Benefits to elite schools and the formation of expected returns to education: Evidence from Mexico City

Ricardo Estrada, Jérémie Gignoux

PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

WORKING PAPER N° 2014 – 06

Benefits to elite schools and the formation of expected returns to education: Evidence from Mexico City

Ricardo Estrada
Jérémie Gignoux

JEL Codes: D83, D84, I21

Keywords: Elite high schools ; Earnings expectations ; Returns to education ; Beliefs formation.

PARIS-JOURDAN SCIENCES ECONOMIQUES
48, BD JOURDAN – E.N.S. – 75014 PARIS
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10
www.pse.ens.fr

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE – ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES
ÉCOLE DES PONTS PARISTECH – ECOLE NORMALE SUPÉRIEURE – INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQU

Benefits to elite schools and the formation of expected returns to education: Evidence from Mexico City

Ricardo Estrada and Jérémie Gignoux*

February 2014

Abstract

We study the effects of admission into elite public high schools in Mexico City on students' expected earnings, arguing these effects provide an indication of the value-added those schools produce. Using data for the centralized and exam-based allocation of students into schools and an adapted regression discontinuity design strategy, we find that admission substantially increases learning achievement, and also the future earnings and returns students expect from a college education, but no effect on the earnings expected with high school education alone. This suggests that students believe that the benefits from their elite education are complements to a college education.

JEL: D83, D84, I21.

Keywords: elite high schools, earnings expectations, returns to education, beliefs formation.

1 Introduction

The benefits to attending selective high schools are in debate; at least, no clear pattern emerges from the literature on the effect of selective secondary schools on academic achievement. 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), but not in others (Lucas and Mbiti (2014) for Kenya). In the United States, results are clearly less optimistic (see Cullen et al. (2006) on Chicago public schools and Abdulkadiroglu et al. (2011) on New York and Boston exam schools) and a similar story goes for the UK (Clark (2010)). These mixed results are at odds with the sustained demand, from parents and students, that generates school selectivity and motivate a more comprehensive investigation of the benefits associated to selective schools.

*Estrada: Paris School of Economics (PSE), 48 boulevard Jourdan, 75014 Paris, restrada@pse.ens.fr. Gignoux: PSE and French National Institute for Research in Agronomy (INRA), 48 boulevard Jourdan, 75014 Paris, gignoux@pse.ens.fr.
We are grateful to François Bourguignon, Denis Cogneau, Julie Cullen, 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, and ZEW Mannheim University, and at the ECINEQ, LACEA, ESPE, EALE, ESSLE and EUDN conferences. 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 whishes to acknowledge funding from CONACYT.

One indication of the value-added that elite schools produce is the causal effect of attending such a school on graduating students' expected earnings. Intuitively, this effect is the difference between the wage that a student graduating from an elite institution expects to have at a given moment in the future and the wage that the same student would have expected to have if she would have attended a non-elite school. Hence, the wage premium of an elite education summarizes the students' valuation of the overall benefits obtained through their education – e.g. augmented skills, prestige of the school's diploma, access to networks, etc. – in terms of expected earnings in the labor market.

In addition to being informative about school quality, the causal effect of elite schools on students' expected earnings matters because, through changes in the expected returns to education, schools can shape educational choices and labor market outcomes. Human capital theory states that schooling decisions depend on the monetary benefits (along with the non-monetary ones) that youth expect from higher attainment. But Manski's (1993) assertion that it is the *heterogeneous* returns perceived by students that influence educational decisions spurred a growing literature which documents the link between perceived returns and educational investments. For instance, in the Mexican context, Attanasio and Kaufmann (2009) find that higher expected returns among high school graduates predicts higher college attendance (see also Jensen (2010) on Dominican Republic, Nguyen (2008) on Madagascar and Arcidiacono et al. (2012) and Zafar (2013) on choice of college major in the U.S.).

There are several ways in which elite schools might affect youths' perceived returns to education. A common concern is that some students – particularly those from a disadvantaged background – may have imperfect information on labor market opportunities and underestimate their real returns to education. In a pioneering study, Jensen (2010) finds that 8th graders in the Dominican Republic significantly underestimate returns to education. Providing information about higher measured returns to a random subset of students leads to higher perceived returns and school completion among these students. Hence, elite schools could give students access to more precise information on the returns to education, either through interactions with peers from more privileged backgrounds or through informational activities at school. In our context, Mexico City, we do not find evidence that high school students underestimate the returns to education. On average, students' perceived returns to college are close to the Mincerian college premium estimated from current workers' earnings. Still, it is possible that some students from disadvantaged backgrounds have biased estimations of their (specific) returns to education and update their beliefs when exposed to a privileged school environment.

Importantly, admission into elite schools can have an impact on students' expected earnings and returns to education even in a model with youth correctly informed on earnings. First, following human capital theory, if these are also better schools, elite schools could boost learning achievement and, in this way, wage expectations. Higher achievement translates into higher earnings expectations if students view scholastic learning (e.g. productivity) as a direct determinant of future earnings or an input for obtaining a more valuable higher education diploma. The same logic applies to the accumulation of non-cognitive skills that enter in the formation of earnings expectations, such as self-confidence or leadership. Note that the perceived returns to college are the difference between the wages expected with a college and a high school education. Hence, in this model, a positive effect of schools on expected returns requires that the expected wage premium of the augmented skills increases with educational attainment.

Second, 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 the expected earnings either directly in the labor market – with access to better paid jobs – or by facilitating the admission into more selective universities. The later mechanism is likely to be limited in our setting, though, as admission into the most prestigious public universities in Mexico is based on standardized exams. Also consistent with a model of asymmetric information in the labor market, elite schools could give students access to formal or informal networks that lead to better employment opportunities. As in the previous model, the effect on the expected returns to college depends on the expected benefits from the school's diploma and networks in the labor market for college and high school graduates.

In this paper, we examine the effect of admission into the National Polytechnic Institute ("Instituto Politécnico Nacional" or IPN), a system of 16 elite high schools of Mexico City Metropolitan area. Compared to other public high schools, IPN schools give students access to a set of enhanced inputs, including higher achieving peers, but also more qualified teachers, smaller class sizes and more IT equipment. IPN admission is also correlated with better outcomes in the short and long run. IPN students score higher on the high school completion scholastic test, are more likely to attend college and have better labor market outcomes.

First, we document that IPN students have higher expected returns to college than students from other public high schools, even when controlling for learning achievement and other individual characteristics. Then, we identify the causal effect of elite schools on expected wages and returns to college using a quasi-experimental setting with a panel database of students followed from application to the high school to 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. (2011) and Pop-Eleches and Urquiola (2013), by identifying the effect of marginal admission to the IPN system rather than to a specific IPN school.

We find that admission to an elite high school system increases substantially both the returns expected from a college education, with point estimates of about 19 percent points on the expected college premium, and scholastic achievement, with a jump of about .3 standard deviations. IPN students benefit from higher school inputs and acquire more valuable skills. However, the observed association between achievement and perceived returns suggests that the higher expected returns cannot stem uniquely from gains in learning achievement. Thus, IPN elite schools affect also perceived returns either by providing information on the returns, improving the access to valuable networks in the labor market for college graduates, or by enhancing non-cognitive skills that determine the perceived returns to college.

We do not find higher effects, though, on the expectations of students from a more disadvantaged background, which suggests that interactions with peers from a different background (or other mechanisms affecting in different ways richer and poorer students) are not explaining our results.[1]

Finally, admission into the IPN system increases substantially the expected earnings conditional on having a college education – with point estimates from 12% to 23% – but has no effect on the earnings expected with only a high school education. This suggests that graduating students believe that the valuable benefits associated to their elite high school education are complementary inputs for a college education.

Our results contribute to the literature on selective high schools by producing evidence on student outcomes beyond academic achievement. In the United States, Cullen et al. (2006) look at several behavior outcomes (like disciplinary incidents and arrest rates) and Abdulkadiroglu et al. (2011) at college admission patterns. Studies in developing countries with a good causal identification focus more in student achievement or schooling completion – as do Dustan et al. (2012), who use the same dataset we have. In an influential paper, Black (1999) uses variation in housing prices around school district boundaries in the United States to identify parental valuation of schools with higher achieving students. We use a different approach to estimate the valuation that students at the end of high school put on the elite education they got – which includes access to higher achieving peers, but also to other improved inputs.

We also provide evidence on the formation of the perceived returns to education, an area which remains poorly understood. In a recent study, Sequeira et al. (2013) finds that receiving a fellowship awarding academic performance increases girls' perceived returns to college in India, apparently by enhancing the perception that effort begets economic success. Also related, Battaglia and Lebedinski (2013) find evidence of role model effects on parental perceived returns in a remedial education program for primary school-age Roma children. Ours is the first paper that presents causal effects of school environments on expected earnings and returns to college.

The rest of the paper is as follows. Section 2 documents the differences in access to school inputs and earnings expectations between IPN and other students. Section 3 discusses the allocation of students into high schools and presents the data. Section 4 describes the empirical strategy. Section 5 reports and discusses the results, and Section 6 concludes.

[1] We nevertheless find in the descriptive analysis that in our context, conditional on academic achievement, students from a disadvantaged background tend to expect – perhaps rationally – slightly lower returns to college than more privileged youth. This is consistent with the observations of Delaney et al. (2011).

2 Mexico City elite high schools and earnings expectations

2.1 IPN elite schools

We study the effect of admission into any of the 16 high schools of the National Polytechnic Institute ("Instituto Politécnico Nacional" or IPN), a system of elite high schools.[2] Admission into IPN schools follows a centralized process shared with most public high schools of Mexico City, based on a common exam (and explicit school choices).[3] IPN schools provide general high school education with a scientific and technical background. IPN is also the main public technological higher education institution in Mexico, but attendance of an IPN high school does not grant access to the IPN higher education institute. Less than 10% of public high school students of Mexico City are able to attend an IPN school.[4]

Admission to an IPN school provides access to a set of enhanced inputs compared to other public high schools. Table 1 gives descriptive statistics for school inputs in IPN and non-IPN schools (based on information from the Mexican census of schools). Being selected among the best applicants, IPN students tend to have school peers who are high-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 in non-IPN schools, and the dispersion of achievement is lower, with an average school-level test score spread of .53 (unconditional) standard deviations versus .62 in other high schools.[5] Table 1 also shows that more IPN teachers have a college degree (86%) than teachers in other schools (81%), and that IPN teachers are more often in a full-time position (29%) than teachers in other schools (13%).[6] Furthermore, 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).

Not surprisingly, IPN students achieve better outcomes in the medium and long run. Table 2 gives regression estimates of correlations between attendance of, and graduation from, an IPN high school and 1) scores in an achievement test at the end of senior high school and 2) higher education and job market outcomes among individuals aged 23 to 35 years old who graduated from a public senior high school in Mexico City. (It uses data from the Enlace test at high school completion and the ENTELEMS labor force survey). IPN students obtain better scores at high school completion than students in non-elite public high schools with a gap in mathematics and Spanish, respectively, of 1.4 and 1 standard deviations.

[2]There are ten institutional systems of senior public high schools in Mexico City Metropolitan Area. IPN along with the "Universidad Nacional Autónoma de México" (UNAM) system stand out from the rest. Our data does not allow us to investigate the effects of attending a UNAM school though because UNAM students do not take the Enlace test at high school completion and thus drop out of our panel.

[3]We describe this process in Section 3.

