London School of Economics and Political Science

Essays in Economics of Education

Andrés Barrios Fernández

July, 2019

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy.

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I declare that my thesis consists of 65,000 words.

Statement of conjoint work

I confirm that Chapter 1 was jointly co-authored with Giulia Bovini and I contributed 50% of this work.

I confirm that Chapter 3 was jointly co-authored with Marin Drlje, Dejan Kovac and Christopher Neilson and I contributed 60% of this work.

Statement of use of third party for editorial help

I can confirm that my thesis was copy edited for conventions of language, spelling and grammar by the LSE Language Center.

Acknowledgments

I am extremely grateful to my supervisors Steve Pischke and Johannes Spinewijn for their insightful comments and invaluable guidance during the Ph.D. Besides my supervisors, I would like to thank Esteban Aucejo, Xavier Jaravel, Camille Landais, Steve Machin, Alan Manning, Sandra McNally, Guy Michaels, Christopher Neilson and Daniel Reck for their support and for many useful comments. I also thank Karun Adusumilli, Diego Battiston, Giulia Bovini, Dita Eckardt, Giulia Giupponi, Felix Koenig, Eui Jung Lee, Claudio Schilter and Vincenzo Scrutinio for their support and useful suggestions.

I am also grateful to the Research Centre of the Ministry of Education, to the Education Quality Agency of Chile, to the Department of Assessment, Evaluation and Educational Records (DEMRE) of the University of Chile and to the Agency for Science and Higher Education of Croatia for granting access to the data used in this thesis.

I acknowledge financial support from the Advanced Human Capital Programme of the National Commission for Scientific and Technological Research of the Government of Chile (CONICYT) and the Department of Economics of LSE.

Abstract

This thesis consists of three chapters that investigate how the institutions and organization of schools affect their performance, and how neighbors and siblings affect human capital investment decisions.

Chapter 1 studies whether the effect of a reform that substantially increased daily instruction time in Chilean primary schools varies depending on school institutions. Focusing on legacy enrollment students and exploiting an IV strategy, it finds that gains are larger in no-fee charter schools than in public schools. Autonomy over personnel decisions emerges as an important institutional feature: to provide additional instruction hours, charter schools rely more on hiring new teachers, and less on increasing the workload of incumbent teachers.

Chapter 2 investigates whether the decision to attend university depends on enrollment of close neighbors. The analysis uses detailed geographic and educational information from Chile and exploits the variation in enrollment generated by the rules that define eligibility for financial aid. The chapter shows that close neighbors have a large and significant impact on university enrollment of younger applicants, suggesting that policies that expand access to university generate spillovers on the peers of their direct beneficiaries. The documented effect is particularly strong among individuals who are more likely to interact and in areas where university attendance is low.

Chapter 3 analyzes how the probability of applying and enrolling in a particular university (or program) changes when an older sibling enrolls in it. In both countries, universities select their students using deferred acceptance admission systems (DA). This chapter exploits thousands of sharp admission cutoffs generated by these systems, and in a fuzzy RD setting shows that older siblings generate significant spillovers on the application and enrollment decisions of their younger siblings. The chapter discusses five classes of mechanisms and presents evidence consistent with information being a relevant driver of the results.

Contents

List of Figures

List of Tables

Chapter 1

It's Time to Learn: School Institutions and Returns to Instruction Time

Andrés Barrios Fernández
LSE

Giulia Bovini
LSE

Abstract

This paper investigates whether the effects of a reform that substantially increased daily instruction time in Chilean primary schools vary depending on school institutions. Focusing on legacy enrollment students and exploiting an IV strategy, we find that longer daily schedules increase reading scores at the end of fourth grade and that the benefits are greater for pupils who began primary education in no-fee charter schools rather than in public schools. We provide evidence that these two types of publicly subsidized establishments, which cater to similar students but differ in the degree of autonomy, expand the teaching input in different ways: in order to provide the additional instruction time, no-fee charter schools rely more on hiring new teachers and less on increasing teachers' working hours than public schools. We also show that additional time at school seems to be more beneficial for schoolchildren of low socio-economic status, who have limited support at home.

Keywords: School autonomy, charter schools, instruction time.
JEL classification: I28, I24, I20.

1.1 Introduction

Given the important role played by schools in the formation of human capital, academics and policymakers have long been interested in understanding what makes a school more effective. While a first branch of the literature focuses on school inputs (i.e. class size, instruction time, teachers), a more recent one highlights the relevance of school institutions and governance on students' performance [Woessmann, 2016]. Indeed, there is active research exploring the effects of granting more autonomy to schools on learning outcomes (see section 1.2). In this paper we study if and why the effects of expanding a specific school input –instruction time– vary across public and charter schools. This paper therefore lies at the intersection of these two literatures, as we address the question of whether the way in which schools are managed and organized affects how they use the additional resources.

