	(0.542)	(0.434)
SD of School Entry Score	0.663	0.516
	(0.129)	(0.0905)
Class Size	41.56	39.59
	(6.254)	(2.782)
Students per PC	9.779	3.521
	(13.62)	(1.133)
Share Teachers with College	0.807	0.856
	(0.125)	(0.133)
Share Full-Time Teachers	0.135	0.287
	(0.184)	(0.0952)
Observations	593	16

Notes: Means of characteristics of students (Panel A) and school inputs (Panel B) for the samples of students admitted to IPN (column (2)) and other (column (3)) public high schools. Other schools do not include UNAM schools. Column 1 reports standard errors, in parenthesis, for a t-test on the equality of means in columns 2 and 3. Source: COMIPEMS 2005 and schools census data.

Table 2: Applicants assigned to IPN and Non-IPN schools by school selectivity

	IPN		Non-IPN		IPN Students
	Schools	Students	Schools	Students	%
31	0	0	307	72,831	0
32-45	0	0	105	31,044	0
46-55	0	0	103	30,582	0
56-65	0	0	64	34,055	0
66-75	8	11,070	22	7,605	59
76-85	5	4,874	3	643	88
86-95	2	2,379	0	0	100
96-99	1	719	0	0	100
Total	16	19,042	604	176,760	

Notes: Distribution of 2005 Comipems applicants assigned to non-UNAM schools.

Table 3: Correlation between College Graduation Outcomes and Student Charactetistics

VARIABLES	(1) College Graduate	(2) IPN College Graduate	(3) UNAM College Graduate	(4) Engineering Graduate
IPN High School Student	-0.0133***	0.0394***	-0.00833***	0.0188**
	(0.00444)	(0.0147)	(0.00258)	(0.00882)
Entry Exam Score	0.0112***	0.00227***	0.00255***	0.00228***
	(0.00348)	(0.000775)	(0.000977)	(0.000851)
Junior HS GPA	0.0247***	0.00535***	0.00292***	0.00633***
	(0.00290)	(0.00130)	(0.000319)	(0.00129)
Private Junior HS	0.0356***	0.00348**	0.00415***	0.00600***
	(0.00366)	(0.00176)	(0.00117)	(0.00180)
At least one parent has High School	0.00948***	0.00162**	0.000748*	0.000463
	(0.00165)	(0.000704)	(0.000440)	(0.000540)
At least one parent is white-collar	0.00967***	0.00111*	0.000689**	0.000365
	(0.00205)	(0.000604)	(0.000336)	(0.000642)
Female	-0.00249**	-0.00165*	0.000516	-0.0120***
	(0.00120)	(0.000894)	(0.000476)	(0.00287)
Observations	190,177	190,177	190,177	190,177
R-squared	0.023	0.034	0.004	0.012
Mean Dependent Variable	0.0465	0.00635	0.00393	0.0115

Notes: OLS estimates of partial correlations between college graduation outcomes and student characteristics. Source: COMIPEMS 2005 and National Registry of Proffesionals 2012-2014. Sample: 2005 Comipems applicants. *** p<0.01, ** p<0.05, * p<0.1

Table 4: RD Design Estimates: Peer Characteristics and School Inputs

VARIABLES	(1)	(2)	(3)
School Entry Score	1.139***	1.128***	1.115***
	(0.140)	(0.0689)	(0.0477)
SD School Entry Score	-2.47e-05	0.0139	0.0103
	(0.0409)	(0.0198)	(0.0138)
Class Size	0.173	-1.692**	-1.646***
	(1.426)	(0.737)	(0.546)
Students per PC	-7.072***	-8.268***	-7.177***
	(1.718)	(0.862)	(0.588)
Teachers with College	0.0962***	0.0785***	0.0692***
	(0.0331)	(0.0179)	(0.0134)
Full Time Teachers	0.177***	0.116***	0.103***
	(0.0411)	(0.0210)	(0.0154)
Observations	462	1,028	1,855
Bandwidth	[2]	[5]	[10]
Clusters	160	401	744
chi2	239.8	818.0	1583
Prob > chi2	0	0	0

Notes: RD design estimates, using the bandwidths of respectively 2, 4 and 8 in the three columns, of the effects of IPN admission on peer characteristics and school inputs, i.e.: school peers' entry scores: mean and standard deviation, of students who graduated from a private junior HS and have a parent with HS education; class size, number of students per computer, share of teachers with a college degree and share of teachers employed full time in the school. All models include cutoff fixed effects interacted with a linear control function. Robust standard errors in parentheses are clustered at the distance to the admission cutoff X school level. Source: COMIPEMS 2005, ENLACE 2008 and 2009, Census of schools. Sample: applicants from Mexico City junior high schools to IPN schools observed in ENLACE survey. *** p<0.01, ** p<0.05, * p<0.1