[4]In 2005, the cohort among whom our data was collected, IPN schools admitted 19,042 applicants of the Mexico City centralized admission process, while other public high schools admitted 210,468. This number includes 34,625 students admitted to UNAM schools, but our analysis will focus on the effects of admission to an IPN versus a non-elite (and non-UNAM) school. Hence, we exclude UNAM students from all figures from here on.

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

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

Along the same lines, IPN graduates are more likely to attend college, with a differential of 34 percentage points compared to a baseline college attendance of 32%, 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%).

A number of studies have examined whether attending an elite school causally explains those better outcomes. Those studies have generally focussed the effects that transit through has on gains in scholastic learning. To account for a broader set of benefits from elite schools, we examine their effects on the wages and returns to college education that students expect after having completed their secondary education.

2.2 Wage expectations differentials across attended systems of high schools

To investigate the effects of attending an elite high school on expected wages and returns to college, we use data from a survey answered by a representative sample of students in the last year of high school, when they take the national Enlace achievement test. The survey gathers information on students' background, schooling experience and expected wages (we provide more details on this Enlace survey below.)[7]

Two questions elicit the information about the earnings expected with given educational attainments. Youth are asked questions about their 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. The answers are given using a pre-codified set of brackets. The questions are the following:

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 earnings brackets for both questions are: i) $4,000 or less; ii) $4,001 to $7,000; iii) $7,001 to $10,000; iv) $10,001 to $15,000; v) $15,001 to $20,000; and vi) more than $20,000.[8] 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" expected earnings, and we can obtain for each individual a point estimate of his expected earnings with high school and college education, as well as of the expected college premium. To obtain a measure of the implied expected college premium, 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, [$3,000–

[7]We present in more details the subjective expectations data we use, its interests and limitations, and conduct a number of consistency checks, in Appendix A1. For more information about the measurement of subjective expectations in developing countries, see surveys by Attanasio, 2009 and Delavande et al., 2011.

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

$4,000] and [$20,000–$27,000]. Then, the implied college premium, or expected returns to college, is given by the difference between log expected college earnings and log expected senior high school earnings:

$$expected\ \ returns = log\ \ (expected\ \ earnings \mid college) - log\ \ (expected\ \ earnings \mid high\ \ school).$$

The online appendix 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 the Federal District 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.

We are primarily interested in differentials between 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. The first result, in Figure 1 Panel A, is that, 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% confidence intervals. Earnings expectations of IPN students are 10% to 20% higher than those of non-IPN students. The differentials are clear across deciles. We observe also that the expected earnings of both groups tend to increase with scholastic achievement.

Second, see also Figure 1 Panel A, there are no clear differences in the earnings expected with only a high school education between IPN and non-IPN students. The IPN students have slightly higher expectations, but the 95% 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, the ones 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 of the two previous results, as shown in Figure 1 Panel B, IPN students tend to expect higher returns of attending college than non-IPN students, again after conditioning on achievement, with a difference of 5% to 15% that is marginally significant in 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.[9]

[9]The Appendix shows similar graphs for the dispersion, measured by the coefficient of variation (CV) of the earnings expectations of students in IPN and non-IPN schools conditional on learning achievements (for each decile). Expectations of earnings with college tend to be less dispersed among students in IPN schools, with a value of the CV lower by about .005 at all deciles except the fourth. Earnings expectations tend to be also less dispersed at higher levels of achievements. There

Table 3 reports OLS regression estimates to test if the observed differentials in expectations between IPN and non-IPN students persist after controlling for family and individual characteristics. The outcomes are logarithms of expected earnings and returns to college. The gaps between IPN and non-elite school students in expected earnings with college and returns to college remain with the controls. The estimates indicate that IPN students tend to expect both higher earnings with college (by 8%) and returns to college (by 6 percentage points). The difference in earnings with only high school is small (2%) 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. Unless it is entirely driven by selection into school systems, this pattern suggests that the attendance of an IPN school has some value on the labor market conditional on obtaining a college degree.

We discussed in the introduction that elite schools could shape students' expected earnings through access to better information, an increase in human capital and a signaling (or access to networks) effect. We can derive testable predictions that are consistent with the two first mechanisms. A pure information-provision mechanism implies both that disadvantaged students underestimate their returns to education and that IPN schools, with more informed peers or direct informational activities, correct this downward bias. We can check then if – given similar gains in achievement – the perceived returns to college of youth with less educated parents increase more than those of their more privileged peers given, exogenous, IPN admission. [10] However, the expectations data do not show evidence of underestimation of returns even among the sample of non-elite high school graduates.

If students are sufficiently informed about the labor-market value of the education they received, gains in earnings expectations should reveal increases in skills (or in awareness of own ability). In a human capital model, the effect of attending an elite high school on expected earnings should be positively associated with the school effect on the accumulation of cognitive or non-cognitive skills. We show in Table 1 that, when compared to other high schools, IPN schools provide to their students a set of enhanced inputs. Whether IPN (or more generally elite) schools can translate higher levels of inputs into enhanced students' skills is ultimately an empirical question that we also investigate.[11]

We return to these predictions in the empirical analysis below.

Now, the selection of students with higher earnings expectations into elite high schools could also

are also smaller gaps in the dispersion of expected earnings with high school and college premium, although the dispersion of expected college premium is lower among IPN students at most deciles.

[10] Although we only have anecdotal evidence on this, IPN schools may engage in activities to provide information to students on the returns to higher education. Also, given their institutional relations with the IPN higher education institute, the administrative and teaching personnel may have information on the job opportunities and wages, notably in the scientific and manufacturing sectors, that could be communicated to their students.

[11] Note that the predicted effect of elite schools on achievement might depend on students' background. On the one hand, attending a more selective school could have adverse effects on skill acquisition for students who could get marginalized either because of lower relative scores (that could make the teaching in those schools less suitable for them) or because they belong to a socioeconomic minority in the school they attend (e.g. from poor neighborhoods). On the other hand, some empirical studies (like Abdulkadiroglu et al. (2011)) have found that gains from attending very selective schools in the United States are larger for (if not restricted to) ethnic minorities.

drive the observed differences, if unobserved characteristics are correlated with the likelihood to attend IPN schools. To address this concern, we use the local natural experiments generated by the process of assignment of students into schools to identify the causal effects of admission into an elite IPN high school.

3 Data

3.1 The Comipems application system

The Metropolitan Commission of Public Senior-Secondary Education Institutions (Comipems) oversees nine of the ten systems of public high schools in Mexico City metropolitan area.[12] Since 1996, Comipems manages the allocation of students to high schools through a single 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 the serial dictatorship algorithm (Pathak, 2011).

The matching process operates, more specifically, in the following way. First, before taking the exam, applicants submit their ranking of preferred school choices (we use the terms of choice set for simplification). 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 schooling options (from 634 options in 2005). They actually submit a list of preferred tracks as some schools offer more than one track. We use the term school as synonymous of track, though, because most schools, notably IPN schools, have only one track at the time of admission. In practice, students receive at school an application package (to take home) in January of their last year of junior high school. Applicants must turn in their registration form (with their ranked sets of chosen schools) in February or March of the same year.

Second, applicants take a common standardized exam in the last weekend of June. All applicants with at least 31 correct answers out of 128 questions in the exam are allowed to register in a Comipems school. [13]

In the third step, students are allocated to schools. In practice, the Mexican Center of Evaluation of Education (Ceneval) uses a computer program that ranks students according to their exam score and proceeds to allocate individuals to the school with available seats they ranked the highest in their choice sets of preferred schools. As some schools are oversubscribed, not all applicants are admitted into their

[12] A recent system of high schools administered by the Federal District government and targeted to low-achieving students does not belong to Comipems. The Metropolitan area includes the Federal District and 22 municipalities from the neighboring State of Mexico.

[13] 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.

preferred option(s). In 2005, only one third of applicants gained admission into their first-choice school.

Finally, students who only chose 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% of the applicants.

Note that, in this system, schools are free to determine the number of seats they open, but afterwards simply apply the admission list produced by Ceneval based on students' stated choices and exam ranks. Neither schools' admission of students nor students' ranking of schools affects the priority that students have in the allocation process. Hence, no complex choice strategies would help students increase their probabilities of being admitted into their top choices. Under the serial dictatorship allocation mechanism, students' dominant strategy is to list their chosen schools in a way that is consistent with their true preferences, assuming students are not constrained in the number of schools they can list (*Pathak (2011)*). This later assumption seems satisfied here as, in 2005, the applicants from Federal District junior high schools submitted 9.2 schooling options on average and only 2% listed the maximum number of options allowed (which is 20).

The combination of the institutional setting and students' preferences produces a set of admission thresholds which are determined ex-post and cannot be predicted beforehand – where an oversubscribed school's admission threshold or cutoff is the exam score of the last student admitted to the school. As students tend to prefer schools with higher inputs, schools with higher admission thresholds offer on average better school inputs. As we discuss below, these admission thresholds generate variation in the allocation of students into schools that can be considered as locally random and exploited for identifying the causal effects of admission to more selective high schools.

3.2 Datasets

The data we use matches information from the 2005 Comipems admission process to information both from the 2008 and 2009 national achievement test (Enlace) of 12th graders of all high schools and from the 2008 and 2009 versions of a questionnaire survey of a representative sample of students who took the same Enlace test. It thus forms a panel dataset in which students are followed from application in 2005 up to graduation from high schools in 2008 or 2009 (unless they drop out before – we document this in depth below). We also match to this panel of students, information on the schools they attend from the 2005 version of the annual census of high schools carried out by the Secretary of Education (called "Formato 911").

The main outcomes of interest are measured in 2008 and 2009, as the cohort of 2005 Comipems applicants graduate from high school. The Mexican Evaluation of Scholastic Achievement of Educational Institutions (Enlace) is a national standardized exam taken, since 2008, by students in the last year of senior-secondary education (12th grade).[14] The purpose of the examination is to evaluate schools, and the educational system as a whole, and hence it has no bearing for students on graduation or university admissions. However, Enlace results are widely reported by media outlets and nongovernmental organizations in Mexico, and are used as the principal input for the creation of school league tables. The publicity provides school agents with incentives to perform better and makes Enlace a medium-stake test.[15]

In parallel to the Enlace exam and on the same day, the Federal Ministry of Education conducts a complementary survey among a random sample of students, gathering information on their individual characteristics, family background, and schooling experience.

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 Ministry of Education ('Secretaría de Educación Pública') data administration teams. 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 datasets contain the following information. The Comipems database includes the submitted listing of choices (tracks and schools), the score at the entry exam, the assignment outcome for all applicants, and also some family background information from a questionnaire attached to the registration form. The Enlace dataset contains the exam scores for all 12th graders. 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. The additional survey conducted at Enlace exam taking provides information on students' background and experience in high school, expected wages conditional on schooling levels and aspirations to pursue further education. Finally, the school census data provides information on the characteristics of schools, in particular: class size, teachers' profile and information technology equipment.

3.3 Samples

The Comipems process concerns most applicants to public senior-secondary schools in the Mexico City Metropolitan area. Excluding students admitted to UNAM schools (as we will from here on), 220,659 eligible applicants took the Comipems exam in 2005 (another 24,999 took the exam, but did not complete junior high on time). 195,802 (88.7%) applicants were allocated to 620 tracks. 2.3% of applicants did not achieve the minimum required score on the exam to access senior secondary education and 9.0% were

[14] Enlace exams are also taken by students at other grades.