Studying how school institutions affect returns to additional instruction time is interesting because many countries are considering devoting, or have already allocated, substantial funds to increasing the amount of time that pupils spend at school.[1] In addition, since time is an inherently limited resource, extended school schedules reduce the amount of time that students can dedicate to other activities. Therefore, the effect on achievement could depend not only on the absolute quality of time use at school, but also on its relative quality with respect to the learning opportunities available outside of school. Hence, we also examine whether the effects of increased instruction time are heterogeneous in terms of students' socio-economic background.

To address these questions, we take advantage of two attractive features of the Chilean educational system. Firstly, we exploit the passing of the Full School Day (*Jornada Escolar Completa*, or FSD henceforth) reform in 1997, which substantially increased daily instruction time in all publicly subsidized (i.e. public and charter) primary and sec-

[1] For instance, since 2003 Germany has begun phasing in all-day schooling and the percentage of pupils attending all-day primary schools has increased from 7.9 per cent in 2005 to 24.2 per cent in 2013 [OECD, 2016a]. Several Latin-American countries have recently transitioned from two-shift schemes, where some grades are taught in the morning and some in the afternoon, to one-shift schemes that feature a longer school day (see section 1.2). As other examples, President Obama in 2009 and Chancellor Osborne in 2016 advocated for longer school days in the US and UK respectively. In the US the National Center on Time and Learning (NCTL) promotes extended school schedules.

ondary schools, whilst leaving the term length and the national curriculum unchanged. The increment was sizable, ranging from 4 to 9 additional instruction hours per week depending on the grade. In grades 1 to 4 it translated into a 26.7 per cent increase in weekly instruction time. Schools could decide when to adopt the longer school day and how to allocate the additional time across subjects. Secondly, we leverage the fact that in Chile there are public and charter schools. Both types of schools are funded through a voucher system, but differ substantially in terms of ownership and in the degree of autonomy that they enjoy: charter schools have more autonomy over staff decisions and over the course offer and content.

We estimate the effect of additional instruction time on achievement by exploiting within-school variation in years of exposure to the FSD across several cohorts of pupils that start grade 1 between 2002 and 2010, and that later take standardized reading and mathematics tests at the end of grade 4. As the availability of longer schedules may affect the composition of pupil intake, we restrict our attention to cohorts of *incumbent* students, i.e. those who started primary education in schools that had yet not adopted the FSD and who may become exposed to it at different stages of their primary education. We further deal with potentially endogenous mobility across schools by instrumenting actual exposure with the exposure a student would have accumulated if she had remained in the school where she initially enrolled.

Our preferred linear specifications show that an additional year of exposure to the FSD raises fourth grade reading scores by 0.024σ. The effect on mathematics scores is smaller (0.007-0.008σ) and non significant.

We then explore heterogeneity by pupils' socio-economic characteristics. The effect of additional instruction time on academic performance is smaller for students from advantaged backgrounds than for the rest of the pupils in our sample. This finding is consistent across different measures of household resources, which include parental education and the availability of books, a computer and a connection to the Internet at home. Although non negligible, the difference in the benefits of additional instruction time for pupils from different backgrounds is not statistically significant. With this caveat about the precision of our estimates in mind, results suggest that returns to additional time at school are higher for pupils who —according to information drawn from time-use surveys— have scarcer support available outside of school. We further provide evidence suggesting that the longer school day is associated with a reduced frequency of homework. As this is likely to particularly benefit students with limited support outside of school, it could be a potential mechanism underlying these findings.

Then, exploiting the features of the Chilean setting, we study how school institutions affect returns to instruction time. In order to avoid capturing differences in the characteristics of students attending different types of schools, we compare public schools to charter schools that do not charge fees. While fee-charging charter schools serve pupils from more advantaged backgrounds, public and no-fee charter schools cater to schoolchildren of similarly lower socio-economic status. We document greater benefits of longer schedules for pupils who start primary education in no-fee charter schools than for those who start in public schools. The difference is large for both subjects —mathematics and reading— but is only statistically significant for the latter. Moreover, it does not decline when, in the same regression specification, we allow for the effect of longer schedules to differ depending on students' socio-economic background; this further suggests that the estimated difference does not reflect heterogeneity in the characteristics of students attending the two types of schools.

