School Choice and Educational Opportunities:

The Upper-Secondary Student-Assignment Process in Mexico City

Maria Elena Ortega Hesles

Richard J. Murnane
Fernando Reimers
Felipe Barrera-Osorio

A Thesis Presented to the Faculty
of the Graduate School of Education of Harvard University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Education

2015

To Arturo

To Elena and the baby on the way

To my Parents and Brothers

Acknowledgements

This dissertation is the result of a long and enriching journey. During this journey, I knocked many doors and crossed paths with people who have uniquely influenced my work. Each one gave me different tools, materials or guidance for the construction of this dissertation. I owe many people for the completion of this but I would need an entire chapter to express my deep appreciation to everyone.

First of all, I would like to extend thanks to the members of my ad hoc committee for their constructive comments, generosity with their time, and efforts in helping me through this process. My deepest gratitude to my advisor and mentor Richard Murnane for his patience, constant support, and invaluable guidance throughout these years. The training I received from him –as a student, an advisee, and a teaching fellow– has shaped me as a person and scholar. I am indebted to Fernando Reimers for his trust, advice, and encouragement not only throughout this project but also over the course of my time at HGSE. It has been a real honor learning from him both inside and outside the classroom. I also want to thank Felipe Barrera-Osorio for his guidance and useful feedback during process. His comments have greatly enriched and improved these papers. Finally, this work would not have been the same without the statistical training, useful feedback, careful editing, and long hours of discussion with John Willett.

This dissertation was possible thanks to the help of many people who listened to me at different stages of the project and who helped me knock the right doors. I want to thank Miguel Herrera, Daniel Hernandez, Carlos Mancera, Sofialeticia Morales, Reyes Tamez, Claudia Uribe, and Lorenzo Gomez Morin for listening to my ideas and guiding me through the system.

I also extend my appreciation to all the people who trusted me and opened their doors. I am deeply grateful to all the members of the COMIPEMS' Technical Committee. I want to especially thank Ana Maria Aceves and Roberto Pena; their support was fundamental for the conception and completion of this project. Jose Guadalupe Fernandez Galarza also provided useful insights regarding the COMIPEMS database. This work would not have been completed without the collaboration of multiple upper-secondary institutions in Mexico and the support of their staff. I want to personally thank Isidro Avila and Ivonne Ramirez (UNAM), Yadira Davila and Emmanuel Vazquez (IPN), Jorge González and Alejandro Hernandez (COLBACH), Maria Elena Araiza (DGB), Alejandro Quintin and Jorge Angeles (DGETI), Catalina Gallardo (DGETA), and Salvador Alvarado, Miguel Angel Cruz and Juan Manuel Uribe (CONALEP). They opened me the doors, answered my never ending calls, emails or letters, and helped me to build this ambitious database. I am also deeply indebted to Matteo Bobba and Veronica Frisancho for their inspiring discussions and for sharing with me their survey data and focus groups recordings that I used as part of the first paper. I also thank Victor Perez who provided research assistance in the first paper.

I appreciate financial support from the HGSE and the Fundacion Mexico en Harvard. In particular, I would like to thank Barbara Randolph for all her help since I first got accepted into the Master's program. Also, I appreciate the resources provided by the Institute for Quantitative Social Sciences (IQQS) at Harvard to conduct research for this dissertation. Also, I want to thank the IQSS and the David Rockefeller Center for Latin American Studies (DRCLAS) for their financial support to participate in different academic conferences.

I am thankful to my incredible friends in Mexico, USA, Sweden, France, and other parts of the world. They believed in me and have supported me professionally, academically, and personally in the years of my doctoral program. Each one has taught me something valuable and helped me get to this stage of my academic journey. At Harvard, I crossed paths with wonderful people who accompanied me at different stages of this process. I want to thank my academic and emotional support group: Maria Cecilia Acevedo, Mylena Aguilar, Anjali Adukia, Robina Bhasin. Eliana Carranza, Gwedolyn Casazza, Annie Chu, Sergio Cardenas, Lorenza de Icaza, Shaun Dougherty, Qian Guo, Rebecca Holcombe, Claudia Ramirez, Santiago Rincon-Gallardo, Anna Rosefsky, Patricia Villarreal, just to mention a few. My last 3 years collecting data and writing this dissertation in Mexico would not have been the same without the encouragement, support, and mental breaks with Eugenia Garduno and Ingrid Garcia. Also, I would like to very specially thank Isabel and

Hector for their friendship but most importantly for becoming my family in Cambridge during my constant visits during these last years.

Needless to say, I would like to express my very profound gratitude to my parents, Maria Elena and Miguel, for teaching me the value of education and paving the road for me to get here. They provided me with unfailing support and love throughout all these years of study. I can finally tell them: I did it! I am finally done with school (although just starting with schools). I also want to thank my brothers, Miguel and Alejandro, who have taught me so many things and who always find a way to make me laugh. They listened and understood all my praises and complaints about life as a graduate student even at 3am. I also want to thank my extended family and family in-law; this accomplishment would not have been possible without their constant encouragement. To my grandfather Mario I need to say that I found out that there are places where the sun in winter does warm a little and others where it doesn't; fortunately on those cold places there is the option of wearing a winter jacket and having heating.

Most importantly, I would like to thank my husband Arturo. His useful feedback, unending support, constant encouragement, quiet patience, and unwavering love has made this possible. He has been by my side always, even during those long days when I hesitated to continue. He has been a crucial piece in this academic journey and my life, having the roles of friend, classmate, teacher, advisor, and partner. No words will suffice to describe how (statistically) significant he has been in all this. TAF.

Finally, I dedicate this thesis to my daughters, Elena and the baby on the way. They have changed the meaning of my life and inspire me to continue improving education. Elena has been a great source of joy, love, and motivation throughout this process.

Table of Contents

School Choice and Educational Opportunities:
The Upper-Secondary Student-Assignment Process in Mexico City

Abstract

Many education systems around the world use a centralized admission process to assign students to schools. By definition, some applicants to oversubscribed schools are not offered admission to their most-preferred school. Thus, one naturally asks whether it makes a difference to applicants' educational opportunities and outcomes which schools they apply to, are offered admission to, and eventually enroll in. Each year in Mexico City, about 300,000 teenagers apply for a seat at one of the nearly 650 public upper-secondary schools. In this centralized, merit-based admission process, applicants are assigned to a school based on entrance examination score and their ranked list of school choices, subject to school capacity constraints.

In this dissertation, I include two papers assessing data from the upper-secondary application cohorts in Mexico City from 2005 to 2009. In the first paper, I find evidence of socio-economic stratification across schools. I also find dissimilarities in the application behavior of individuals according to their socio-economic background, even for those with high achievement levels. Based on qualitative and quantitative data from a small sample of applicants, I suggest that in addition to differences in economic resources, asymmetries in access to information might help to explain disparities in the application behavior of individuals from different socio-economic backgrounds.

In the second paper, I capitalize on the natural experiment created at each oversubscribed public upper-secondary school in Mexico City by the imposition of exogenous admission cut-off scores. Using a regression-discontinuity design, I estimate that, on average, upper-secondary applicants who score just above the admission threshold for a more competitive school (i.e. a school with higher cut-off score and higher average examination scores) have lower probability of graduating on time and within 5 years than do applicants who scored just below the admission threshold. Given the high take-up rates of the offers of admission, I find that the effects for enrollment in a more competitive school are only slightly larger than they are in their analogous reduced-form estimates. In addition, I show that effects differ across the distribution of admission cut-off scores and for applicants with selected socio-demographic characteristics who scored just above the admission threshold.

School Choice and Educational Opportunities:

The Upper-Secondary Student-Assignment Process in Mexico City

Introduction

Recently, school choice has become a topic broadly discussed in the education and economics literature. Arguments behind the design of school-choice programs include the introduction of market mechanisms to remedy inefficiencies and the expansion of students' access to public schools, while making education systems more equitable (Musset, 2014). However, in settings where demand for schools exceeds supply, it is necessary to design mechanisms to decide who is offered admission to a specific school and who is not admitted (Abdulkadiroğlu et al., 2012). That decision is often defined by schools' priorities over applicants and capacity constraints. In large education markets, centralized admission processes allocate students to schools usually based on their revealed preferences over schools, and either through a lottery or a merit system (Lucas & Mbiti, 2012). In particular, academic merit-based systems should provide high-achieving students with the opportunity to attend more selective schools, regardless of their social background or feeder school (Ajayi & Telli, 2013). Thus, such systems might play an important role in reducing inequality and promoting mobility in contexts with large variation in family income and school quality.

Thus, improving the equality of educational opportunities is one of the implicit roles of school-choice programs. Equality of educational opportunities implies, among other things, that family characteristics should not have a strong influence on the access, transition, and persistence through schooling. Unfortunately, in many countries, educational opportunity has not expanded enough to close the education gaps among different social groups and promote educational mobility.

In my dissertation, I explore issues of equality of educational opportunities in terms of access to, and completion of, upper-secondary education in Mexico City. For almost two decades, nine public upper-secondary school subsystems[1] in Mexico City have been implementing a centralized merit-based student-assignment mechanism, in order to reduce the inefficiencies and costs of the old decentralized application and admission system. Each year over 300,000 applicants compete for a seat at one of the nearly 650 public upper-secondary school options in the Metropolitan Area. Admission to schools is based on applicants' stated preferences over options and their score on a standardized entrance examination, subject to schools' capacity constraints. Given the set-up of assignment algorithm, schools with higher cut-off scores are always associated with higher-achieving peers.

In the first paper, I explore the role of socio-economic background on access to upper-secondary education. Using administrative data from the 2005-2009 application cohorts, I examine whether student stratification across school subsystems is related to social background. I also inspect whether the application behavior of students differs by socio-economic level, particularly among high-achievers.[2] In the second paper, I exploit the exogenous variation in the placement offer, at each oversubscribed school, for applicants of equivalent skills and preferences around the cut-off for admission. I adopt a regression-discontinuity design in combination with instrumental-variables estimation to estimate the

[1] A subsystem is a group of schools that share the same curricula, institutional organization, and rules.
[2] High-achiever is defined as an individual who scored in the top 10 percent of the entrance-examination distribution, for his or her application cohort.

causal impact for the marginal applicant, of the offer of admission to, and enrollment in, a more competitive school on subsequent graduation outcomes.

My descriptive analysis in the first paper suggests that there is stratification of students by socio-economic background across the school subsystems. In addition, I observe differences in the application behavior of individuals according to their socio-economic (SES) background, even for those with similar scores in the entrance examination. Consistent with findings from other settings, I also find that high-achieving students from low-SES families in Mexico City have, on average, lower education aspirations and apply to less selective upper-secondary schools than similar students from high-SES households. I hypothesize that, in addition to differences in economic resources, asymmetries in access to information related to the admission process might help in explaining disparities in the application behavior of individuals from different socio-economic backgrounds. This differential application behavior can translate eventually into higher SES stratification across schools and lower educational mobility.

My causal analysis in the second paper shows that, on average, applicants who are offered admission to a more competitive school option and are near the cut-off for admission –lie at the bottom of that school's ability distribution– have on average a 3.1 percentage points lower probability of graduating on time, and 1.1 percentage points lower probability of graduating within 5 years, from public upper-secondary education relative to applicants who scored just below the admission cut-off. These magnitudes represent 6.6 percent of the average on-time graduation rate of 47 percent, and 2 percent of the average 5-years graduation rate of 57 percent. Considering the fact that some admitted applicants never registered, I show that the causal effect of enrollment is slightly larger, at -3.2

percentage points on the probability of timely graduation and a -1.2 percentage points on the probability of graduation, for those at the margin of admission. I also find that effects differ across the distribution of admission cut-off scores and for applicants with selected socio-demographic characteristics who scored just above the admission threshold. My results are robust to a variety of sensitivity analyses.

Though these two papers address different questions and have certain limitations, both indicate that students from less privileged backgrounds tend to face greater barriers to access and graduation from upper-secondary education than their more privileged counterparts. This could potentially have important long-term academic, economic and social implications for this group. Further research is needed in this area, particularly regarding targeted interventions designed to address the financial, academic and informational barriers faced by teenagers of different social and cultural backgrounds in Mexico City.

References

Abdulkadiroglu, A., K., Che and Y. Yasuda (2012). "Expanding ‘Choice’ in School Choice," unpublished working paper.

Ayaji, K. and H. Telli (2013). “Imperfect Information and School Choice in Ghana,” unpublished working paper.

Lucas A. and I. Mbiti (2012). "The Determinants and Consequences of School Choice Errors in Kenya," American Economic Review 102(3): 283-88.

Musset, P. (2012), “School Choice and Equity: Current Policies in OECD Countries and a Literature Review”, OECD Education Working Papers, No. 66, OECD Publishing. http://dx.doi.org/10.1787/5k9fq23507vc-en

Reimers, F. (2000). Educational opportunity and policy in Latin America. In F. Reimers (Ed.), Unequal schools, unequal chances: The challenges to equal opportunity in the Americas. Cambridge, Massachusetts: Harvard University, 54-111.

The Role of Socio-Economic Background and Information in the Upper-Secondary Admission Process in Mexico City

The Role of Socio-Economic Background and Information in the Upper-Secondary Admission Process in Mexico City

Since 1996, admission to upper-secondary education in Mexico City has followed a centralized merit-based process. Each year over 300,000 applicants compete for a seat at one of the nearly 650 public upper-secondary school options. Admission to schools is based on applicants' stated preferences over options and their score on a standardized entrance examination, subject to the number of seats available in each school. This system was adopted to reduce the inefficiencies and costs generated by the old decentralized application and admission system. Although many people regard this process as "fair," others have raised questions about its role in reducing opportunities for educational mobility given the strong association that exists between social background and academic performance.

Centralized admission systems, like the one in Mexico City, are being implemented in other settings where demand for some schools exceeds supply.[3] Although the particular student-allocation process may differ from system to system, it is usually based on applicants' revealed preferences over schools, combined with a lottery, priority ordering,[4] or merit system (Lucas & Mbiti, 2012; Abdulkadiroğlu & Sönmez, 2003). Under an academic merit-based system, admission to a particular school is assumed to depend on students' academic performance and not directly on other factors such as social background

[3] Ghana (Ajayi, 2013), Romania (Pop-Eleches & Urquiola, 2013), and Trinidad and Tobago (Jackson, 2010) to mention a few.
[4] Some school systems give priority based on: attendance at feeder schools, residence in the school's walk zone, and sibling attendance at the school (Abdulkadiroğlu et al., 2006).

or feeder school. However, there is an intrinsic problem when measuring academic merit via test scores since they often correlate with students' socio-economic background.

In contexts with large variation in family background and school quality, academic merit-based systems aim to provide an opportunity for high-achieving students to attend more selective schools, irrespective of their socio-economic background (Ajayi & Telli, 2013). But, this is not always the case. Conditional on achievement, students from more disadvantaged social backgrounds often apply to less selective schools than those from more favorable backgrounds (Ajayi, 2012; Hoxby & Avery, 2012).

In this paper, I explore the relevance of socio-economic background on access to upper-secondary education, under a centralized merit-based admission system. I focus on the case of upper-secondary education in Mexico City, which has implemented a centralized merit-based admission process for almost two decades. Using administrative data from the 2005-2009 application cohorts, I first examine whether student stratification across school subsystems[5] is related only to performance on the entrance examination or also to other dimensions, like social background. For this, I investigate the composition of the student body for each school subsystem. Second, I explore whether the application behavior of students differ according to the socio-economic background of their families. In particular, I explore whether high-achievers, who are the ones that should benefit from a merit-based system, show different application behaviors than their socio-economic status would lead one to anticipate.

[5] A subsystem is defined as a group of schools that share the same curricula, institutional organization, and rules. In Mexico City, there are 9 subsystems that participate in the student-assignment process.

My descriptive analysis indicates that there is ability –and socio-economic– stratification across schools, with more resources being allocated to schools that incorporate mostly high-achieving students from the top quartiles of the socio-economic distribution. I also find differences in the application behavior of individuals from distinct socio-economic backgrounds. Consistent with findings from other settings, I also observe that high-achieving students from low socio-economic families in Mexico City have, on average, lower educational aspirations and apply to less selective upper-secondary schools than do students from high socio-economic households. Based on qualitative and quantitative data from a small sample of recent upper-secondary applicants, I then propose that, in addition to economic resources, asymmetries in the access and quality of information related to the admission process, and available choices, might help in explaining disparities in the application behavior of individuals from different socio-economic backgrounds.

The evidence I present here is descriptive rather than causal. Therefore, I cannot explain why low-income students tend to apply to less selective schools. Further extensions of the research –that move beyond the associations and explore causal relations through quasi-experimental designs or the methods of instrumental-variables estimation– are desirable. However, the descriptive evidence that I present here documents that upper-secondary applicants in Mexico City are making different schooling choices based on their social background. It also suggests that imperfect information may play an important role in reinforcing school stratification and reproducing initial inequalities.

In the rest of the paper, I proceed as follows. In the next section, I review the existing literature on the relationship between social background and schooling decisions. I then describe the institutional background of my study, and continue by describing my dataset. Then, in the following section, I describe the results of my descriptive analysis and hypothesize that imperfect information could help to explain them. In the last section, I conclude with a discussion of the findings and recommendations for further research.

Background and Context

A growing body of evidence has shown that socio-demographic characteristics, particularly income, play a major role in predicting the schools to which students apply and in which they eventually enroll (Belasco & Trivette, 2012). Research consistently finds that, conditional on other variables, socio-economic background is positively associated with students' academic performance, educational aspirations, and education outcomes (Perna & Thomas, 2006). In other words, individuals from higher socio-economic backgrounds are less likely to fall behind and drop out, and are more likely to have aspirations that promote school persistence. In the case of Mexico, evidence from research indicates that there is a clear relationship between the social origin of students and their access, transition, and persistence in upper-secondary education (Tirado, 2004; Solis et al., 2013).

Recent studies of access to education reveal that individuals of low socio-economic (SES) status disproportionately apply to less selective and less resourced schools, regardless of their academic performance. That is, many high-achieving students from low-SES families do not maximize their educational opportunities and fail to apply to selective schools for which they have the sufficient academic skills to attend. This is particularly troubling in settings where low-income students have larger returns to attending selective schools than high-income students. This "under-match" phenomenon has been particularly evident in the context of college applications in the U.S. (Avery & Kane, 2004; Bowen et al., 2009; Hoxby & Avery, 2012) but few researchers have focused on the case of admission to secondary education in developing countries (Ajayi, 2011). For instance, de Janvry et al. (2012) report that students with lower GPAs and less-educated parents in Mexico City are less likely to apply to elite schools.

Among other things, school-choice models aim to promote access to schools for all students, regardless of their social background (Musset, 2014). However, some researchers argue that school choice can reinforce initial disparities because it increases sorting of students between schools based on their socio-economic status (Hsieh & Urquiola, 2006; Ajayi & Telli, 2013). Given the association between academic performance and socio-economic background, MacLeod and Urquiola's (2009) model predicts that, if admission is merit-based, competition in school-choice systems will lead to socio-economic stratification.

School-choice models assume that applicants have perfect and complete information when making schooling decisions. School choice is difficult when applicants have imperfect information about schools and their own academic performance. In fact, research suggests that imperfect information may help in explaining why some high-achieving applicants from low-income households make suboptimal schooling decisions (Avery & Hoxby, 2012). Information is particularly relevant in Mexico City's upper-secondary admission process because applicants must list their preferred schools before knowing their score in the entrance examination, and because there is a large set of school options to choose from.

Recent research indicates that providing students and families with information about schools, or assistance with the admission process, can lead to changes in their schooling choices, particularly for more disadvantaged students (Hastings & Weinstein, 2008; Jensen, 2010; Bettinger et al., 2012; Hoxby & Avery, 2012). For the Mexican case, Bobba and Frisancho (2014) found that providing students with individualized information about their own academic ability changed the application decisions of upper-secondary applicants who received positive feedback.

Institutional Background

Mexico is a country with a very extensive and complex education system. During the last decades, Mexico has made considerable progress towards an increase of the enrollment rate at basic education[6] and decrease of illiteracy (SEP, 2005). This expansion was accompanied by the diversification of education services to provide for a 'more adequate' education, and according to the needs and demands of specific groups (Schmelkes, 2000). Thus, different types of basic and upper-secondary education services coexist in the country. For instance, the upper-secondary level is divided into 3 tiers: general or academic, technological or bivalent, and technical or vocational.[7] However, research suggests that the segmentation of education services in basic education is providing unequal educational opportunities and maintaining social segregation (Treviño, 2006; Cardenas, 2009). In upper-secondary education there is some evidence of pronounced

[6] Basic education includes the primary (grade 1 to 6) and the lower-secondary (grades 7 to 9) levels.
[7] Each of these tiers has different programs and curricula. This makes difficult to transfer between tiers. See Appendix A for a more detailed description of these tracks.

disparities in academic achievement, availability of resources, school conditions, and teachers' profile between different types of schools (INEE, 2012).

One of the biggest barriers to improve education outcomes in Mexico is found at the upper-secondary level (i.e. grades 10 to 12), which has recently been made mandatory. While primary and lower-secondary coverage rates have been increasing and currently exceed 90 percent, coverage in upper-secondary education remains relatively low. The increase in coverage and expansion of supply in basic education implies that a higher percentage of students have the academic requirements to enter upper-secondary education. Despite the advances, upper-secondary education has the largest dropout rates in the system. The dropout rate was 14.5 percent per year in 2012, which means that in the course of the three years covered by this level nearly 45 percent of students that entered into this level eventually dropped out (Bentaouet Kattan & Szekely, 2014).

The educational challenges in Mexico City are different from those in the rest of the country. While the average schooling in the country is less than complete lower-secondary education (i.e. 9 years), Mexico City has among the highest average coverage and absorption rates in upper-secondary education. Mexico City's upper-secondary level represents a very interesting and particular case study within the country. It has the highest transition rates from lower to upper-secondary (90%) in the country, but also the highest repetition (40%) and dropout rates (13% per year) in upper-secondary education. These indicate that, in contrast with the national average, many teenagers enter upper-secondary education in Mexico City but a substantial proportion fall behind or never graduate.

The Upper-Secondary Admission Process in Mexico City

The Metropolitan Area of Mexico City[8] has the largest education market in the country and has been implementing a merit-based student-assignment mechanism in the upper-secondary public-school system since 1996 (COMIPEMS, 2012). Before that, the decentralized application and admission systems were characterized by congestion and wait lists in some schools, and vacant seats in others. Students had to apply to several schools simultaneously and then choose their most-preferred school that had accepted them. Applicants from lower-income families were at greater disadvantage because they had fewer financial resources to cover the costs of applying to many schools (COMIPEMS, 2008).

In 1996, nine upper-secondary school subsystems from the Federal District and the State of Mexico created COMIPEMS,[9] a commission to coordinate institutional efforts and respond to the increasing demand for upper-secondary education (COMIPEMS, 2012). The main task of this commission was to create a centralized and competitive admissions process, currently known as the CIEMS.[10] Through the CIEMS process, applicants pay a single administrative fee for the process and list up to 20 school options they are interested in attending.[11] Applicants are assigned to public schools based on a designed and explicit interplay between their score in an entrance examination, a ranked list of up to 20 preferred school options, and the predetermined capacity constraints of schools. In Mexico City, each

[8] The Metropolitan Area includes the Federal District and 22 municipalities of the State of Mexico. See Appendix A for the geographical distribution of school options by tier.
[9] COMIPEMS are the initials in Spanish for the Metropolitan Commission of Upper-Secondary Public Institutions. In Appendix A I include a list of the subsystems that integrate COMIPEMS.
[10] CIEMS is the Spanish acronym for Concurso de Ingreso a la Educación Media Superior.
[11] Applicants have to rank schools or school/track combinations. In the case of technical and some technological schools, applicants must indicate their desired track or area of specialization (for example, accounting). For simplicity, I use the term “school” to refer to school/track combinations.

year about 300,000 teenagers apply for a seat at one of the nearly 650 public upper-secondary school options from three different tracks (i.e. general, technological and technical). There are enough seats in the system for all applicants, with about 25 percent of the options being undersubscribed schools. The most demanded schools within the process are those affiliated with two large public universities, UNAM and IPN, which tend to fill all their seats early in the assignment process.[12]

The student-assignment process proceeds as follows.[13] In any given academic year, the process starts in late January with a public announcement of the CIEMS on the Internet and local newspapers (COMIPEMS, 2012). Interested candidates access the registration materials through their current lower-secondary school, local information centers, or the COMIPEMS website.[14] Such information is public and free of cost. By the beginning of March, applicants must complete all the required documents, including the ranked list of up to 20 school options. Registration has a fixed cost of about $25 USD regardless of the number of schools listed. Applicants must go to a nearby registration center, where they receive a study guide for the examination and an identification voucher necessary for taking the examination.

The entrance examination takes place at the end of June. It is composed of 128 multiple-choice questions from 10 different subjects or content areas. The raw scores in

---

[12] UNAM stands for the National Autonomous University of Mexico. IPN stands for National Polytechnic Institute. Admission to UNAM schools translates into eventual transfer to UNAM for students who graduate within 4 years and have a GPA of at least 7 out of 10.

[13] See Appendix C.1 for the timeline of the process.

[14] Materials include registration forms, a calendar, and a manual with the list of schools, their location and specialization fields, if applicable. A complete list of the registration materials is presented in Appendix C.2. The COMIPEMS website provides additional information on school locations and past cut-off scores for each school option.

these subjects are then added to give the total score. Once the entrance examinations have been centrally graded by a computer, COMIPEMS's board meets for the assignment process. Applicants who scored less than 31 points, or have not graduated from lower-secondary education, are excluded from further consideration.[15]

During the first round, at the end of July, each school reports the maximum number of seats available and all qualified applicants are ranked from highest to lowest by their examination score. A computer algorithm is used to assign ranked applicants to their most-preferred school with open seats, once their turn arrives. That is, seats are allocated down the student ranking: the top scorer is assigned to his first choice, the second highest scorer gets his most-preferred choice among schools with open seats, and so on. If several applicants with the same score compete for the last seats at a specific school, a representative from that institution must decide immediately whether he opens new seats to accept all tied applicants, or rejects all of them.[16] The assignment process continues until all applicants are allocated, except for those who only request options with cut-off scores above their own score. Thus, the cut-off score for each school option, each year, is set equal to the entrance exam score of the applicant who fills the last seat. Final school-assignment decisions are published by the end of July.