Table 5: RD Design Estimates: Student Achievement

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Mathematics			Language		
Admitted	0.290**	0.296***	0.272***	0.0590	0.0531	0.0380
	(0.132)	(0.0918)	(0.0679)	(0.168)	(0.104)	(0.0678)
Observations	517	946	1,753	517	946	1,753
R-squared	0.195	0.159	0.189	0.101	0.089	0.116
Bandwidth	2	4	8	2	4	8
Clusters	175	350	674	175	350	674
Mean Non-Admitted	0.269	0.236	0.162	0.159	0.129	0.0558

Notes: RD design estimates of the effects of IPN admission on achievement in Mathematics (columns (1)-(3)) and Spanish (columns (4)-(6)) – those outcomes are standardized test-scores at Enlace high-school completion exam. All models include cutoff fixed effects interacted with a linear control function. Robust standard errors in parentheses are clustered at the distance to the admission cutoff X school level. Source: COMIPEMS 2005, ENLACE 2008 and 2009. Sample: applicants from Mexico City junior high schools to IPN schools observed in ENLACE survey. *** p<0.01, ** p<0.05, * p<0.1

Table 6: RD Design Estimates: Expected Earnings and Returns to College

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	Earnings HS			Earnings College			College Premium		
Admitted	0.0551	-0.0413	0.0302	0.209**	0.105**	0.156***	0.154	0.147**	0.125***
	(0.110)	(0.0627)	(0.0437)	(0.0910)	(0.0519)	(0.0370)	(0.109)	(0.0614)	(0.0422)
Observations	517	946	1,753	517	946	1,753	517	946	1,753
R-squared	0.035	0.032	0.019	0.054	0.056	0.032	0.058	0.049	0.023
Bandwidth	2	4	8	2	4	8	2	4	8
Clusters	175	350	674	175	350	674	175	350	674
Mean Non-Admitted	8.868	8.822	8.817	9.573	9.556	9.571	0.705	0.734	0.753

Notes: RD design estimates of the effects of IPN admission on expected wages with high school (columns (1)–(3)) and college (columns (4)–(6)), and on expected relative returns to college (columns (7)–(9)). Outcomes are logs of expected earnings with college and high school (columns (1)–(6)) and log of ratio of expected earnings with college and high school (columns (7)–(9)). All models include cutoff fixed effects interacted with a linear control function. Robust standard errors in parentheses are clustered at the distance to the admission cutoff X school level. Source: COMIPEMS 2005, ENLACE 2008 and 2009. Sample: applicants from Mexico City junior high schools to IPN schools observed in ENLACE survey. *** p<0.01, ** p<0.05, * p<0.1

Table 7: RD Design Estimates: College Attendance Intentions

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Intentions Undergraduate			Intentions Graduate		
Admitted	0.0910	0.0525	0.0427*	0.135	0.169***	0.157***
	(0.0588)	(0.0352)	(0.0258)	(0.106)	(0.0621)	(0.0432)
Observations	517	946	1,753	517	946	1,753
R-squared	0.114	0.054	0.041	0.107	0.058	0.051
Bandwidth	2	4	8	2	4	8
Clusters	175	350	674	175	350	674
Mean Non-Admitted	0.905	0.896	0.877	0.581	0.587	0.578

Notes: RD design estimates of the effects of IPN admission on declared intentions to attend college (columns (1)-(3)) and graduate school (columns (4)-(6)). All models include cutoff fixed effects interacted with a linear control function. Robust standard errors in parentheses are clustered at the distance to the admission cutoff X school level. Source: COMIPEMS 2005, ENLACE 2008 and 2009. Sample: applicants from Mexico City junior high schools to IPN schools observed in ENLACE survey. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 8: RD Design Estimates: College Graduation

VARIABLES	(1) College Graduate	(2) IPN College Graduate	(3) UNAM College Graduate	(4) Engineering Graduate
Admitted	0.00763	0.0153***	-0.0104*	0.0192***
	(0.0128)	(0.00527)	(0.00593)	(0.00556)
Observations	3,014	3,014	3,014	3,014
R-squared	0.016	0.023	0.021	0.013
Bandwidth	4	4	4	4
Clusters	664	664	664	664
Mean Non-Admitted	0.0520	0.00461	0.00789	0.00921