While survey evidence suggests that public and charter schools allocate the additional instruction time across subjects in a similar way, we uncover a significant difference in how public and no-fee charter schools adjust total contract and teaching hours to provide longer schedules: no-fee charter schools rely more on hiring new teachers and less on increasing work hours per teacher than public schools, as shown by the evolution of the number of teachers and their contract hours after the adoption of the FSD. It therefore appears that the higher degree of autonomy enjoyed by charter schools over staff decisions allows them to adjust the teaching input in a different way. Moreover, we show that public school teachers display a lower degree of satisfaction with the FSD scheme than their colleagues in charter schools. If extended teachers' working hours translate into a lower quality of time use at school, this could be one mechanism underlying the documented heterogeneity. The literature has found that charter schools are typically associated with better learning outcomes. Our findings suggest that in the context of a large-scale expansion of a school input (i.e. instruction time), charter schools may be able to adjust other inputs (i.e. teaching hours) in a more effective way.

The rest of this paper is organized as follows. Section 1.2 reviews the related literature. Section 1.3 describes the Chilean education system and the FSD reform. Section 1.4 presents the identification strategy. Section 1.5 describes the data and the sample. Section 1.6 discusses the main findings. Section 1.7 concludes.

1.2 Related Literature

This paper speaks to the growing literature that examines the effects of school institutions on pupils' performance. Several papers focus on newly founded or converted charter schools in the US. Studies on oversubscribed charter schools exploit the fact that admission depends on a lottery and document a positive effect both on school performance [Abdulkadiroğlu et al., 2011, Dobbie et al., 2011] and on medium-term non academic outcomes [Dobbie and Fryer Jr, 2015]. Dobbie and Fryer Jr [2013] and Angrist et al. [2013] highlight that high-performing charter schools are characterized by certain practices and features, among which there is increased instruction time. Along this line, Baude et al. [forthcoming] show that the quality of charter schools in Texas improved between 2001 and 2011 driven largely by the adherence to *No Excuses*-style curricula. Abdulkadiroğlu et al. [2016] analyze charter takeovers (i.e. formerly public schools converted into charter schools) and also report positive effects on achievement of grandfathered students. Eyles et al. [2017] and Eyles and Machin [2018] examine how the conversion of some English community schools into academies —autonomous, state-funded education establishments not subject to local authority control— affects achievement of legacy enrolled pupils. The former work studies several post-2010 episodes of conversion involving already high-performing primary schools and does not find significant effects on achievement. The second assesses the first round of conversion of mostly under-performing secondary schools in the 2000s, documenting instead a positive impact on test scores.

Our work is also related to the literature on the relationship between instruction time and achievement.[2] A set of papers studies the effect of the number of school days prior to standardized tests on performance, by exploiting either unplanned school closures due to adverse weather conditions [Marcotte, 2007, Marcotte and Hemelt, 2008, Hansen, 2011, Goodman, 2014] or changes in term dates and/or test dates [Sims, 2008, Agüero and Beleche, 2013, Aucejo and Romano, 2016]. These studies find positive, although in some cases modest, effects. While they leverage small variations in the number of school days, we focus on substantial and permanent changes to the length of the school day. Varying the length rather than the number of school days may have different consequences on student achievement. For example, while the former entails a re-organization of daily routines, the latter does not.

Starting from Lavy [2015], recent studies examine the effect of instruction time on achievement and the drivers of its effectiveness by using cross-country PISA data and exploiting within-pupil variation in subject-specific classroom hours. Lavy [2015] finds that a one-hour increase of weekly subject-specific instruction time raises scores by 0.06σ and that school and student characteristics matter: the effect is larger for schools that enjoy more autonomy and for pupils from disadvantaged backgrounds. Rivkin and Schiman [2015] further highlight that productivity of instruction time depends positively on the quality of the classroom environment, as captured by student disruption and student-teacher interactions. Cattaneo et al. [2017] focus their attention on Switzerland and document that students in more demanding school-tracks enjoy greater benefits. Also these studies leverage a source of variation different from ours. Different allocations of weekly instruction time across subjects do not necessarily entail a change in the length of the school day. Students do not have to re-arrange their daily routine or reduce the time for activities carried out outside of school, nor do schools need to operate for more hours.