During the second round, applicants who did not get a seat at one of their listed choices, but meet the requirements, have the opportunity to select a school that did not fill

[15] The minimum score restriction was eliminated in 2013. Schools affiliated to UNAM and IPN also require a minimum GPA of 7 out of 10 in lower-secondary education.
[16] The representative only has information regarding the number of tied applicants and their score but not about their personal characteristics.

their seats. Finally, applicants must complete the paperwork at their assigned school in order to register. Applicants are only allowed to register at their assigned school.[17]

In Table 1, I present the main statistics summarizing the 2005-2009 assignment processes. I observe that over those years the number of options and applicants increased. Although applicants can fill in up to 20 educational options, under 3 percent completed all of their preference portfolio. On average, applicants listed 9 options. Not all of the applicants that registered in March took the exam in June - four percent of them did not.[18] Of those who took the entrance test, about 11 percent were not eligible for assignment either because they did not graduate from lower-secondary education or because they had a score of under 31 points in the examination. About 83 percent of eligible applicants were assigned during the first stage of the process and around 40 percent of them were offered admission to their top-choice. The proportion of applicants who choose a school with open seats during the second round has decreased over time. There were about 10 percent of eligible applicants each year who were not offered admission to any of their listed schools and did not choose a school in the second round of the process. On average, 20 percent of the applicants who demonstrated their willingness to continue their studies by registering in the assignment process were left out of the public upper-secondary system.

[17] If applicants want to switch to another school they would have to take the exam again the following year.
[18] There might be different reasons why some registered students do no take the examination, such as not graduating from lower-secondary education, moving to another state, working full-time, enrolling into a private school or just not following up the process.

Dataset

In this paper, I use data from three sources. First, I use administrative data from the set of applicants who registered for the COMIPEMS admission process that took place between 2005 and 2009. These five cohorts include over 1.5 million applicants and about 3,200 school/track options. The dataset contains rich information on applicants' background, entrance examination score, ranking of up to 20 school options, assigned school, and responses to a context survey. Background information includes participants' gender, age, zip code, lower-secondary school attended and GPA. In addition, the context survey gathers information about the students' family structure, parental education, durable goods, family income, financial aid, and other related socio-economic and family questions. These data were made available to me through the COMIPEMS technical committee. My analytic sample includes 1,186,248 applicants who were offered admission, during the first or second round of the process, to any of the 9 subsystems that participated in the admission process from 2005 to 2009.[19]

Second, to explore access to sources of information and application behaviors, I draw on survey data from a sample of 9th grade students that participated in the 2013 upper-secondary admission process. This survey was administered to a random sample of 15 lower-secondary schools of the Federal District and State of Mexico, during the pilot stage of the randomized experiment conducted by Bobba and Frisancho (2014). This 45-minute self-administered survey took place in the second part of May 2013, once students had

[19] I acknowledge there can be a sample bias if those who did not answer certain survey questions were systematically different from those that did. A comparison between the unrestricted and the restricted samples showed very little differences in the average baseline characteristics, which leads me to believe that missing values are random.

already registered for the admission process and were close to taking the examination. Among other things, the survey included questions about students' preferences, expected score in the entrance examination, and about the knowledge of the application process and sources of information. To have more data regarding their background, I merged this database with COMIPEMS's administrative databases (i.e. student survey, list of preferences, actual score, and school assignment).[20] My sample has 1,526 lower-secondary students who participated in the 2013 admission process.

Third, to have a better understanding of applicants' experience during the CIEMS process, I analyze qualitative data from 15 focus groups conducted as part of the previously mentioned pilot study (Bobba & Frisancho, 2014). Within each of the sampled lower-secondary schools, twelve students attending 9th grade were randomly chosen and invited to participate in the focus groups. Most of these students agreed to participate. The focus group sessions lasted between 60 and 90 minutes each, and were either video- or audio-recorded. The discussions were led by a trained qualitative researcher. The purpose of the focus groups was to understand students' perceptions and experiences regarding the upper-secondary admission process. Both, the survey data and qualitative records from this sample were provided to me by Bobba and Frisancho, who were the lead researchers of the project.

Although data I analyze here comes from different application cohorts, the comparison between different cohorts (2005-2009) shows that the main patterns and

[20] 92% of the students in the sample were found in the official COMIPEMS data, when merging databases using the Curp and/or in some cases name and school. This discrepancy is most likely due to the result of participation decisions in the CIEMS and could proxy the dropout rate between lower- and upper-secondary education in the sample.

characteristics of applicants are relatively stable over time.[21] Therefore, the main conclusions of my analysis are likely to be time-insensitive.

Findings

In this section, I explore questions of whether –in the centralized merit-based admission system implemented in Mexico City– there is (1) evidence of student stratification, by ability and social background across school subsystems, and (2) evidence of students' differential application behavior according to their socio-economic background, particularly among high-achieving students. I propose that imperfect and asymmetric information are possible mechanisms to help explain differences in the application behavior of students from different socio-economic backgrounds.

Composition of the Student Body, by School Subsystem

Previous research has shown that, in Mexico, segmentation of education services in basic education leads to unequal educational opportunities and to the maintaining of social segregation (Treviño, 2006; Cardenas, 2009). This suggests that the reproduction of initial disparities begins early on the school trajectories of children in Mexico. In upper-secondary education, there is some aggregate evidence of pronounced disparities in academic achievement, availability of resources, school conditions, and teachers' profiles among types of schools (INEE, 2012).

[21] Statistics disaggregated by application cohort can be provided upon request.

To explore student stratification by ability and other possible dimensions in the merit-based system under analysis, in Table 2 I present summary statistics on the student body composition of each subsystem. Since there is not much variation between cohorts and for the sake of space, I only show aggregate statistics for the period 2005-2009. I have divided the table according to the track (i.e. general, technological or technical) to which each subsystem belongs. For instance, schools in the subsystem SS1 belong to the general education track.[22] Each column corresponds to a different subsystem. The only exception is SS5, which includes schools from both the general and technological tracks, and which I divided into SS5a and SS5b.

As indicated by their high average entrance scores, SS1 and SS6 are the most selective subsystems. Both are subsystems affiliated to public universities. Regarding the academic characteristics of the student body, I observe differences in the average entrance scores of students admitted to each subsystem and some stratification by their GPA in lower-secondary education. This is to be expected given the merit-based nature of the student-allocation system. I also see differences in schooling trajectories, with applicants assigned to more selective schools having failed fewer courses during lower-secondary education. I notice differences in the studying and reading habits of students according to the selectivity of the subsystem. For instance, applicants admitted to the most selective subsystem (SS1) reported studying on average 6.53 hours per week compared to 4.37 hours among those admitted to the least selective subsystem (SS9).

[22] By request of the COMIPEMS committee, I labelled the subsystems as SS# instead of indicating their actual names.

Although gender composition is relatively stable across subsystems, there is a gender bias that favors males in SS6 and SS9, probably related to the specialization fields that they offer. On average, students admitted to all subsystems have the same age. The percentage of students from indigenous families is very low in all subsystems but particularly among the most selective ones (i.e. SS1, SS2 and SS6). About 94% of applicants who are offered admission to the upper-secondary level lived in the Metropolitan Area of Mexico City at the time of registration. In addition, I show that subsystems with more schools located along the border of the city (i.e. SS5, SS8 and SS9) are more likely to attract non-residents of the area.

In Table 2b, I observe that, on average, students offered admission to more selective schools have higher educational aspirations than those attending less selective schools. That is, most of the students in elite schools want to continue into higher education. I must note that differences in this dimension, between tiers, might reflect to some extent the different focus of each tier. However, the pattern that I just described is also valid for comparisons of subsystems within tiers and which have the same focus. Overall, this suggests that, as expected, the admission mechanism is indeed acting to stratify students by ability, academic habits and effort.

In terms of characteristics of the household head, there is little or no variation in average age across subsystem. However, students admitted to the more selective subsystems tend to have more-educated parents, when compared to those admitted into the less selective subsystems. In particular, approximately half of the students admitted into the three most selective subsystems (i.e. SS1, SS2 and SS6) have parents with at least an upper-secondary education. This contrasts to the 20 percent of students, whose parents have this level of qualification, that were admitted into the least selective subsystems (i.e. SS8 and SS9). Related to this, I note that a higher percentage of the parents of students in elite schools have white-collar jobs, compared to those whose children attend non-elite institutions.

Finally, regarding the socio-economic background and cultural capital of the family, I detect a relatively clear student stratification. That is, applicants who are admitted into the most selective school subsystems come from more favorable family backgrounds than those admitted into the least selective subsystems. Students who are admitted to elite schools not only had a higher family income per capita and a higher probability of having attended a private lower-secondary school, but they also participated in more cultural activities (like theater, cinema and museum) than students admitted to the least selective schools. For example, the per capita monthly family income of someone in SS1 is more than twice the per capita family income of an applicant assigned to SS8 or SS9.

Overall, the description that I have presented here provides evidence of student stratification by socio-economic and cultural capital. In addition, I note that resources allocated to the different subsystems seem to reinforce existing disparities within the student population, with more resources being spent per student in the more selective subsystems. For instance, in 2013, the average expenditure per student in the SS1 schools was 40,827 pesos while, in the SS4 and SS9 schools, it was 17,253 and 14,177 pesos, respectively. That is, on average, the most selective subsystem invested –in each student– more than twice the money spent by the least selective subsystem, which is attended by more-disadvantaged students. It is not clear whether the application process is responsible for discriminating against individuals from more vulnerable contexts, or whether there are market inefficiencies that lead vulnerable students to make different choices regarding their education, or both.

Student Application Behavior, by Socio-Economic Background

As I previously described, research has shown a strong positive relationship between educational outcomes and social background. Individuals from lower socio-economic backgrounds are more likely to fall behind and drop out, and are less likely to have high educational aspirations. Evidence has shown that, conditional on achievement, students from more disadvantaged social backgrounds often apply to less selective and resourced schools than those from more favorable backgrounds (Ajayi, 2012; Hoxby & Avery, 2012).

In the setting under analysis, Tirado (2004) showed that a student's social-cultural background is the characteristic that has the strongest association with the score in the upper-secondary entrance examination, followed by the characteristics of the school. Similarly, Villa (2007) pointed out that scores on the upper-secondary entrance examination are positively correlated with family income. De Janvry et al. (2012) reported that, in Mexico City, students with lower GPAs and less-educated parents are less likely to apply to upper-secondary elite schools. In this section, I explore whether applicants have different application behaviors, in terms of their socio-economic status. I also investigate whether high-achieving students, those who are supposed to most benefit from an academic merit-based system, have similar application behaviors regardless of their socio-economic background.

For this analysis, I estimate the socio-economic (SES) status of a student's family using principal-components analysis (PCA). I created a SES index using data from all the cohorts simultaneously, so that its values were equitable across cohorts.[23] I estimated the values of a SES index based on other characteristics of the family, including: the self-reported number of persons per room in the family, access to services (like sewerage, hot water, telephone, Internet) at home, and the availability of durable goods (like a car, computer, DVD, mobile phone, books). Then, I divided applicants from each cohort into quartiles according to their value on the SES index. I define as high-SES and low-SES, respectively, as the top and bottom quartile of the SES index distribution.

In Table 3, I present descriptive statistics for the 2005-2009 applicants, according to their SES level. In each column, I report sample means and standard deviations (in parenthesis) for all students in that particular quartile. For most of the variables, I found striking similarities across SES groups. In Panel A, I do not observe important differences between respondents in the SES quartiles in terms of their demographic characteristics, such as gender and age. About half of the applicants are males and average age is 15.5 years. I see that the proportion of indigenous population applying to upper-secondary education in Mexico City is small (2.5%) but slightly larger among the bottom SES quartile (6%). As expected, there is a positive relationship between the schooling of the household head and the socio-economic level. Students in the top quartile have, on average, parents with 11.7 years of schooling compared to 7.7 years among those in the bottom quartile.

[23] Although the COMIPEMS survey includes self-reported family income, responses to this question can only be made within a series of coarse income bins, and many students fail to respond. The SES index which I estimated has a strong correlation with the self-reported family income (0.681).

Regarding the academic characteristics of applicants (Panel B in Table 3), I notice a relatively constant GPA of 7 (out of 10) in lower-secondary education across socio-economic quartiles. But, consistent with evidence from this particular setting, I find that there is a positive relationship between scores on the entrance examination and socio-economic background. That is, the average entrance score increases –from 58.3 to 69.6 points– as one moves from a lower to a higher quartile in the SES distribution. I detect that the top quartile's applicants have a significantly higher likelihood of having attended a private lower-secondary school than students in the rest of the SES distribution. On average, 18 percent of high-SES applicants attended a private school, compared with only 2 percent of low-SES applicants.

In Panel C, I present summary statistics on a selection of variables related to the students' application behavior, and to the results of the school assignment. I do not observe important differences across quartile. I note that applicants from all quartiles rank, on average, nine schooling options. In fact, only 2 percent of all applicants rank 20 options, which is the maximum number of options allowed. Applicants are assigned, on average, to their third option and about 25 percent of them are offered admission to their first-choice, regardless of socio-economic status. Distance from their lower-secondary school to their first-choice school is about the same (8 km) for all groups.

However, I do detect differences in the selectivity of the preferences portfolio, calculated as the mean of the cut-off scores of all the school options listed by an applicant.[24] Applicants in the top-quartile seem to have school choice portfolios with an average selectivity of 70 points, compared to a 61 points average selectivity of the portfolio for

[24] The higher the cut-off scores of the options selected, the higher the selectivity of the portfolio.

those in the first quartile. Applicants in the top quartile are more likely to select an 'elite' school as their first choice of school than those from the most disadvantaged families. Finally, as I mentioned before, the dominant strategy under this assignment mechanism is to rank schools according to their cut-off score, otherwise it would be like wasting an option.[25] Listing a less selective option above a more selective option may reflect a poor understanding of process. Given the exogeneity and uncertainty of the cut-off scores, I define an application 'mistake' as listing a first-choice school that has a cut-off score that is at least 10 points below the cut-off score of the second-option. I observe that about 20 percent of applicants in the bottom quartile make this 'mistake' while only 16 percent in the top quartile do.

High-achievers

Given the strong relationship between family socio-economic status and student achievement, one could argue that the patterns that I just described are reflecting that association. That is, applicants from more disadvantaged contexts tend to have lower achievement levels than students living in more privileged families, who therefore receive higher quality educational opportunities. To address this concern, an interesting extension to this analysis is to compare the application behavior of high-achieving students among socio-economic groups. I define as 'high-achieving' an individual who scored in the top 10 percent of the entrance examination distribution, for his or her application cohort. In a similar exercise for the US, Hoxby and Avery (2012) found that low-income students'

[25] It is not possible to be assigned to a more selective option after being rejected from a less selective option.

college-application behavior differ greatly from that of their high-income counterparts who have similar achievement levels.

In Table 4, I show that high-SES applicants are over-represented in the high-achieving student population. However, 28 percent of the high-achievers come from the bottom and second quartiles. I observe that among high-achievers, those from higher-SES families perform slightly better in the entrance examination than do the rest of the distribution, but the difference is very small (97.5 versus 95.6 points). I notice that there is a higher proportion of males (60%) than females in the high-achieving group. I report that, on average, the household head has a higher level of schooling among high-achievers in the top quartiles. In fact, an average parent in the top quartile has completed one year of college while an average parent in the bottom quartile has just completed lower-secondary education.

As expected, this group of high-achieving students is very homogenous in terms of academic achievement (Panel B) but differs in their attendance at private lower-secondary school. Almost one-third of high-achievers in the top quartile come from a private institution, while only one-tenth in the bottom quartile do.

Regarding application behavior, in Panel C of Table 4, I see that high-achievers also rank, on average, 9 schooling options. Given their high scores, they are therefore usually offered admission to their first-choice of school. Applicants in the bottom quartile are slightly less likely to select an elite school as their most-preferred option than those in the top quartile, even though they are equally prepared to attend this kind of institution. Although high-achievers do not list technical schools as their top-option and favor general-track schools, applicants in the first and second quartiles are slightly more likely, than those in the other two SES quartiles, to select a technological school as their top-option. This pattern might reflect to some extent different educational and work aspirations across quartiles, with lower-SES students favoring options that facilitate early transition into the job market.

Even though applicants do not know cut-off scores in advance, they have access to cut-off scores from the previous 2-3 years. Given the relative stability of the selectivity of some schools, one possible application mistake is to incorrectly rank their schools based on the schools' relative selectivity (Lucas & Mbiti, 2012). I define this mistake as listing a first option that has a cut-off score that is at least 10 points below the cut-off score of their second listed option. I observe that high-achievers from the lowest-two quartiles tend to choose less selective portfolios of school options than the rest of the SES distribution; with those in the top quartile having on average a slightly more ambitious portfolios. For instance, the average cut-off score for options listed by students in the top quartile is 77, while it is 70 for those in the bottom quartile. This group of talented students make slightly fewer "application mistakes" than the rest of the population. Nevertheless, I still find that individuals in the bottom quartile are more likely to make this kind of mistake and list as a top-choice an option that has a cut-off score which at least 10 lower than their second-choice, suggesting some misunderstanding about the admission mechanism.

Summing up, the descriptive evidence that I have presented here indicates that there are similarities across SES quartiles among high-achievers in terms of general application behavior, like the number of options listed. However, there are also some differences in terms of the composition of the preference portfolio according to their socio-economic background; with those in the bottom quartile being more cautious regarding their schooling choices than those in the top quartile. On average, low-SES applicants have a less selective portfolio and tend to favor non-elite technological or technical schools. In addition, high-achievers from low-SES families tend to underestimate their performance in the entrance examination more than those from high-SES families. Finally, students in the bottom quartile are more likely to make application mistakes by ranking first a school with a cut-score that is at least 10 points below the cut-score of the second-option. Therefore, it seems that, at the application stage, there are important differences in the behavior of high-SES and low-SES students with similar achievement levels. In an academic merit-based system like the one studied here, high-achieving applicants from different socio-economic backgrounds could have similar educational paths if they made similar application choices.

The Role of Information

Recently, research has turned to investigate the reasons why, conditional on academic achievement, students from more disadvantaged backgrounds have less ambitious education aspirations than their high-income counterparts (Hastings et al., 2006; Hoxby, 2009; Hoxby & Avery, 2012). Some researchers propose that market inefficiencies, like information asymmetries and imperfect information, could help explain the dissimilarities in application behavior followed by college applicants from different social backgrounds (Hastings & Weinstein, 2008; Hoxby & Avery, 2012; Ajayi & Telli, 2013). According to Lavecchia et al. (2014), too little information and too many options can lead to suboptimal educational outcomes, especially among low-income students. They argue that helping students and parents navigate the processes with an abundance of information or resources, such as simplifying how information is conveyed, can lead to better decisions and improved outcomes. Lee and Fitzgerald (1996) propose that sufficient and accurate information about the quality of educational options is fundamental for a rational choice. Bobba and Frisancho (2014) show that providing students in Mexico City with positive feedback about their performance in a mock examination induced changes in their ranked list of schools.

Growing evidence from the literature indicates that families gain information on schools from a wide variety of sources, both formal and informal; each one with its own strengths and weaknesses (Hall, 2009). Formal sources of information reach many people simultaneously. Nevertheless, subjective opinions from close networks about schools can be at least as important as formal sources (Brasington & Hite, 2012). Some researchers argue that high-SES families are endowed with an informational advantage from social networks that could explain differences in school-choice strategies (Allen et al., 2014). In particular, Schneider et al. (2000) suggest that middle- and high-income families have social networks of high-quality information, while low-income families have access to social networks that provide low-quality data. In the case of Mexico City, Dustan (2014) finds evidence that positive academic experiences of older siblings in a particular school, means that applicants are more likely to prefer that school over other choices. He argues that applicants prefer schools for which they have better information.

In what follows, I use quantitative and qualitative data collected in May 2013, as work previous to the randomized experiment conducted by Bobba and Frisancho (2014). Based on feedback from the focus groups, I describe the main take-aways from the perceptions and experiences of students regarding the upper-secondary admission process.

In my presentation, I emphasize the role that formal and informal sources of information play in the decision process. In particular, I characterize three types of students, based on their application behavior and their access to sources of information. I then use survey data to explore in more detail whether perceptions, examination preparation strategies, and access to sources of information related to the upper-secondary admission process differ by student socio-economic level.

As I mentioned above, I merged these survey data with administrative data from COMIPEMS. I estimated the students' SES quartile in a similar way to before, based on the entire SES distribution of the application cohort. A limitation of the data analyzed here is that the focus group and survey processes were conducted in a small sample of schools, with students from slightly higher SES families than the cohort's average.[26] However, even though the survey and focus groups samples might not be representative of the full population of applicants, I argue that my findings shed some light on interesting patterns and possible mechanisms to explain why individuals from different socio-economic backgrounds have different applicant behaviors.

Perception and Experience of Applicants in the Admission Process

In general, 9th grade students who participated in the focus groups perceived the upper-secondary admission process as "fair" because it is based on an objective and homogenous measure of student achievement. They acknowledged the relevance of the CIEMS admission for the fulfillment of their educational and professional plans. This,

[26] See Table 5 for a comparison of selected characteristics between the survey sample and the universe of 2013 applicants.

combined with the perception of a "long" and "complicated" process, translated into a feeling of stress among most of the participants. Students expressed their confusion and uncertainty when having to select, very early in the process, 20 school options from a "very large set of options." Some mentioned that just going over the list of 600 options was "overwhelming" and that choosing options was something too difficult to do without any guidance. In addition, their feelings of stress was augmented by the generalized perception that their lower-secondary schools had not prepared them well for the entrance examination and that additional extra-curricular preparation was needed to have a good performance.

According to participants, each lower-secondary school has a person or teacher in charge of helping 9^{th} grade students with the application process. However, in most of the sampled schools, students mentioned that the role of this "counselor" was limited to the administrative part of the process (i.e. documents, important dates, etc.) and, to a lesser extent, on providing "useful" information (e.g. type of options and characteristics of the subsystems) that could help them make a better-informed decision about their portfolio of options. I found out that, in some schools, the counselor had an active role, including administering mock or practice tests.

In some cases, I detected that participants received confusing or inadequate advice from teachers or counselors, like "choose the nearest school to avoid transportation fees," "select schools with low demand to guarantee a place," or "choose a low-demand option first and then a more selective option." This not only reflects disparities in the quality of the information received but also indicates that some school counselors do not fully understand the admission process.

In the focus groups, I noticed that a reduced group of participants had incorrect beliefs about the assignment mechanism. They believed that, in addition to the score in the entrance examination and preferences, the process considers other factors like age, distance to school, sibling, and their performance in the national standardized test (i.e. ENLACE). Confusion may arise from the fact that some of these factors are relevant for the primary and lower-secondary school assignment.

Since the information and guidance that students received at school seemed limited, most of the participants acknowledged looking for external sources of information. Formal sources included reading the manuals, reading the study guide, checking the COMIPEMS website, attending preparation courses, and visiting schools, among others. Nevertheless, informal sources, such as social networks, were mentioned among the most frequent sources of information and advice. In particular, siblings and family members, who had participated already in the process, were considered the most useful source of information. Recommendations provided by social networks also differed widely and tended to be founded on subjective perceptions and past experiences. The "quality and reputation" of the schools were mentioned on multiple occasions as reasons to select one school over another, but very few participants were able to explain what features these terms were based on.

Overall, in analyzing the data from the focus groups, I realized that –across sampled schools– there were notable differences in the degree of knowledge that students possessed about the admission process, the sources of information, and the quality of the advice they received. Although all of the participants were aware of the administrative process and timeline, some of them were uncertain as to which schools to choose and how to prepare for the entrance examination.

Types of Applicants

Regarding the application experiences of 9th grade students and their access to different sources of information, I categorized three types of applicants based on information from the focus groups. The first type included students whose parents actively engaged in the admission process and who had access to multiple sources of both formal and informal information. They had parents who had completed upper-secondary education, knew close friends and relatives who had participated in the process and had attended elite upper-secondary schools. They also had the resources to pay for private tutoring or preparation courses. Students from this group received good guidance regarding the admission process at their lower-secondary school. In general, applicants from this group targeted selective schools and were confident about their outcomes.

The second type of students had some degree of parental support and access to informal sources of information. In some cases, they were the first generation in their families to attend upper-secondary education so they looked for teachers' and friends' advice. Students in this group constructed a preference portfolio with a combination of selective and non-selective schools, from both general and technological tiers. They studied for the examination alone or in groups, usually with the help of a printed study guide. Overall, they had ambitious educational aspirations but were cautious when constructing their portfolio.

The third type of applicants had little parental involvement and low educational aspirations. They only wanted to attend upper-secondary education in order to enter the labor market in a better position. Students in this group registered for the centralized admission process but listed a group of low-demand options –mostly technical schools– near their homes, and in which they thought they could secure a seat. In contrast with the other groups, they spent little time looking for external sources of information and preparing for the examination. In summary, students from this group did not pay a lot of attention to the admission process and constructed a portfolio with mainly safe-bets and nearby options.

To some extent, the application behavior of students appeared to be closely related to their socio-economic background and their access to sources of high-quality information. In the next section, I explore these relationships in more detail.

Differences in the Access to Information, by Socio-Economic Background

Qualitative data from the focus groups suggested that, consistent with the literature, access to formal and informal sources of information about the upper-secondary admission process in Mexico City may differ by socio-economic background. Using data from a small sample of lower-secondary education students in 2013, I examined the variation in their perceptions about their admission chances, as well as their access to different sources of information, by socio-economic background. My analysis had two main caveats. First, my findings are derived from a small convenience sample, containing an over-representation of students in the top SES quartiles. Therefore, my findings do not reflect accurately what is happening in the population, but might shed light on interesting patterns present in specific socio-economic sectors. Second, although the survey asked respondents about their access to different sources of information, I cannot assess the quality of the feedback participants received from each source. That is, I can distinguish the proportion of the applicants who reported asking teachers for guidance, but I cannot assess whether they received high- or low-quality information.