Notes: RD design estimates of the effects of IPN admission on having a college degree (column (1)), having a college degree and being an IPN college graduate (column (2)), having a college degree and being an UNAM college graduate (column (3)), having a college degree and being a college graduate from an Engineering degree (column (4)). All models include cutoff fixed effects interacted with a linear control function. Robust standard errors in parentheses are clustered at the distance to the admission cutoff X school level. Source: Registro Nacional de Profesionistas 2012-2015. Sample: applicants from Mexico City junior high schools to IPN schools. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Earnings differentials by graduation major

VARIABLES	(1) ln Monthly Wage	(2) ln Hourly Wage
Engineering Major	0.150***	0.0848**
	(0.0428)	(0.0425)
Female	-0.0888***	-0.00213
	(0.0326)	(0.0310)
Age	0.333***	0.204***
	(0.0713)	(0.0660)
Age squared	-0.00489***	-0.00278**
	(0.00123)	(0.00114)
Observations	4,447	4,447
R-squared	0.122	0.096
State of birth FE	Yes	Yes

Notes: OLS estimates of partial correlations between ln wage and individual characteristics. Source: National Employment Surveys (ENOE) from 3rd quarter of 2006 to 2nd quarter of 2010. Sample: employed 23-35 years-old individuals with a college education living in localities of more than 100,000 inhabitants of Mexico City and State of Mexico. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 10: Long-Run Outcomes and IPN Graduation

VARIABLES	(1) Attended College	(2) Active in Labor Market	(3) Unemployment	(4) ln Hourly Wage	(5) ln Hourly Wage	(6) ln Hourly Wage	(7) Engineer
IPN Graduate	0.343***	0.133***	-0.0673**	0.506***	0.588***	0.214*	0.163***
	(0.0518)	(0.0389)	(0.0312)	(0.0848)	(0.104)	(0.117)	(0.0625)
Female	-0.0612	-0.191***	-0.0240	-0.164**	-0.134*	-0.111	-0.170***
	(0.0399)	(0.0297)	(0.0251)	(0.0665)	(0.0734)	(0.103)	(0.0588)
Age	0.00345	-0.0148	-0.0799	0.119	0.324*	-0.0777	0.168
	(0.00631)	(0.0747)	(0.0639)	(0.172)	(0.185)	(0.277)	(0.142)
Age Squared		0.000297	0.00148	-0.00163	-0.00532	0.00241	-0.00276
		(0.00132)	(0.00113)	(0.00303)	(0.00325)	(0.00489)	(0.00249)
Mexico City Born	0.00803	-0.0277	-0.0656*	0.183**	0.0917	-0.00475	-0.189**
	(0.0543)	(0.0404)	(0.0342)	(0.0924)	(0.0942)	(0.194)	(0.0786)
Observations	568	568	475	292	179	113	231
R-squared	0.081	0.097	0.024	0.171	0.197	0.165	0.118
Sample	All	All	All	All	Only HS	College	College
Mean Non-IPN	0.324	0.812	0.0773				0.187

Notes: OLS estimates of partial correlations between IPN graduation and: college attendance (column (1)), labor market participation (column (2)), unemployment (column (3)), hourly wages (column (4)–(6)), engineering degree (column (7)). The sample comprises respondents from the module ENTELEMS of the National Labour Force Survey of the third quarter of 2008. Estimation is restricted to individuals 23 to 35 years old who graduated from a public high school from Mexico City and the State of Mexico in localities larger than 100,000 inhabitants. Graduates from UNAM high schools are excluded. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 11: Balance of Covariates at Baseline Assignment to Schools and end of High School