A number of papers exploit instead reform-induced variation in instruction time. Pischke [2007] and Parinduri [2014] study the effects of exceptionally short or long school

[2]The early studies mostly focus on term length and report modestly positive to insignificant effects. These studies rely either on variation in term length between and within US states over time [Rizzuto and Wachtel, 1980, Card and Krueger, 1992, Grogger, 1996, Betts and Johnson, 1998, Eide and Showalter, 1998] or on cross-country differences [Lee and Barro, 2001, Wößmann, 2003]. A review of studies conducted in the 1985-2009 period can be found in Patall et al. [2010].

years due to country-level reforms of school calendars that leave the curriculum unchanged.[3] More similarly to us, Huebener et al. [2017] and Lavy [forthcoming] examine reforms that increase weekly instruction hours in Germany and Israel, respectively. They both find a positive effect on achievement. The former documents a larger gain for high-performing students, while the latter does not find evidence of differential benefits across pupils from different socio-economic backgrounds. Battistin and Meroni [2016] and Meroni and Abbiati [2016] study an expansion of mathematics and reading instruction time in lower secondary schools in southern Italy, documenting positive effects on mathematics test scores, concentrated among high-achieving disadvantaged pupils. Recently, Figlio et al. [2018] show that extending the school day and providing additional literacy instruction time in low-performing schools in Florida have a positive effect on reading test scores.

Similarly to Chile, several other Latin American countries have switched from a two-shift scheme —where some grades are taught in the morning and some in the afternoon— to a one-shift scheme, substantially lengthening the school day. The impact has been evaluated in a series of reports.[4] Findings are mixed, suggesting that how the reform of school schedules is implemented and how additional instruction time is used play important roles in shaping returns.

Two papers study the effect of the FSD reform in Chile on achievement. Bellei [2009] focuses on performance at grade 10 in 2001 and 2003, adopting a difference-in-differences approach. Berthelon et al. [2016] explore the effect on early literacy skills at grade 2. Based on one year of observations (2012), they instrument exposure to the FSD with the local availability of schools offering longer schedules. Both papers find positive and significant effects on performance. We focus on all cohorts that start primary school between 2002 and 2010 and examine the effect of the FSD on their academic performance in a different grade (grade 4). In addition, we propose a different identification

[3]The former studies the short 1966-67 West German school year and documents an increase in repetition rates in primary school as well as a reduction in enrollment to higher secondary school tracks, but no effects on earnings and employment. The latter examines the long 1978-79 Indonesian school year and reports a reduction in repetition rates and improved educational attainment, with positive effects also on wages and on the probability of working in the formal sector.

[4]Cerdan-Infantes and Vermeersch [2007] on Uruguay, Almeida et al. [2016] on Brazil and Hincapie [2016] on Colombia.

strategy to assuage concerns about student endogenously sorting into schools offering the FSD. Furthermore, we examine how schools' types and students' characteristics affect returns to longer schedules and explore some mechanisms that could explain the heterogeneous effects.

1.3 Institutional Setting

1.3.1 The Chilean School System

The Chilean school system features two education cycles: primary education (grades 1-8) and secondary education (grades 9-12). Standardized tests called SIMCE assess pupils' knowledge and skills in core subjects at the end of various grades. The testing frequency is highest in fourth grade, with tests taking place every year since 2005.[5]

Education is provided by three types of schools: public schools, charter schools and non subsidized private schools. Public schools are free and are funded through student vouchers.[6] Charter schools are private, but they are publicly subsidized through the voucher system as well. Since 1994 charter schools can also charge tuition fees, but the size of the voucher decreases as tuition fees increase. Non subsidized private schools are funded only through tuition fees and are usually substantially more expensive than charter schools.

The FSD reform applies to public and charter schools, which serve more than 90 per cent of the students attending regular programs in the school system.[7] Despite both being publicly subsidized, they are different in how they are managed and regulated. Public schools are either managed by the Municipal Department of Education (*DAEM*) or by non-profit Municipal corporations.[8] The working conditions are regulated by a

[5]Fourth graders were also tested in 1999 and 2002.

[6]During the 1980s the Chilean school system experienced a major transformation that transferred the administration of public schools from the Ministry of Education to Municipalities. Furthermore, the funding system was changed by introducing a voucher that could be used both in public and charter schools.

[7]This figure excludes education for adults, education for students with specific disabilities and other types of special programs.

[8]While the director of the *DAEM* is usually a teacher appointed by the Municipality, corporations are led by a board of directors who do not need to be teachers and whose president is the major of the Municipality.

labor code specific to education professions.[9] Charter schools are private organizations and, accordingly, the working conditions of teachers are regulated by the private sector labor code.[10] Different regulations translate into charter schools having greater autonomy and flexibility in the management of the teaching staff, in terms of recruiting, dismissal and compensation policies. They also enjoy more responsibility and freedom over the design of the curriculum. In Appendix section 1.A we discuss in more detail the main regulatory differences between public and charter schools. We also provide further evidence when exploring the differential effect of the FSD by type of school in sub-section 1.6.3.