[15] Students enrolled in UNAM schools do not take Enlace and, hence, are not considered in our analysis.

not admitted to any of their listed options and did not use the second-stage process to select a school with remaining seats. The pool of applicants includes students currently enrolled at junior high schools in the Federal District (34.0%) and in the participating localities from the State of Mexico (40.0%), but also youth who graduated in previous years (23.7%), who attend schools in other localities (1.7%) and students of special adult-education institutions (0.6%).

For estimating the effects of IPN admission, we focus on a restricted sample of 2005 Comipems applicants a) who graduated in the academic year 2004-2005 from a junior high school of the Federal District and b) who applied to at least one IPN school and were either admitted or rejected from such school.

We exclude returning-to-school and special adult-education students because they have lower chances of completing senior-secondary schooling and likely different responses to IPN admission. Similarly, as 15 of the 16 IPN high schools are located within the Federal District, we omit students from the neighboring State of Mexico. These restrictions reduce the sample to 75,137 applicants.

The restriction to eligible applicants of IPN schools is required because our identification strategy compares the outcomes of admitted and rejected IPN applicants and it reduces the sample to 18,523 applicants at the Comipems exam (baseline) – 7,783 admitted IPN applicants and 10,740 rejected. We explain formally our definition of rejected IPN applicants in Section 4. Intuitively, a rejected IPN applicant is one who, with a higher entry exam score, would have been allocated to an IPN school.[16]

Only a share of the IPN eligible applicants (51.4%) can be matched to the Enlace test in 2008 or 2009 (more on attrition below). This leads to samples of 9,119 students at Enlace test-taking, and 6,358 students (34.3%) at the Enlace complementary survey conducted among a random sub-sample of the test-takers. Note that applicants can opt for a private school after the admission results, or transfer to this system during their senior-secondary education. However, only 4.1% of the applicants in our sample take the Enlace test in a private high school.[17]

In practice, as we explain below, our estimators only use observations "close" to the IPN admission cutoffs.

3.4 Attrition

In our sample, attrition between the end of junior high and high school could be caused particularly by dropping out from school, but also by imperfect matching and Enlace turnout rates.

First, dropping out from high school is a widespread phenomenon in Mexico. In 2010, only 45% of the 25 to 34-year-old population in the country had completed a high school education. The same year,

[16]In principle, we omit 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, which use only observations "close" to the IPN admission cutoffs.

[17]Private schools do not belong to Comipems and manage admission decisions independently.

according to OECD data, the completion rate among all Mexican students admitted in high school three years before was 52%. As dropping out is correlated with family background, completion rates should be higher among students in Mexico City and notably IPN students. According to administrative statistics, the ratio of graduates in 2007 (regardless of the year of admission) to new entrants in 2004 was 62% in IPN high schools. As stated above, we observe 51.4% of the 18,253 IPN applicants in the Enlace dataset at the end of high school (45.7% in 2008 and 5.7% in 2009). This suggests that dropping out explains the bulk of the attrition between entry and completion of high school and our 2005 and 2008/09 data.

Second, we estimate that matching errors (due to incorrect or imprecise information on CURP, name and/or birth date) could contribute to overall attrition in our sample by up to 6 points. In the 2008 Enlace-exam database, we observe 13,453 students in Comipems schools who are not matched in the Comipems data. If all of them were eligible applicants from the 2005 Comipems process, matching errors would explain 6 points of attrition. However, some of those observations should correspond to students who did not take the Comipems exam in 2005, e.g. students admitted in 2004 or before or who migrated into Mexico City after completing the first year of senior-secondary education in another region (those school changes are possible).

Third, some 12th grade students, while still enrolled, could not show up at the Enlace exam. Exam participation might be affected by random absenteeism; or, more worrisome, some principals, looking to achieve a higher mean Enlace score in their school, could have encouraged, or at least not deterred, lower-performing students to skip school the day of the exam. If we approximate the Enlace exam turnout rate by the ratio of students who took the Enlace exam in Comipems schools in 2008 over the number of students enrolled in grade 12th reported in the school census at the beginning of the 2007–2008 academic year, the average Enlace turnout rate is estimated at 90%. Although dropouts during the year could lead to overestimate the Enlace turnout, this suggests that incomplete exam turnout could explain up to 5 points of attrition.

To summarize, our back-of-the-envelop calculations suggest that the attrition rate of 48.6% between the beginning and end of senior secondary education in our sample can be roughly decomposed in the following way: at least 38% due to school drop-outs, up to 6% to matching errors, up to 5% to incomplete Enlace turnout, and 3% due to other factors (such as more than one grade repetition). Beyond the overall composition of attrition, we are interested in verifying that there is no relationship between the treatment of interest, i.e. admission to an IPN school, and attrition, measured by Enlace exam taking. Before that, we need to present our identification strategy.

4 Empirical strategy

Our empirical strategy is based on a regression discontinuity (RD) design and provides estimates of the effects of admission into the elite IPN system. We focus on the effects of admission to a school system rather than a single school, thus departing from related studies of school effects, to make the interpretation of the estimated effects more straightforward. The identification of a system-level parameter in an admission process done at the school level requires some adjustments to the usual RD design strategy, notably in the definition of admission cutoffs.

4.1 Discontinuities in admission to the IPN school-system

To explicit 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.

Note that the ranking of schools in students' choice sets can differ from a ranking of schools by selectivity, so that 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 of schools or other school features could generate a ranking of schools that is not strictly decreasing in admission cutoffs. 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. However, 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.[18] One 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 $c_1 > c_2 > \ldots > c_{20}$.

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

[18] We insist that this simplification excludes all choices of schools which are more selective and less preferred than any other choice, and not only than the school the student was allocated to. Thus it does not introduce any selection above or below the c_K cutoff.

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.

However, one 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[19]).

To see this, assume that the choice set C_i of student i is such that the first L schools are all IPN schools and the next ones are all non-IPN schools, i.e.:

$$C_i = \left\{ \underbrace{c_1, ..., c_L}_{\in \Theta^{IPN}}, \underbrace{c_{L+1}, ..., c_{20}}_{\notin \Theta^{IPN}} \right\} \tag{1}$$

where Θ^{IPN} denotes the set of 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 a choice set C_i, one can define a cutoff score for admission to the IPN system as the lowest cutoff of the chosen IPN schools:

$$c_{IPN} = min \ \left\{ c(S_k), \ \ S_k \in C_i \cap \Theta^{IPN} \right\} \tag{2}$$

In practice, for a student admitted to a school of 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 one 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 the one of that first preferred IPN school.[20]

Note that, 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

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

[20]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 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 doesn't track UNAM students until high school completion.

choice set. Two variations can occur which however do not affect the way IPN admission cutoffs are obtained in practice. First, students can list one or several non-IPN schools before any IPN school in their choice sets, so that:

$$C_i = \left\{ \underbrace{c_1, ..., c_J}_{\notin \Theta^{IPN}}, \underbrace{c_{J+1} ..., c_L}_{\in \Theta^{IPN}}, \underbrace{c_{L+1}, ..., c_{20}}_{\notin \Theta^{IPN}} \right\} \tag{3}$$

with J non-IPN schools preferred to $L - J$ IPN ones. Only 0.1% of students have such simplified choice sets in our data. Our definition of IPN admission cutoffs accommodates such a choice set, pointing at c_L as the relevant cutoff. The only difference is that, with a score above c_J, the student would be allocated to a selective non-IPN school. We delete from our dataset the observations of students with such choice sets and scores above c_J as they cannot serve to identify the effects of admission to the IPN system (they do not receive the treatment and would not receive it even with a counterfactually higher score). Second, a slightly more tricky setting occurs when the set of IPN schools is disjoint with non-IPN schools inserted between IPN schools, i.e.:

$$C_i = \left\{ \underbrace{c_1 ..., c_I}_{\in \Theta^{IPN}}, \underbrace{c_{I+1}, ..., c_J}_{\notin \Theta^{IPN}}, \underbrace{c_{J+1} ..., c_L}_{\in \Theta^{IPN}}, \underbrace{c_{L+1}, ..., c_{20}}_{\notin \Theta^{IPN}} \right\} \tag{4}$$

Such choice sets are also rare in our data, representing about 7% of students admitted to IPN schools. In this case, one can still identify some intervals of scores over which the student is admitted to the IPN system ($s_i \in [c_L, c_J[$ or $s_i \geq c_I$), but there are now two IPN admission cutoffs and 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: starting from the school the student was allocated to, the relevant IPN cutoff would be either c_I or c_L.

In the online appendix we show in detail the IPN admission cutoffs for the sample of IPN applicants and the schools to which students are allocated in case they score below the cutoffs. The IPN admission cutoffs are in average 75.5 exam points (with a standard deviation of 7.37) and range from 66 to 99. Around 36% of the students who were not admitted to the IPN system were allocated to a school system with a technical background (DGETI and Conalep), 41% to a school system with a general education focus (Colegio de Bachilleres, DG Bachillerato and SE Edomex), and 26% to the second-stage admission process.[21]

The admission process then generates discontinuities in the relation between the students' score 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.

[21] Among those allocated to the second-stage process, 39% chose a general-education high school (Colegio de Bachilleres), 27% a technical-oriented school (DGETI or Conalep) and 33% did not register in any school.

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 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, the top panel of Figure 2 plots the distribution of the distance of students' scores from the relevant IPN admission cutoffs. No discontinuities are apparent at the cutoff (zero) in that distribution, which 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 3 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 described above make it possible to identify the effects of admission to a school of the IPN system using a RD design strategy. There are students who stated some choices C which lead to the same IPN admission cutoffs and achieved slightly different scores above and below those cutoffs, so that the ones with scores above get admitted to an IPN school and the others not. We can then identify the effect of IPN admission, for the group of students with choices that lead to the same IPN admission cutoffs, by comparing the outcomes of those who got admitted to the ones who were rejected. The variations in treatment near the cutoffs are as good as random and thus can serve to unveil a causal effect of admission on given outcomes (Imbens and Lemieux, 2008 and Lee and Lemieux, 2010). As applicants cannot locally self-select into the IPN system, their expected potential outcomes if 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. For each subsample of students with a same IPN admission cutoff, the effects of admission into the IPN system will then be given by the discontinuities in later outcomes at the cutoff. The discontinuities thus provide a set of local experiments which allow us to identify the effects of IPN admission at different admission cutoff scores.

Note that we do not need all treatment and comparison students to have stated the exact same choice sets for preferred schools. Because whether students' scores fall above or below the cutoffs is locally random, 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.

Furthermore, 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 schools system will present a discrete jump at the cutoff (from zero below) but may not be complete above it for two reasons.

First, some students may drop out before or during senior-secondary schooling. Selective 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. Such attrition would reflect in a discontinuity, at the admission cutoff, in the distribution of the Comipems score among students completing senior-secondary education. The bottom panel of Figure 2 plots the distribution of the distance of students' scores from their IPN admission cutoffs for the sample of students who completed high school: the distribution is very similar to the one at baseline (in top panel) indicating no differential attrition in our sample. We conduct complementary tests of the presence of differential attrition in Section 4.3.

Second, some students can decide not to attend the public school they were allocated to through the Comipems system and opt instead for a private school, or some could change school after the first or second year of high school. In the online appendix, 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 some will drop out (see below), the great majority of IPN admittees who complete senior-secondary schooling do attend an IPN school until completion. The high but incomplete take-up does not threaten empirical identification, but conducts to an intent-to-treat interpretation of the parameters of interest.