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	Assignment			End of Senior HS			Enlace Survey		
Junior HS GPA	-0.0853	-0.0352	-0.0177	0.136	0.128	0.170***	0.127	0.155	0.246***
	(0.0936)	(0.0582)	(0.0389)	(0.134)	(0.0795)	(0.0594)	(0.155)	(0.0952)	(0.0748)
Private Junior HS	-0.0120	-0.00963	-0.00566	0.0470	0.0497	0.0346	0.0904**	0.0601**	0.0333
	(0.0469)	(0.0269)	(0.0193)	(0.0563)	(0.0329)	(0.0219)	(0.0377)	(0.0283)	(0.0210)
Female	-0.0552	-0.0168	-0.0405*	-0.100	-0.0294	-0.00204	-0.129	0.00219	0.00729
	(0.0551)	(0.0330)	(0.0243)	(0.0830)	(0.0525)	(0.0371)	(0.104)	(0.0619)	(0.0433)
Number of Choices	0.356	-0.211	0.00909	0.826	-0.0616	0.112	0.598	-0.294	-0.259
	(0.579)	(0.348)	(0.242)	(0.698)	(0.427)	(0.285)	(0.792)	(0.496)	(0.318)
IPN Top 10	-0.149	-0.156	-0.0513	0.259	0.110	0.120	0.342	0.176	0.177
	(0.305)	(0.182)	(0.125)	(0.330)	(0.233)	(0.170)	(0.351)	(0.239)	(0.170)
High Demand Top 10	0.198	-0.0777	-0.0244	0.586	0.278	0.262	0.718	0.396	0.370*
	(0.327)	(0.173)	(0.129)	(0.400)	(0.244)	(0.174)	(0.458)	(0.280)	(0.198)
Parent with HS	-0.0218	-0.0453	-0.0221	0.0816	0.0596	0.0298	0.0640	0.0667	0.0472
	(0.0515)	(0.0315)	(0.0246)	(0.0786)	(0.0468)	(0.0345)	(0.0941)	(0.0594)	(0.0406)
Parent White Collar	0.0490	-0.00660	-0.00519	0.126	0.0558	0.0275	0.0863	0.0383	0.0113
	(0.0541)	(0.0317)	(0.0229)	(0.0781)	(0.0456)	(0.0325)	(0.0864)	(0.0527)	(0.0368)
Intentions Grad School	0.123**	0.0216	0.0183	0.175**	0.0372	0.0613	0.0751	0.00626	0.0444
	(0.0483)	(0.0326)	(0.0257)	(0.0696)	(0.0530)	(0.0397)	(0.0864)	(0.0607)	(0.0462)
Observations	1,633	2,989	5,502	710	1,314	2,517	516	946	1,763
Bandwidth	2	4	8	2	4	8	2	4	8
Clusters	340	659	1285	229	445	880	176	352	677
chi2	13.04	4.857	5.002	13.91	7.346	13.44	15.42	14.14	23.95
Prob > chi2	0.161	0.847	0.834	0.125	0.601	0.144	0.0799	0.118	0.00438

Notes: RD design estimates, using the bandwidths of respectively 2, 5 and 10 in the three columns, of the discontinuities associated with IPN admission in a set of covariates, i.e. junior high school GPA, attendance of a private junior high school, gender, number of school choices submitted in the COMIPEMPS allocation process, number of IPN schools and high demand schools in top 10 choices, high school graduate and white collar parent, and aspirations to graduate schools. All models include cutoff fixed effects interacted with a linear control function. Robust standard errors in parentheses are clustered at the distance to the admission cutoff X school level. Source: COMIPEMS 2005 and ENLACE 2008 and 2009. Sample: (1–3) applicants from Mexico City junior high schools to IPN schools observed at COMIPEMS test taking and (4–6) Enlace test taking.

Table 12: RD Design Estimates: Enlace-test and survey taking at end of high school

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES		Test taker			Survey taker	
Admitted	0.000777	-0.0464	-0.0485*	0.0409	0.0283	-0.00252
	(0.0692)	(0.0395)	(0.0282)	(0.0812)	(0.0451)	(0.0309)
Observations	1,636	2,992	5,506	1,636	2,992	5,506
R-squared	0.031	0.020	0.010	0.031	0.026	0.018
Bandwidth	2	4	8	2	4	8
Clusters	340	659	1285	340	659	1285
Mean Non-Admitted	0.462	0.464	0.475	0.308	0.313	0.307

Notes: RD design estimates, using the bandwidths of respectively 2, 5 and 10 in the three columns, of discontinuities associated with IPN admission in the probabilities of taking the Enlace exam (columns (1)–(3)) and the Enlace Survey (columns (4)–(6)) in 2008 or 2009. All models include cutoff fixed effects interacted with a linear control function. Robust standard errors in parentheses are clustered at the distance to the admission cutoff X school level. Source: COMIPEMS 2005. Sample: applicants from Mexico City junior high schools to IPN schools observed at COMIPEMS test taking.*** p<0.01, ** p<0.05, * p<0.1

Table 13: Non-IPN Elite Schools – Peer Characteristics and School Inputs

VARIABLES	(1)	(2)	(3)
School Entry Score	0.716***	0.687***	0.701***
	(0.135)	(0.0736)	(0.0497)
SD School Entry Score	-0.0584*	-0.0658***	-0.0733***
	(0.0330)	(0.0178)	(0.0126)
Share Students have Parent with HS	0.133***	0.121***	0.113***
	(0.0246)	(0.0143)	(0.00923)
Share Students from Private Junior HS	0.0434***	0.0393***	0.0392***
	(0.00895)	(0.00548)	(0.00344)
Class Size	-2.789**	-2.762***	-2.927***
	(1.304)	(0.740)	(0.515)
Students per PC	4.711***	3.522***	2.609***
	(1.452)	(0.828)	(0.571)
Teachers with College	0.00969	0.0268*	0.0194*
	(0.0245)	(0.0147)	(0.0106)
Full Time Teachers	-0.173***	-0.109***	-0.0738***
	(0.0386)	(0.0225)	(0.0154)
Observations	233	442	748
Bandwidth	2	4	8
Clusters	230	431	724
chi2	87.58	167.7	300.5
Prob > chi2	0	0	0