1.3.2 The FSD Reform

In 1997 the Chilean government decided to increase daily instruction time in all publicly subsidized primary and secondary schools (i.e. public schools and charter schools), whilst leaving the term length and the national curriculum unchanged.[11] The Full School Day (*Jornada Escolar Completa*, or FSD henceforth) reform aimed at improving the quality of education and reducing inequality in learning outcomes.[12] The reform envisaged a substantial increase of instruction hours (which last 45 minutes). Specifically, in primary schools 8 instruction hours were added in grades 1 to 6 and 5 hours in grades 7 and 8; in secondary schools, grades 9 and 10 experienced a 9-hour increase of instruction time per week.[13] In grades 1 to 4, this translated into a 26.7 per cent increase of weekly instruction time. As a result, in 2015 the length of school days in Chilean primary schools was the highest among OECD countries, when considering total compulsory instruction time [OECD, 2016b].

Schools could choose when to implement the FSD.[14] The deadline was initially set to 2002. However, it was later extended and differentiated by type of school and student: 2007 for all public schools and for charter schools catering to disadvantaged pupils,

[9] This is called *Estatuto de los Profesionales de la Educación*.
[10] This is called *Código del trabajo*.
[11] Increasing daily instruction time is not mandatory in grades 1 and 2.
[12] The Law 19494 that introduced the FSD was published on January 25, 1997.
[13] In grades 11 and 12, 6 instruction hours were added in the scientific-humanities track, while 4 hours were added in the vocational track.
[14] Schools could also adopt the FSD in different grades at different times, but they were mandated to ensure that pupils who started attending the longer school day in a given grade would then also be exposed in all following grades.

2010 for the rest of charter schools. Yet, by 2013 —the last year of data available to us— there were still schools operating under the old scheme. Figure 1.1 illustrates the pattern of adoption of the FSD between 1997 and 2013 for primary schools. For every year, it shows the number of schools, as well as the share of public and charter schools, that had adopted the policy by that year. The two types of schools display similar patterns of adoption, although a larger share of public schools had implemented the FSD by 2013.[15]

By the time the reform was announced many schools were operating a two-shift scheme: some grades were taught in the morning and some in the afternoon. The increased instruction time and the longer school day required a change to a one-shift scheme, where all pupils attend school from the morning to mid-afternoon. Table 1.1 illustrates the daily school schedules with and without the FSD, inclusive of time devoted to breaks. Without the FSD pupils spend at least 5 hours per day at school. The typical morning shift runs from 8.00 to 13.00, while the typical afternoon shift runs from 14.00 to 19.00. Under the FSD students spend at least 7.08 hours per day at school. If the school adopts the FSD from Monday to Friday, the typical school day starts at 08.00 and ends at 15.05. If the school adopts the FSD on 4 days and the shorter school day on the remaining one, the typical longer school day starts at 8.00 and ends at 15.45.[16]

The passage from a two-shift to a single-shift scheme implies that pupils have lunch at school. For most students, however, this did not translate into a substantial change in the nutritive content of their diet. First, Chile had virtually no problem of infant malnutrition when the FSD reform was passed.[17] Second, students from disadvantaged backgrounds could also have lunch at school under the short school day scheme. The main difference between the short and the long school day is therefore the increase of instruction time, which requires adjusting the teaching input.

Table 1.2 reports weekly instruction hours per subject with and without the FSD for grades 1 to 4. It shows that the legislated increase in instruction time was not allotted

[15]By 2013 around 12 per cent of primary schools were still operating without the FSD.

[16]The minimum hours of daily instruction are prescribed by the law. Schools can freely choose the time at which the school day starts. The daily schedules in Table 1 are built assuming that the full school day and the morning shift start at 8.00, while the afternoon shift starts at 14.00.

[17]In 2000 only 2.9 per cent of children aged 0-6 suffered from malnutrition and only 0.3 per cent suffered from moderate or serious malnutrition (Mönckeberg B. [2003]).

to specific subjects, but rather allocated to the so-called "Free Choice time", which schools could decide how to use. Therefore, schools had considerable freedom over the organization of the FSD, the only constraint being the approval by the Ministry of Education of a pedagogical plan that described the use of the additional time.

We do not observe how each school allocates the additional time across subjects. However, we can provide some evidence based on a survey administered in 2005 to investigate the use of time in fifth grade at 387 urban primary schools that had already implemented the FSD at that point.[18] Drawing on this, Table 1.3 reports the allocation of weekly instruction time across curricular subjects, both for all schools (columns 1-2) and distinguishing between public (columns 3-4) and charter (columns 5-6) schools. "Core Time" excludes "Free Choice Time". It shows that schools devote a substantial portion of "Free Choice Time" to core subjects. Among those, more hours are allocated to Spanish than to mathematics.[19] A small fraction of additional instruction time is dedicated to other subjects; the remaining portion of "Free Choice time" is distributed among various extra-curricular activities (not reported in the table for brevity). Charter schools devote slightly less additional time to Spanish and mathematics. However, the allocation of additional instruction time across subjects is similar in public and charter schools. The only significant differences emerge with regards to foreign languages and religion, to which charter schools devote more of the additional instruction time.