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 be estimated is:

$$Y_i = \alpha + \tau W_i + \beta d_i + \gamma d_i.W_i + \epsilon_i \tag{5}$$

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 (5) 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. For estimating this parameter, we aggregate all students with different IPN admission cutoffs into a single sample, incorporate cutoff fixed-effects in Equation (5), and estimate the model:

$$Y_{ic} = \alpha + \tau W_i + \delta_c + \sum_c 1\{c\}.\left[\boldsymbol{\beta_c} d_i + \boldsymbol{\gamma_c} d_i.W_i\right] + \epsilon_i \tag{6}$$

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 $\boldsymbol{\beta_c} d_i$ and$\boldsymbol{\beta_c} d_i + \boldsymbol{\gamma_c} d_i.W_i$ are now 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: larger bandwidths will increase precision by using more data but may induce more bias by relying more on the linear extrapolation on the two sides of the cutoff. The Imbens-Kalyanaraman (IK) optimal bandwidth (Imbens and Kalyanaraman, 2011) is a particular solution of this tradeoff obtained by minimizing the mean squared error; in our data, the IK bandwidth is about 2 and we consider this as our benchmark. We also use two larger bandwidths of 5 and 10 that allow higher precision. For the larger bandwidth of 10, in order to reduce extrapolation error, we replace the linear control function by a quadratic one. In addition, when estimating Equation (6), we cluster the standard errors at the level of the senior-secondary schools to which students are admitted to allow for common unobserved shocks at this level.

For those estimates, the sample of 18,523 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, 5 and 10 points, we have samples of respectively 1,359, 3,206 and 6,356 observations, and of 6,358, 471, 1,121 and 2,204 observations when restricting to students taking the Enlace survey.

Several points should be noted on those estimates. First, the estimated parameter τ captures the local effects of IPN admission for marginal IPN admittees. 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. As already mentioned, although it captures the effect of admission (students cannot be offered to choose between two Comipems schools after the allocation process), it has an intent-to-treat interpretation because students can opt for a private school, study in another region or leave the education system.

Second, the important feature of our strategy is that 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 admittees were allocated, and a control group of students who were not admitted to an IPN school. Our estimates differ from those proposed by several related studies, notably Abdulkadiroglu et al. (2011), and Pop-Eleches and Urquiola (2013), who consider school-specific cutoffs and focus on the marginal effects of admission to a more selective school compared to admission to a less selective school (which can be another elite 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, the differences in schooling environments and received inputs are more pronounced between school systems than between schools. This approach thus provides more variation for identifying the effects of school inputs.

4.4 Validity of the RD design at assignment and differential attrition

We follow Lee and Lemieux (2010) to analyze the validity of the regression discontinuity design, in a process analog to assess whether the randomization was carried out properly in a randomized experiment. 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.

We now 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 not able to sort above the IPN cutoffs. Table 4 (columns 1-3) gives the results of three sets of RD estimates, using the bandwidth of respectively 2, 5 and 10, for a set of baseline covariates (similar to the one used before) as dependent variables. We also test the joint significance of discontinuities in the full set of covariates using seemingly unrelated regressions (SUR). We do not find any evidence that any groups of students sort above the IPN admission threshold, and the results are consistent across specifications. With one exception, the admission coefficient is not statistically significant at the 10% level in any of the nine regressions reported in each column. The Chi-square test for the discontinuity indicator being zero in all equations takes high p-values of .607, .803, and .903 using the three bandwidths. This is strong evidence in support of the locally random-like variation of assignment to treatment.

We discussed in Section 3 the presence of attrition in our sample, due mainly to dropping out from school, between the beginning and end of senior-secondary education. We thus need to investigate the robustness of the RD design to this attrition at Enlace survey-taking. Figure 2 already provided visual evidence rejecting selective attrition linked to admission into the IPN system. We formally test the presence of such attrition using RD estimates of the effect of IPN admission on the probabilities of taking the Enlace exam and answering the Enlace survey. Results are reported in Table 5. The magnitude of the point estimates is negative but small, of -0.0398 using the bandwidth of 2, and we cannot reject that IPN admission has zero effect on the probability of taking the Enlace exam. The students of our sample admitted to IPN schools are thus not more or less likely than the non-admitted ones to drop out before completing high school. There is no clear-cut prediction of the relationship between IPN admission 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. The lack of effects in our data suggests that overall some countervailing effects balance. A similar result was obtained by Pop-Eleches and Urquiola (2013) using Romanian data.[22]

Now, even with no effect of admission into the IPN system on the probability of taking the Enlace exam on average, some heterogeneous effects across specific groups could remain possible. This would happen if, for example, IPN schools are 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 conditional on taking the Enlace exam. Table 4 (columns 4-6) gives the results of RD estimates (again for the three bandwidths), together with the SUR joint-significance test, of differences at the cutoff in the same characteristics used to evaluate the balance of covariates at Comipems assignment. With the bandwidths of 2 and 5, none of the admission coefficients in the nine equations estimated is statistically significant at the 10% level. With the bandwidth of 10, only one discontinuity, for junior high school GPA, is found statistically significant at the 5% level. However, the SUR tests reject the joint significance for the three bandwidths with p-values of respectively .405, .684 and .196. Thus we fail to reject that all discontinuities are jointly equal to zero, and overall students below and above the admission thresholds are not systematically different at Enlace taking.

These checks confirm that discontinuities in outcomes at completion of high school can be causally attributed to the locally exogenous allocation of students, generated by the exam-based allocation process, to schools of the IPN system.

[22]This is in contrast with the results recently obtained by Dustan et al. (2012) who find that admission to IPN schools increases the probability to drop out. The difference can be explained by the samples used. Our sample is restricted to Comipems applicants from schools of the Federal District, while Dustan et al. (2012) include applicants from neighboring municipalities. When admitted to IPN schools, the later may be more likely to drop out because all IPN schools (but one) are located in the Federal District.

4.5 Admission into IPN and school inputs received

Students who scored above the cutoffs and were admitted to an IPN high school are exposed to a different school environment and benefit from different school inputs than students who were rejected. Figure 4 plots the averages of a series of school characteristics of by distance to the IPN admission cutoff and shows that students scoring just below and above the cutoffs will experience rather different schooling environments during the following high school years. The graphs suggest that IPN students tend to be taught by more qualified and experienced teachers, in smaller class sizes, and have access to more computers, and probably other school inputs.

To confirm this, Table 6 gives RD design estimates of the effects of admission to an IPN school on several measures of the inputs received in senior-secondary schools. Consistently with Figure 4, the estimates indicate that students admitted to IPN schools have on average peers with test scores at the entry Comipems exam about 1.1 standard deviations higher, but not more or less variable, are taught by teachers who are 8–10% more likely to have graduated from college (20% of teachers of non-admitted students are not college graduates) and 11–17% more likely to work full time in the school (21% of teachers of non-admitted students are in a full-time position). They also benefit from smaller classes by .1 to 1.7 students (although the estimates with the bandwidths of 2 and 10 are not statistically significant; non-admitted students have in average class sizes of 42 students), and from much better access to computers with about four students per computer (versus 11 for non-admitted students). This confirms that IPN admittees are exposed to a set of enhanced inputs.

5 Results

5.1 Higher earnings expectations

Panels A to C of Figure 5 plot the local averages of the main outcomes we investigate, the expected earnings with a high school and college education and the associated 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 7 (columns 1-6) reports the RD design estimates of the effect of IPN admission on the earnings expected with a college and a high school education. Admission into the IPN system increases the expected earnings associated to a college education. The estimates of the three local linear regressions show large increases from 12–23% (all statistically significant at the 5% level). In comparison, average expected earnings below the cutoff are about $14,000 pesos (around 1,900 USD). In contrast, we do not find any effect on the wages expected conditional on staying with a high school education. The magnitude of the point estimates is rather small and the sign changes across specifications, from 1% with the bandwidth of 2 to -4%/-6% with the larger ones, and the effects are never statistically significant effect at the 10% level.

Table 7 (columns 7-9) presents the effect of IPN admission on the returns to college derived from the expected earnings with a college and a high school education. The regression results confirm that IPN admission increases the expected returns to college. Compared to average expected returns of about 75% below the cutoff, the point estimates indicate a large effect of about 18/22 percentage points, and are very robust to the bandwidth selection. 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 1% level and the point estimate remains unchanged in the smallest bandwidth, suggesting that bias is negligible in the larger bandwidths estimates.

Thus, our results show that IPN schools increase the earnings expected with a college education, but not with only a high school degree, which in turn increases the expected returns to college. This suggests that IPN students believe that only if they attend college can capture the large value-added obtained in their elite education.

5.2 Higher levels of skills

For understanding the mechanisms through which elite schools affect students' earnings expectations, one must account for the potential effect of elite schools on learning. Panel D of Figure 5 provides visual evidence of a discontinuity at the cutoff for achievement in mathematics, and the estimates in Table 8 confirm that admission into the IPN system increases substantially achievement in mathematics at high school completion. Again, the local regressions show large effects – .30 to .34 standard deviations and statistically significant at the 5% level – for all three bandwidths. On the other hand, we do not find a significant effect on achievement in Spanish, which may reflect the emphasis of IPN schools in scientific and technical fields.

Other studies find estimates of similar magnitude to the ones we observe for achievement in mathematics: Jackson (2010) (.3-.5 standard deviations) working on selective secondary schools in Trinidad and Tobago, and Pop-Eleches and Urquiola (2013) (.5-.15 standard deviations) in Romania.. In contrast, studies in developed countries (e.g. Cullen et al. (2006), Abdulkadiroglu et al. (2011), and Bui et al. (2011) for the United States) find no effects or a positive effect restricted to minority students. For instance, Abdulkadiroglu et al. (2011) find gains in Boston exam schools (of .17 standard deviations) for Blacks and Hispanics in English. The gains in achievement in mathematics associated with IPN admission lie thus in the upper part of the range identified in the literature, which confirms the capacity of this elite school system to combine inputs in a meaningful way. Also, the higher gains in mathematics – as opposed to language – are consistent with the findings in the literature on charter schools in the U.S. (see for example Dobbie and Fryer (2013)).

Regarding the formation of earnings expectations, we outlined that, in a human capital model, an increase in skills (productivity) would induce an increase in the expectation of future earnings. So the observation of simultaneous gains in cognitive skills and expected wages with college is compatible with a human capital model in which students believe (maybe correctly) that the labor markets for college and high school graduates reward differently cognitive skills. This explanation is consistent with the descriptive evidence presented in Figure 1, which suggests that students expect no wage gains from higher scholastic achievement if they stay with a high school degree.

However, the gains in scholastic skills (in mathematics) are not sufficient to explain the higher wages – conditional on college attendance – expected by IPN students. Going back again to results in Section 2.2, even after controlling for learning achievement and several individual characteristics, IPN students expect higher wages if they go to college than non-IPN ones. In addition, the slope of the relationship between achievement and expected returns (disregarding potential selection) seems too low for the .3 standard deviation increase in mathematics achievement to explain the approximately 20% increase in expected returns to college – the correlation coefficient between Enlace test scores in mathematics and expected returns to college in the 10 bandwidth sample amounts to 2 percentage points in expected returns for a 1 standard deviation increase in mathematics achievement.