Notes: RD design estimates, using the bandwidths of respectively 2, 5 and 10 in the three columns, of the effects of admission to non-IPN selective high schools on peer characteristics and school inputs, i.e.: school peers' entry scores: mean and standard deviation, of students who graduated from a private junior HS and have a parent with HS education; class size, number of students per computer, share of teachers with a college degree and share of teachers employed full time in the school. All models include cutoff fixed effects interacted with a linear control function. Robust standard errors in parentheses are clustered at the individual level. Source: COMIPEMS 2005, ENLACE 2008 and 2009, Census of schools. Sample: applicants from Mexico City junior high schools to non-IPN selective high schools observed in ENLACE survey. *** p<0.01, ** p<0.05, * p<0.1

Table 14: Non-IPN Elite Schools – Outcomes

VARIABLES	(1) End of SHS	(2) End of SHS	(3) College Premium	(4) Wage SHS	(5) Wage College	(6) Math Score	(7) Language Score	(8) Intentions College	(9) Intentions Grad School
Admitted	0.00234	0.00412	0.0760	-0.00643	0.0696	-0.0115	-0.00988	0.0809	0.0850
	(0.0525)	(0.0486)	(0.0974)	(0.0951)	(0.0881)	(0.137)	(0.151)	(0.0755)	(0.0915)
Observations	2,011	2,011	444						
R-squared	0.009	0.011	0.029	0.044	0.029	0.115	0.159	0.044	0.049
Bandwidth	4	4	4	4	4	4	4	4	4
Clusters	299	299	160						
Mean Non-Admitted	0.402	0.232	0.790	8.796	9.586	0.130	0.187	0.876	0.614

Notes: RD design estimates, using the bandwidth of 4, of the effects of admission to non-IPN selective high schools on students' outcomes, i.e. completion of high school (column (1)), expected returns to college and wages (columns (2)–(4)) and scores at ENLACE exam (columns (5)–(6)), aspirations to attend undergraduate and graduate school (columns (7)–(8)). All models include cutoff fixed effects interacted with a linear control function. Robust standard errors in parentheses are clustered at the individual level. Source: COMIPEMS 2005, ENLACE 2008 and 2009, Census of schools. Sample: applicants from Mexico City junior high schools to non-IPN selective high schools observed in ENLACE survey. *** p<0.01, ** p<0.05, * p<0.1

Table 15: RD Design Estimates: Treatment Effects by Parental Education

VARIABLES	(1) College Premium	(2) Wage SHS	(3) Wage College	(4) Math Score	(5) Language Score	(6) Aspires College	(7) Aspires Grad School
Admitted	0.130*	-0.0301	0.0999	0.303***	-0.0865	0.0708	0.155**
	(0.0687)	(0.0663)	(0.0629)	(0.101)	(0.120)	(0.0436)	(0.0717)
Parent has high school	0.000203	0.0392	0.0394	-0.0700	-0.0932	0.0947***	0.0615
	(0.0459)	(0.0524)	(0.0470)	(0.0673)	(0.0807)	(0.0284)	(0.0502)
Interaction	0.0392	-0.0432	-0.00406	0.00189	0.245**	-0.0526	0.0141
	(0.0700)	(0.0686)	(0.0598)	(0.105)	(0.114)	(0.0371)	(0.0697)
Observations	946						
R-squared	0.050	0.034	0.059	0.163	0.097	0.070	0.063
Bandwidth	4	4	4	4	4	4	4
Clusters	350						
Mean Non-Admitted	0.725	8.811	9.536	0.238	0.150	0.862	0.565

Notes: RD design estimates, using the bandwidth of 4 and interaction terms, of heterogeneities, by parental education in the effects of IPN admission on expected returns to college and wages (columns (1)–(3)) and scores at ENLACE exam (columns (4)–(5)). All models include cutoff fixed effects interacted with a linear control function. Robust standard errors in parentheses are clustered at the distance to the admission cutoff X school level. Source: COMIPEMS 2005, ENLACE 2008 and 2009. Sample: applicants from Mexico City junior high schools to IPN schools observed in ENLACE survey. *** p<0.01, ** p<0.05, * p<0.1