To further investigate the effect of the FSD on schools' time use, we rely on data that reports, at the school-class-year level, the list of subjects taught.[20] Appendix Figure 1.C1 shows, in an event study framework that collapses information at the school-grade-year level, the evolution of the total number of subjects, as well as subjects related to specific disciplinary areas, around the adoption of the FSD.[21] Following the

[18]The survey was administered by the Studies Directorate of the Sociology Faculty at the Catholic University of Chile (*DESUC*) and a report based on it was written by Ruz Pérez and Madrid Valenzuela [2005].

[19]Spanish also features more instruction time also under the shorter school day.

[20]This data is available at `http://datos.mineduc.cl/dashboards/19923/bases-de-datos-de-planes-y-programas-de-estudios-anos-2002-al-2016/`.

[21]The event study specification reads:

$$Y_{gst} = \gamma_g + \eta_s + \theta_t + \lambda_{sg} + \phi_{st} + \mu_{gt} + \sum_{\rho=-5}^{-2} \beta^\rho \mathbb{1}(p_{gst} = \rho) + \sum_{\rho=0}^{4} \beta^\rho \mathbb{1}(p_{gst} = \rho) + \varepsilon_{gst} \quad (1.1)$$

introduction of a longer school day in a given grade, the number of subjects taught in that grade increase by a small but statistically significant amount, up to almost 0.1 four years after the implementation of the policy. The increase is driven by the fact that there are more subjects related to foreign languages as well as more tutorials and workshops. While the number of subjects is an imperfect measure of how schools use the additional time, as they could simply increase the hours devoted to each subject, this provides further evidence that the longer school day translated into an increase of instruction time.

Augmenting instruction time and lengthening the school day generated additional operational costs, which were funded through an increase in the baseline vouchers, by 25-50 per cent depending on the grade (Table 1.2).[22] Some schools also had to expand their infrastructure in order to switch to the single-shift scheme. Infrastructure-related costs were funded through *ad-hoc* additional resources, which were allocated through public tenders organized by the Ministry of Education and its regional offices. Priority was usually granted to schools catering for students from lower socio-economic backgrounds.[23]

1.4 Empirical Strategy

In order to study whether increased instruction time and a longer school day affect achievement, we exploit the fact that we observe several cohorts of pupils starting primary education in a given school in the 2002-2010 period and then taking a standardized test at the end of grade 4 – possibly in a different establishment – over the period 2005-2013. Since we can follow the entire school career of each one of such students, we can compute actual years of exposure to the FSD by the end of grade 4

g, s and t index the grade, the school and the year, respectively. $p_{gst} = t - E_{gs}$ is the distance (in years) from the event, which is the introduction of the FSD in grade g of school s. Controls consist of grade (γ_g), school (η_s) and year (θ_t) fixed effects, as well as their interactions. The FSD is adopted in event-year 0 and coefficients β^ρ show how different the number of subjects taught is in event-year ρ relative to event-year -1, which is taken as the reference year. Standard errors are clustered at the school-grade level.

[22] The final amount that a school receives through student vouchers also depends on its location, size, and other characteristics. We report the increase of the baseline voucher, because this was the change common to all schools.

[23] Yet schools serving pupils from higher socio-economic backgrounds were less likely to need infrastructure-related investments.

as $ExpFSD4_i = \sum_{j=1}^{4} d_i^j$, where d^j is a dummy that takes value 1 if the pupil is ever exposed during grade j to the FSD.[24]

We then estimate the following specification:

$$Y_{ist} = \eta_s + \theta_t + \beta ExpFSD4_{ist} + \gamma X_{ist} + \delta Z_{st} + \varepsilon_{ist} \tag{1.2}$$

Y_{ist} is the test score of student i who starts primary school in school s in year t and then takes the standardized test at the end of grade 4. η_s is a set of school fixed effects that account for time-invariant heterogeneity across schools; θ_t is a set of year fixed effects that control for common unobserved year-specific shocks. In the richest specifications, we also include a set of controls at the student and at the school level. Specifically, X_{ist} is a vector of student characteristics measured in first grade, which include: gender, age, attendance rate and end-of-year status (i.e. promotion to second grade or retention in first grade). Z_{st} averages student characteristics contained in X_{ist} at the school level. It also includes enrollment and average class size in first grade. Standard errors are clustered at the school level.