The unexplained difference could stem in part from additional skills not captured by scholastic achievement. In particular, elite schools could enhance non-cognitive skills such as future-orientedness, self-confidence, leadership, etc. We do not have a comprehensive set of measures for students' non-cognitive skills at the end of high school. We construct though two indexes – using principal component analysis – for students' self-organization and attitude to school using related questions from the Enlace survey. The conditional expectation functions of both variables on the entry exam score are plotted in Panels F and G from Figure 5. Table 9 reports the corresponding parametric estimates for the effect of IPN system admission in these two proxies for non-cognitive skills.

We find a positive effect of admission to the IPN system on students' attitude to school. The point estimates go from .17 to .25 standard deviations and are statistically significant (at least at the 10% level) in two of the three bandwidths (the exception is the bandwidth of 2 exam points). There is no clear effect on self-organization, though. The point estimates have a positive sign, but their magnitude is smaller – from .03 to .18 standard deviations – and in none of the three bandwidths are statistically different from zero at the 10% level. Again, for non-cognitive skills to explain the effects of IPN admission on earnings expectations, we would need the skills to be relevant only conditional on college attendance.

Thus, the higher expected returns to college of IPN students could derive from the acquisition of additional skills that students value conditional on getting a college education. This would occur if the skills that elite high schools produce make college attendance more profitable either in terms of learning (complementarities in skill production) or earnings (complementarities in labor market productivity). Along these lines, we reported in the descriptive analysis that students seem to believe that the return to scholastic achievement in the labor market for high school graduates is small or insignificant, and the opposite in the labor market for college graduates. However, the gains in scholastic skills seem insufficient to explain the increase in expected earnings and returns to college, a gap that may be explained in part from a better development in IPN schools of other skills, including non-cognitive ones.

5.3 Disadvantaged students and no evidence of imperfect information

We investigate now if students from a disadvantaged background benefit also from admission to the IPN system and if there might be updating from imperfect information about the returns to college among these students. In short, we do not find evidence of heterogeneities by family background, and notably students from more disadvantaged backgrounds do not increase their earnings expectations any more than other students. For instance, the estimates in Table 10, with interaction terms, show that the effects do not differ by parental education in any of the considered outcomes: youth with more educated parents benefit in terms of learning achievement, and expected earnings and 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 secondary level.

Thus, students from a disadvantaged background benefit as much as their more privileged peers when admitted to IPN schools. More specifically, attendance of IPN schools do not increase the expected returns to college of disadvantaged youth more than others. We had already mention that the elicited perceived returns in our sample are close to the estimated returns using actual wages from current workers in Mexico City (see details in online appendix). Together, both pieces of evidence count against the story that disadvantaged youth correct their downward-biased perceived returns to education when exposed to better information in the IPN schools. Hence, an information channel cannot explain the observed gains in expected returns.

5.4 Higher schooling aspirations

For getting a sense of the long-term effects on students' trajectories associated with admission to an IPN school, one would ideally want to observe individuals' later outcomes in college (if they attend it) and on the labor market. Our data does not allow tracking students after high school completion, but the Enlace survey contains information on the educational attainment students aspire to reach.

The visual evidence in Panels H and I of Figure 5 is again suggestive of positive effects. Table 11 gives the results of RD design estimates of the effects of IPN admission on aspirations for undergraduate (first three columns) and graduate (last three) degrees. Although the estimates with the smaller bandwidth of 2 shows a positive effect of 8 percentage points on the probability to aspire to an undergraduate education, the effect is not statistically significant and is not robust to the larger bandwidths (the point estimate decreases to 2 points). The already high share of students who aspire for such education, about 90%, below the cutoff, may explain this result.

However, there is evidence for a positive effect on the probability to aspire to a graduate education, with increases of 11 percentage points with the smaller bandwidth of 2 and about 18-19 points with the two larger bandwidths, the latest two being significant at the 1% level. This effect is to be compared with a share of about 57% aspiring to a graduate education. Thus, the effects of admission to an IPN school on the financial gains from a college education seem to translate into higher educational aspirations.

While self-reported aspirations to reach a given attainment are likely an over-estimation of actual enrollment decisions (and completion rates) at that schooling level, the declared intention to achieve a certain schooling level may be informative of the likelihood of achieving the previous schooling level. Hence the increase in aspirations for a graduate education eventually translate into a higher likelihood of enrollment at the undergraduate level.

Besides, Attanasio and Kaufmann (2009), in an observational study of the relationship between perceived returns and attainment among a sample of Mexican students find that a one percentage point increase in the expected college premium is associated with a .2 to .4 percentage point higher probability (higher for boys) of attending college. These results are an average for the entire population and are descriptive, so do not compare exactly with our estimates of the local average treatment effect (LATE) for the population of IPN applicants around the admission cutoffs. But causal effects of such a magnitude would imply that the increase in earnings expectations we found could translate into a higher probability of college attendance of up to 4 to 8 percentage points. This estimate seems in the lower range of the magnitude, but consistent with the effects we obtain on aspirations for a graduate diploma and the observed gap in college attendance of 34 percentage points between IPN and non-elite-school graduates – reported in Table 2. This suggests that elite schools can affect decisions to enroll at college and have long-lasting consequences on individual outcomes.

5.5 Robustness checks

We examine the robustness of our results by adding a vector of baseline covariates as explanatory variables in our econometric model. As in the case of experimental designs with random assignment to treatment and control groups, in the absence of selective attrition (the main concern here) or manipulation of treatment assignment, the inclusion of baseline covariates should not affect the LATE estimates (Lee and Lemieux, 2010).

We include controls for a vector of junior high school GPA deciles fixed effects, schooling in a private junior high school, gender, parental education (indicator for whether at least one parent has senior secondary education), and parental occupation (indicator for whether at least one parent has a white collar job). This specification aims in particular to control for junior high school GPA and past attendance at a private junior high school, the only two variables for which there were some marginally significant (at the 10% level) differences between IPN admittees and rejected candidates at follow-up. In this specification, we are comparing the outcomes of students within the same decile of junior high school GPA or past schooling background that were marginally admitted and rejected from the IPN system. We also include three control variables for the set of school choices stated during the Comipems allocation process (although the local randomness of the IPN admission cutoffs should be orthogonal to those choices).

Table 12 reports the RD design estimates with the controls for the set of outcomes examined above for the bandwidth of 5. We obtain a similar set of results as before, with statistically significant effects of IPN admission on expected college premium, earnings expected with a college education, learning achievement and aspirations for a graduate education, but no effect on earnings expected with an high school education. The magnitudes of the coefficients across the specifications are also very similar to the ones reported in Tables 7 to 8 and 11. The results from our former RD design estimates thus are very robust to the inclusion of those controls.

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

For understanding better what inputs make a difference on students' expected earnings and returns to college and particularly if peers are a crucial input, we investigate the benefits from admission into non-IPN elite schools. As IPN schools are not the only selective ones in Mexico City, our empirical strategy can serve also to estimate the effects of admission into other selective public high schools. For this purpose, 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 -we present stacked results for this set of admission cutoffs to schools in different systems. In the online appendix we plot the average of the same school characteristics by distance to the cutoff and Table 13 provides the corresponding RD estimates. Admission into non-IPN elite high schools also provides access to peers with higher (by more than .75 standard deviations) average achievement, but not to systematically better other school inputs: although class sizes tend to be slightly smaller (by 2-3 students per class), those schools do not provide better access to computers and their teachers work full time less often than in other schools. So, these selective schools provide fewer inputs than the IPN ones.

In the online appendix we plot the averages of the main outcomes on the left and right of the cutoff for admission into non-IPN elite schools, and Table 14 gives the corresponding RD estimates. Admission into such schools has no effect on learning achievement at high school completion and does not either affect the returns students expect from a college education. The point estimates for the effects on wages expected with college are positive (of 6 to 12 percentage points), but these estimates are not statistically significant. Similarly, positive point estimates are found for the effects on aspirations for further education (undergraduate or graduate), but again the effects are not statistically significant. Thus, non-IPN elite schools also provide access to scholastically better peers, but do not generate the same gains as IPN schools. Our results suggest that IPN schools provide some specific inputs that do affect both skills and the returns students expect from college, and that those benefits do not merely stem from interactions with more carefully selected peers.[23]

6 Conclusions

In this article, we argue that the causal effects of attending specific schools on wage expectations provide a measure of the economic value that students receive from these schools (e.g. in form of augmented skills, signaling and access to networks) – given that students are sufficiently informed about the wages they will be able to obtain in the future.[24] We use this framework to examine the benefits from attending the IPN, a system of elite public schools in Mexico City metropolitan area. In this context, although expected wages are high, reported expectations of returns to college compare well with basic econometric estimates of wage returns to college, suggesting that lack of information is not prevalent.

We exploit data from a natural experiment to identify the effect of attending an elite school system on expected wages and returns to college education. The natural experiment arises from the allocation of students into high schools based on a centralized exam, common to all public schools of the area; we compare the outcomes of students just above and below IPN admission cutoff scores. From a methodological standpoint, while other studies use a related RD design framework (such as Abdulkadiroglu et al. (2011), and Pop-Eleches and Urquiola (2013)) , our strategy departs by identifying the effect of admission to an elite school system compared to another school system, rather than admission to more and less selective schools. This modified parameter of interest allows a more straightforward interpretation.

We find that admission to an elite school system increases substantially the earnings and returns students expect from a college education, with point estimates of the effect of IPN admission on the expected college premium of about 19 percentage points. Students who attend IPN schools thus attach economic benefits to the elite education they receive. But these students expect such benefits only if attending college – we find no effects on wages expected with only a high school degree. So the benefits that this elite school system produce seem to be more – or only – valuable in the labor market for college graduates or as a complement for a college education. While we do not find any statistically significant effect of IPN admission on the intention to attend college – already around 90% of students in the comparison group aspire to proceed to this level of schooling – we do find that students admitted to elite schools aspire to a graduate degree more often. The benefits that IPN schools generate could explain a share of the substantial differentials, that we document, in the college and labor market long-run

[23] We also examined heterogeneities by IPN school selectivity, splitting the set of IPN admission cutoffs into two subsets corresponding to less and more selective IPN schools. These results (available upon request) show that the gains stem mainly from admission to less selective IPN schools.

[24] While estimating the effects of elite education on earnings would be an obvious alternative, it can be done only after many years and requires long-run panel data that is not generally available.

outcomes between IPN graduates and graduates of other non-elite public high schools.

Our findings contribute to the debate on elite schools by using a broader measure of the economic benefits from attending such a school. Supporting our interpretation of gains in expected wages as an indication of school value-added, we find that admission to the IPN system increases learning achievement in mathematics at high school completion by about .3 standard deviations. This gain in learning confirms the capacity of the elite school system to combine inputs in a meaningful way.

We are also interested in the way school inputs shape student outcomes. Specificly, we investigate if the access to higher achieving peers can explain the gains we observe. The IPN elite schools we study provide to their students not only such peers, but also a number of better inputs, such as more qualified teachers and better access to information technologies. Those inputs, and maybe other unobserved ones, do make a difference. We compare the effects of IPN and other equally 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, suggesting that, beyond peers, the IPN schools provide specific inputs that matter. This finding echoes with recent studies (for instance Abdulkadiroglu et al. (2011)) that observe that schools only providing interactions with peers who are more carefully selected and scholastically rigorous have almost no effect on students' learning.