I present, in Table 6, summary statistics that describe students' own perceptions of their admission chances and the strategies they followed to prepare for the examination. In each column, I reported means and standard deviations (in parenthesis) for all sampled students in that particular quartile. For simplicity, I focused on the top and bottom quartiles of the SES distribution, high- and low-SES respectively. In Panel A, I observed a statistically significant difference between low- and high-SES students regarding their perceptions of the quality of the lower-secondary school they attended; with those in the top quartile having a better perception of their school, on average. I saw that, on average, high-SES students had a slightly better perception of their academic performance and expected to obtain a higher score in the entrance examination than low-SES students. Comparing their anticipated scores with their actual scores, I noticed that both groups of students overestimated their score on the entrance examination. The overestimation tended to be significantly higher among low-SES students. Related to this, a statistically significantly lower proportion of students in the bottom quartile listed a university-affiliated school (i.e. an elite school) among their preferences, compared to those in the top quartile. Overall, this suggests that surveyed lower-secondary students had a limited knowledge of quality of the school they attended and about their own ability relative to the entire distribution of upper-secondary applicants.

In Panel B of Table 6, I present statistics for different strategies adopted by surveyed students when preparing for the examination. In each column, I reported means and standard deviations (in parenthesis) for all sampled students in that particular SES group. I noticed that low-SES students were statistically significantly more likely to study alone and less likely to attend exam preparation courses than high-SES students. Among those that attended courses, the average cost paid by students in the top quartile ($3,700 pesos) was around 1,500 pesos more than that paid by students in the bottom quartile. Attendance at preparation courses can help explain, to some extent, disparities in the access to information about their own ability and in guidance received when ranking schools during the admissions process. In fact, many courses have a money-back guarantee policy, if course participants were not offered admission to one of their five highest-ranked choices.

Finally, I found no statistically significant differences in the percentage of students who reported studying with friends, family or a private tutor. This indicates that, while high-SES students tended to receive a more formal training and a lot of practice exercises in the preparation courses, low-SES students prepared alone and in a less structured way for the examination.

Next, in Table 7, I compare access to different sources of information between low- and high-SES applicants and found striking similarities. Although I detected no statistically significant differences across groups, I identified some interesting patterns that are consistent with the content of the conversations in the focus groups. As I expected, the application manual was the most-consulted source of formal information, with visits to the subsystems' exposition in January of each year the least exploited resource (Panel A). In general, there was little difference in access to formal sources of information, which indicates that administrative information of the CIEMS process is reaching 9th grade students effectively, independently of their social background. In fact, low-SES students used COMIPEMS' materials at a slightly, although not statistically significant, higher rate than did high-SES students. The main difference regarding formal sources of information seems to be that a larger proportion of high-SES students, than low-SES students, visited upper-secondary schools to inform their choices. Therefore, those in the top quartile have, on average, more direct school-based information than those in the bottom quartile.

Regarding informal sources of information (Panel B in Table 7), I did not observe any statistically significant difference by SES level. Consistent with conversations during the focus group meetings, family members that had already participated in the admission process was the most important informal source of information. High-SES students had slightly higher chances of knowing someone that had already participated in the process. On average, only 25 percent of the surveyed applicants reported talking with their parents about the admission process; even though parents were required to approve and sign their ranked list of school options. Almost 40 percent of the surveyed applicants, regardless of the SES quartile, looked for information about their options on the Internet. Although I did not detect any statistically significant difference in their access to these informal sources of information, I did see that social networks are an important part of the admission process. At this point, the question remains as to whether the quality of the information received by students from different social backgrounds is similar.

Conclusion

In this paper, I have explored barriers to access to upper-secondary education for the case of the centralized upper-secondary education admission system implemented in Mexico City. First, I showed that, in this academic merit-based system, there existed student stratification, by ability and socio-economic background. This suggests that the student-allocation process in the system might be reinforcing initial disparities and hindering educational mobility. For instance, many high-SES students accessed selective upper-secondary public schools through lower-secondary private schools, making competition disadvantageous given their family and academic background.

Second, I demonstrated that the similarities and differences in the application behavior of high-achieving students in Mexico City according to their socio-economic background. In an academic merit-based system like the one studied here, high-achievers from different socio-economic backgrounds could have similar educational paths if they made similar application choices. Although on average they list the same number of options, the composition of their preference portfolio differs. Many high-achieving students from low-SES families did not apply to selective schools, despite being qualified for admission. This pattern might reflect to some extent different educational and work aspirations across quartiles, with lower-SES students favoring options that facilitate early transition into the job market. Dissimilarities in application behavior could be explained by multiple factors such as the availability of economic resources, the opportunity for attending private school and taking preparation courses, and access to information networks. I propose that, in addition to economic resources, imperfect information and asymmetries in the access to high-quality information may explain why some students from disadvantaged families made different schooling choices. In summary, I find that the benefits of a merit-based system are more likely to promote equality of educational opportunities when there is perfect information among all the participants.

I believe that more research is needed in this area, particularly in the investigation of the causal relationship between students' SES and their application behaviors, in order to design targeted interventions to promote educational mobility under merit-based systems. Based on the analysis presented in this paper, I believe that possible policy implications to level the playing field across SES groups might include: (1) changing the timeline of the process so that students apply to schools after they have taken the entrance examination and know their results, (2) offering exam preparation courses and counselling to low income students, to help them level the playing field, and (3) create mechanisms to standardize quality of information such as more online systems, systems aided by artificial intelligence, and a 01-800 number to answer questions.

References

Abdulkadiroglu, A., K., Che and Y. Yasuda (2012). "Expanding 'Choice' in School Choice," Working paper.

Ajayi, K. F. (2011). "School choice and educational mobility: Lessons from secondary school applications in Ghana," University of California-Berkeley Working Paper.

Ayaji, K. and H. Telli (2013). "Imperfect Information and School Choice in Ghana," unpublished working paper.

Belasco, A. S., & Trivette, M. J. (2012). "Aiming low: estimating the scope and predictors of postsecondary undermatch," Working paper, College Transitions.

Bentaouet Kattan, R. and M. Szekely (2014). "Dropout in upper secondary education in Mexico: patterns, consequences and possible causes," Policy Research working paper 7083, World Bank Group.

Bobba, M., and V. Frisancho, V. (2014). "Learning About Oneself: The Effects of Signaling Academic Ability on School Choice," Working paper.

Bowen, W., M. Chingos and M. McPherson (2009). Crossing the Finish Line: Completing College at America's Public Universities. Princeton: Princeton University Press.

Brasington, D. M. and D. Hite (2012). "School choice and perceived school quality," Economics Letters 116(3): 451-453.

Cardenas, S. (2009). “Is the Class Schedule the Only Difference between Morning and Afternoon Shift Schools in Mexico?” Doctoral dissertation, Harvard Graduate School of Education.

COMIPEMS (2008). “Informe del Concurso de Ingreso a la Educación Media Superior de la Zona metropolitana de la Ciudad de México (1996-2005),” Ceneval Publication.

COMIPEMS (2012). “Informe del Concurso de Ingreso a la Educación Media Superior de la Zona metropolitana de la Ciudad de México (1996-2010),” Ceneval Publication.

de Janvry, A., A. Dustan and E. Sadoulet (2012). “The Benefits and Hazards of Elite High School Admission: Academic Opportunity and Dropout Risk in Mexico City,” Unpublished Working Paper, University of California at Berkeley.

Dustan, A. (2013). “Peer Networks and School Choice under Incomplete Information,” Working paper, University of California, Berkeley.

Hall, C. (2009). “Does making upper secondary school more comprehensive affect dropout rates, educational attainment and earnings? Evidence from a Swedish pilot scheme,” Working paper No. 2009: 9, IFAU-Institute for Labour Market Policy Evaluation.

Hastings, J., T. Kane, and D. Staiger (2006). “Preferences and Heterogeneous Treatment Effects in a Public School Choice Lottery,” NBER Working Papers No. 12145, National Bureau of Economic Research.

Hastings, J. and J. Weinstein (2008). "Information, School Choice, and Academic Achievement: Evidence from Two Experiments," The Quarterly Journal of Economics 123(4): 1373-1414

Hastings, J., C. Neilson, and S. Zimmerman (2012). “The Effect of School Choice on Intrinsic Motivation and Academic Outcomes,” NBER Working Paper No. 18324, National Bureau of Economic Research.

Hoxby, C. and C. Avery (2012). “The Missing "One-Offs": The Hidden Supply of High-Achieving, Low Income Students,” NBER Working Paper No. 18586, National Bureau of Economic Research.

INEE (2012). La Educacion Media Superior en Mexico. Instituto Nacional para la Evaluacion de la Educacion, Mexico.

Jackson, K. (2010). "Do Students Benefit From Attending Better Schools?: Evidence From Rule-based Student Assignments in Trinidad and Tobago," The Economic Journal, 120(549): 1399–1429.

Lavecchia, A., H. Liu and P. Oreopoulos (2014). “ Behavioral Economics of Education: Progress and Possibilities.” NBER Working Paper 20609, National Bureau of Economic Research.

Lucas A. and I. Mbiti (2012). "The Determinants and Consequences of School Choice Errors in Kenya," American Economic Review 102(3): 283-88.

Musset, P. (2012), “School Choice and Equity: Current Policies in OECD Countries and a Literature Review”, OECD Education Working Papers, No. 66, OECD Publishing.

Pallais, A. (2013). “Small Differences that Matter: Mistakes in Applying to College,” NBER Working Paper No. 19480, National Bureau of Economic Research.

Perna, L. W., and S. L. Thomas (2006). “A framework for reducing the college success gap and promoting success for all,” In National Symposium on Postsecondary Student Success: Spearheading a Dialog on Student Success, July.

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

Solis, P., E. Rodriguez Rocha and N. Brunet (2013). “Origenes sociales, instituciones y Decisiones Educativas en la Transicion a la Educacion Media Superior: El Case del Distrito Federal,” Revista Mexicana de Investigacion Educativa 18(59):1103-1136.

Tirado, F. (2004). “Perfiles del EXANI-I. Evaluación de la educación en México,” In Indicadores del EXANI-I, CENEVAL, Mexico, pp 97-148.

Trevino, E. (2006). “Evaluacion del Aprendizaje de los Estudiantes Indígenas de America Latina”. Revista Mexicana de Investigación Educatival 11(28): 225-268.

Table 1: Universe of upper-secondary applicants, by cohort

	Application Cohort				
	2005	**2006**	**2007**	**2008**	**2009**
Number of schools	305	308	312	315	321
Number of Options (School-track combinations)	634	650	657	663	676
Applicants	287,886	298,291	296,778	303,224	317,603
Average number of options	8.8	9.2	9.3	9.5	9.5
Filled 20 possible options	2%	2%	3%	3%	3%
Took examination	280,283	287,814	290,425	290,237	304,709
Average rank of option assigned	3.0	3.3	3.4	3.5	3.6
Average score	62.2	63.8	63.6	64.7	60.8
Eligible for assignment	249,886	257,287	259,372	258,038	271,576
Assigned	230,010	236,503	238,527	239,001	241,669
1st Round	207,042	212,236	216,717	221,277	226,911
1st option	45%	41%	40%	40%	38%
2nd option	14%	14%	14%	14%	14%
3rd option	10%	10%	10%	10%	10%
4th option	8%	8%	8%	8%	9%
5th-10th option	21%	24%	24%	25%	26%
11th-20th option	2%	3%	3%	3%	4%
2nd Round	22,581	22,959	18,773	16,886	13,279
Not eligible	30,397	30,407	31,053	32,198	33,132
Without lower secondary certificate	25,416	25,802	25,212	26,915	25,881
Less than 31 points	4,981	4,605	5,841	5,283	7,251

Table 2a: Summary statistics describing the composition of the student body for the 2005-2009 application cohorts, by school subsystem

	General					Technological				Techincal	All
	SS1	SS2	SS3	SS4	SS5a	SS6	SS7	SS5b	SS8	SS9	
Subsystem's characteristics											
Average score of admitted applicants on the entance examination	89.23	85.29	70.65	63.66	61.62	88.36	59.09	59.5	51.62	51.51	66.14
Expenditure per student (2013 pesos)	40,827	…	…	17,253	…	…	…	…	…	14,177	…
Demographic characteristics											
Male	48%	50%	40%	48%	39%	65%	49%	48%	44%	51%	49%
Age	15.29	15.11	15.66	15.70	15.24	15.25	15.71	15.33	15.42	15.68	15.49
Indigenous	1%	1%	2%	3%	4%	2%	4%	5%	6%	5%	4%
Lives in the metropolitan zone	95%	94%	95%	94%	93%	95%	93%	92%	92%	93%	94%
Travel distance to to school (km)	9.65	7.30	6.25	7.55	6.26	11.19	8.07	6.68	8.42	7.59	8.00
Schooling trajectories											
GPA lower-secondary	8.64	8.68	8.01	7.92	8.16	8.51	7.86	8.05	7.91	7.73	8.09
Failed at least one course in lower-secondary education	17%	17%	33%	35%	27%	19%	36%	29%	35%	39%	30%
Attended a general lower-secondary	68%	82%	68%	65%	71%	63%	61%	69%	67%	64%	66%
Attended a technical lower-secondary	29%	15%	28%	29%	23%	35%	31%	24%	25%	27%	28%
Attended a telesecundaria lower-	1%	2%	1%	2%	4%	1%	3%	5%	5%	5%	3%
Attended an adult education	1%	0%	2%	3%	1%	1%	2%	1%	2%	3%	2%
Observations	172,635	3,283	10,059	203,930	187,274	101,772	208,959	82,006	4,356	211,852	1,186,126

Table 2b: Summary statistics describing the composition of the student body for the 2005-2009 application cohorts, by school subsystem

	General					Technological				Techincal	All
	SS1	**SS2**	**SS3**	**SS4**	**SS5a**	**SS6**	**SS7**	**SS5b**	**SS8**	**SS9**	
Studying Habits											
Study time (hrs per week)	6.53	5.83	5.39	5.03	4.96	6.17	4.71	4.71	4.47	4.37	5.17
Reading time (hours per week)	4.63	4.09	3.99	3.88	3.71	4.20	3.64	3.58	3.64	3.49	3.87
Books read (per year)	4.54	4.21	3.43	3.37	3.62	4.01	3.16	3.51	3.42	3.04	3.57
Educational Aspirations											
Want to study higher education	93%	88%	83%	78%	73%	86%	62%	60%	57%	53%	72%
Want to attend a private university	1%	2%	1%	1%	1%	1%	1%	1%	1%	1%	1%
Expect to complete upper-secondary education	0%	1%	1%	2%	3%	0%	3%	4%	5%	5%	2%
Expect to complete vocational education	6%	15%	18%	20%	21%	17%	25%	25%	25%	25%	20%
Expect to complete college	29%	35%	40%	39%	41%	29%	35%	34%	32%	32%	34%
Expect to complete graduate studies	65%	53%	42%	39%	33%	57%	28%	26%	25%	22%	38%
Observations	172,635	3,283	10,059	203,930	187,274	101,772	208,959	82,006	4,356	211,852	1,186,126

Table 2c: Summary statistics describing the composition of the student body for the 2005-2009 application cohorts, by school subsystem

	General					Technological				Techincal	All
	SS1	SS2	SS3	SS4	SS5a	SS6	SS7	SS5b	SS8	SS9	
Characteristics of the Household head											
Age	43.22	42.45	42.66	42.20	41.34	42.46	41.75	41.18	41.15	41.47	41.97
Schooling (years)	11.68	11.40	10.22	9.53	8.99	10.89	8.82	8.43	8.13	8.33	9.50
Completed at least upper-secondary	56%	50%	40%	33%	27%	49%	26%	22%	19%	20%	33%
Unemployed	5%	4%	6%	6%	6%	5%	7%	7%	7%	7%	6%
White-Collar job	28%	28%	18%	15%	13%	21%	11%	10%	10%	9%	15%
Socio-economic characteristics											
Currently working	4%	4%	6%	8%	5%	4%	8%	6%	6%	8%	6%
Attended private lower-secondary	19%	17%	6%	5%	4%	14%	3%	2%	2%	2%	6%
Family income (pesos per month)	7,510	6,575	5,754	5,240	4,735	6,549	4,558	4,261	4,032	4,132	5,255
Per capita family income (pesos per month)	1,782	1,491	1,358	1,184	1,040	1,517	1,010	920	860	902	1,188
Family receives Oportunidades	1%	3%	1%	2%	5%	2%	3%	6%	5%	4%	3%
Has computer or laptop at home	65%	54%	53%	46%	39%	61%	39%	36%	32%	34%	45%
Has Internet access at home	36%	24%	26%	21%	16%	31%	16%	13%	13%	13%	21%
People per room	1.97	1.95	2.23	2.24	2.32	2.08	2.37	2.42	2.54	2.44	2.27
Cultural capital											
Went to theater or cinema (per year)	3.67	2.84	3.20	2.79	2.17	3.25	2.34	1.89	1.80	2.08	2.59
- At least once	92%	86%	90%	87%	80%	90%	83%	77%	75%	80%	84%
Visited museum (visit per year)	3.28	1.93	3.04	2.62	1.91	2.97	2.23	1.68	1.81	2.03	2.39
- At least once	93%	79%	91%	88%	79%	91%	84%	76%	79%	81%	85%
Number of books at home	119	101	82	70	61	95	58	54	58	51	72
Observations	172,635	3,283	10,059	203,930	187,274	101,772	208,959	82,006	4,356	211,852	1,186,126

Table 3: Sample means (and standard deviations) of selected characteristics of upper-secondary education applicants, by socio-economic quartile

	Family Socio-economic Status			
	Quartile 1 (bottom)	Quartile 2	Quartile 3	Quartile 4 (top)
Panel A: Socio-demographic characteristics				
Male	0.456	0.498	0.514	0.522
	(0.498)	(0.500)	(0.500)	(0.500)
Age	15.580	15.669	15.639	15.444
	(1.412)	(1.555)	(1.627)	(1.315)
Indigenous	0.063	0.030	0.012	0.009
	(0.242)	(0.171)	(0.139)	(0.122)
Schooling of household head (years)	7.729	8.684	9.789	11.712
	(3.575)	(3.485)	(3.716)	(4.045)
Panel B: Academic characteristics				
GPA lower-secondary (out of 10)	7.262	7.139	7.139	7.327
	(2.473)	(2.508)	(2.542)	(2.497)
Score in entrance examination	58.538	59.890	64.101	69.643
	(18.032)	(18.216)	(18.864)	(19.731)
Attended private lower-secondary school	0.018	0.022	0.047	0.177
	(0.132)	(0.146)	(0.212)	(0.382)
Panel C: Application characteristics				
Number of options listed	9.222	9.297	9.362	9.269
	(3.611)	(3.641)	(3.736)	(3.838)
Selectivity of the portfolio	60.83	62.13	65.63	69.95
	(13.276)	(13.197)	(13.073)	(12.881)
Rank of the assigned option	3.247	3.331	3.377	3.304
	(2.846)	(2.919)	(2.991)	(3.003)
Assigned to first option	0.259	0.257	0.267	0.291
	(0.438)	(0.437)	(0.443)	(0.454)
First-choice elite school	0.527	0.555	0.638	0.746
	(0.499)	(0.497)	(0.480)	(0.435)
Distance to first option (km)	8.292	8.442	8.214	8.268
	(8.539)	(8.300)	(8.209)	(8.041)
Application mistakes	0.194	0.187	0.175	0.162
	(0.396)	(0.390)	(0.380)	(0.369)
Observations	354,038	376,016	367,675	368,542

Note: ‘Selectivity of the preferences portfolio’ calculated as the mean of the cut-off scores of all the school options listed by an applicant. ‘Application ‘mistake’ implies listing a first-choice school that has a cut-off score that is at least 10 points below the cut-off score of the second-option.

Table 4: Sample means (and standard deviations) of selected characteristics of upper-secondary education high-achieving applicants, by socio-economic quartile

	Family Socio-economic status			
	Quartile 1 (bottom)	Quartile 2	Quartile 3	Quartile 4 (top)
Panel A: Demographic characteristics				
Male	0.576	0.596	0.600	0.579
	(0.494)	(0.491)	(0.490)	(0.494)
Age	15.425	15.576	15.543	15.283
	(1.457)	(1.623)	(1.576)	(1.070)
Indigenous	0.029	0.012	0.006	0.004
	(0.169)	(0.109)	(0.079)	(0.059)
Schooling of household head (years)	9.586	10.321	11.395	13.569
	(3.754)	(3.728)	(3.795)	(3.945)
Panel B: Academic characteristics				
GPA lower-secondary (out of 10)	8.579	8.528	8.432	8.494
	(1.533)	(1.447)	(1.516)	(1.462)
Score in entrance examination	95.588	95.571	96.358	97.459
	(6.550)	(6.482)	(6.885)	(7.378)
Attended private lower-secondary school	0.082	0.069	0.121	0.345
	(0.275)	(0.253)	(0.327)	(0.475)
Panel C: Application characteristics				
Number of options listed	9.607	9.539	9.513	9.149
	(3.861)	(3.760)	(3.861)	(3.955)
Selectivity of the portfolio	69.69	69.51	72.27	76.92
	(12.464)	(12.416)	(11.533)	(10.268)
Rank of the assigned option	1.171	1.167	1.205	1.254
	(0.610)	(0.542)	(0.737)	(0.821)
Assigned to first option	0.822	0.820	0.814	0.795
	(0.383)	(0.384)	(0.389)	(0.404)
First-choice elite school	0.828	0.806	0.857	0.917
	(0.378)	(0.396)	(0.356)	(0.276)
Distance to first option (km)	9.623	9.769	9.159	8.470
	(8.785)	(8.685)	(8.256)	(7.656)
Application mistakes	0.138	0.140	0.127	0.116
	(0.345)	(0.347)	(0.333)	(0.320)
Observations	3,619	4,771	7,228	14,733

Note: High-achieving applicant is defined as having an entrance score at or above the 90th percentile.

Table 5: Means (and standard deviations) describing applicants' selected characteristics, in the survey sample and the full sample

	Sample	
	Survey	Full
Panel A: Demographic characteristics		
Male	0.498	0.509
	(0.500)	(0.500)
Age	15.112	15.550
	(0.520)	(1.532)
Indigenous	0.022	0.039
	(0.147)	(0.194)
Working status	0.858	0.828
	(0.349)	(0.378)
Panel B: Family characteristics		
Living in Mexico City	1.000	0.964
	(0.000)	(0.186)
Schooling of household head (years)	6.758	6.129
	(3.039)	(3.242)
SES quartile	2.808	2.501
	(1.090)	(1.117)
Panel C: Academic characteristics		
Score in entrance examination	79.950	67.799
	(22.685)	(24.188)
Attended private lower-secondary school	0.105	0.071
	(0.307)	(0.257)
GPA (lower-secondary)	7.824	7.490
	(1.865)	(2.152)
N	1,593	310,163

Notes: Survey sample includes all 9th grade students that responded the survey administered by Bobba & Frisancho (2013). Full sample corresponds to the universe of applicants that participated in the 2013 COMIPEMS admission process.

Table 6: Sample means (and standard deviations) of selected characteristics of of 9th grade surveyed applicants, by socio-economic level

	Low-SES	High-SES	Difference	
	(1)	(2)	(1) - (2)	
Panel A: Perceptions				
Evaluate your school (out of 10)	7.940	8.248	-0.308	***
	(1.434)	(1.274)	[-2.61]	
Auto-evaluate your academic performance (out of 10)	7.822	8.002	-0.180	*
	(1.171)	(1.041)	[-1.88]	
Expected score in entrance examination	96.545	104.535	-7.991	***
	(17.391)	(14.811)	[-5.66]	
Difference between expected and real score	25.070	18.434	6.635	***
	(22.262)	(17.726)	[3.71]	
Included an university-affiliated school in their portfolio	0.741	0.892	-0.151	***
	(0.439)	(0.310)	[-4.37]	
Panel B: Examination Preparation				
Study Alone	0.948	0.878	0.070	***
	(0.222)	(0.328)	[3.05]	
Study with friends	0.259	0.245	0.015	
	(0.440)	(0.431)	[0.38]	
Study with private tutor	0.594	0.557	0.037	
	(0.492)	(0.497)	[0.84]	
Study with parents	0.569	0.500	0.069	
	(0.497)	(0.501)	[1.56]	
Attend a preparation course	0.525	0.756	-0.231	***
	(0.501)	(0.430)	[-5.15]	
Cost of the course (pesos)	2067.7	3711.4	-1643.6	***
	(1,573.5)	(2,527.0)	[-6.58]	
Observations	213	452		

Note: *** p<0.01, ** p<0.05, * p<0.1. Difference of means in italics, t-statistics in squared brackets. Low-SES and high-SES applicants are defined, respectively, as the ones from the bottom and top quartile of the SES distribution for the application cohort.

Table 7: Sample means (and standard deviations) for selected sources of formal and informal of information for 9th grade surveyed applicants, by socio-economic level

	Low-SES	High-SES	Difference	
Panel A: Formal Information				
Application Manual	0.931	0.919	0.012	
	(0.255)	(0.274)	[0.54]	
Workbook	0.785	0.738	0.047	
	(0.412)	(0.440)	[1.33]	
Expo Subsystems	0.305	0.271	0.034	
	(0.461)	(0.445)	[0.87]	
Website	0.758	0.705	0.052	
	(0.430)	(0.456)	[1.40]	
Past cut-off scores	0.620	0.610	0.010	
	(0.487)	(0.488)	[0.23]	
School location tool	0.667	0.587	0.079	
	(0.473)	(0.493)	[1.95]	
Mock examination	0.403	0.324	0.079	
	(0.493)	(0.469)	[1.54]	
School visit	0.653	0.719	-0.066	*
	(0.477)	(0.450)	[-1.71]	
Panel B: Informal Information				
Family member participated in COMIPEMS	0.753	0.800	-0.047	
	(0.433)	(0.400)	[-1.21]	
Parent or sibling participated in COMIPEMS	0.395	0.416	-0.021	
	(0.490)	(0.494)	[-0.46]	
Talked with parents and siblings about the process	0.232	0.252	-0.019	
	(0.423)	(0.434)	[-0.54]	
Talked with extended family members	0.136	0.095	0.041	
	(0.344)	(0.294)	[1.46]	
Talked with school teachers or counselors	0.197	0.202	-0.005	
	(0.399)	(0.402)	[-0.14]	
Talked with classmates or friends	0.051	0.048	0.003	
	(0.220)	(0.213)	[0.16]	
Used study guide	0.117	0.139	-0.023	
	(0.322)	(0.347)	[-0.80]	
Looked for information in the Internet	0.388	0.380	0.008	
	(0.488)	(0.486)	[0.196]	
Observations	213	452		

Note: *** p<0.01, ** p<0.05, * p<0.1. Difference of means in italics, t-statistics in squared brackets. Low-SES and high-SES applicants are defined, respectively, as the ones from the bottom and top quartile of the SES distribution for the application cohort.