By including first-grade school fixed effects, specification (1.2) leverages variation in exposure to the FSD by the end of grade 4 across cohorts of students who enrolled in the same establishment. It therefore exploits the fact that, depending on whether the school adopted longer schedules within our sample period (i.e. by 2013) and on the exact year of adoption, adjacent cohorts of enrollees could experience a different exposure to the FSD before taking the test. This source of variation can be used to estimate the causal effect of the FSD on learning outcomes if cohorts of pupils are not systematically different along characteristics that are not taken into account in specification (1.2) and that correlate both with years exposure to increased instruction time and with achievement.

Given the staggered adoption of the FSD across schools, a first concern is that parents would factor the availability of the longer school day into their preferences about the

[24]The treatment is therefore more precisely defined as the number of grades attended at least once under the FSD scheme by the end of grade 4. Throughout the paper, we use the term years of exposure to the FSD for brevity. Moreover, the two definitions are exactly equivalent for non repeaters, who constitute the largest majority of the sample (88 per cent).

school in which to enroll their children. This could affect the composition of pupil intake, possibly along dimensions that our set of controls cannot account for. According to parent surveys administered alongside the test in 2005, the FSD was the most important reason for enrolling their child in a given school for only 1.97 per cent of parents. Proximity to home (29.84 per cent), the presence of a relative in the school (18.61 per cent) and the school's prestige (18.48 per cent) were cited as the most important determinants of school choices.[25]

Nonetheless, we address this concern by restricting our analysis to *incumbent* pupils. This means that we only consider pupils who enroll in first grade in a given school before such school adopts the FSD. As an example, if a school adopts the longer school day in 2007, we discard students who start primary education in that school in 2007 or later. Cohorts who enrolled before 2007, on the other hand, made their decision before the introduction of the FSD and possibly became exposed to it at some point in their school career. For *incumbent* students who never repeat a grade the range of the treatment variable ($ExpFSD4_{ist}$) is 0-3, as exposure to the FSD can start as early as grade 2. For repeaters, on the other hand, the range is 0-4; the variable takes value 4 in cases when the school adopts the FSD in the year when the pupil is repeating first grade. Furthermore, restricting the sample to *incumbent* students implies that first-grade controls included in specification (1.2) are observed before the adoption of the FSD and hence are pre-determined with respect to the treatment. The focus on legacy enrollment cohorts also characterizes recent studies on the effects of charter takeovers in the US and of academy conversions in England [Abdulkadiroğlu et al., 2016, Eyles et al., 2017, Eyles and Machin, 2018]. This restriction attenuates identification issues related to unobserved changes in pupil intake, the more so the less parents can anticipate the exact year in which a school will adopt the FSD.

Students can move across schools and in Chile school transfers are indeed a common phenomenon; in our master sample (described in section 1.5) around 35 per cent of students change school between grades 1 and 4. Pupils who enroll in first grade in the same establishment and in the same year can therefore experience a different exposure

[25] The cost (9.35 per cent) and the ethical values (8.95 per cent) of the school follow in the ranking. The presence of the FSD is ranked seventh among fifteen options.

to the FSD by the time they reach grade 4. Furthermore, if mobility across schools is influenced by the availability of longer schedules, student-level actual exposure to the FSD could be correlated with other unobserved determinants of achievement. To mitigate this concern, in our preferred specification we instrument actual exposure to the FSD with the exposure a student would have accumulated had she never transferred from the school where she attended first grade. The instrumental variable is therefore $PotExpFSD4_i = \sum_{j=1}^{4} d_s^j$, where d_s^j is a dummy variable that takes value 1 if student i would have ever been exposed to the FSD in grade j, had she remained in school s, where she started first grade.[26]

When discussing results in section 1.6, we will show that the instrument is relevant, as there is a positive and statistically significant relationship between "potential" exposure and actual exposure to the FSD by grade 4. By relying on this instrument we aim to isolate and exploit the variation in actual exposure that is not affected by possibly endogenous mobility decisions of *incumbent* pupils after first grade. Moreover, we assume that "potential" exposure is not systematically correlated with unobserved determinants of achievement and affects fourth grade test scores only through its impact on actual exposure.