Our results also shed light on the formation of the perceived the returns to college, which are important determinants of decisions to enroll, as a growing literature shows. The finding of simultaneous gains in skills and expected returns to college confirms that students update their expectations as they accumulate more human capital. But the gains in scholastic skills seem insufficient to explain the increases in wage expectations, and some partial evidence suggests that IPN schools also increase non-cognitive skills. Although we do not provide direct evidence of it, IPN schools could also generate economic gains through signaling or access to networks. On the other hand, and differently from other studies (e.g. Jensen (2010), and Battaglia and Lebedinski (2013)), our results do not lend support to a pure informational or role model mechanism in the formation of perceived returns to college. In particular, we find no higher effects on the expectations of students from more disadvantaged background, the ones most likely to have inordinately low expectations due to lack of information.

These results have a potential use to policymakers: information on expected wages – and other expectations regarding future employment, for example – could be used systematically as outcomes to evaluate school quality, as a complement to measures of scholastic learning. In addition, while much attention has been devoted to the effects of different school environments on scholastic learning, our results suggest that schools may help students build other (cognitive or non-cognitive) skills that have value either on the labor market or in a college education. Identifying more precisely those skills and what features in the environment of elite schools produce them is key for the analysis of educational policies.

References

Abdulkadiroglu, Atila, Joshua D. Angrist, and Parag A. Pathak, "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools," NBER Working Papers 17264, National Bureau of Economic Research, Inc July 2011.

Arcidiacono, Peter, V. Joseph Hotz, and Songman Kang, "Modeling college major choices using elicited measures of expectations and counterfactuals," *Journal of Econometrics*, 2012, *166* (1), 3–16.

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

_ **and Katja Kaufmann**, "Educational Choices, Subjective Expectations, and Credit Constraints," NBER Working Papers 15087, National Bureau of Economic Research, Inc July 2009.

Battaglia, Marianna and Lara Lebedinski, "The Curse of Low Aspirations: Remedial Education and Perceived Returns to Education of Roma People," Working paper, Bocconi University 2013.

Black, Sandra E., "Do Better Schools Matter? Parental Valuation Of Elementary Education," *The Quarterly Journal of Economics*, May 1999, *114* (2), 577–599.

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," NBER Working Papers 17089, National Bureau of Economic Research, Inc May 2011.

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

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.

Delaney, Liam, Colm Harmon, and Cathy Redmond, "Parental education, grade attainment and earnings expectations among university students," *Economics of Education Review*, 2011, *30* (6), 1136–1152.

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.

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, "The Benefits and Hazards of Elite High School Admission: Academic Opportunity and Dropout Risk in Mexico City," Working Papers, UC Berkeley 2012.

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," Working Papers 14-03, University of Delaware 2014.

Manski, Charles F., "Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?," in "Studies of Supply and Demand in Higher Education" NBER Chapters, National Bureau of Economic Research, Inc, July 1993, pp. 43–60.

Nguyen, Trang, "Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar," Working paper, Massachusetts Institute of Technology 2008.

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.

Sequeira, Sandra, Johannes Spinnewijn, and Guo Xu, "Rewarding Schooling Success and Perceived Returns to Education: Evidence from India," Working Papers, LSE 2013.

Zafar, Basit, "College Major Choice and the Gender Gap," *Journal of Human Resources*, 2013, *48* (3), 545–595.

Tables

Table 1: Student Characteristics and School Inputs

VARIABLES	(1) Other Schools	(2) IPN Schools
Panel A: Student Characteristics		
Junior high GPA	-0.0636	0.711
	(0.982)	(1.011)
Private junior high	0.0313	0.118
	(0.174)	(0.323)
Female	0.517	0.333
	(0.500)	(0.471)
No. Choices submitted	8.873	9.347
	(3.580)	(3.772)
IPN Schools in top 10 choices	0.659	4.196
	(1.266)	(2.329)
High Demand Schools in top 10 choices	2.551	3.305
	(2.219)	(2.041)
At least one parent has high school	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 and standard deviations (in parentheses) of characteristics of students (Panel A) and school inputs (Panel B) for the samples of students admitted to IPN (2nd column) and other (1st column) public high schools. Other schools do not include UNAM schools. Source: COMIPEMS 2005 and schools census data.

Table 2: Short-Run and Long-Run Outcomes and IPN Graduation

VARIABLES	(1) Math Score	(2) Language Score	(3) Attended College	(4) Active in Labor Market	(5) Unem-ployment	(6) ln Hourly Wage	(7) ln Hourly Wage	(8) ln Hourly Wage
IPN Graduate	1.382***	0.989***	0.342***	0.135***	-0.0622**	0.509***	0.587***	0.214*
	(0.0829)	(0.0766)	(0.0516)	(0.0388)	(0.0312)	(0.0852)	(0.104)	(0.116)
Female	-0.280***	0.175***	-0.0614	-0.190***	-0.0231	-0.145**	-0.119*	-0.111
	(0.0138)	(0.0153)	(0.0399)	(0.0297)	(0.0252)	(0.0661)	(0.0719)	(0.103)
Age			0.00354	-0.0166	-0.0866	0.195	0.361**	-0.0794
			(0.00628)	(0.0747)	(0.0640)	(0.168)	(0.181)	(0.267)
Age Squared				0.000323	0.00158	-0.00289	-0.00593*	0.00244
				(0.00132)	(0.00113)	(0.00297)	(0.00319)	(0.00472)
Observations	81,415	81,415	568	568	475	292	179	113
R-squared	0.236	0.104	0.081	0.096	0.016	0.159	0.192	0.165
Sample	Enlace	Enlace	Entelems	Entelems	Entelems	Entelems	Only HS	College
Mean Non-IPN	-0.158	-0.101	0.324	0.812	0.0773			

Notes: OLS estimates of partial correlations between IPN graduation and several short and long-run outcomes, including: scores at high-school completion achievement test (columns (1)-(2)), college attendance (column (3)), labor market participation (column (4)), unemployment (column (5)), hourly wages (column (6)-(8)). Columns (1)-(2): The sample comprises respondents from the 2008 and 2009 ENLACE achievement survey. Columns (3)-(8): The sample comprises respondents from the module ENTELEMS attached to 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 the Federal District (D.F.) 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 3: Partial Correlations: Earnings Expectations and IPN Attendance

	(1)	(2)	(3)
VARIABLES	Implied Premium	Earnings HS (ln)	Earnings College (ln)
IPN student	0.0569***	0.0194	0.0764***
	(0.0123)	(0.0132)	(0.0125)
High school score	0.0485***	-0.0256***	0.0229***
	(0.00600)	(0.00526)	(0.00455)
Entry exam Score	-0.00524	0.0353***	0.0301***
	(0.00554)	(0.00554)	(0.00479)
Junior high GPA	0.00625**	-0.0115***	-0.00520*
	(0.00247)	(0.00314)	(0.00267)
Private junior high	0.0211*	0.0367***	0.0578***
	(0.0109)	(0.0110)	(0.0106)
At least one parent has high school	0.0162***	0.0174***	0.0336***
	(0.00603)	(0.00663)	(0.00512)
At least one parent is white-collar	0.0199***	-0.0188**	0.00112
	(0.00685)	(0.00778)	(0.00590)
Asset Index	0.00897***	0.0190***	0.0280***
	(0.00260)	(0.00318)	(0.00258)
Female	0.0823***	-0.0483***	0.0340***
	(0.00579)	(0.00617)	(0.00633)
Indigenous Origin	-0.0119	-0.0126	-0.0246**
	(0.0125)	(0.0124)	(0.0113)
Constant	0.681***	8.816***	9.497***
	(0.00648)	(0.00655)	(0.00597)
Observations	44,064	44,064	44,064
R-squared	0.026	0.009	0.037

Notes: OLS estimates of partial correlations between IPN school attendance (and individual characteristics) and expected college premium (Column (1)), wages with senior high-school (Column (2)) and college (Column (3)) degrees. Outcomes are logs of expected earnings with college and high school and ratio (returns). The sample includes 12th graders in public high schools of Mexico City Metropolitan Area who answered the Enlace Survey in 2008. Robust standard errors in parentheses are clustered at the school level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4: Balance of Covariates at Baseline Assignment to Schools and end of High School

EQUATION	VARIABLES	(1) Assignment [2]	(2) Assignment [5]	(3) Assignment [10]	(4) End [2]	(5) End [5]	(6) End [10]
JHS GPA	Admitted	-0.144	0.0111	0.0448	0.0275	0.112	0.158**
		(0.101)	(0.0494)	(0.0526)	(0.138)	(0.0771)	(0.0797)
Private JHS	Admitted	-0.0421	-0.0171	-0.00900	0.0357	0.0340	0.0443
		(0.0425)	(0.0238)	(0.0227)	(0.0624)	(0.0304)	(0.0316)
Female	Admitted	-0.0768	0.00537	0.00215	-0.128	-0.0337	-0.0461
		(0.0611)	(0.0329)	(0.0376)	(0.0997)	(0.0471)	(0.0548)
Number of Choices	Admitted	0.710	-0.171	-0.145	0.767	-0.208	-0.0679
		(0.527)	(0.301)	(0.302)	(0.786)	(0.407)	(0.425)
IPN in Top 10	Admitted	-0.0730	-0.226	-0.171	0.417	-0.0178	0.0197
		(0.308)	(0.160)	(0.170)	(0.306)	(0.184)	(0.198)
High Demand in Top 10	Admitted	0.225	-0.0774	-0.00817	0.418	0.269	0.227
		(0.268)	(0.171)	(0.184)	(0.439)	(0.210)	(0.242)
One Parent has SHS	Admitted	0.341	-0.00162	0.153	0.617	0.0982	0.150
		(0.309)	(0.178)	(0.183)	(0.412)	(0.220)	(0.239)
One Parent is White Collar	Admitted	0.338	0.120	0.223	0.0666	0.0493	-0.00310
		(0.333)	(0.171)	(0.181)	(0.523)	(0.232)	(0.252)
	Observations	1,359	3,206	6,356	631	1,541	3,142
	Control Fn	Linear	Linear	Quadratic	Linear	Linear	Quadratic
	Clusters	208	277	344	153	216	287
	chi2	7.291	5.251	3.496	9.253	6.217	11.71
	Prob > chi2	0.506	0.730	0.900	0.321	0.623	0.165

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. All models include cutoff fixed effects fully interacted with the control function. Robust standard errors in parentheses are clustered at the school level. Source: COMIPEMS 2005 and ENLACE 2008 and 2009. Sample: (1-3) regular applicants from D.F. junior high schools to IPN schools observed at COMIPEMS test taking and (4-6) Enlace test taking.