Appendix

Appendix A. Subsystems and Institutions that integrate the COMIPEMS

Table A.1. Subsystems and Institutions that participate in COMIPEMS

Model	Subsystem		Institutions
General	UNAM	Universidad Nacional Autonoma de Mexico	CCH, ENP
	UAEM	Universidad Autónoma del Estado de Mexico	Texcoco
	COLBACH	Colegio de Bachilleres	COLBACH
	DGB	Dirección General de Bachillerato	CEB
	SE	Secretaría de Educación del Estado de México	COBAEM
Technical	CONALEP	Colegio Nacional de Educación Profesional Técnica	DF, State of Mexico
Technological (General + Technical)	SE	Secretaría de Educación del Estado de México	CECYTEM
	DGETI	Dirección de Educación Tecnológica Industrial	CBTIS, CETIS
	DGETA	Dirección de Educación Tecnológica Agropecuaria	CBTA
	IPN	Instituto Politécnico Nacional	CET, CECYT

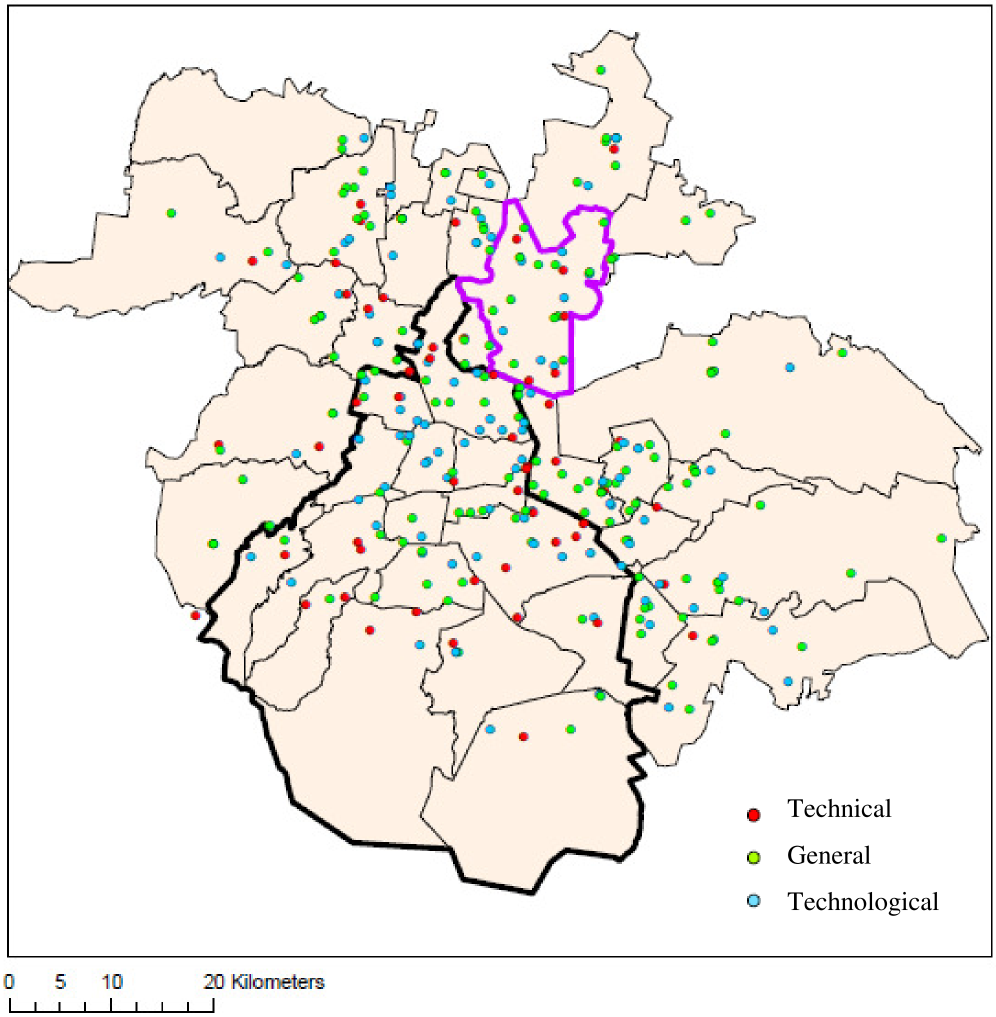
The main characteristics of each subsystem:

- General: prepares students with general knowledge to continue studying for post-secondary education. It is offered in three-year programs. Graduates receive a certificate that is necessary to enter higher education. There are several modalities: standard, open and distance, and mixed.

- Technological: covers the curriculum of general education to prepare students for post-secondary education. In addition, it gives students the opportunity to receive vocational training. This bivalent character allows graduates to continue into higher education and gives them also a technical diploma at a semi-professional level. Programs last three years.

- Technical: trains students to become professional technicians and gives them the option to continue into tertiary education. Programs are completed in three years.

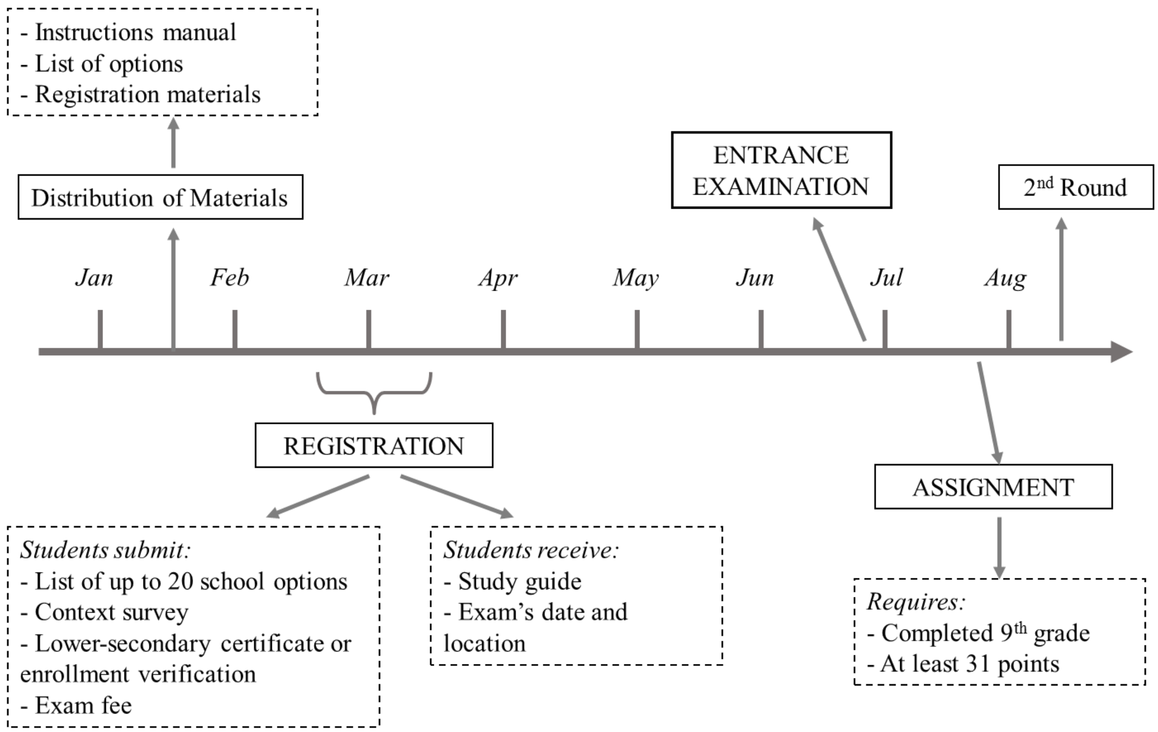
Appendix B. Geographic Distribution of School Options



Source: Avitable, Bobba and Pariguana (2015). "High-School Track Choice and Financial Constraints: Evidence From Urban Mexico," Working paper.

Appendix C: The COMIPEMS Admission Process

C.1. Timeline



C.2. Registration Materials

The registration materials that are distributed to individuals who want to attend a public upper-secondary school in Mexico City include:

a) A manual with a detailed explanation of the process and the directory of available school options, including each option's cut-off scores from previous years.

b) A registration form to fill with personal information and the list of up to 20 preferred school options they were willing to attend, ordered by preference.

c) A context survey seeking information on the students' demographic, academic and socio-economic characteristics.

d) A bank deposit slip to pay the fee for sitting the standardized entrance examination (approximately 25 U.S. dollars).

Appendix D: Definition and coding description of the variables

Table D.1. Definition and coding of the variables for the statistical analysis

Variable	Definition	Values
MALE	Applicant is a male	yes=1, no=0
INDIGENOUS	Applicant reported speaking an indigenous language	yes=1, no=0
AGE	Age of applicant at baseline	14 to 25
WORK	Applicant was working for a wage at baseline	yes=1, no=0
HH_SCHOOLING	Years of education for the highest level of education completed by the household head at baseline	no education=0, primary=6, lower-secondary=9, upper-secondary=12, university=17, graduate=20
RESIDENT	Applicant lives in the metropolitan area of Mexico City	yes=1, no=0
SES_INDEX	Principal components index of access to durables goods (like car, computer, dvd) and services (like sewage, telephone, internet) at home	-6 to 6
SES_quartile	Quartile of the SES distribution to which applicant's family belongs according to the SES_INDEX	1 to 4
SCORE	Score on the standardized entrance examination	0 to 128
GPA	GPA obtained in lower-secondary education	6 to 10
PRIVATE_LOWERSEC	Applicant attended a private lower-secondary school	yes=1, no=0

Just Making the Admission Cut-off:

The Impact of the Offer of Admission to, and Enrollment in, a More Competitive School on the Graduation Outcomes of Upper-Secondary Applicants in Mexico City

Just Making the Admission Cut-off:

The Impact of the Offer of Admission to, and Enrollment in, a More Competitive School on the Graduation Outcomes of Upper-Secondary Applicants in Mexico City

In education markets with oversubscribed schools, not all the applicants can be assigned to their most-preferred school. Therefore, important policy questions concern whether it matters to applicants' subsequent educational outcomes which school they are offered admission to, and enroll in. The empirical difficulties in identifying the causal effect of admission to, and enrollment in, a more competitive[27] school lie in the fact that students may self-select schools based on the latter's observed and unobserved characteristics. Nevertheless, using quasi-experimental designs that capitalize on plausibly exogenous variation in school admission among students, some researchers have provided credible estimates of the effect on subsequent educational outcomes of having access to a better quality, or more selective, secondary school.[28]

In this paper, using such a quasi-experimental discontinuity design, I estimated the causal impact on subsequent upper-secondary graduation outcomes, of an offer of admission to, and enrollment in, a more competitive school for an applicant at the margins of being accepted into that school, using data from Mexico City. In particular, I explore the effect on 3-year on-time graduation and graduation within 5 years. These outcomes are relevant because they are closely related to other education outcomes and subsequent labor market outcomes, such as earnings.[29] I am able to use this methodology because, since

[27] Throughout the paper I use the term "more competitive" to define a school with a higher entrance cut-off score than the comparison. By definition, it is always the case here that a school with a higher cut-off score is more-preferred and has higher-achieving peers than a school with a lower cut-off score, see Figure 1.
[28] Peer ability is often used as a proxy of school quality.
[29] Aguilar and Ortega (2006) estimated that the marginal return of completing upper-secondary education on labor income was about 27% in 2002.

1996, the Metropolitan Area of Mexico City has been implementing a centralized merit-based admission mechanism in its upper-secondary public school system. As I explain in more detail later, applicants are offered admission to their preferred school option subject to the school's capacity constraints, with priority given to students with higher scores on a standardized entrance examination.[30] The data available include examination scores, the students' school preferences, and the school assignment resulting from the application of this merit-based admission rule.

My identification strategy exploits arguably exogenous variation in the placement offer, at each oversubscribed school, among applicants of equivalent skills and preferences at the margins of the cut-off for admission. In my research, I adopt a regression-discontinuity design –in combination with instrumental-variables estimation– to estimate the causal impact of both the offer of admission to, and enrollment in, a more competitive school (i.e. higher cut-off score) for applicants just below, and just above, a school's admission threshold. Thus, I compare the graduation outcomes of students closely on either side of an admission threshold, as a measure of the impact of gaining admission to a school with a higher cut-off score and thus, being immersed in a school environment with higher-achieving peers (see Figure 1) and often higher expenditure per student.

I found that, on average, applicants who scored just above the admission threshold, and thereby effectively received an offer to attend a more competitive school, had a 3.1 percentage points lower probability of graduating within 3 years (on time) and a 1.1 percentage points lower probability of graduating within 5 years, relative to applicants who

[30] The school-assignment mechanism is a serial dictatorship with applicants ordered by admission score. It gives the incentives to reveal their true preference ranking.

scored just below the admission threshold, and did not receive the offer of admission to the more competitive school.[31] These effects are therefore estimates of an intent-to-treat effect local to the population at the cut-off. In addition, I use the offer of admission as an instrument for enrollment in the more competitive school, and estimate local treatment-on-the-treated (LATE) effects that are very similar (i.e. -3.2 and -1.2 percentage points, respectively) to their analogous reduced-form estimates. My results are robust to several sensitivity analyses that I describe. Overall, in this particular setting, it seems that it is better in terms of graduation outcomes to be an average-ranked student in a school with lower- or average-achieving peers, than the lowest-ranked student in a more competitive school with higher-achieving peers. However, I acknowledge that in some selective schools the costs in terms of graduation outcomes for the marginal student should be weighed against the benefits in terms of academic preparation for those who eventually graduate.

My main findings only pertain to the marginal applicant offered admission to a more competitive school. It is possible that the effects are different for specific demographic groups, and at different parts of the cut-off score distribution. Thus, in an extension to the baseline analysis, I show that the negative impact of the offer of admission to a more competitive school for the marginal applicant is larger for males, applicants from low socio-economic status families, and in public lower-secondary schools than for females and those from higher-SES families and those who attended private lower-secondary schools. I also find that the magnitude of the effects I detected differ according to the location of a school's cut-off score in the overall admission cut-off distribution. In

[31] These effects represent a 6% and 2% decrease, respectively, among applicants near the eligibility cut-off. The average probability of graduating on time is 47%, while probability of graduating within 5 years is 57%.

particular, the negative causal effects of the offer of admission tend to disappear as one moves to the admissions decisions of the less selective schools. Finally, I show that the magnitude of the estimated effects on graduation outcomes are larger when focusing only on the offer of admission to, and enrollment in, a first-choice school than when considering all the choices.

There are two important caveats to my analysis. First, I cannot generalize my findings to applicants who score far from the admission cut-off score, to not-oversubscribed public schools, or to private schools. Second, given the available data, I cannot effectively disentangle the particular channels that are driving my findings. Peers may not be the only thing that changes with admission to a more competitive school. Applicants who score above the cut-off for a particular school might be exposed to different schooling environments (i.e. peers, teachers, resources, etc.) than those who scored just below.

I have organized the remainder of this paper as follows. In the next section, I begin with a review of the existing literature on the impact of the offer of admission to, and enrollment in, a better quality –or more selective– school on the students' subsequent educational outcomes. I also describe the student-allocation mechanism implemented in the admissions process in Mexico City. I end this section by posing my specific research questions. In the following section, I then detail my research design, describing my dataset and identification strategy. I next present my results and describe the potential threats to validity of my inferences. I conclude with a discussion of the findings and opportunities for further research.

Background and Context

In many contexts, it is not always possible to assign students to their most-preferred school because of capacity constraints. So different student-allocation mechanisms have been devised and implemented to complete the assignment process. The design of the student-assignment mechanism thus becomes fundamental in settings where demand for schools exceeds supply (Abdulkadiroğlu & Sönmez, 2003). Commonly, centralized admission processes allocate students to schools based on their revealed preferences over schools and, in some cases, schools' priority ordering[32] and maximum capacity (Abdulkadiroğlu & Sönmez, 2003; Klijny & Haeringerz, 2007).

School choice has become a topic discussed broadly in education (Abdulkadiroğlu & Sönmez, 2003).[33] The underlying rationale for allowing students to choose their schools is that it is hypothesized that they will benefit from being able to enroll in a school that best matches their interests and abilities. The offer of admission to one's preferred school may then play a role in the student's engagement in academic work, which in turn is associated with the decision to persist, or drop out of, school (Rumberger, 2011). For instance, Hastings et al. (2012) suggest that the opportunity to attend a first-choice school improves intrinsic motivation for students, even before they arrive at their new school, and has subsequent positive effects on test scores.

---

[32] Some school systems give priority based on: attendance to feeder schools, residence in the school's walk zone, and sibling attendance at the school (Abdulkadiroğlu et al., 2006).

[33] The school choice problem is closely related to the college admissions problem introduced by Gale and Shapley (1962). In the college admissions problem students have preferences over schools and vice versa, while in the school choice problem schools are simply objects to be consumed by students (Abdulkadiroglu & Sonmez, 2003).

Researchers have linked differences in school quality and student achievement to ultimately differences in labor market outcomes (Card & Krueger, 1992; Grogger, 1996; Murnane et al., 2000), suggesting that attending a better quality school may have important long-term effects. The empirical difficulties in identifying the causal effect of admission to, and enrollment in, a 'better' school lie in the fact that students may self-select schools based on unobserved characteristics; some of which may also affect their educational outcomes directly (Manning & Pischke, 2006; Imbens & Wooldridge, 2009). For instance, children may attend different schools because they have different preferences or levels of motivation (Jackson, 2010). Despite these selection biases, some researchers have estimated the impact of the offer of admission to, and enrollment in, a better quality or more selective secondary school by relying on plausibly exogenous variation in school admission created by the action of external student-allocation mechanisms.

Prior studies have applied a regression-discontinuity design (RDD) to compare the average outcomes of students just below, and just above, an exogenously-set admissions cut-off, arguing that such students are equivalent in all respects other than falling on one side, or the other, of the cut-off. In this way, Pop-Eleches and Urquiola (2013) exploited the admission cut-offs of almost 2,000 secondary schools in Romania and found that, on average, students just above the cut-off for a school or track with higher-achieving peers scored 0.02 to 0.10 standard deviations higher on the graduation test than those who just missed the cut-off. Using survey data, they found evidence that students who barely made it into more selective schools perceived themselves as weaker than their classmates and felt marginalized. De Hoop (2011) estimated the impact of admission to competitive elite public-secondary schools in Malawi and found that marginal students, who were not assigned to a public school, were more likely to retake the primary-school examination. He also found that marginal students selected into 'high-quality' public secondary schools, defined in terms of peer ability and school characteristics, were less likely to drop out or switch to another school than those assigned to lower quality schools.

In contrast, for the case of Mexico City, Dustan (2010) estimated that the offer of admission to a set of 'elite' upper-secondary schools had a 0.19 standard deviations positive impact on the 12th grade examination for those at the margin of admission. Analyzing the same subsample for Mexico City, de Janvry et al. (2012) found that the marginal applicant admitted to an IPN-affiliated school had on average a 7.7 percentage points lower probability of taking the 12th grade examination.[34] They also found that, students who just passed the IPN's school admission threshold and took the 12th grade examination had, on average, 0.12 standard deviations higher scores than those who did not pass the threshold to enter an elite school.

Although more complex analytically, other studies have sought to estimate the causal effect of enrollment in a first-choice or selective secondary school on student outcomes, rather than focusing simply on the impact of the offer of admission. For instance, using lottery assignment to oversubscribed schools and instrumental-variables estimation, Cullen et al. (2006) found no impact of attending a first-choice public school in Chicago on a student's subsequent test scores. Similarly, Hastings et al. (2006) showed that, in a school district of North Carolina, winning a lottery to attend a first-choice school led to no gain in

[34] IPN stands for National Polytechnic Institute. Although authors interpret "not taking" the national standardized examination as dropout, this is a low-stakes test and there is anecdotal evidence of self-selection of students in certain schools into taking the test. In addition, students attending UNAM-affiliated schools do not take the standardized 12th grade test. UNAM stands for Universidad Nacional Autonoma de Mexico and it is the largest public university in the country.

test scores, on average. However, in contrast to this, in a follow-up study, Hastings et al. (2012) detected gains on average test scores –across all subjects– among lottery winners who attended their first-choice school, compared to those who lost the lottery. Capitalizing on a regression-discontinuity design (RDD), Dobbie and Fryer (2011) found that attending an examination school (i.e. elite public high schools) in New York City had little impact on SAT scores and college enrollment, but increased the rigor of the high-school courses taken. Dobbie and Fryer (2014) found that exposure to higher-achieving and more homogeneous peers has little impact on college enrollment, college graduation, or college quality.

Outside the U.S., Clark (2007) found that, using a RDD, attendance at selective secondary schools in the United Kingdom did not improve test scores four years later for those at the margin of admission, but had positive effects on course-taking and university enrollment. Relying also on a RDD, Jackson (2010) studied the student-allocation mechanism in Trinidad and Tobago based on an algorithm using student preferences and standardized test scores. Jackson found that students attending schools with higher-achieving peers had positive effects on performance in the end-of-secondary education examination, and on earning the prerequisites for admission to tertiary education. Using a similar methodology, Lucas and Mbiti (2014) studied the causal effect of attending an elite government school on student progression and test scores in secondary school. They showed that attending a national school resulted in exposure to a higher quality and more diverse peer group in a better resourced school. However, they found little evidence of positive impacts on test scores and no impact on the probability of timely progression through secondary school for the marginally admitted student. In this tradition, in my own research, I capitalize on the exogenous variation in school placement created by the student-assignment mechanism in Mexico City to estimate the causal effect of both the offer of admission to, and enrollment in, a more competitive school, on the probability of graduating and graduating on time, among those applicants at the admission cut-off.

Previous research has also found evidence that the impacts of the offer of admission to, and attendance at, a selective or better school differ among demographic subgroups.[35] For instance, Hastings et al. (2006) detected positive effects of attendance at a first-choice school among white and high-income students who won the lottery, but no effects for African-Americans and low-income students. Jackson (2010) reported that, at the admission cut-off, the effect for girls of attending a school with higher-achieving peers on examination performance was twice the effect for boys. Clearly, such differences by group may not be universal, but are perhaps unique to the specific settings, systems and students. In this paper, I explore possible heterogeneity in the effects of an offer of admission into, and enrollment in, a more competitive school, by selected demographic characteristics.

A shortcoming of earlier research is that it has focused mostly on the impact of the offer to, or attendance at, a selective school on student test performance but few have investigated the impact on student graduation outcomes for schools of different selectivity levels. Graduation outcomes are relevant because they are closely related to other education outcomes and subsequent labor market outcomes, such as earnings. Although researchers recognize that the effects of school quality can be driven by different factors (such as applicant's motivation and interest, peer composition, teacher attention, school resources,

[35] A "selective school" often refers to an elite school or a school with very high average peer scores. A "better school" is often defined as a school with higher average peers scores and/or higher average resources.

match quality), most studies have focused on peer quality as the main mechanism through which admission to a selective school ultimately impacts students' outcomes (Abdulkadiroğlu et al., 2014). Overall there is no consensus in the literature about the impact on subsequent educational outcomes of attending selective or better schools for the marginal applicant. On the one hand, students could benefit from the interaction with higher-achieving peers (Duflo et al., 2010). On the other hand, the effect of the interacting environment with higher-achieving peers might depend on where he or she falls in the ability distribution and on teacher attention to low-performing students (Cicala et al., 2011). The student-assignment mechanism implemented in Mexico City gives me the opportunity to estimate the impact of the offer of admission to, and enrollment in, a more competitive school on subsequent graduation outcomes of the student, which I argue may be more closely related to labor market outcomes than test scores. This particular setting allows me to estimate the net impact of attending a more competitive school relative to a school with a lower cut-off score. However, given the available data, in this paper, I cannot effectively separate the particular channels that are driving my findings.

The Upper-Secondary Admission Process in Mexico City

The Metropolitan Area of Mexico City[36] has the largest education market in the country and has been implementing a merit-based student-assignment mechanism in the upper-secondary public school system for more than a decade (COMIPEMS, 2008). Before that, teenagers who wanted to attend this education level had to apply to several schools simultaneously and then withdraw from all but the most-preferred school that had accepted

[36] The Metropolitan Area includes the Federal District and 22 municipalities of the State of Mexico.

them. This meant that often certain schools were left with vacant seats while many applicants were not admitted to any of the schools to which they applied. In this inefficient context, applicants from lower-income families were at greater disadvantage because they had fewer financial resources to cover the costs of applying to many different schools (COMIPEMS, 2008).

In 1996, nine upper-secondary school subsystems from the Federal District and the State of Mexico created a commission, known as COMIPEMS,[37] to coordinate institutional efforts and respond to the increasing demand for upper-secondary education (COMIPEMS, 2012). The main task of this commission was to create a centralized competitive admissions process, currently known as the CIEMS.[38] Through the CIEMS process, applicants are assigned to public schools based on a designed and explicit interplay between their score on an entrance examination, a ranked list of up to 20 preferred school options,[39] and the pre-determined capacity constraints of schools.

The student-assignment process proceeds as follows.[40] In any given academic year, the process starts in late January with a public announcement of the CIEMS on the Internet and local newspapers (COMIPEMS, 2012). Interested candidates access the registration materials through their current lower-secondary school, local information centers, or the

---

[37] COMIPEMS are the initials in Spanish for the Metropolitan Commission of Upper-Secondary Public Institutions. In Appendix A I include a list of the subsystems that integrate COMIPEMS.
[38] CIEMS is the acronym for Concurso de Ingreso a la Educación Media Superior, which translates to Upper-Secondary Education Entrance Competition.
[39] Applicants have to rank schools or school/track combinations. In the case of technical and some technological schools applicants must indicate their desired track or area of specialization (for example, accounting). Therefore, an applicant can list multiple tracks from the same school but each track will have a different cut-off score. For simplicity, I use the term "school" to refer to school/track combinations.
[40] See Appendix B.1 for the timeline of the process.

COMIPEMS website.[41] By the beginning of March, applicants must complete all the required documents, including the ranked list of up to 20 school options. Registration has a fixed cost of about $25 USD regardless of the number of schools listed. At the registration center, applicants receive a study guide for the test and an identification voucher necessary for taking the test. The entrance examination takes place at the end of June.

Once the entrance tests have been centrally graded by a computer, COMIPEMS's board meets for the assignment process. Applicants who scored less than 31 points, or have not graduated from lower-secondary education, are excluded from further consideration.[42] During the first round, all qualified applicants are ranked by their examination score, from highest to lowest. Then, a computer algorithm is used to assign ranked applicants to their most-preferred school with open seats, once their turn arrives. That is, seats are allocated down the student ranking: the top scorer is assigned to his first choice, the second-highest scorer then gets his most-preferred choice among schools with open seats, and so on. If several applicants with the same score compete for the last seats at a specific school, a representative from that institution must decide immediately whether he opens new seats to accept all tied applicants, or rejects all of them.[43] The process continues until all applicants are assigned, except for those who only request options with cut-off scores above their own score. Thus, the cut-off score for each school option, each year, is set equal to the entrance examination score of the applicant who fills the last seat. Final school-assignment decisions are published by the end of July.

[41] Materials include registration forms, a calendar, and a manual with the list of schools, their location and specialization fields, if applicable. A complete list of the registration materials is presented in Appendix B.2.
[42] The minimum score restriction was eliminated in 2013. Schools affiliated to UNAM and IPN also require a minimum GPA of 7 out of 10 in lower-secondary education.
[43] The representative only has information regarding the number of tied applicants and their score but not about their personal characteristics.

During the second round, applicants who did not get a seat at one of their listed choices, but meet the requirements, have the opportunity to select a school that did not fill their seats. Finally, applicants must complete the paperwork at their assigned school in order to register. Applicants are only allowed to register at their assigned school.[44]

The merit-based admission system implemented in Mexico City provides exogenous variation in the offer of admission, and ultimately enrollment. Two applicants with the same preferences over schools could be assigned to different schools for having a one-point difference on the entrance examination. This allows the estimation of the causal impact on subsequent graduation outcomes, among applicants immediately on either side of the cut-off score for a specific oversubscribed school. It is such comparisons that are the topic of this paper.

Specific Research Questions

Like Dustan (2010) and de Janvry et al. (2013), I use data from the upper-secondary student-assignment process in Mexico City.[45] While they analyze the impact of the offer of admission to a limited sample of 'elite' schools on the 12th grade test scores and the probability of taking such tests,[46] I focus on the impact on ultimate high-school graduation of both the offer of admission to, and enrollment in, a more competitive school using a

[44] If applicants want to switch to another school they would have to take the exam again the following year.
[45] Estrada and Gignoux (2014) also use COMIPEMS data combined with a 12th grade survey in a subsample of schools to estimate the effect of the offer of admission to an IPN-affiliated (elite) upper secondary school on expected earnings at the end of secondary and higher education.
[46] They consider taking the ENLACE 12th grade test as a proxy for graduating, not necessarily on time. Some bias can arise from students who do not take the exam if they are systematically different from those that do take the low-stakes ENLACE test. Also, UNAM-affiliated schools, listed as the top-choice by about 30 percent of the applicants, do not take the ENLACE test so they are excluded from their sample.

large sample of oversubscribed schools, both elite and non-elite. I exploit the exogenous variation generated by the student-assignment mechanism implemented in Mexico City to estimate the causal effect on the probability of graduating and graduating on time, at the threshold of admission, of the offer of admission to and enrollment in a school with a higher cut-off score and hence, higher-achieving peers. Specifically, I address the following research questions:

(1) Does the offer of admission to a more competitive upper-secondary public school increase the probability that an applicant at the margin of being admitted will graduate within three years (on time) and within five years?

(2) Does enrollment in a more competitive upper-secondary public school increase the probability that the individual at the margin of being admitted will graduate from upper-secondary education within three years (on time) and within five years?

Research Design

Due to the sorting of students into schools it is often challenging to estimate the impact of an offer of admission to, or enrollment in, a chosen upper-secondary school. In this paper, I deal with this self-selection problem by exploiting the way in which individuals are offered admission to public upper-secondary schools in Mexico City; according to the value of an observable "forcing" variable, relative to an exogenously defined cut-off point (Shadish et al., 2002). The source of variation comes from comparing the outcomes of applicants, who have similar characteristics and preferences, but are offered admission to different schools because they scored just below, or just above, the cut-off for admission to a specific school. To address my two research questions concerning the impact of the offer of admission to, and ultimately enrollment in, the chosen school, I use both a "sharp" and a "fuzzy" regression-discontinuity design (RDD), respectively.[47]

The idea behind the RDD is that the probability of being offered admission to a more competitive school, and hence attending a school with higher achieving peers, changes discontinuously at the cut-off score of that particular school. The key identifying mechanism underlying the RDD, which ensures estimates of treatment effect are unbiased and can be interpreted causally, is that the cut-off on the admissions forcing variable must be assigned exogenously and cannot be subject to strategic responses or manipulation by participants (Shadish et al., 2002). Under this continuity assumption, individuals close to the admission threshold, and on either side of it, will be equal in expectation prior to treatment, so that differences in their average outcomes post-treatment can be interpreted causally (Murnane & Willett, 2011). I argue that these conditions prevail in the Mexico City data because applicants do not know their examination score nor the schools' exact cut-offs when submitting their ranked list of preferences five months before the examination. Therefore, I can attribute any discontinuous jumps in the outcomes, on average, at the threshold, to the increase in the probability of being offered admission to a more competitive school. The main caveat of this approach is the localness of the treatment that only yields internally valid estimates for the marginally admitted applicant.

[47] While in the "sharp" RDD the forcing variable perfectly predicts the treatment status (i.e. probability of treatment jumps from 0 to 1), in the "fuzzy" RDD the treatment status is only partially determined by the forcing variable (i.e. the probability of treatment changes by less than 1) (Jacob et al., 2012).

Using data from multiple cohorts and oversubscribed schools, each of which establishes its own cut-off score on the admissions examination yearly, I constructed 'sharp samples' for oversubscribed schools following Abdulkadiroglu et al. (2014).[48] I re-center the forcing variable for each applicant relative to the cut-off score of his respective sharp-sample school, such that the new variable will express the difference between the applicant's examination score and the minimum score required to be offered admission at that specific school. As I mentioned, my data contains many cut-offs, one for each oversubscribed school. Therefore, I follow the approach used in Pop-Eleches and Urquiola (2013) to combine all these discontinuities into a single analysis, pooled across applicants and oversubscribed schools, where the capacity constraint is binding, into one discontinuity with suitable fixed effects for the combination of cohort and application risk sets (i.e. School-School-Cohort).[49] I describe this approach in more detail later.

Data

To address my research question, I draw on a linked dataset that allows me to track the graduation of students who participated in the public upper-secondary assignment process in Mexico City between 2005 and 2009. My analysis incorporates two sources of data.

---

[48] The "sharp sample" for school S is the sample for applicants for whom offers of admission are sharp in the sense of being deterministically linked with S's forcing variable. See Appendix C for a more detailed explanation of the construction of the sharp samples and the structure of the data.

[49] That is, dummies for interaction of cohort and applicant preferences ordering over schools (i.e. school pair combinations: School-School-Cohort). For example, when estimating the effect of enrolling in school A over school B, I pooled applicants whose ranking are A, B, C, with applicants whose ranking is A. B, D but some scored just above the cut-off for A while others scored just below the cut-off for A and got assigned to B. For more detail, see Appendix C.

I use primary data from the set of applicants who registered for the COMIPEMS admission processes between 2005 and 2009. These five cohorts include over 1.5 million applicants and about 2,200 oversubscribed school/track options. I chose to focus on these particular cohorts because they include the students for whom I am able to measure on-time graduation (i.e. graduating within three years). I can also observe graduation within five years for the 2005 to 2007 cohorts. The dataset contains rich information on applicants' background, entrance-test score, ranking of up to 20 school options, assigned school, and responses to a context survey. Background information includes participants' gender, age, zip code, lower-secondary school attended and GPA. In addition, the context survey gathers information about the students' family structure, parental education, durable goods, family income, financial aid, and other related socio-economic and family questions. The data were made available to me through the COMIPEMS technical committee.

I augment the COMIPEMS database with student administrative records from different upper-secondary subsystems, again for the 2005-2009 applicant cohorts. Each upper-secondary subsystem keeps administrative records of students' enrollment and graduation from its schools, thereby providing information on my outcomes of interest, on-time upper-secondary graduation and graduation.[50] I merged the existing databases using the COMIPEMS registration number and/or the national identification number (known as CURP).

[50] One limitation of the data is that I do not observe graduation outcomes for applicants assigned to the State of Mexico's subsystem. This subsystem represents about 20% of the total universe of applicants assigned during the first round.

Analytic Sample

I include, in my analytic sample, applicants who participated in the 2005 to 2009 upper-secondary admission's processes in Mexico City and who, if they had graduated on time would have done so between 2008 and 2012. I limit my sample to applicants who were assigned during the first round of the process to oversubscribed schools that did not require specifying a specialization field.[51] I further limit my sample to individuals that were taking the examination for the first time (i.e. non-repeaters) and who were age 25 or younger when taking the entrance exam.[52]

As I mentioned before, I constructed a 'sharp sample' for each pair of oversubscribed schools in the analytic sample. Each applicant serves as an observation at least for the sharp sample of its assigned school (see Appendix C).[53] Given that each applicant can be in more than one sharp sample, in the analysis I adjusted the standard errors of my parameter estimates to account for the clustering of observations within student.

My analytic sample includes students applying to 305 oversubscribed schools that did not require the explicit selection of a specialization field. In the pooled sharp samples, I have approximately 505,000 observations on 336,000 applicants who were assigned, between 2005 and 2009, during the first round to those oversubscribed schools. For

---

[51] This subsample includes schools from 5 subsystems (i.e. IPN, UNAM, UAEM, DGB, COLBACH), excluding the SE subsystem for which students' graduation records are not available. To assess the potential bias of omitting applicants assigned to the SE subsystem, I used the 12th grade standardized test as a proxy of graduation. Results are not very different from the ones presented here. Also, results for other estimations, including those that consider the rest of the subsystems and each school/track combination as a separate school have a slightly smaller magnitude, but same direction and statistical significance, than the ones for the analytical sample.

[52] I limit my sample to first-time applicants because applicants who decide to retake the test in the following year may have different unobserved characteristics, like motivation, than applicants who did not retake the entrance examination.

[53] For instance, an applicant who ranked school A first but did not qualify there, is also in the sharp sample for school B if B is his second choice and he is assigned there.

comparability reasons across statistical specifications, I drop applicants with missing baseline covariate information and use the remaining sample of students in my discontinuity analysis.[54] Thus, my discontinuity sample consists of about 272,000 observations assumed to fall just below, or just above, their respective admissions threshold.[55]

Descriptive Statistics

In Table 1, I present a set of summary statistics for the four different samples. In column 1, I report the average values of the included variables for all the applicants. About 70 percent of them were assigned during the first round to both oversubscribed and non-oversubscribed schools. In column 2, I display the average values of the same variables for the same applicants in a sample that I have restricted to first-time applicants under the age of 25 who were assigned during the first round of the process and who listed oversubscribed schools. In column 3, I report mean values on the same variables for my analytic sample comprised of applicants with non-missing values for the selected covariates. Finally, in column 4, I present the average values for those same variables for applicants assumed to fall within my ultimate preferred bandwidth of 9 points (that is, those that fall within 9 points either side of their respective cut-off points on the admissions-score forcing variable). I report standard deviations of all variables in parentheses. Note that there are explicit differences in averages between the full sample of upper-secondary applicants and

[54] I acknowledge that this can pose a threat to internal validity if those who did not answer certain questions were systematically different from those that did. A comparison between the unrestricted and the restricted samples showed very little differences in the average baseline characteristics, which leads me to believe that missing values are random. I discuss this further in the Threats to Validity section.
[55] Considering a bandwidth of 9 points obtained through the cross-validation (CV) method.

the other samples. That is expected given that not all applicants are assigned to an oversubscribed school during the first round.

Those that are not offered admission to an oversubscribed school tend to have lower scores and come from more vulnerable families. For instance, the full sample has a greater percentage of low-SES applicants and household heads with fewer years of schooling than the other subsamples. Despite the expected differences with the full sample, the rest of my samples (columns 2 to 4) considering only oversubscribed schools are comparable, in average value, on most of the presented variables.

In my analytic sample, in column 3, I observe that 48 percent of applicants are males, the average age is 15.2 years, only 2 percent are indigenous, the average GPA in lower-secondary education is 8.4 (out of 10), 12 percent attended a private lower-secondary school, and 92 percent of them live in the Metropolitan Area of Mexico City. In addition, I notice that the household head has on average 11 years of schooling (i.e. incomplete upper-secondary), 42 percent come from families below the median income but only 3.3 percent of them work and 1.6 percent receive a conditional cash transfer (i.e. Oportunidades-Progresa). Regarding their school preferences (Panel B), applicants in the analytic sample listed, on average, 9 school options and were assigned to their third option. The average score in the entrance test was 78 points (out of 128). Finally, in Panel C of Table 1, I show the average values for the two outcomes of interest: about 46.7 percent of the applicants graduate on time ultimately, while 57.0 percent of them graduate within five years of starting upper-secondary education.

Measures

For each individual in my analytic sample, I create the outcome variable, GRADONTIME, a dichotomous indicator of whether the applicant graduated from the assigned school within three years (=1; 0 otherwise).[56] I also created the outcome variable, GRAD, a dichotomous indicator of whether the applicants from cohort 2005 to 2007 graduated from the assigned school within five years (=1; 0 otherwise). I will refer to these outcomes as Y.[57] These outcomes have clear policy relevance because they are closely related to other education outcomes and subsequent labor market outcomes. In my regression-discontinuity analyses, my forcing variable (Murnane & Willett, 2011) is the upper-secondary admission-examination score (out of 128), SCORE, that was used to determine assignment each year. I center the forcing variable, SCOREc, with respect to the admissions cut-off of the relevant school in a specific year. That is, applicants' scores on the forcing variable are expressed in terms of their distance from the respective sharp sample cut-off (j) such that the marginal admitted student has a SCOREc of zero.

My question predictors include the dichotomous variables OFFER and ENROLL, for my two research questions respectively. OFFER is a dichotomous variable that records whether an applicant scored above the relevant cut-off score and, therefore, received an admission offer (=1; 0 otherwise). ENROLL is a dichotomous indicator recording whether the applicant enrolled subsequently in his or her assigned school (=1; 0 otherwise).

[56] Applicants can only register at their assigned school; otherwise they would have to participate again in the CIEMS the following year.
[57] There is a positive relationship between the cut-off score for a school and the probability of graduation or on-time graduation. The best fit for this relationship is a quadratic polynomial.

To increase the precision of my estimates, I include selected background covariates that I hypothesize to be related to my outcome (Bloom, 20012; Jacob et al., 2012).[58] They include indicators of gender, age, indigenous language, lower-secondary GPA and working status, as well as indicators of parental education, socio-economic status,[59] living outside the metropolitan area, family is recipient of conditional cash transfer (Progresa-Oportunidades), and version of the examination[60] taken. In my statistical models, I refer to the complete covariates as the vector **X**. Additionally, as I mentioned before, I include dummies for the combination of cohort and application risk sets; that is, school-school-cohort fixed effects, as described in Appendix C.

Empirical Strategy

To address my research questions, I use a standard regression-discontinuity design (RDD) approach, based on the assumption that the relationships between the probability of upper-secondary school graduation, and on-time graduation, with my centered forcing variable are "locally" linear within a bandwidth of the respective cut-off point. I implement a RDD using local linear regression (LLR),[61] in combination with a cross-validation

---

[58] I list and define these covariates in Appendix D.

[59] I created a principal components index of access to durable goods (like car, computer, DVD) and services (like sewerage, telephone, Internet) at home. This shows positive correlation with reported family income. Family income is not used here because not everyone reported it.

[60] Applicants who list a UNAM-affiliated school as their first choice take UNAM's version of the examination, while the rest take CENEVAL's. These examinations are supposed to be "technically equivalent" in content and difficulty, and to have a high degree of reliability, validity, lack of bias, and equity (COMIPEMS, 2012). Empirical analysis, not shown here, suggests that the general results are not sensitive to pooling across examination versions within the same cohort. In the analyses, I include a dummy of the UNAM version of the exam.

[61] In this case, a probit or logit specification yields marginal effects similar to those obtained under a linear-probability model.

bandwidth determined by the methods of Imbens and Lemieux (2008).[62] As a robustness test, I check the sensitivity of my results to different bandwidth choices.

Under my design, applicants may have received an offer of admission depending on their placement relative to the exogenous admission cut-off at a specific school. Applicants, who scored just below the cut-off for a particular school and have the same preference ordering,[63] provide an adequate counterfactual for those who scored just above. Any difference in their outcomes can be attributed to the fact they have access to different schools.

I first address the question of the impact of the offer of admission to a more competitive school (or intent-to-treat) on the probability of graduating on time, or graduating, by fitting the following "reduced-form" regression:

$$P(Y_{ij}=1)=\pi_o+\pi_1 OFFER_{ij}+\pi_2 SCOREc_{ij}+\pi_3 SCOREc\times OFFER_{ij} \quad [1]$$
$$+X_i^{'}\delta+\varphi_j+\varepsilon_{ij}$$

for the i^th^ student in the sharp sample j, where φ_j is the vector of school-school-cohort fixed effects and ε_{ij} is the residual. As I mentioned before, I adjust the standard errors for the clustering of observations within individual as a student may appear in more than one sharp sample. I include the interaction between the treatment and the forcing variable to allow

[62] With a discrete forcing variable it is not possible to compare outcomes in very narrow bins, like the optimal bandwidth IK by Imbens & Kalyanaraman (2012), just to the right and the left of the threshold (Lee & Card, 2008). Based on the cross-validation criterion, I chose a bandwidth of 9. I use the term bandwidth to indicate the width of the analytic window used for estimation.
[63] I define “same preference ordering” as having the same ranking as pertaining to the two choices at issue. For example, in comparing the impact of enrolling in school A or B, I pooled together applicants whose rankings are A,B, C, with applicants whose ranking is A, B, D. For more detail see Appendix C.

the slope to differ at the two sides of the cut-off. In the model, all parameters have their usual interpretations. Note that I have re-centered the forcing variable to express the difference between each applicant's examination score and the minimum score required to be offered admission at a specific school and in a specific year. Because of that, the critical parameter is the slope associated with question predictor OFFER (π_1), referred to as the local intent-to-treat (ITT) effect (Bloom, 2012). This parameter expresses the causal effect of being offered admission to a more competitive school (i.e. a school with higher cut-off score and higher-achieving peers).

Not all the applicants who score above the cut-off enrolled in their assigned school; some decided to re-take the exam the following year. Given the design of the allocation mechanism under study, it is not possible for an applicant to enroll in a school to which he or she was not offered admission. Consequently, the magnitude of the discontinuity at the threshold may be reduced because a proportion of the group offered admission did not register (Bloom, 2012). Thus, to estimate the causal impact of enrollment (i.e. take-up of the admission offer), I use a standard instrumental-variable (IV) approach to model the relationship between on-time graduation (or graduation) and enrollment following a "fuzzy" regression-discontinuity design (Murnane & Willett, 2011; Jacob et al., 2012). In the first stage of the IV analysis, I predict the probability of attending a more competitive school using a student's position relative to his admission cut-off as an instrument in a local-linear probability model, as follows:

$$P(\text{ENROLL}_{ij}=1)=\alpha_o+\alpha_1\text{OFFER}_{ij}+\alpha_2\text{SCOREc}_{ij}+\alpha_3\text{SCOREc}\times\text{OFFER}_{ij} \quad [2a]$$
$$+X_i^{'}\theta_1+\varphi_{1j}+\delta_{ij}$$

for the i^th^ student in the sharp sample j. Where, δ_{ij} is the first-stage residual and φ_j is the vector of school-school-cohort fixed effects. In the second stage of the IV approach, I fit the following regression model:

$$P(Y_{ij} = 1) = \beta_o + \beta_1 \hat{ENROLL}_{ij} + \beta_2 SCOREc_{ij} + \beta_3 SCOREc \times OFFER_{ij} + X_i^{'}\theta_2 + \varphi_{2j} + \varepsilon_{ij} \quad [2b]$$

where $\hat{ENROLL}$ is the predicted value (i.e. the exogenous part) of the first-stage outcome for each applicant. Because of my inclusion of these predicted values in place of the nominal (and potentially endogenous) question predictor, ENROLL, I adjust the second-stage standard errors of the parameter estimates using standard methods (Murnane & Willett, 2011; Wooldridge, 2002).[64] In the model, the parameter of interest is β_1, referred to as the local treatment-on-the-treated (TOT). It represents the causal effect of enrolling in a more competitive school on the probability of graduating, or graduating on time, from upper-secondary education for applicants at the margin of admission (Bloom, 2012).

The selection of bandwidth, or the window around the cut-off, is an important decision in the application of the RDD approach. The closer you get to the cut-off, the more similar the individuals are, but the less power you have; the further away you move from the cut-off point, you gain in power, but you lose on similarity of groups. Nevertheless, a wider bandwidth makes it more difficult to model the functional form of the relationship between the forcing variable and the outcome, as one must consider non-linear functional forms. Although I chose the cross-validation bandwidth of 9 points as my preferred

[64] In practice I use the ivregress 2sls and xtivreg commands available in the Stata software, which yields the adjusted second-stage standard errors.

bandwidth, as a robustness test I check the sensitivity of my results to different bandwidth choices.

Results

In this section I present my empirical results. I begin by examining the soundness of the proposed identification strategy. First, I show that the discontinuity in the offer of admission is not associated with discontinuous changes in baseline covariates at the threshold for admission. I also show that within defined choice sets applicants cannot manipulate their position with respect to the cut-off score. Then, I turn to the main results. I present the results from the “sharp” and “fuzzy” regression-discontinuity designs, pooling together the cut-offs of oversubscribed schools under analysis. Finally, I re-estimate the effects of the offer of admission for selected demographic groups and at different parts of the cut-off score distribution.

Internal Validity

The rationale behind the regression-discontinuity design is that, if the offer of admission to a more competitive school changes discontinuously at the threshold, then the causal effect of the offer of admission can be identified at the cut-off even when applicants' scores are related systematically to factors that affect graduation and on-time graduation. The main identification assumption of this methodology is that individuals cannot manipulate their position with respect to the cut-off point, such that the cut-off point should work as a lottery for those close to it. Individuals immediately on either side of the cut-off are, on average, equivalent on observed and unobserved characteristics and differ only in position relative to the threshold. To confirm that the regression-discontinuity design provides internally valid estimates, I verify that two key conditions are met. First, a valid regression-discontinuity design requires that the determination of school-specific thresholds are made independently of the examination scores obtained by applicants (Bloom, 2012). In this sense, a potential threat arises when applicants are able to manipulate their position relative to the cut-off score or when schools choose cut-offs so they can admit specific applicants. Second, the RDD methodology requires that the distribution of baseline characteristics is the same for applicants whose scores place them just on one, or the other, side of the cut-off (Bloom, 2012; Murnane & Willett, 2011). Thus, an important potential threat to validity may be that there are statistically significant differences in baseline characteristics between applicants just below, and just above, the cut-off score.

First, I argue that the cut-off score for each school is unknown a priori by both applicants and schools, leaving little room for manipulation. Applicants do not know their examination score when ranking their options, while the cut-off scores are not determined until the allocation algorithm is run and ties are broken.[65] Examinations include only multiple-choice questions that are marked centrally by a computer, limiting the possibility of manipulating the scores. I inspect the density of the forcing variable graphically in order to rule out a discontinuity in the number of observations around the admission cut-off.[66] In Figure 1, I display a histogram of the applicants at each admission score on the re-centered

[65] Decisions regarding tied applicants are based solely on scores and the number of additional seats each school could open. The school administrators have no information on the characteristics of tied applicants and have to decide whether they accept or reject all tied applicants.
[66] McCrary (2008) proposed to test directly for any manipulation of the forcing variable. This works well when the running variable is continuously distributed, but is inconsistent when the running variable is discrete (Lee & Card, 2007; Frandsen, 2014).

forcing variable, for all schools in my sample. The density is relatively smooth in the neighborhood of the cut-off. In the presence of manipulation, I would have expected to see a big lump just above the cut-off. Therefore, endogenous sorting, or manipulation, does not seem to be a major concern.