Table 5: RD Design Estimates: Enlace-taking at end of high school

VARIABLES	(1) End of HS	(2) End of HS	(3) End of HS	(4) Surveyee	(5) Surveyee	(6) Surveyee
Admitted	-0.0398	-0.0362	-0.0496	0.0356	0.0210	0.00386
	(0.0633)	(0.0339)	(0.0375)	(0.0706)	(0.0392)	(0.0400)
Observations	1,359	3,206	6,356	1,359	3,206	6,356
R-squared	0.028	0.015	0.014	0.035	0.015	0.020
Bandwidth	[2]	[5]	[10]	[2]	[5]	[10]
Control Fn	Linear	Linear	Quadratic	Linear	Linear	Quadratic
Clusters	208	277	344	208	277	344
Mean Non-Admitted	0.487	0.507	0.509	0.337	0.349	0.337

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 fully interacted with the control function. Robust standard errors in parentheses are clustered at the school level. Source: COMIPEMS 2005. Sample: regular applicants from D.F. junior high to IPN schools observed at COMIPEMS test taking.*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6: School Inputs

EQUATION	VARIABLES	(1)	(2)	(3)
School Entry Score	Admitted	1.125***	1.136***	1.117***
		(0.110)	(0.0657)	(0.0671)
SD School Entry Score	Admitted	-0.00736	0.00801	0.00878
		(0.0319)	(0.0209)	(0.0221)
Class Size	Admitted	-0.109	-1.714**	-1.273
		(1.238)	(0.779)	(0.856)
Students per PC	Admitted	-6.918***	-7.921***	-7.698***
		(1.311)	(0.870)	(0.849)
Teachers with College	Admitted	0.0979***	0.0818***	0.0888***
		(0.0276)	(0.0185)	(0.0201)
Full Time Teachers	Admitted	0.173***	0.112***	0.115***
		(0.0380)	(0.0221)	(0.0263)
Observations		467	1,107	2,176
Bandwidth		[2]	[5]	[10]
Control Fn		Linear	Linear	Quadratic
Clusters		123	175	232
chi2		302.8	916.1	806.5
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 IPN admission on received school inputs, i.e.: school peers' entry scores: mean and standard deviation, class size, number of students per computer, share of teachers with a college degree, share of teachers employed full time in the school. All models include cutoff fixed effects fully interacted with the control function. Robust standard errors in parentheses are clustered at the school level. Source: COMIPEMS 2005, ENLACE 2008 and 2009, Census of schools. Sample: regular applicants from D.F. junior high to IPN schools observed in ENLACE survey. *** p<0.01, ** p<0.05, * p<0.1

Table 7: RD Design Estimates: Expected Earnings and Returns to College

VARIABLES	(1) Earnings HS [2]	(2) Earnings HS [5]	(3) Earnings HS [10]	(4) Earnings College [2]	(5) Earnings College [5]	(6) Earnings College [10]	(7) Expected return [2]	(8) Expected return [5]	(9) Expected return [10]
Admitted	0.0113	-0.0610	-0.0361	0.231**	0.116**	0.152***	0.220*	0.177***	0.189***
	(0.123)	(0.0559)	(0.0621)	(0.0920)	(0.0470)	(0.0493)	(0.123)	(0.0604)	(0.0608)
Observations	470	1,115	2,184	470	1,115	2,184	470	1,115	2,184
R-squared	0.024	0.031	0.022	0.036	0.047	0.034	0.040	0.040	0.028
Control Fn	Linear	Linear	Quadratic	Linear	Linear	Quadratic	Linear	Linear	Quadratic
Clusters	125	179	236	125	179	236	125	179	236
Mean Non-Admitted	8.863	8.815	8.790	9.561	9.555	9.564	0.698	0.740	0.774

Notes: RD design estimates, using the bandwidths of respectively 2, 5 and 10 in the three columns, of the effects of IPN admission on expected wages with high school (columns (1)-(3)) and college (columns (4)-(6)), and on expected returns to college (columns (7) - (9)). Outcomes are logs of expected earnings with college and high school (columns (1)-(6)) and ratio of logs of expected earnings with college and high school (columns (7) - (9)). All models include cutoff fixed effects fully interacted with the control function. Robust standard errors in parentheses are clustered at the school level. Source: COMIPEMS 2005, ENLACE 2008 and 2009. Sample: regular applicants from D.F. junior high to IPN schools observed in ENLACE survey. *** p<0.01, ** p<0.05, * p<0.1

Table 8: RD Design Estimates: Student Achievement

VARIABLES	(1) Math Score	(2) Math Score	(3) Math Score	(4) Language Score	(5) Language Score	(6) Language Score
Admitted	0.300**	0.343***	0.313***	0.0530	0.0790	0.0828
	(0.138)	(0.0856)	(0.0948)	(0.170)	(0.0983)	(0.0920)
Observations	470	1,115	2,184	470	1,115	2,183
R-squared	0.181	0.145	0.210	0.088	0.072	0.127
Bandwidth	[2]	[5]	[10]	[2]	[5]	[10]
Control Fn	Linear	Linear	Quadratic	Linear	Linear	Quadratic
Clusters	125	179	236	125	179	236
Mean Non-Admitted	0.285	0.262	0.143	0.181	0.127	0.0820

Notes: RD design estimates, using the bandwidths of respectively 2, 5 and 10 in the three columns, of the effects of IPN admission on achievement in Mathematics (columns (1)-(3)) and Spanish (columns (4)-(6)). Outcomes are standardized test-scores at Enlace high-school completion exam. All models include cutoff fixed effects fully interacted with the control function. Robust standard errors in parentheses are clustered at the school level. Source: COMIPEMS 2005, ENLACE 2008 and 2009. Sample: regular applicants from D.F. junior high to IPN schools observed in ENLACE survey. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 9: RD Design Estimates: Non-Cognitive Skills

VARIABLES	(1) Self Organization	(2) Self Organization	(3) Self Organization	(4) Attitude to School	(5) Attitude to School	(6) Attitude to School
Admitted	0.180	0.0340	0.0578	0.169	0.253**	0.242*
	(0.220)	(0.126)	(0.139)	(0.228)	(0.119)	(0.125)
Observations	454	1,069	2,091	454	1,070	2,097
R-squared	0.047	0.020	0.028	0.040	0.027	0.032
Bandwidth	[2]	[5]	[10]	[2]	[5]	[10]
Control Fn	Linear	Linear	Quadratic	Linear	Linear	Quadratic
Clusters	121	177	234	122	177	234
Mean Non-Admitted	-0.0769	-0.00694	0.00408	-0.0898	-0.0340	0.0166

Notes: RD design estimates, using the bandwidths of respectively 2, 5 and 10 in the three columns, of the effects of IPN admission on two indexes of non-cognitive skills: self-organization (columns (1)-(3)) and attitude to school (columns (4)-(6)). Outcomes are indexes based on principal component analysis and are normalized with mean 0 and standard deviation 1. All models include cutoff fixed effects fully interacted with the control function. Robust standard errors in parentheses are clustered at the school level. Source: COMIPEMS 2005, ENLACE 2008 and 2009. Sample: regular applicants from D.F. junior high to IPN schools observed in ENLACE survey. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 10: RD Design Estimates: Treatment Effects by Student Characteristics

VARIABLES	(1) College Premium	(2) Wage HS	(3) Wage College	(4) Math Score	(5) Language Score
Admitted	0.170***	-0.0331	0.137***	0.349***	0.0288
	(0.0597)	(0.0552)	(0.0497)	(0.0930)	(0.106)
One Parent has High School	0.0731	-0.000287	0.0728	-0.0533	-0.0679
	(0.0639)	(0.0653)	(0.0580)	(0.100)	(0.0923)
Interaction	0.0339	-0.0872	-0.0533	-5.35e-05	0.130
	(0.0467)	(0.0539)	(0.0439)	(0.0841)	(0.0868)
Observations	1,115	1,115	1,115	1,115	1,115
R-squared	0.057	0.047	0.055	0.149	0.085
Bandwidth	[5]	[5]	[5]	[5]	[5]
Control Fn	Linear	Linear	Linear	Linear	Linear
Clusters	179	179	179	179	179
Mean Non-Admitted	0.749	8.775	9.524	0.277	0.138

Notes: RD design estimates, using the bandwidth of 5 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 fully interacted with the control function. Robust standard errors in parentheses are clustered at the school level. Source: COMIPEMS 2005, ENLACE 2008 and 2009. Sample: regular applicants from D.F. junior high to IPN schools observed in ENLACE survey. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 11: RD Design Estimates: Aspirations to Further Schooling

VARIABLES	(1) College	(2) College	(3) College	(4) Grad School	(5) Grad School	(6) Grad School
Admitted	0.0808	0.0245	0.0262	0.110	0.177***	0.191***
	(0.0604)	(0.0326)	(0.0386)	(0.113)	(0.0519)	(0.0545)
Observations	470	1,115	2,183	470	1,115	2,183
R-squared	0.110	0.042	0.034	0.095	0.056	0.047
Bandwidth	[2]	[5]	[10]	[2]	[5]	[10]
Control Fn	Linear	Linear	Quadratic	Linear	Linear	Quadratic
Clusters	125	179	236	125	179	236
Mean Non-Admitted	0.902	0.882	0.885	0.573	0.586	0.591

Notes: RD design estimates, using the bandwidths of respectively 2, 5 and 10 in the three columns, of the effects of IPN admission on declared aspirations to attend college (columns (1)-(3)) and graduate school (columns (4)-(6)). All models include cutoff fixed effects fully interacted with the control function. Robust standard errors in parentheses are clustered at the school level. Source: COMIPEMS 2005, ENLACE 2008 and 2009. Sample: regular applicants from D.F. junior high to IPN schools observed in ENLACE survey. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 12: RD Design Estimates: Robustness to Controls for Covariates

VARIABLES	(1) Expected Return	(2) Expected Wage HS	(3) Expected Wage College	(4) Score Math	(5) Score Language	(6) Aspires College	(7) Aspires Grad School
Admitted	0.196***	-0.0924	0.103**	0.336***	0.0161	0.0266	0.149**
	(0.0615)	(0.0563)	(0.0503)	(0.0846)	(0.105)	(0.0318)	(0.0577)
Observations	1,115	1,115	1,115	1,115	1,115	1,115	1,115
R-squared	0.186	0.161	0.188	0.324	0.260	0.243	0.243
Bandwidth	[5]	[5]	[5]	[5]	[5]	[5]	[5]
Control Fn	Linear	Linear	Linear	Linear	Linear	Linear	Linear
Clusters	179	179	179	179	179	179	179
Baseline Covariates	YES	YES	YES	YES	YES	YES	YES
Mean Non-Admitted	0.740	8.815	9.555	0.262	0.127	0.882	0.586

Notes: RD design estimates, using the bandwidth of 5 and controlling for a set of baseline covariates, of the effects of IPN admission on previous outcomes. Covariates include a vector of dummies for junior high-GPA deciles, dummies for graduation from a private junior high, gender, whether at least one parent has high school education, whether at least one parent has a white-collar occuppation and whether the student took the UNAM version of the entry exam, the number of IPN schools submitted in their top 10 choices, the number of highly selective schools submitted in their top ten choices, the number of schools submitted in their choice set and an index for household assets. All models include cutoff fixed effects fully interacted with the control function and the baseline covariates. Robust standard errors in parentheses are clustered at the school level. Source: COMIPEMS 2005, ENLACE 2008 and 2009. Sample: regular applicants from D.F. junior high to IPN schools observed in ENLACE survey. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 13: Non-IPN Elite Schools – School Inputs

EQUATION	VARIABLES	(1)	(2)	(3)
School Entry Score	Admitted	0.726***	0.754***	0.714***
		(0.148)	(0.108)	(0.109)
SD School Entry Score	Admitted	-0.0618*	-0.0663***	-0.0781***
		(0.0335)	(0.0253)	(0.0269)
Class Size	Admitted	-2.560*	-2.590***	-2.917***
		(1.426)	(0.844)	(0.855)
Students per PC	Admitted	5.762***	2.502*	1.247
		(1.967)	(1.322)	(1.383)
Teachers with College	Admitted	0.0284	0.0134	0.0150
		(0.0317)	(0.0190)	(0.0181)
Full Time Teachers	Admitted	-0.209***	-0.0592*	-0.0347
		(0.0521)	(0.0326)	(0.0325)
Observations		206	559	1,229
Bandwidth		[2]	[5]	[10]
Control Fn		Linear	Linear	Quadratic
Clusters		54	102	156
chi2		56.30	65.29	56.67
Prob > chi2		2.54e-10	0	2.13e-10