As an additional test, I look for potential discontinuities in several of the background characteristics. I examine the distribution of selected covariates to reveal that no major discontinuities existed at the cut-off that might contribute to my results. I re-fit the hypothesized RDD model [1] for alternative bandwidths and without additional controls, while treating each of the baseline characteristics as a dependent variable. In Table 2, I present the different estimates for the slope parameter associated with the offer of admission for each covariate and for different choices of bandwidth. Each row in Table 2 represents the estimate of a different baseline covariate regressed on the offer of admission for different choices of bandwidth. In these analyses, the parameter of interest is the estimate associated with the offer of admission. In most of the cases I fail to reject the null hypothesis that the parameter associated with the offer of admission is equal to zero, in the population, for the selected covariates, except for schooling of the household head.[67] To further test the assumption of baseline equivalence, I regress the offer of admission on a set of applicants' characteristics. Across all window widths, I fail to reject null hypothesis that the covariates jointly explain the variation in the offer of admission. Based on these results, I claim that applicants around the threshold are, on average, equal in expectation

[67] I attempt to address this issue by controlling for observable characteristics of the applicants.

such that any differences in the outcomes of interest result from discontinuities in the offer of admission.[68]

Causal Effect of the Offer of Admission on Graduation Outcomes

In Figure 2, I provide a graphical representation of the bivariate relationship between the forcing variable, SCOREc, and my outcomes of interest. In this graph, I plot the sample mean of outcome against the re-centered forcing variable. By visual inspection, it appears that the offer of admission to a more competitive school has led to a discrete decrease on the probability of graduating and graduating on time. On average, applicants just above the cut-off (i.e. offered admission) appear to have lower probability of timely graduation than applicants just below the cut-off (Figure 2a). The effect on graduation within 5 years is also negative but slightly smaller (Figure 2b). The difference between these plots suggest that some students are held back 1 or 2 grades but eventually graduate.

I confirm the previous trend by fitting a local-linear probability model [1]. In Table 3, I present the intent-to-treat (or reduced-form) estimates of the parameter of interest, under alternative model specifications. As a robustness test, I check the sensitivity of results to a variety of window widths, in addition to the CV bandwidth (9 points), with each of the columns containing the results of fitting the same model in a slightly different window width. Going from left to right, the columns present models fitted in progressively wider windows, starting with 3 points in column 1 and ending with 15 points in column 5.

[68] The lack of a statistically significant break for the covariates, around the admission threshold, is similar to the finding of no difference in mean values in a randomized control trial (Lee and Lemieux, 2010).

I observe little fluctuation in the impact of OFFER. This suggests that my results are, in fact, robust to the choice of bandwidth.

The estimates in Panels A and B of Table 3 represent the causal effect, for applicants at the margin of admission (within the bandwidth noted in the column heading), of receiving an offer of admission to a more competitive school on the subsequent probability of graduating on time from upper-secondary education, controlling for baseline characteristics and school-school-cohort fixed effects. Similarly, estimates in Panel C and D correspond to the effect, for the marginal applicant, of the offer of admission to a more competitive school on upper-secondary graduation within 5 years.

My preferred specification uses a CV bandwidth of 9 points (column 3 in Table 3), controls for baseline characteristics and school-school-cohort fixed effects (Panels B and D). It also includes the two-way interaction between the treatment and the forcing variable (OFFER×SCOREc). I interpret the parameter estimates as indicating that, on average, applicants at the margin of admission to a more competitive school are 3.1 percentage points less likely to graduate on time and 1.1 percentage points less likely to graduate within five years, relative to similar applicants who just did not receive an offer to that particular school in the same year. For the analytic sample, these represent a 6.5 percent decrease below the mean on-time graduation rate of 46 percent, as well as a 2 percent of the mean on-time graduation rate of 57 percent. The estimates are statistically significant at conventional levels (p-value<0.01) and are relatively stable across different bandwidths.

In the previous analyses, I assumed a linear relationship between my outcomes of interest and my forcing variable. However, given the sensitivity of RD estimates to non-linear specifications, Imbens & Lemieux (2008) suggest choosing a bandwidth within which the relationship between the outcome and the forcing variable is locally linear. As I illustrate in Figure 3, the relationship between the outcomes and the forcing variable seems linear within the selected bandwidth. To assess the sensitivity of my findings to the linearity assumption, however, in my reduced-form regression I include polynomial specifications of SCOREc and two-way interactions between each polynomial term and OFFER. In Table 4, I report the point estimates for the different non-linear specifications (up to quartic) within my selected 9 points bandwidth, controlling for baseline covariates and school-school-cohort fixed effects. Although my estimates fluctuate a little, I find that the direction and statistical significance of my findings are preserved across the different specifications. In addition, the terms that describe the polynomial specifications of the relationship are not statistically significant suggesting that within the selected bandwidth the relationship is locally linear.

Causal Effect of Enrollment on Graduation Outcomes

Since not all applicants enroll in their assigned school, the jump in the probability of enrollment around the cut-off is smaller than unity. To account for this, I follow an instrumental-variables approach and use the offer of admission to instrument for actual enrollment in a more competitive school. I argue that this variable meets the two conditions needed for valid instrument: exogeneity and relevance.

First, although I cannot formally test the exogeneity condition in my model,[69] throughout this paper I have mentioned that this particular student-allocation mechanism

[69] To test this condition, it is necessary to have an over-identified model. That is, a specification with a larger number of instruments than endogenous regressors.

creates an exogenous variation in the placement offer at an oversubscribed school among applicants of equivalent skills at the cut-off for admission. That is, the variation in enrollment captured by the offer of admission at the threshold is exogenous.

Second, I argue that the offer of admission is related to the variation in enrollment in a more competitive school. In Figure 4, I illustrate this discrete discontinuity on the probability of enrollment at the cut-off as a function of the applicant's admission score. I indicate the cut-off score for receiving the offer of admission (SCOREc$\geq$0) with a dashed vertical line at SCOREc=0. Given the design and enforcement of the student-allocation mechanism under analysis, it is not surprising that enrollment to the left of the admission cut-off score is 0. Although there is a substantial jump to the right of the cut-off score, the proportion of the applicants who enroll at their assigned school does not rise to one. On average, 95% of the applicants offered a seat take the offer, which indicates a very high compliance rate.[70] This is why in this particular case the intent-to-treat (ITT) are so close to the treatment-on-the-treated (TOT) estimates.

My first-stage estimates confirm that receiving an offer of admission results in a substantial jump in the probability of enrollment at the cut-off for admission. In Table 5, I report the first-stage estimates for the difference in the average probability of enrollment in a more competitive school between those who scored just below and just above the admission threshold, for alternative choices of bandwidth. Going from left to right, the columns present models in progressively wider bandwidths, starting with 3 points in column 1 and ending with 15 points in column 5. The coefficient on the parameter of

[70] Take-up rates vary by school subsystem. The opportunity cost of not registering is high because applicants who do not enroll in their assigned school cannot register to attend another public school and need to re-take the entrance test the following year.

interest, OFFER, is statistically significant at conventional levels for the different choices of bandwidth. In addition, the large F-statistic (>50) associated with this parameter suggests that the offer of admission is a strong instrument for predicting the probability of enrollment of an applicant even within the narrowest bandwidth. The magnitude and direction of my estimates at the cut-off are comparable across the different choices of bandwidth.

My instrumental-variable estimates of the causal impact of enrolling in a more competitive school also indicate a negative effect on graduation and on-time graduation for applicants that were just offered admission relative to those who just missed the admission cut-off. In Table 6, I present the local treatment-on-the-treated (TOT) estimates for different choices of bandwidth. The parameters of interest in my preferred specification, with bandwidth of 9, covariates, and school-school-cohort fixed effects (Panels B and D), have negative and statistically significant coefficients for on-time graduation and graduation. They indicate that, on average, an applicant who enrolled in a more competitive school is 3.2 percentage points less likely to graduate on time, and 1.2 percentage points less likely to graduate from upper-secondary education than a similar applicant student who scored just below the admission threshold. Given the high compliance rates, IV estimates are only slightly larger than their analogous reduced-form estimates.

Heterogeneous Effects

As discussed in the literature review, some previous studies have found evidence of differential effects of admission to a selective school, by students' characteristics such as gender, race and income level. In this section, I explore possible interactions between the parameter of interest and selected demographic characteristics. I also examine whether the effects differ according to where the schools' cut-offs are located in the admission cut-off score distribution. Finally, I investigate whether the effects of the offer of admission to, and enrolment in, a more competitive school are different for those at the margin of admission to their top-choice school (i.e. their most-preferred option).

In Table 7, I explore the effects of possible interactions between the predictor of interest and selected characteristics of applicants (CHARACT), such as: gender, socio-economic status, working status, type of lower-secondary school attended and lower-secondary GPA.[71] For the sake of space and simplicity, in this table, I only present the reduced-form estimates for observations within my preferred bandwidth (i.e. 9 points). I fitted the following general "reduced-form" regression:

$$P(Y_{ij}=1)=\pi_o+\pi_1 OFFER_{ij}+\pi_2 SCOREc_{ij}+\pi_3 SCOREc\times OFFER_{ij} \quad [3]$$
$$+\pi_4 CHARACT\times OFFER_{ij}+X_i^{'}\delta+\varphi_j+\varepsilon_{ij}$$

In column 1, I interact OFFER with MALE and find a negative relationship. On average, males that scored just above the cut-off to a specific school have an estimated probability of graduating on time that is 1.8 percentage points less, and have an estimated probability of graduating within five years that is 1.6 percentage points less, than similar females that just cleared the threshold for admission. That is, while the effect for both males and females is different from zero, I find larger negative effects for males than for females of being a low-scoring student in a more competitive school. In column 2, I examine

[71] I do not find a statistically significant interaction with other covariates such as indigenous, working status, and receiving Oportunidades.

whether the effect differs for applicants who work for a wage compared to those that do not work. Though I estimate a negative relationship, it is not statistically significant for either of the graduation outcomes. In column 3, I interact OFFER with a binary indicator of low socio-economic status and find a negative relationship. Thus, I estimate that marginal applicants who come from low-SES households are at greater risk of falling behind and not graduating than similar applicants who scored just above the cut-off but come from high-SES families.

Regarding the academic characteristics of applicants, I find that marginal applicants, who attended a private lower-secondary school, have a higher estimated probability of graduating and graduating on time (column 4). Finally, in column 5, I display the results of interacting OFFER with the GPA. I find that for those at the margin of admission, a one-point increase in GPA (out of 10) is associated with an increase of 1 percentage points in the estimated probability of graduating on time and 0.9 percentage points in the probability of graduating within 5 years. This indicates that high- and low-ability students close to the threshold of a specific school are affected differently. All these interactions, except for the one with working status, were statistically significant at conventional levels (p-value<0.05). They confirm that, at the margin, the effect of the offer of admission on timely graduation, and graduation, differs according to certain demographic and academic characteristics of applicants. Taken together, these results suggest that the negative effect of scoring just above the threshold, and receiving an offer of admission to a more competitive school, is larger for the marginal applicants who come from more vulnerable contexts (i.e. low-SES, public lower-secondary school, lower GPA).

The results that I have present so far pool together students applying for admission to all oversubscribed schools. Given that the effects may differ depending on the selectivity of the school, I next examine how the impact of the offer of admission on timely graduation differs according to where the school cut-offs are located in the admission cut-off distribution. I subdivide the sharp samples according to their quartile in the admission cut-off distribution, with the most selective school being included in the top interquartile.[72] Given this structure, the higher the admission cut-off score, the more selective a school is (and the higher the average peer achievement is).

In Table 8, I present the results for school choices in each of the interquartiles. I showed that the magnitude and statistical significance of the effect of the offer of admission on timely graduation and graduation gets larger as one moves to a more selective school (i.e. a school with a higher cut-off score). Among schools in the top quartile, I estimate that the offer of admission to a more competitive school reduces the probability of graduating and graduating on time, for the marginal applicant, by 4.8 and 7.5 percentage points respectively. In contrast, I do not find any effect of the offer of admission for those at the margin of admission to a school in the two bottom quartiles.[73] When focusing only in the schools in the top 5 percent of the cut-off score distribution (not shown here), I detect a negative effect of the offer of admission of about 7.3 percentage points on graduation and 10.1 percentage points on timely graduation. Therefore, it appears that most of the impact

---

[72] The interquartiles correspond to having a cut-off score above 82 (q4), a cut-off score between 72 and 82 (q3), a cut-off score between 61 and 72 (q2), or a cut-off below 61 (q1).

[73] When focusing on admission to IPN-affiliated schools, regarded as elite schools, de Janvry et al. (2013) found an increase of 7.7 percentage points in risk of dropout, proxied by the probability of taking the 12th grade standardized test.

of the offer of admission on graduation and graduation on time is driven by the most selective schools (i.e. in the top quartile of the admission cut-off distribution).

Finally, I explore whether the effects of the offer of admission to, and enrollment in, a more competitive school might be different for those at the margin of admission to their top-choice school (i.e. their most-preferred option). I restrict the analysis to applicants who scored just below and just above their top-choice school. On average, 43 percent of the applicants in the analytic sample were offered admission to their most-preferred school. In Table 9, I present the results analogous to those in Table 2 but for focusing on the offer of admission to one's first-, second- or third-choice. I observe that applicants who scored just above the cut-off and were offered admission to their top-choice school have, on average, an estimated probability of graduating on time that is 5.9 percentage points less, and have an estimated probability of graduating within five years that is 3.6 percentage points less, than similar applicants who scored just below the threshold for that same school. Thus, effects for the top-choice school were particularly large. Effects on timely graduation and graduation are smaller, but still negative, for the offer of admission to one's second-choice school. I did not find statistically significant effects on graduation outcomes when estimating the impact of being just admitted or denied admission to one's third-choice school.

Threats to Validity

In this section, I discuss the main threats to internal and external validity of the inferences that can be drawn from my results.

The regression-discontinuity methodology has strong internal validity for inferences at the cut-off point for admission, providing its assumptions are met. Under this type of design, internal validity can be threatened if: (1) applicants can manipulate their position relative to the cut-off score or if schools choose cut-offs to admit particular applicants; and (2) there are statistically significant differences in baseline characteristics between applicants just below, and just above, the cut-off score. I have already shown evidence and argued that neither of them seems to pose a major threat in this particular setting. In addition, I must note that the process of student allocation implemented in Mexico City does not allow for crossovers of students between schools, which could bias the results. Applicants who were assigned during the first round cannot easily transfer to a different school that has no available seats. If they want to transfer to a school that still has seats open, they must meet the minimum cut-off score determined during the first round and go through a very time-consuming bureaucratic process. During the years under analysis, the only way to transfer to a different school was to re-take the entrance examination the following year.

Since not everyone answered all the questions of the background survey, one potential threat to internal validity of my findings concerns the potential for selection bias arising from those who did not answer them all. Comparing the unrestricted and the restricted samples I find very little difference in the average baseline characteristics, which leads me to believe that missing values are random.[74] When analyzing the balance of the covariates for applicants just below and just above the cut-off for admission, I found that

[74] Tables available upon request.

there are no systematic differences between the two groups for alternative choices of bandwidth.

As in any regression-discontinuity design, the limited external validity is the main limitation of the findings presented in this study. Despite the internal validity of my estimates, I was only able to make causal claims for individuals at the margins of being offered a seat at an oversubscribed upper-secondary public school in Mexico City that did not require specifying a specialization field in the academic years under analysis. Given my sample restrictions, my inferences apply to applicants who took the entrance examination for the first time and were less than 25 years old when taking the examination. This implies that I cannot extrapolate my findings on how the offer to, and enrollment in, more competitive schools impacts subsequent upper-secondary graduation for students with very high, or very low, admission scores or who have taken the examination multiple times. In sum, unless stronger assumptions are imposed, my findings cannot then be generalized to applicants who score far from the admission cut-off, or to not-oversubscribed public schools, or private schools, or different places or years.

Another limitation to the external validity of my findings comes from the fact that I have no administrative records of enrollment and graduation for students attending schools from one particular subsystem in the State of Mexico. In contrast to other large subsystems, this subsystem has many small schools spread along the border of the Metropolitan Area. Many of those schools are not oversubscribed and tend to have low to medium cut-off scores. When creating the sharp samples, I omitted applicants that were offered admission to a school from this subsystem. For those assigned to any other subsystem under analysis, I considered their next most-preferred option to which they qualified and that was not from the subsystem with missing data. To assess the potential bias, I re-estimated the reduced-form model using the year a student took the 12th grade standardized examination as a proxy for graduation (low-stakes attached to the test).[75] Reduced-form estimates from the regression, which uses the year of taking the 12th grade examination as a proxy for graduation and on-time graduation among applicants from the subsystem with missing data, are slightly smaller (-0.7 and -2.2 percentage points, respectively) than those for the sample with complete administrative data, but they have the same direction and statistical significance.

Discussion

My results indicate that applicants who barely passed the cut-off to a more competitive school (i.e. higher cut-off score and higher achieving peers) in Mexico City are less likely to graduate and to graduate on time than similar applicants who just missed the cut-off admission and were offered admission to, and enrolled in, a less competitive school. Taken together, my findings suggest that the most vulnerable students (i.e. males, students from low-SES families, those enrolled in public lower-secondary schools, and those with relatively low GPAs) who just make it into selective or elite schools (i.e. top quartile of the cut-off distribution), are at greater risk of not graduating. Overall, in this particular setting it seems that it is better, in terms of graduation outcomes, to be an

[75] Taking the exam does not necessarily imply a student graduates, just that he or she is enrolled in 12th grade. There is anecdotal evidence of self-selection into the test. That is, some schools might ask their lower-performing students to not attend the day of the test to inflate school mean score. This might bias the estimates downward. Students attending UNAM-affiliated schools do not take the test.

average-ranking student in a school with lower- or average-achieving peers than the lowest ranking student in a more competitive school with higher-achieving peers.

As I mentioned before, I am only able to detect the net effect of the offer of admission to, and enrollment in, a school with a higher cut-off score and hence, higher achieving peers. However, peer composition may not be the only component of the educational production function that changes at the cut-off (Dobbie & Fryer, 2014). Other observable (i.e. resources, teacher quality) and unobservable (i.e. curriculum, academic rigor, parental involvement) characteristics of the schools may differ.

There is no consensus in the literature regarding the role of peers for the marginal students. On the one hand, peers and social interactions are important for the formation of skills and human capital (Sacerdote, 2011). On the other hand, being among the lowest-ranking students within a school, relative to the other peers, could have a negative effect (Cicala et al., 2011). First, marginal students are likely to have a lower class rank than they otherwise would have, so their confidence and motivation could be reduced by being at the bottom of the class (Barros & Salcedo, 2009). Thus, students at the bottom of the ability distribution at a given school will then be more likely to fail courses, fall behind, and eventually drop out of school. Second, teachers may devote less time and attention to the least academically skilled students in the classroom. If teachers teach to the mean of the class, students at the bottom of the ability distribution in a specific school will be exposed to more academic rigor than they would otherwise face. That is, instruction may be far above the level the marginal student can understand.

In a similar case study to the one presented here, Pop-Eleches and Urquiola (2013) studied behavioral responses of students who participated in the Romanian centralized admission process and found that those who scored just above the cut-off to a ‘better’ school perceived themselves, on average, as weaker relative to their peers and felt marginalized. For the case of college students in Colombia, Saavedra (2009) noted that marginal applicants struggled to get through graduation relative to rejected applicants who attended less selective schools.

Below, I describe a potential mechanism that may undergird my results and discuss some policy implications of my findings.

Reference-group Perspective

As I mentioned before, the underlying rationale for allowing students to choose their schools is that they will benefit from being able to enroll in a school that best matches their interests and abilities. My results are consistent with the model proposed by Cicala et al. (2011) which suggests that the effect of the interacting environment with higher-achieving peers might depend on where a student falls in the ability distribution. That is, marginal applicants, who just passed the cut-off to gain admission to their top-choice school, are usually the ones at the bottom of the ability distribution and the ones that tend to have a harder time keeping up with the rest of the class. On average, marginal students have a lower class rank than they would have had in another school and might be less competitive in subsequent outcomes “even if their absolute level of achievement is unchanged” (Dobbie & Fryer, 2014, p.59).

This explanation is also consistent with sociology models showing that if relative ability matters, then admission into selective schools could adversely impact children who scored close to the cut-off. It relates to the sociology literature regarding relative deprivation and social comparison dynamics, such as the frog-pond model (Davis, 1966) or the big-fish-little-pond effect (Marsh & Parker, 1984; Marsh 1987). Davis (1966) applied the reference-group perspective to the study of college students and examined the relationships between academic achievement, GPA and career choice. He argued that an undergraduate's career aspirations were based in part on his assessment of his academic performance relative to that of his peers, concluding with the saying: “It is better to be a big frog in a small pond than a small frog in a big pond.” Similarly, the big-fish-little-pond effect (Marsh 1987) proposed that a student will have a lower academic self-concept and lower educational occupational aspirations in an academically selective school than in a non-selective school. Overall, this strand of literature points out that students form self-concepts about their academic abilities by a social comparison process, particularly with classmates. This might provide some insight on a potential mechanism driving the effects I found.

Policy Considerations

Although a merit-based student mechanism in upper-secondary education in Mexico City may be an efficient way to assign students to schools, in terms of costs and logistics, the negative effects of admission to a more competitive school on graduation outcomes for the marginal student calls into question the strategies implemented to keep on track and retain the lowest-ranking students within each school. Most of the current programs aimed at decreasing upper-secondary dropout in Mexico are designed to provide financial aid or economic incentives directly to students (like Oportunidades, Probems, and Prepa Si). Few and isolated efforts are focused on the supply side, such as providing counselling, tutoring and remedial education for those students who are falling behind. For instance, a school's resources that are spent on students who take longer to graduate may be better used in early programs to prevent grade retention. Recognizing that the most vulnerable students at the bottom of the ability distribution for a given school, particularly among the most selective schools, are at greater risk of falling behind and not graduating is an initial step to design better programs that target and support these groups.

Marginal students to a more competitive school can be divided into three groups: those who graduate on time, those who take more than 3 years to graduate but eventually do, and those who never graduate. Findings from earlier research suggest that there may be positive gains on achievement and subsequent educational or labor outcomes from admission to a more selective school (Dobbie & Fryer, 2014; Estrada & Gignoux, 2014; de Janvry et al., 2013; Jackson, 2010). In some cases, these gains may compensate for the negative effects of not graduating on time (i.e. having been held back) from a selective or elite school since, in the end, these students successfully receive a regular upper-secondary diploma from a selective school. However, the trade-off is not so clear for those who attend non-selective schools or those who attend a selective school but never graduate. In terms of the labor market, the most pervasive effects may be for the marginal applicants who, despite having high entrance scores and attending a selective school, never graduate from upper-secondary education. These school dropouts are the ones who in the long-term would be more affected, in terms of labor income, from the negative effect of scoring just above the admission threshold to a more competitive school and not graduating.

My findings show that there are some students with relatively high admission scores, and who are barely admitted into selective schools, that are facing difficulties graduating at all. Thus, the upper-secondary education system is losing high-ability students that could have probably graduated from a less competitive school. It would be desirable to design programs to prevent high-ability students from dropping out and which facilitates transfer to other schools without requiring them to re-take the entrance examination and repeat all the courses. In settings with imperfect information like this, allowing students to make post-enrollment transfers in order to complete their studies could alleviate the risks of receiving poor school matches and reduce the probability of dropping out of the system. Recent efforts in Mexico have focused on designing protocols to facilitate the transfer of academic credits between institutions, but there are still many challenges for those who want to switch to another school.

Conclusion

In this paper, I explore whether it matters to applicants' subsequent educational outcomes to which school they are offered admission to, and they enroll in. This is a difficult question to answer mainly because students self-select into schools of different characteristics. I address this obstacle by exploiting the student-allocation mechanism implemented in upper-secondary education in Mexico City. This mechanism creates a series of regression-discontinuity quasi-experiments that allow me to examine the average effects, for the marginal applicant, of the offer of admission, and enrollment in, a more competitive school. It also enables me to explore the heterogeneity in effects for different subgroups and at different points of the admission-score distribution.

I find that being offered admission to a more-preferred and higher-achieving upper-secondary school in Mexico City decreases the probability of graduating on time and of graduating within five years for individuals at the margins of being offered a seat at an upper-secondary public school in Mexico City, in the designated academic years. This suggests that applicants, who were otherwise equal in expectation at the end of 9th grade, had statistically-different graduation outcomes in upper-secondary education as a result of admission to different schools. I show that the estimates for the impact of enrollment are very similar given the high compliance rate. In addition, my results indicate that the negative effect of scoring just above the threshold is larger for the marginal applicants who come from more vulnerable contexts (i.e. low-SES, public lower-secondary school, lower GPA) and are offered admission to a selective school (i.e. top quartile of the school cut-off distribution). The negative impact of being among the worst-performing students relative to the other peers in the class could operate through different channels. However, at this point I do not have enough information to explore the channels that are driving these effects. Further research is needed in this area.

While the magnitude and direction of the effects presented here may be particular to the case of Mexico City, additional work is needed to examine whether these negative effects exist in settings with different student allocation systems and educational institutions. Future work in this particular context might also undertake an examination of whether admission to a more competitive school has effects on longer-term outcomes, such as college attendance and labor income.

References

Abdulkadiroglu, A., J. Angrist, and P. Pathak (2011). "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools." Econometrica 84(1): 137-196

Abdulkadiroglu, A. and T. Sönmez (2003). “School choice: A mechanism design approach,” American Economic Review (93): 729–747.

Barros, R. and A. Salcedo (2009). “The Effect of College Quality on Academic Performance: Evidence from Mexico,” Unpublished Working Paper, Stanford University.

Bloom, H. S. 2012. “Modern Regression Discontinuity Analysis,” Journal of Research on Educational Effectiveness 5 (1): 43-82.

Cameron, C. and P. Trivedi (2005). Supplement to Microeconometrics: Methods and Applications. New York, NY: Cambridge University Press.

Clark, D. (2007). “Selective Schools and Academic Achievement," IZA Discussion Paper No. 3182.

COMIPEMS (2012). “Informe del Concurso de Ingreso a la Educación Media Superior de la Zona metropolitana de la Ciudad de México (1996-2010),” Ceneval Publication.

Cullen, J., B. Jacob, and S. Levitt. (2006). “The Effect of School Choice on Student Outcomes: Evidence from Randomized Lotteries,” Econometrica, 74(5):1191-1230.

de Hoop, J. (2010), “Selective Secondary Education and School Participation in Sub-Saharan Africa: Evidence from Malawi,” Tinbergen Institute Discussion Papers 10-041/2, Tinbergen Institute, April.

de Janvry, A., A. Dustan and E. Sadoulet (2012). “The Benefits and Hazards of Elite High School Admission: Academic Opportunity and Dropout Risk in Mexico City,” Unpublished Working Paper.

DOF (2012). “Decreto por el que se declara reformado el párrafo primero; el inciso c) de la fracción II y la fracción V del artículo 3o., y la fracción I del artículo 31 de la Constitución Política de los Estados Unidos Mexicanos (Obligatoriedad de la Educación Media Superior),” Diario Oficial de la Federación (February 9th, 2012).

Dobbie, W. and R. Fryer (2011). "Exam High Schools and Academic Achievement: Evidence from New York City," NBER Working Papers 17286, National Bureau of Economic Research.

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

Dustan, A. (2010). “Have Elite Schools Earned their Reputation?: High School Quality and Student Tracking in Mexico City," Unpublished Working Paper, Berkeley.

Estrada, R. and J. Gignoux (2014). "Benefits to elite schools and the formation of expected returns to education: Evidence from Mexico City," CEPREMAP Working Papers 1407, CEPREMAP.

Frandsen, B. (2014). “Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design When the Running Variable is Discrete,” Unpublished Working Paper.

Gale, D. and S. Shaphley (1962). “College Admissions and Stability of Marriage,” The American Mathematical Monthly, 69(1): 6-15.

Hastings, J., T. Kane, and D. Staiger (2006). “Preferences and Heterogeneous Treatment Effects in a Public School Choice Lottery,” NBER Working Papers No. 12145, National Bureau of Economic Research.

Hastings, J. and J. Weinstein (2008). "Information, School Choice, and Academic Achievement: Evidence from Two Experiments," The Quarterly Journal of Economics, 123(4): 1373-1414

Hastings, J., C. Neilson, and S. Zimmerman (2012). “The Effect of School Choice on Intrinsic Motivation and Academic Outcomes,” NBER Working Paper No. 18324, National Bureau of Economic Research.

Imbens, G. and Lemieux, T. (2008). "Regression discontinuity designs: A guide to practice," Journal of Econometrics 142(2): 615-635.

Imbens, G and K. Kalyanaraman (2012). "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," Review of Economic Studies 79(3): 933-959.

Jackson, K. (2010). "Do Students Benefit From Attending Better Schools?: Evidence From Rule-based Student Assignments in Trinidad and Tobago," The Economic Journal, 120(549): 1399–1429.

Jacob, R., P. Zhu, M. Somers, H. Bloom (2012). “A Practical Guide to Regression Discontinuity,” MDRC Working Paper, MDRC.

Haeringer, G. & F. Klijn (2008). "Constrained School Choice," Working Papers 294, Barcelona Graduate School of Economic.

Lucas A. and I. Mbiti (2012). "The Determinants and Consequences of School Choice Errors in Kenya," American Economic Review 102(3): 283-88.

McCrary, J. (2008). “Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test.” Journal of Econometrics, 142(2): 698-714.

Manning, A. and Pischke, J. (2006). "Comprehensive versus Selective Schooling in England in Wales: What Do We Know?" NBER Working Papers12176, National Bureau of Economic Research.

Murnane, R. and J. Willett (2011). Methods matter: Improving causal inference in educational and social science research. New York, NY: Oxford University Press.

Pathak, P. and J. Sethuraman (2011). “Lotteries in student assignment: An equivalence result,” Theoretical Economics 6: 1–17.

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

Saavedra, J. (2009). “The Returns to College Quality: Evidence from Colombia,” unpublished Working Paper, Harvard University.

Shadish, W., T. Cook and D. Campbell (2002). Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Boston: Houghton Mifflin Company.

Figure 1. Scatterplot of the bivariate relationship between school's cut-off score (School cut-off score) and the average score of applicants offered admission to that particular school (Peer mean score). Restricted to oversubscribed schools in the analytic sample.

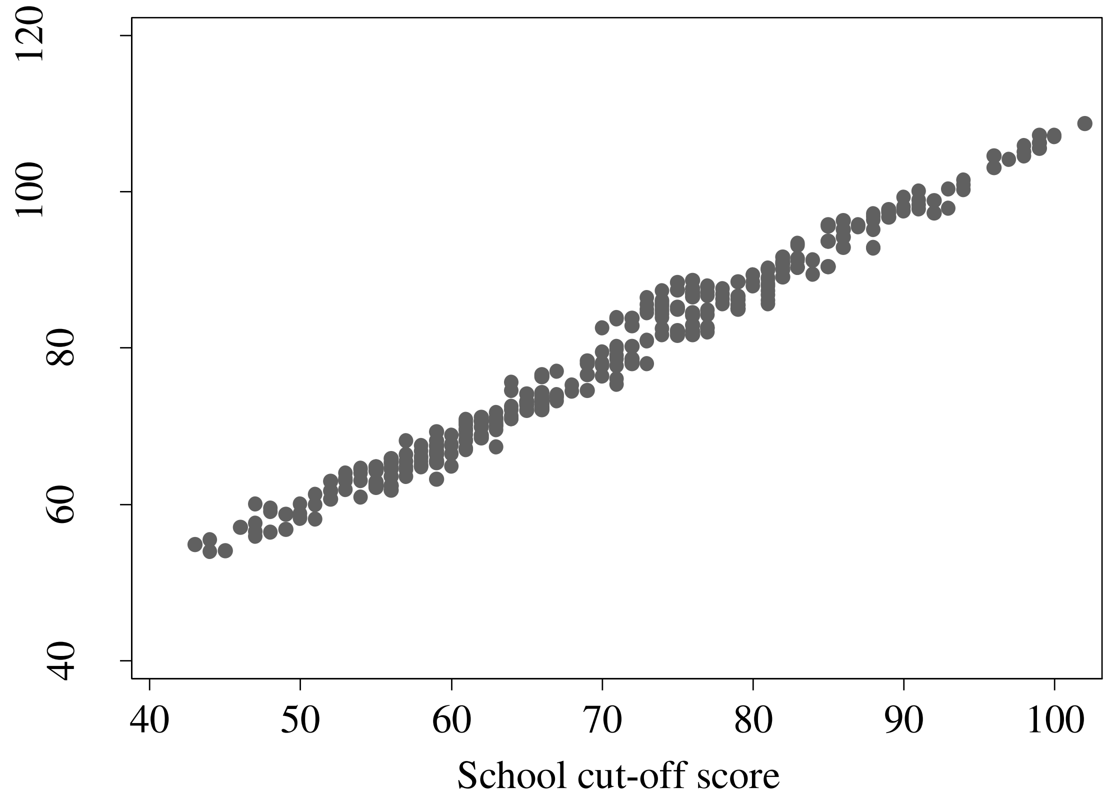
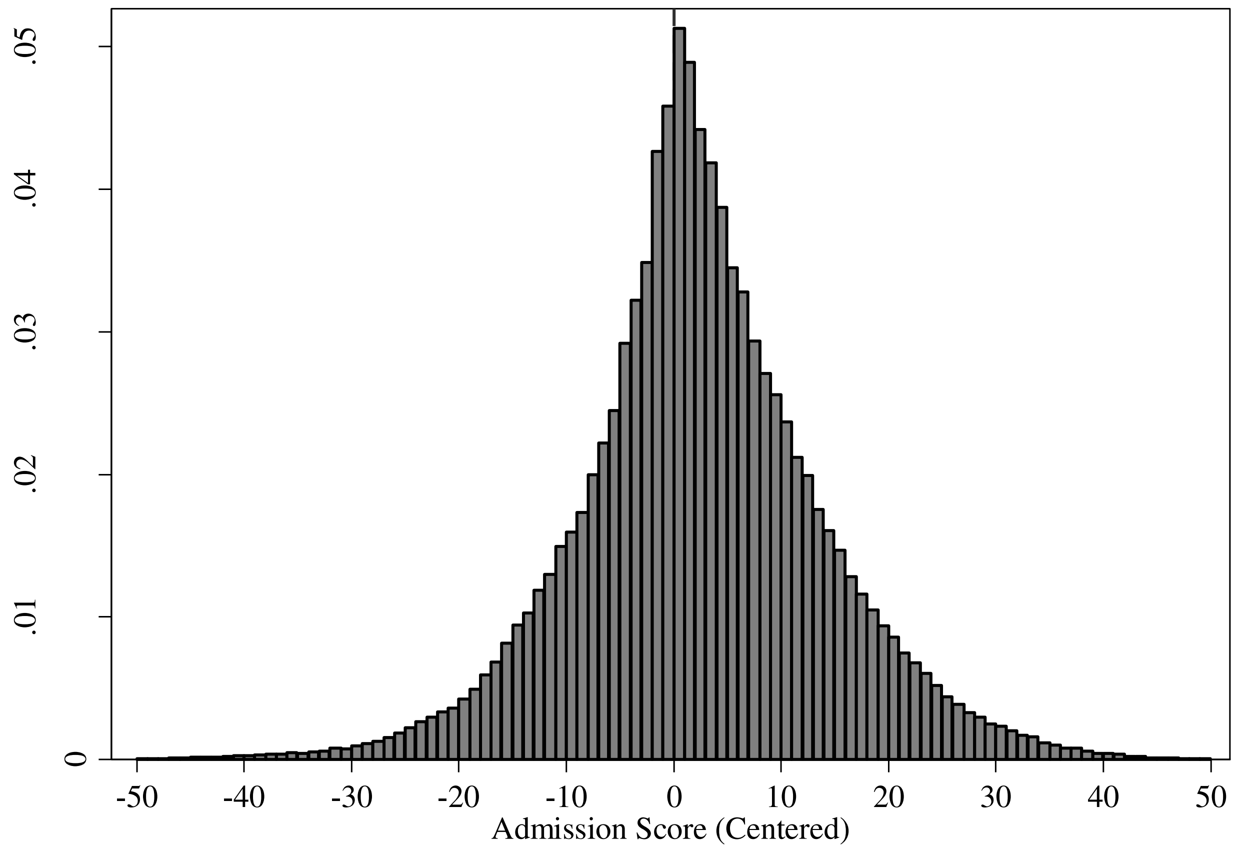


Figure 2. Sample distribution of the admission-score forcing variable relative to the admissions cut-off for a more competitive school.



Note: The score of each applicant is re-centered relative to the relevant threshold. Sharp sample for oversubscribed schools for which it was not necessary to specify a specialization field are pooled together.

Figure 2. Sample bivariate “reduced-form” relationship between the outcomes of interest GRADONTIME and GRAD and the forcing variable, SCOREc, displaying the observed discontinuity in each outcome, centered on the cut-off score (n=505,015)

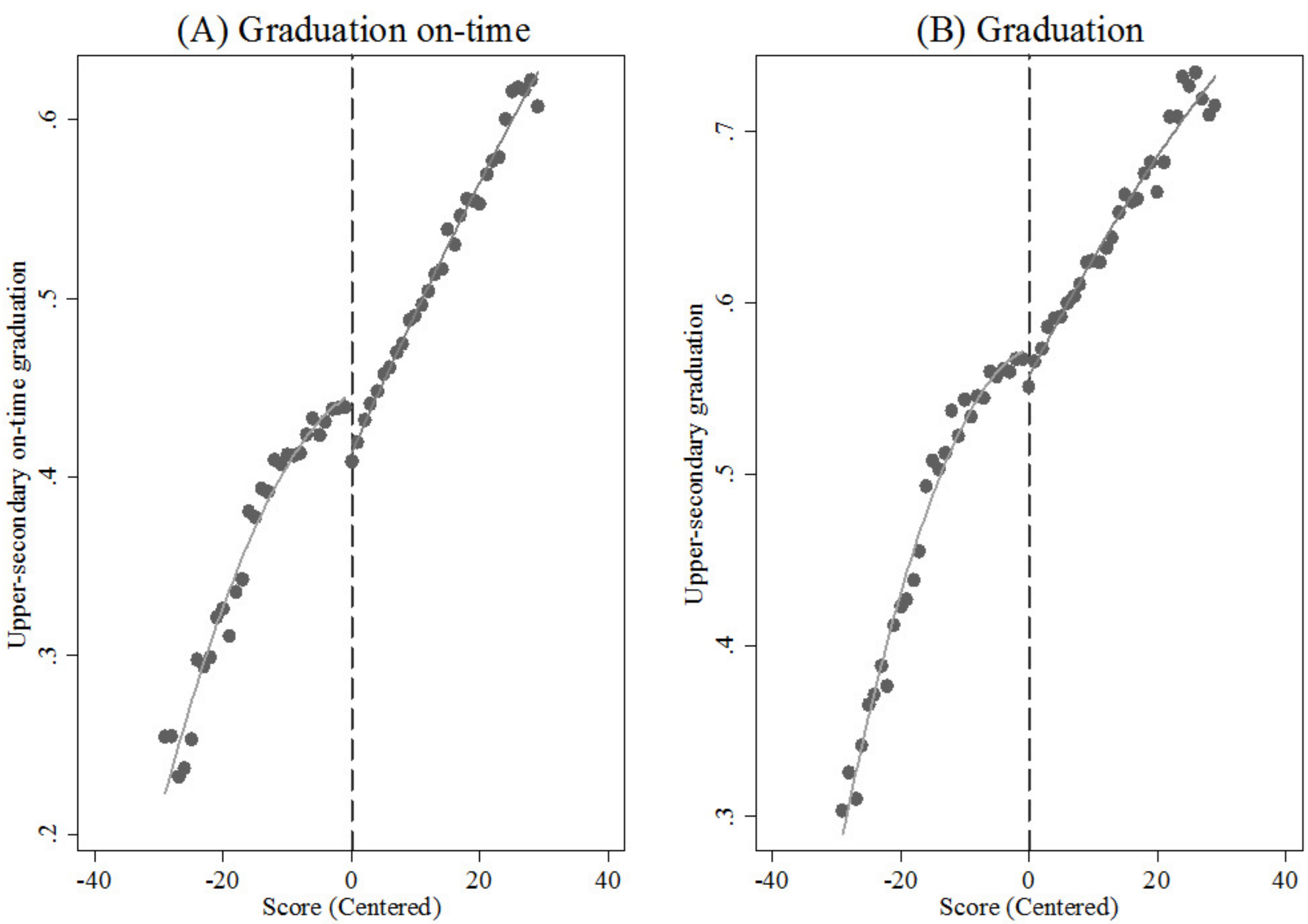


Figure 3. Sample bivariate “reduced-form” relationship between the outcomes of interest GRADONTIME and GRAD and the forcing variable, SCOREc, displaying the observed discontinuity in each outcome, centered on the cut-off score, within a 9 points bandwidth around the respective cut-off (n=271,655)

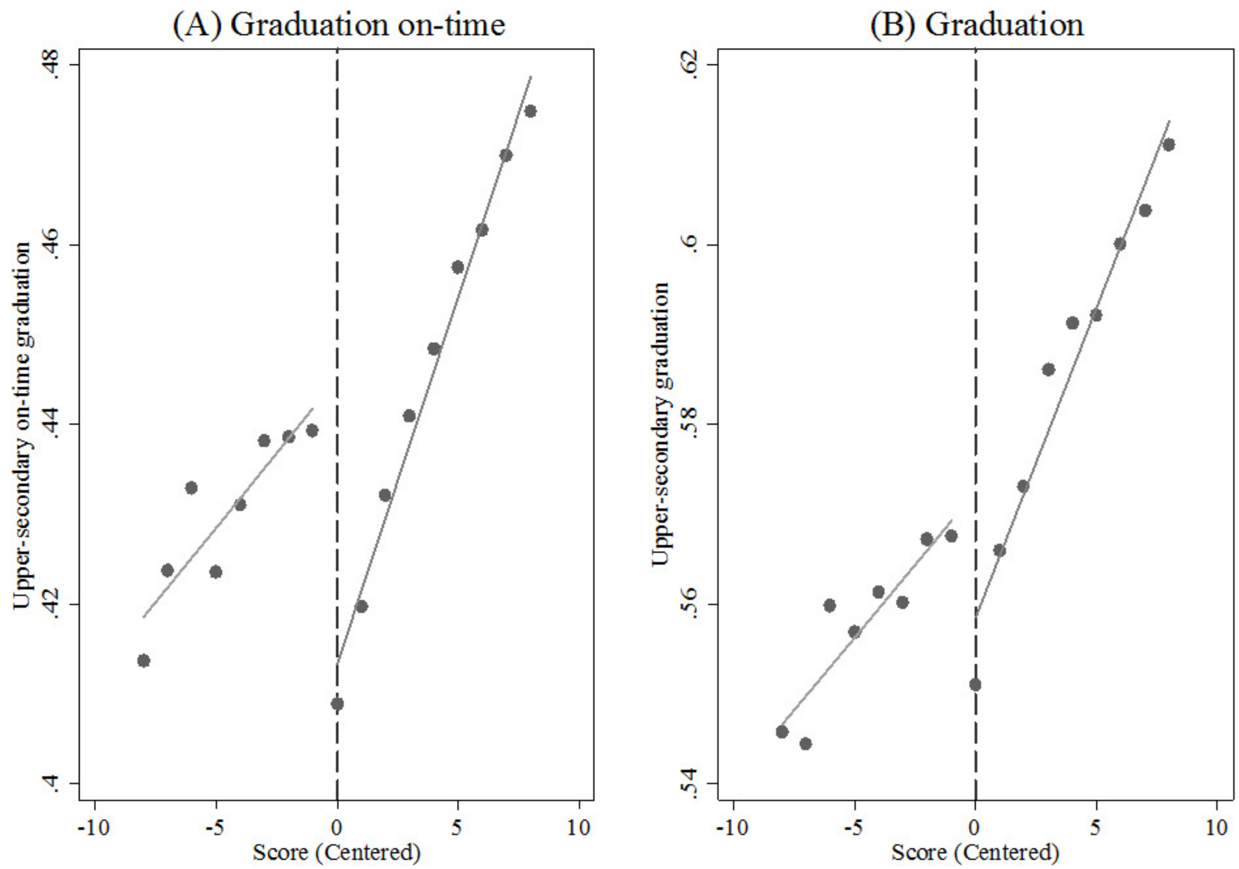


Figure 4. Enrollment in more-competitive oversubscribed school as a function of the re-centered admission score, in the analytic sample. (n=505,015)

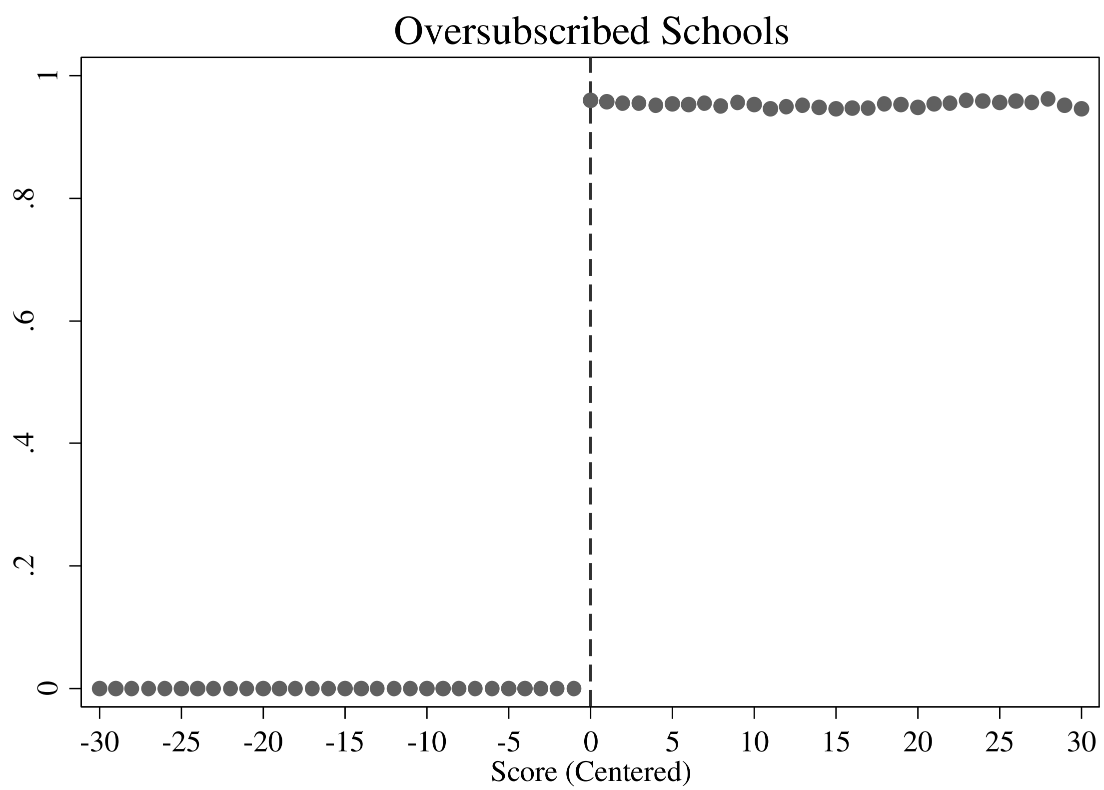


Table 1. Sample means (and standard deviations) describing applicants' selected characteristics, in different samples

	Sample			
	Full	Restricted	Analytic	Discontinuity
Panel A. Demographic and Socio-economic characteristics, at baseline				
Male	0.503	0.502	0.495	0.483
	(0.500)	(0.500)	(0.500)	(0.500)
Age (years)	15.636	15.315	15.203	15.193
	(1.583)	(0.923)	(0.720)	(0.705)
Indigenous	0.031	0.017	0.020	0.019
	(0.174)	(0.131)	(0.141)	(0.137)
Resident of Mexico City	0.776	0.891	0.915	0.920
	(0.417)	(0.312)	(0.225)	(0.219)
GPA in lower-secondary (out of 10)	7.212	8.332	8.379	8.391
	(2.518)	(0.848)	(0.837)	(0.815)
Private lower-secondary school	0.065	0.118	0.119	0.121
	(0.246)	(0.323)	(0.323)	(0.326)
Working	0.059	0.030	0.034	0.033
	(0.236)	(0.170)	(0.182)	(0.178)
Schooling of household head (years)	9.499	10.786	10.786	10.896
	(4.031)	(4.150)	(4.150)	(4.121)
SES index	0.004	0.397	0.430	0.471
	(1.575)	(1.597)	(1.693)	(1.687)
Low-SES family	0.551	0.436	0.440	0.429
	(0.497)	(0.496)	(0.496)	(0.495)
Oportunidades recipient	0.025	0.015	0.017	0.016
	(0.156)	(0.121)	(0.130)	0.126
Panel B. School preferences, at baseline				
Number options listed	9.283	9.935	9.996	10.295
	(3.712)	(3.817)	(3.800)	(3.809)
Rank of assigned option	3.373	3.580	3.557	3.659
	(2.972)	(2.888)	(2.872)	(2.822)
Score on entrance test	61.937	77.841	78.292	78.420
	(20.652)	(15.106)	(15.077)	(12.002)
Panel C. Graduation Outcomes				
On-time graduation (3 years)	0.423	0.451	0.468	0.454
	(0.494)	(0.498)	(0.499)	(0.498)
Graduation (5 years)	0.493	0.552	0.570	0.562
	(0.500)	(0.497)	(0.495)	(0.496)
Observations	1,503,782	588,622	505,015	271,655

Notes: The analytic and discontinuity sample include observations with no missing covariates. The discontinuity sample considers a bandwidth of 9 points. Some characteristics have missing values in the full and restricted sample.

Table 2. Estimated average difference in selected background characteristics (with the respective robust standard errors of the difference, in parentheses), at the cut-off used to determine the offer of admission.

	Bandwidth				
	3	6	9	12	15
Male	0.0011	-0.0006	-0.0021	-0.0022	-0.0030
	(0.0086)	(0.0047)	(0.0037)	(0.0032)	(0.0029)
Age	0.0091	0.0002	-0.0006	0.0002	0.0005
	(0.0120)	(0.0065)	(0.0053)	(0.0046)	(0.0042)
Indigenous	0.0014	-0.0007	-0.0007	-0.0005	-0.0000
	(0.0024)	(0.0013)	(0.0010)	(0.0009)	(0.0008)
GPA lower-secondary	0.0030	-0.0017	0.0049	0.0051	0.0043
	(0.0132)	(0.0072)	(0.0058)	(0.0050)	(0.0045)
Work	0.0012	0.0005	-0.0003	-0.0003	-0.0006
	(0.0031)	(0.0017)	(0.0013)	(0.0012)	(0.0010)
Private lower-secondary school	-0.0073	-0.0060*	-0.0041*	-0.0028	-0.0013
	(0.0054)	(0.0030)	(0.0024)	(0.0021)	(0.0019)
HH's schooling	-0.0776	-0.0797**	-0.0767**	-0.0727***	-0.0612***
	(0.0690)	(0.0377)	(0.0300)	(0.0260)	(0.0235)
SES index	-0.0358	-0.0143	-0.0188	-0.0191*	-0.0170*
	(0.0288)	(0.0157)	(0.0125)	(0.0108)	(0.0098)
Oportunidades	0.0013	0.0001	-0.0000	0.0002	0.0004
	(0.0021)	(0.0011)	(0.0009)	(0.0008)	(0.0007)
Resident	0.0043	0.0036*	0.0029	0.0020	0.0018
	(0.0038)	(0.0021)	(0.0018)	(0.0014)	(0.0013)
Observations	105,343	201,068	271,655	325,481	365,684

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors adjusting for clustering of students. Parameter estimates were obtained with a local-linear regression using an edge kernel with the listed bandwidth, and no additional covariates included. All models included school-cohort fixed effects.