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 received school inputs, i.e.: school peers' entry scores: mean and standard deviation, class size, number of students per computer, share of teachers with a college degree, share of teachers employed full time in the school. All models include cutoff fixed effects fully interacted with the control function. Robust standard errors in parentheses are clustered at the individual level. Source: COMIPEMS 2005, ENLACE 2008 and 2009, Census of schools. Sample: regular applicants from D.F. junior high to non-IPN selective high schools observed in ENLACE survey. *** p<0.01, ** p<0.05, * p<0.1

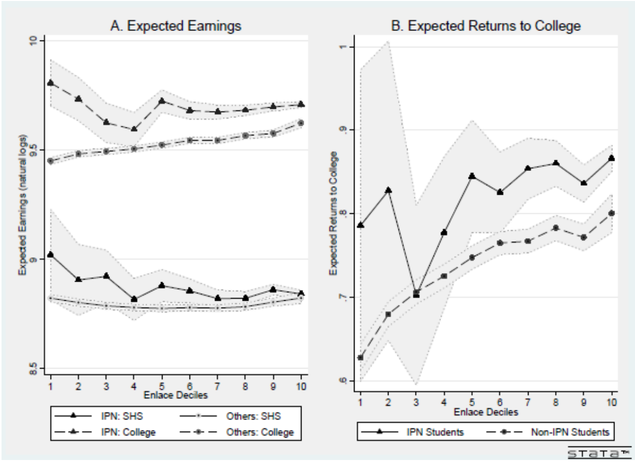
Table 14: Non-IPN Elite Schools – Outcomes

VARIABLES	(1) End of HS	(2) College Premium	(3) Wage HS	(4) Wage College	(5) Math Score	(6) Language Score	(7) Aspires College	(8) Aspires Grad School
Admitted	-0.0287	0.106	-0.0451	0.0612	0.134	0.0203	0.0707	0.0246
	(0.0436)	(0.0844)	(0.0875)	(0.0812)	(0.125)	(0.143)	(0.0588)	(0.0813)
Observations	2,152	563	563	563	563	563	562	562
R-squared	0.012	0.020	0.022	0.021	0.081	0.081	0.030	0.022
Bandwidth	[5]	[5]	[5]	[5]	[5]	[5]	[5]	[5]
Control Fn	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear
Clusters	2022	540	540	540	540	540	539	539
Mean Non-Admitted	0.464	0.783	8.765	9.548	0.165	0.201	0.894	0.647

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 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 fully interacted with the control function. Robust standard errors in parentheses are clustered at the individual level. Source: COMIPEMS 2005, ENLACE 2008 and 2009, Census of schools. Sample: regular applicants from D.F. junior high to non-IPN selective high schools observed in ENLACE survey. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

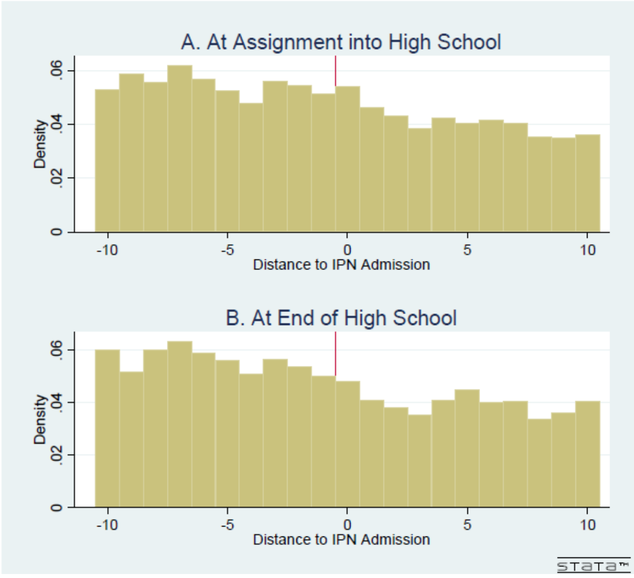
Figures

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



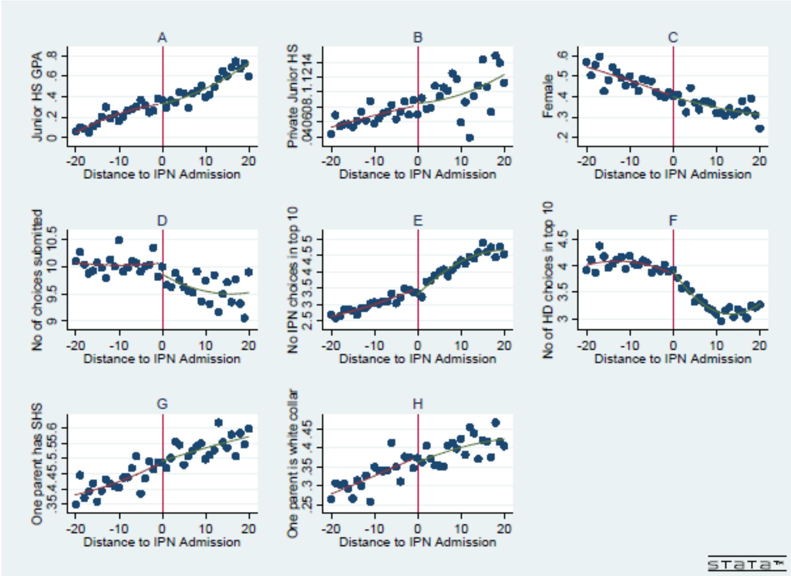
Notes: Conditional means, by decile of high school completion (ENLACE) achievement test, of expectations of earnings returns to college (Panel A) (with 95% confidence interval) and wages expected with high school and college degrees (Panel B) of students of IPN and other public high schools. Source: ENLACE 2008 and 2009 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 and end of high school



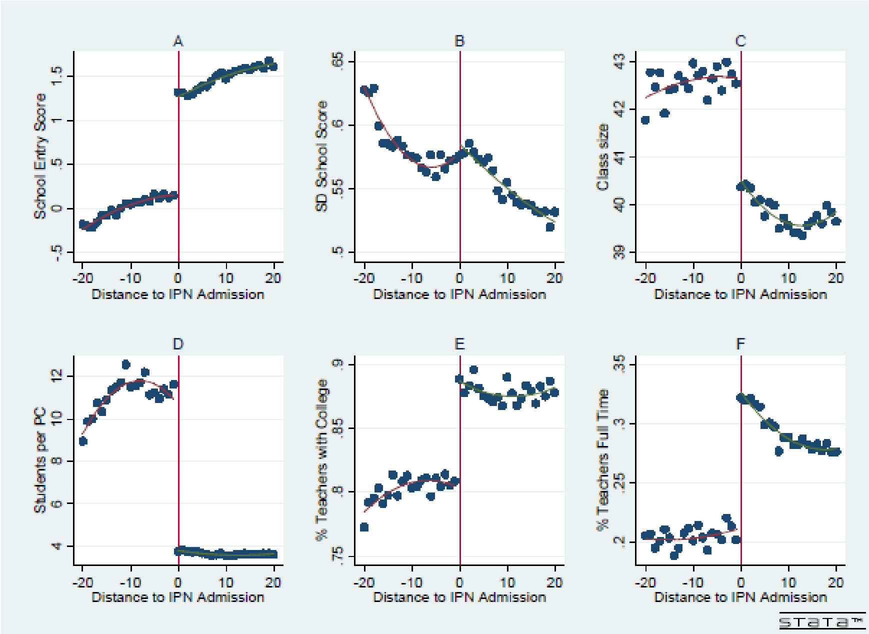
Notes: Distribution of the distance of students' scores at the COMIPEMS entry exam and from their cutoff scores for admission into the IPN system. Source: COMIPEMS 2005 school choices and entry exam scores. Sample: (a) regular applicants from Federal District junior high schools to COMIPEMS public high schools and (B) individuals from the same group observed at ENLACE high school completion achievement test.

Figure 3: Means of individual characteristics by distance from IPN admission cutoffs



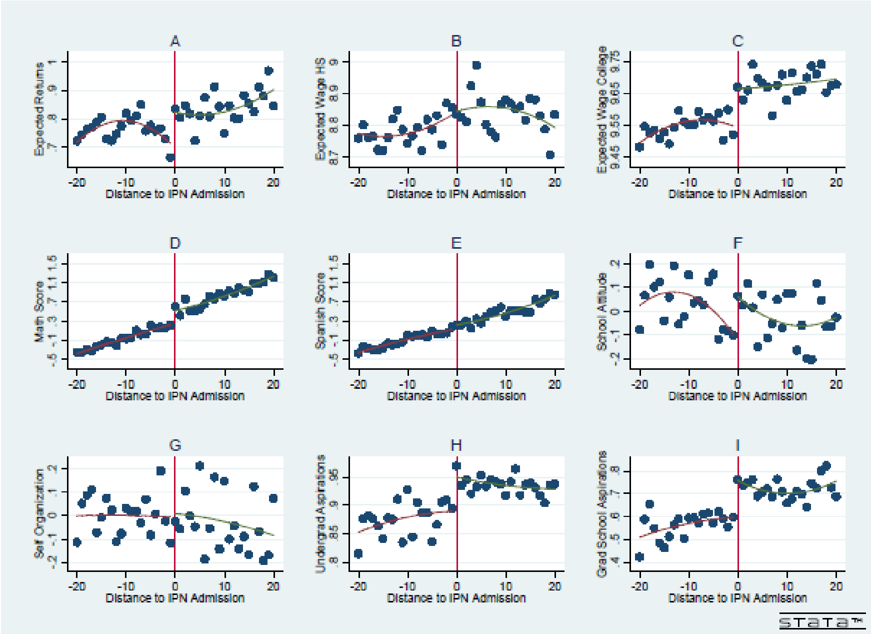
Notes: Conditional means of measures of students' characteristics (junior high school GPA, attendance of a private junior high school, gender, number of school choices submitted in the COMIPEMPS allocation process, number of IPN schools and high demand schools in top 10 choices, high school graduate and white collar parent) by distance of IPN entry exam test score from IPN admission cutoff. Source: COMIPEMS 2005. Sample: regular applicants from Federal District junior high to IPN schools.

Figure 4: Means of school inputs by distance from IPN admission cutoffs



Notes: Conditional means of measures of school inputs (school peers' entry scores: mean and standard deviation, 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: regular applicants from D.F. junior high to IPN schools observed at ENLACE high school completion achievement test.

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



Notes: Conditional means of measures of students' outcomes (expected returns to college, expected wages with high school and college degree, scores in Mathematics and Spanish at ENLACE high school completion exam, indexes of attitude toward school and self-organization, aspiration for an undergraduate and graduate diploma) by distance of IPN entry exam test score from IPN admission cutoff. Source: COMIPEMS 2005, ENLACE 2008 and 2009. Sample: regular applicants from D.F. junior high to IPN schools observed at ENLACE high school completion achievement test (and, for all outcomes except test scores, responding the ENLACE survey) .