Table 3. Estimated parameter estimates (with approximate p-values and associated robust standard error) summarizing the effect of an offer of admission to a more competitive school on the outcomes of interest, by the value of the forcing variable, without and with covariates

	Bandwidth				
	3	6	9	12	15
Panel A: On-time graduation, No controls					
OFFER	-0.0300***	-0.0324***	-0.0296***	-0.0290***	-0.0280***
	(0.0086)	(0.0047)	(0.0037)	(0.0032)	(0.0029)
OFFER×SCOREc	0.0101*	0.0064***	0.0050***	0.0041***	0.0038***
	(0.0056)	(0.0017)	(0.0010)	(0.0007)	(0.0005)
R-2 statistic	0.078	0.060	0.055	0.054	0.053
Observations	105,343	201,068	271,655	325,481	365,684
Panel B: On-time graduation, With controls					
OFFER	-0.0306***	-0.0321***	-0.0309***	-0.0302***	-0.0290***
	(0.0080)	(0.0043)	(0.0034)	(0.0030)	(0.0027)
OFFER×SCOREc	0.0107**	0.0079***	0.0061***	0.0055***	0.0052***
	(0.0052)	(0.0015)	(0.0009)	(0.0006)	(0.0005)
R-2 statistic	0.219	0.202	0.198	0.196	0.196
Observations	105,343	201,068	271,655	325,481	365,684
Panel C: Graduation, No controls					
OFFER	-0.0156	-0.0126**	-0.0095**	-0.0088**	-0.0081**
	(0.0112)	(0.0061)	(0.0049)	(0.0042)	(0.0038)
OFFER×SCOREc	0.0064	0.0054**	0.0033***	0.0024***	0.0018***
	(0.0073)	(0.0021)	(0.0012)	(0.0009)	(0.0007)
R-2 statistic	0.085	0.068	0.063	0.061	0.061
Observations	61,312	117,565	158,005	188,611	211,107
Panel D: Graduation, With controls					
OFFER	-0.016	-0.0124**	-0.0109**	-0.0105***	-0.0099***
	(0.0106)	(0.0058)	(0.0046)	(0.0040)	(0.0036)
OFFER×SCOREc	0.0067	0.0065***	0.0044***	0.0035***	0.0030***
	(0.0069)	(0.0020)	(0.0012)	(0.0008)	(0.0006)
R-2 statistic	0.188	0.171	0.167	0.166	0.166
Observations	61,312	117,565	158,005	188,611	211,107

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors adjusted for the clustering of students. Parameter estimates were obtained with a local-linear regression analysis using an edge kernel with the listed bandwidth. All models included school-fixed effects and an intercept. Additional controls included: male, age, indigenous, lower-secondary GPA, private lower-secondary school, working status, years of education of the household head, SES index, living in the metropolitan area of Mexico City, and receiving Oportunidades-Progresa (CCT).

Table 4. Parameter estimates (with approximate p-values and standard errors) summarizing the fitted probability of graduating and graduating on-time as a function of the offer of admission and the forcing variable, for selected non-linear specifications of the forcing variable, with controls.

	Specification of the Forcing Variable			
	(1) Linear	(2) Quadratic	(3) Cubic	(4) Quartic
Panel A: On-Time Graduation, With Controls				
OFFER	-0.0309***	-0.0306***	-0.0248***	-0.0177
	(0.0034)	(0.0057)	(0.0095)	(0.0176)
OFFER×SCOREc	0.0061***	0.0120***	0.0199**	0.0266
	(0.0009)	(0.0034)	(0.0096)	(0.0247)
OFFER×SCOREc^2		-0.0003	0.0017	0.0096
		(0.0004)	(0.0027)	(0.0115)
OFFER×SCOREc^3			0.0002	0.0005
			(0.0002)	(0.0019)
OFFER×SCOREc^4				0.0001
				(0.0001)
Bandwidth	9	9	9	9
R-2 statistic	0.198	0.198	0.198	0.198
Observations	271,655	271,655	271,655	271,655
Panel B: Graduation, With Controls				
OFFER	-0.0109**	-0.0117	-0.0134	-0.0201
	(0.0046)	(0.0075)	(0.0125)	(0.0233)
OFFER×SCOREc	0.0044***	0.0110**	0.0111	-0.0048
	(0.0012)	(0.0045)	(0.0127)	(0.0329)
OFFER×SCOREc^2		-0.0005	-0.0015	-0.0039
		(0.0006)	(0.0036)	(0.0152)
OFFER×SCOREc^3			0.0000	-0.0017
			(0.0003)	(0.0025)
OFFER×SCOREc^4				-0.0000
				(0.0002)
Bandwidth	9	9	9	9
R-2 statistic	0.167	0.167	0.167	0.167
N	158,005	158,005	158,005	158,005

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors adjusted for the clustering of students. Parameter estimates were obtained with a local-linear regression analysis using an edge kernel with the listed bandwidth. All models included school-fixed effects and an intercept. Additional controls included: male, age, indigenous, lower-secondary GPA, private lower-secondary school, working status, years of education of the household head, SES index, living in the metropolitan area of Mexico City, and receiving Oportunidades-Progresa (CCT).

Table 5. Estimated parameters (with approximate p-values and robust standard errors, in parentheses) describing the first-stage effects of being offered admission to a more competitive school on the probability that a student enrolls, without and with covariates.

	Bandwidth				
	3	6	9	2	15
Panel A: Enrollment, No Controls					
OFFER	0.9613***	0.9601***	0.9595***	0.9591***	0.9590***
	(0.0015)	(0.0011)	(0.0009)	(0.0009)	(0.0008)
R-2 statistic	0.892	0.904	0.905	0.905	0.905
Observations	105,343	201,068	271,655	325,481	365,684
Panel B: Enrollment, With Controls					
OFFER	0.9605***	0.9595***	0.9591***	0.9588***	0.9588***
	(0.0015)	(0.0011)	(0.0009)	(0.0009)	(0.0008)
R-2 statistic	0.894	0.906	0.907	0.907	0.907
Observations	105,343	201,068	271,655	325,481	365,684

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors adjusted for the clustering of students. Parameter estimates were obtained with a local-linear regression analysis using an edge kernel with the listed bandwidth. All models included school-fixed effects and an intercept. Additional controls included: male, age, indigenous, lower-secondary GPA, private lower-secondary school, working status, years of education of the household head, SES index, living in the metropolitan area of Mexico City, and receiving Oportunidades-Progresa (CCT). All the F-statistic associated with the excluded instrument, not reported here, have very large values (>50).

Table 6. Instrumental-Variables estimates of the causal effect of attending a more competitive school on the probability that a student, at the margins of admission, graduates within 3 or 5 years.

	Bandwidth				
	3	6	9	12	15
Panel A: On-time graduation, No controls					
ENROLL	-0.0312***	-0.0338***	-0.0309***	-0.0303***	-0.0292***
	(0.0089)	(0.0050)	(0.0041)	(0.0036)	(0.0033)
R-2 statistic	0.035	0.036	0.035	0.036	0.037
Observations	105,343	201,068	271,655	325,481	365,684
Panel B: On-time graduation, With controls					
ENROLL	-0.0319***	-0.0335***	-0.0322***	-0.0315***	-0.0303***
	(0.0083)	(0.0046)	(0.0038)	(0.0033)	(0.0031)
R-2 statistic	0.151	0.149	0.149	0.150	0.150
Observations	105,343	201,068	271,655	325,481	365,684
Panel C: Graduation, No controls					
ENROLL	-0.0162	-0.0131**	-0.0099*	-0.0092**	-0.0085**
	(0.0116)	(0.0065)	(0.0053)	(0.0047)	(0.0043)
R-2 statistic	0.021	0.022	0.023	0.021	0.022
Observations	61,312	117,565	158,005	188,611	211,107
Panel D: Graduation, With controls					
ENROLL	-0.0167	-0.0129**	-0.0118**	-0.0109**	-0.0103**
	(0.0110)	(0.0062)	(0.0050)	(0.0044)	(0.0041)
R-2 statistic	0.111	0.110	0.111	0.112	0.112
Observations	61,312	117,565	158,005	188,611	211,107

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors adjusting for clustering at the student level in parentheses. The R-squared correspond to the specifications with controls. Parameter estimates were obtained with a local linear regression using an edge kernel with the listed bandwidth. All models include school-fixed effects. Additional controls include: male, age, indigenous, lower-secondary GPA, private lower-secondary school, working status, years of schooling of the household head, SES index, and Oportunidades-Progresa (CCT).

Table 7a. Reduced-form estimates of the offer of admission on the outcomes of interest including interactions of the offer with selected demographic characteristics of the applicant, with a bandwidth of 9 points.

	Specification				
	(1)	(2)	(3)	(4)	(5)
Panel A: On-time graduation, With controls					
OFFER	-0.0227***	-0.0309***	-0.02649***	-0.0331***	-0.1163***
	(0.0038)	-0.0036	(0.0037)	(0.0035)	(0.0154)
MALE	-0.0115***				
	(0.0032)				
MALE×OFFER	-0.0172***				
	(0.0031)				
WORK		-0.0219***			
		(0.0025)			
WORK×OFFER		-0.0126			
		(0.0099)			
LOW_SES			-0.0031		
			(0.0031)		
LOW_SES×OFFER			-0.0107***		
			(0.0032)		
PRIV_SEC				-0.0515***	
				(0.0050)	
PRIV_SEC×OFFER				0.0188***	
				(0.0050)	
GPA					0.2309***
					(0.0019)
GPA×OFFER					0.0102***
					(0.0018)
Bandwidth	9	9	9	9	9
R-2 statistic	0.198	0.198	0.198	0.198	0.198
Observations	271,655	271,655	271,655	271,655	271,655

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors adjusting for clustering at the student level in parentheses. Parameter estimates were obtained with a local linear regression using an edge kernel with the listed bandwidth. All models include school-fixed effects. Additional controls include: male, age, indigenous, lower-secondary GPA, private lower-secondary school, working status, years of schooling of the household head, SES index, and Oportunidades-Progresa (CCT).

Table 7b. Reduced-form estimates of the offer of admission on the outcomes of interest including interactions of the offer with selected demographic characteristics of the applicant, with a bandwidth of 9 points.

	Specification				
	(1)	(2)	(3)	(4)	(5)
Panel B: Graduation, With controls					
OFFER	-0.0043	-0.0109***	-0.0041	-0.0144***	-0.0852***
	(0.0049)	(0.0046)	(0.0049)	(0.0046)	(0.0217)
MALE	-0.0036				
	(0.0042)				
MALE×OFFER	-0.0139***				
	(0.0041)				
WORK		-0.0499***			
		(0.0111)			
WORK×OFFER		0.0014			
		(0.0117)			
LOW_SES			-0.0022		
			(0.0041)		
LOW_SES×OFFER			-0.0153***		
			(0.0041)		
PRIV_SEC				-0.1057***	
				(0.0068)	
PRIV_SEC×OFFER				0.0312***	
				(0.0068)	
GPA					0.1845***
					(0.0026)
GPA×OFFER					0.0089***
					(0.0025)
Bandwidth	9	9	9	9	9
R-2 statistic	0.167	0.167	0.167	0.167	0.167
Observations	158,005	158,005	158,005	158,005	158,005

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors adjusting for clustering at the student level in parentheses. Parameter estimates were obtained with a local linear regression using an edge kernel with the listed bandwidth. All models include school-fixed effects. Additional controls include: male, age, indigenous, lower-secondary GPA, private lower-secondary school, working status, years of schooling of the household head, SES index, and Oportunidades-Progresa (CCT).

Table 8. Reduced-form estimates of the offer of admission on the outcomes according to the location cut-off in the admission of the school cut-off score distribution, with a bandwidth of 9 points.

	Specification			
	(1)	(2)	(3)	(4)
	Interquartile 1	Interquartile 2	Interquartile 3	Interquartile 4
Panel A: On-time graduation, With controls				
OFFER	-0.0006	0.0127	-0.0108**	-0.0754***
	(0.0075)	(0.0088)	(0.0054)	(0.0058)
Bandwidth	9	9	9	9
R-2 statistic	0.200	0.187	0.176	0.177
Observations	23,971	47,317	98,631	101,736
Panel B: Graduation, With controls				
OFFER	0.0011	0.0161*	-0.0161**	-0.0476***
	(0.0113)	(0.0096)	(0.0071)	(0.0076)
Bandwidth	9	9	9	9
R-2 statistic	0.160	0.147	0.152	0.159
Observations	10,397	33,723	56,986	56,899

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors adjusting for clustering at the student level in parentheses. Parameter estimates were obtained with a local linear regression using an edge kernel with the listed bandwidth. All models include school-fixed effects. Additional controls include: male, age, indigenous, lower-secondary GPA, private lower-secondary school, working status, years of schooling of the household head, SES index, living in the metropolitan area of Mexico City, and receiving Oportunidades-Progresa (CCT). The interquartile 1 the sharp samples of schools with cut-of score of less than 61, the interquartile 2 includes schools with cut-off score between 61 and 72, the interquartile 3 corresponds to schools with cut-off scores between 72 and 82, and the interquartile 4 include schools with cut-off scores above 82 points.

Table 9. Reduced-form estimates of the offer of admission on the outcomes according to the ranking of the offer, with a bandwidth of 9 points.

	Subsample		
	First-choice	Second-choice	Third-choice
Panel A: On-time graduation, No controls			
OFFER	-0.0595***	-0.0325***	-0.0206***
	(0.0066)	(0.0081)	(0.0772)
OFFER×SCORE	0.0061***	0.0119	0.0112
	(0.0016)	(0.0185)	(0.0192)
R-2 statistic	0.082	0.125	0.178
Observations	110,309	34,490	14,368
Panel B: On-time graduation, With controls			
OFFER	-0.0562***	-0.0234*	0.0437
	(0.0060)	(0.0622)	(0.0786)
OFFER×SCORE	0.0068***	0.0053	0.0159
	(0.0015)	(0.0187)	(0.0191)
R-2 statistic	0.234	0.264	0.307
Observations	110,309	34,490	14,368
Panel C: Graduation, No controls			
OFFER	-0.0356***	-0.0264**	-0.0183**
	(0.0082)	(0.0189)	(0.0086)
OFFER×SCORE	0.0040**	-0.0013	0.0186
	(0.0020)	(0.0252)	(0.0256)
R-2 statistic	0.097	0.134	0.192
Observations	65,755	20,452	8,618
Panel D: Graduation, With controls			
OFFER	-0.0323***	-0.0139	0.0034
	(0.0077)	(0.0862)	(0.1052)
OFFER×SCORE	0.0051***	-0.0023	0.0167
	(0.0019)	(0.0244)	(0.0247)
R-2 statistic	0.205	0.229	0.287
Observations	65,755	20,452	8,618

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors adjusting for clustering at the student level in parentheses. The R-squared correspond to the specifications with controls. Parameter estimates were obtained with a local linear regression using an edge kernel with the listed bandwidth. All models include school-fixed effects and a constant. Additional controls include: male, age, indigenous, lower-secondary GPA, private lower-secondary school, working status, years of education of the household head, SES index, living in the metropolitan area of Mexico City, and receiving Oportunidades-Progresa (CCT).

Appendix

Appendix A. Subsystems and Institutions that integrate the COMIPEMS

Table A.1. Subsystems and Institutions that participate in COMIPEMS

Model	Subsystem		Institutions
General	UNAM	Universidad Nacional Autonoma de Mexico	CCH, ENP
	UAEM	Universidad Autónoma del Estado de Mexico	Texcoco
	COLBACH	Colegio de Bachilleres	COLBACH
	DGB	Dirección General de Bachillerato	CEB
	SE	Secretaría de Educación del Estado de México	COBAEM
Technical	CONALEP	Colegio Nacional de Educación Profesional Técnica	DF, State of Mexico
Technological (General + Technical)	SE	Secretaría de Educación del Estado de México	CECYTEM
	DGETI	Dirección de Educación Tecnológica Industrial	CBTIS, CETIS
	DGETA	Dirección de Educación Tecnológica Agropecuaria	CBTA
	IPN	Instituto Politécnico Nacional	CET, CECYT

The main characteristics of each subsystem:

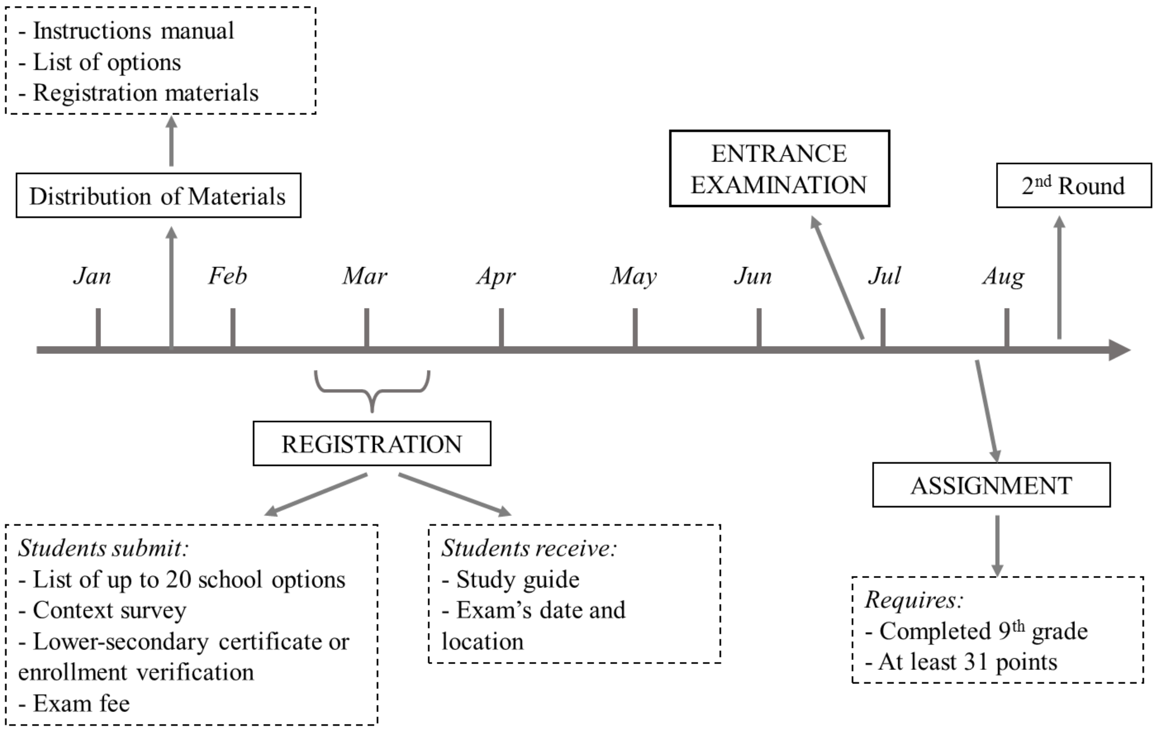
- General: prepares students with general knowledge to continue studying for post-secondary education. It is offered in three-year programs. Graduates receive a certificate that is necessary to enter higher education. There are several modalities: standard, open and distance, and mixed.

- Technological: covers the curriculum of general education to prepare students for post-secondary education. In addition, it gives students the opportunity to receive vocational training. This bivalent character allows graduates to continue into higher education and gives them also a technical diploma at a semi-professional level. Programs last three years.

- Technical: trains students to become professional technicians and gives them the option to continue into tertiary education. Programs are completed in three years.

Appendix B: The COMIPEMS Admission Process

B.1. Timeline



B.2. Registration Materials

The registration materials that are distributed to individuals who want to attend a public upper-secondary school in Mexico City include:

a) A manual with a detailed explanation of the process and the directory of available school options, including each option's cut-off scores from previous years.

b) A registration form to fill with personal information and the list of up to 20 preferred school options they were willing to attend, ordered by preference.

c) A context survey seeking information on the students' demographic, academic and socio-economic characteristics.

d) A bank deposit slip to pay the fee for sitting the standardized entrance examination (approximately 25 U.S. dollars).

C. Construction of Sharp Samples

In the setting under analysis, applicants to public upper-secondary schools are ranked based on their score on an entrance examination. Applicants may qualify for schools (i.e. be scored above the respective school's cut-off score) that they listed among their preferences, but not be offered admission if they have already received an offer of admission at a more preferred school.

According to Abdulkadiroğlu et al. (2014), the "sharp sample" for school S is the sample for whom offers of admission are sharp in the sense of being linked deterministically with S's forcing variable. Thus, the sharp sample for school S implies the union of different sets of applicants, those who ranked S and qualified, and those who ranked S and did not qualify.

In this particular case, considering the application risk sets, I constructed a "sharp sample" for each oversubscribed school and included fixed effects for each pair of schools combination in a given cohort (i.e. School-School-Cohort fixed effects), not necessarily the same preference set. For example, I grouped applicants that prefer A over B but some scored just above the cut-off for A and some scored just below the cut-off for A and got assigned to B.

Below, I explain the procedure I followed to construct the sharp samples, using a simple example.

Assume that, in 2006, with 3 applicants and 3 schools:

- Applicant 1 ranks schools in descending order: A, B, C
- Applicant 2 ranks schools in descending order: A, B, C
- Applicant 3 ranks schools in descending order: B, A, C

where cut-off(A)> cut-off(B) > cut-off(C).

Imagine that Applicant 1 is offered admission to school A, Applicant 2 is offered admission to B and Applicant 3 cleared the admission cut-off score for school C. Thus, all students prefer schools A and B over school C, but Applicant 3 prefers School B over A, instead of A over B. Then, given how the assignment algorithm works, it is not possible that Applicant 3 receives an offer of admission to School A after being rejected from B, because cut-off(A)>cut-off(B). If he gets rejected from School B, he will then be considered for School C, which is his next most-preferred school with available seats.

In this scenario, School B is the next most-preferred option to which Applicant 1 would qualify if rejected from School A. Thus, Applicant 1 will be included in the sharp sample for School A with a positive centered-score ($SCORE_1$-cut-off(A)>0). To control for the applicant's preference ordering, I include as a covariate in my regression models a dichotomous predictor to represent the applicant's School-School-Cohort combination A_B_2006, which indicates that for Applicant 1, School A is preferable to B in 2006. Similarly, School C will be the next most-preferred option to which Applicant 2 would qualify, so he will be considered in the sharp sample for School B with a positive centered-score ($SCORE_2$-cut-off(B)>0), and will have a corresponding value for the dichotomous covariate B_C_2006.

In addition, School B is the "most informative qualification" for Applicant 3. That is, the school with the lowest cut-off score that rejected him or the more-preferred to which he would have been more likely to be admitted. Then, Applicant 3 will be included in the sharp sample for School B with a negative centered-score ($SCORE_3$-cut-off(B)<0), and will include a corresponding value for the dichotomous covariate B_C_2006.[76] Analogously, School A is the most informative qualification for Applicant 2 so he will be included in the School A sharp sample with a negative centered-score ($SCORE_2$-cut-off(A)<0), and a corresponding value for the dichotomous covariate A_B_2006.

To summarize:[77]

- Applicant 1 is included in the sharp sample for School A with SCOREc>0;

  applicant's preference is reflected in a value of 1 on the dichotomous covariate: A_B_2006; 0 otherwise.

---

[76] As I mentioned before, although student 3 could be included in the sharp sample for A, there would be no applicant with the school-school-cohort fixed effect B_A_2006 that would be assigned to school A. Therefore, by construction, a more competitive school (i.e. higher cut-off score) is always a more preferred school with higher-achieving peers.

[77] Another approach is to include each applicant in the sharp sample of each school that they ranked. However, the strategy I follow is a more refined way to construct the sharp samples while controlling for the applicant's preference ordering. Ultimately, my analyses are restricted to individuals around the cut-off so those applicants that score away from the cut-off for a particular school are unlikely to change the results.

- Applicant 2 is included in the sharp sample for school A with SCOREc<0;

  applicant's preference is reflected in a value of 1 on the dichotomous covariate: A_B_2006; 0 otherwise.

- Applicant 2 is also included in the sharp sample for school B with SCOREc>0;

  applicant's preference is reflected in a value of 1 on the dichotomous covariate: B_C_2006; 0 otherwise.

- Applicant 3 is in the sharp sample for school B with SCOREc<0:

  applicant's preference is reflected in a value of 1 on the dichotomous covariate: B_C_2006; 0 otherwise.

As I mentioned, I follow the approach used in Pop-Eleches and Urquiola (2013) and combine the analysis of all discontinuities into a single analysis. Thus, I pooled all sharp samples into the analysis of a single discontinuity, and included suitable fixed effects to describe the combination of cohort and application risk sets (i.e. school-school-cohort).

Finally, since an applicant can be present in more than one sharp sample, I adjust the obtained standard errors for the clustering of records within the individual.

D. Definition and coding description of the variables

Table D.1: Definition and coding of the variables for the statistical analysis

Variable	Definition	Values
SCORE	Score on the standardized entrance examination	0 to 128
SCOREc	Distance of the score with respect to the cut-off (centered at 0)	-60 to 60
OFFER	Applicant was offered admission to a more competitive school in the sharp sample	yes=1, no=0
ATTEND	Applicant attended his or her assigned school	yes=1, no=0
GRADONTIME	Applicant graduated from upper-secondary education within 3 year of the offer of admission	yes=1, no=0
MALE	Applicant is a male	yes=1, no=0
INDIGENOUS	Applicant reported speaking an indigenous language	yes=1, no=0
AGE	Age of applicant at baseline	14 to 25
WORK	Applicant was working for a wage at baseline	yes=1, no=0
HH_SCHOOLING	Years of education for the highest level of education completed by the household head at baseline	no education=0, primary=6, lower-secondary=9, upper-secondary=12, university=17, graduate=20
GPA	GPA obtained in lower-secondary education	6 to 10
RESIDENT	Applicant lives in the metropolitan area of Mexico City	yes=1, no=0
SES_INDEX	Principal components index of access to durables goods (like car, computer, dvd) and services (like sewage, telephone, internet) at home	-6 to 6
LOW_SES	Applicant has SES below the median	yes=1, no=0
OPORTUNIDADES	Applicant's family receivee the Oportunidades-Progresa conditional cash transfer	yes=1, no=0
PRIVATE_LOWERSEC	Applicant attended a private lower-secondary school	yes=1, no=0

VITA

Maria Elena Ortega Hesles

1999 – 2005	Insituto Tecnologico Autonomo de Mexico Distrito Federal, Mexico	B.S. Jun 2005
2004 – 2006	Research Analyst Banco de Mexico Distrito Federal, Mexico	
2006 – 2007	Master in Education Harvard University Graduate School of Education Cambridge, MA	Ed.M. May 2007
2007 – 2013	Doctor of Education Candidate, Harvard Graduate School of Education Cambridge, MA, U.S.A.	
2008 – 2009	Teaching Fellow Harvard University Graduate School of Education Cambridge, MA	
2009 – 2011	Teaching Fellow Harvard University, Faculty of Arts and Sciences Cambridge, MA	
2011 – 2015	Graduate Student Affiliate Harvard Institute for Quantitative Social Science Cambridge, MA	