A remaining concern is that the timing of adoption may depend on past performance. For example, if schools switch to the longer school day after they observe a cohort of pupils faring particularly poorly at the test, our estimates may simply capture mean-reversion effects. In general, there can be concerns about confounding effects of underlying school-specific trends in test scores. We show in section 1.5 that there are no visible clear trends in reading and mathematics scores in the years preceding the switch to longer schedules. Another concern is that other events may take place at the school around the time of FSD adoption, which could also affect learning outcomes in the following years. We discuss and address these further issues in Appendix

[26]To build this "potential" measure of exposure we also assume that the student would have never repeated, as we do not observe the pattern of repetitions in this counterfactual school career. Therefore, the range of the instrumental variable for all *incumbent* pupils is 0-3. The first stage regression specification reads:

$$ExpFSD4_{ist} = \eta_{0,s} + \theta_{0,t} + \beta_0 PotExpFSD4_{ist} + \gamma_0 X_{ist} + \delta_0 Z_{st} + \epsilon_{ist} \tag{1.3}$$

section 1.B, where we show that findings remain similar when we restrict our attention to pupils starting primary school in establishments that did not receive public funds for expanding infrastructure, when we control for another policy targeting disadvantaged students implemented in 2008 and when we focus our attention only in cohorts of students not exposed to such policy.

1.5 Data and Sample

We link several administrative and survey datasets on account of unique school, student and teacher identifiers.

Data on achievement in fourth grade comes from a nationwide standardized low-stakes test (SIMCE test) designed by the Education Quality Agency (*Agencia de Calidad de la Educación*).[27] It is administered at the end of the school year and is marked by external examiners, therefore leaving little room for test score manipulation. Individual records on performance in the test are available for fourth grade students in 1999, 2002 and then with a yearly frequency from 2005 onward. We restrict our attention to the 2005-2013 waves of the test. The reason being that we can follow students' school careers only for cohorts who start primary school from 2002 onward; this is necessary both to correctly identify *incumbent* students (i.e. pupils who enroll in first grade in a school that has not yet adopted the FSD) and to compute actual exposure to the FSD for students who move across schools between grades 1 and 4. 2013 is the last year of data available to us. We use pupil-level test scores in the reading and mathematics sections of the test as our measure of achievement. Scores are standardized by year and subject to have a mean equal to 0 and a standard deviation equal to 1. Alongside the test, surveys are administered to students and their parents, as well as to teachers. Based on questions that are consistent across all waves of the parent survey, we recover a rich set of information on pupils' backgrounds as of grade 4 that we use to study heterogeneity by students characteristics. Based on teacher surveys, we provide evidence on the frequency of homework assignments in schools with and without the FSD.

[27] While the stakes are low because the test does not impact a student' final evaluation, schools care about it because school-level average scores are publicly available for consultation.

The second source of information is the register of pupils enrolled in the school system over the period 2002-2013, which is maintained by the Ministry of Education. Besides gender and date of birth, for every school year it records information about the school that the student attends, the attendance rate and the end-of-year status (i.e. promotion to the next grade or retention in the same grade). We also have access to the register of educational establishments, from which we recover the location (an urban or rural area) and the administrative status of the school (public, charter or non subsidized private). A companion dataset records the year of adoption of the FSD at the school-grade level over the period 1997-2013. Based on these sources, we reconstruct the school career from grade 1 to grade 4 of every student who started primary school between 2002 and 2010 and took the fourth grade test between 2005 and 2013; we then compute the actual years of exposure to the FSD by the end of grade 4, as well as the exposure a student would have experienced had she never transferred from her first grade school. We also retrieve the set of first grade student- and school-level characteristics that we include in the richest regression specification. In order to distinguish charter schools with and without tuition fees, we rely on a dataset maintained by the Ministry of Education that records all the subsidies that schools received from the government over the 2005-2013 period. Since charter schools that charge tuition fees receive reduced subsidies, we can distinguish them from schools that do not charge tuition fees.[28]

We also exploit the information contained in the register of teachers, which is available for the period 2003-2013. We draw on this dataset to study how no-fee charter and public schools adjust the number of teachers and their working hours after the adoption of the FSD. We also rely on the 2005 Longitudinal Teachers Survey (*Encuesta Longitudinal Docente*) to investigate differences in teachers' opinions on the FSD;[29] based on the 2015 Time -Use Survey (*Encuesta Nacional de Uso de Tiempo*), we examine if academic support outside of school varies across pupils from different socio-economic

[28]We classify a charter school as a no-fee charter school if it never charged fees between 2005 and 2013. For 2.39 per cent of charter schools attended in first grade by students belonging to the master sample we do not find information about the tuition fees in the dataset.

[29]The Longitudinal Teachers Survey was implemented by the Microdata Center of the University of Chile.

backgrounds.[30]

Finally, we digitized from primary sources the list of schools that received additional funds to expand their infrastructure when lengthening the school day; we parsed the releases of the Official Journal (*Diario Oficial*) published by the Interior Ministry over the period 1997-2004 and searched for the outcomes of all public tenders through which *ad-hoc* resources for infrastructures were assigned.[31] Based on this, we create a dataset that records, for every school, the year in which resources were disbursed and the amount received, if any. Since 2008, students from disadvantaged backgrounds are granted additional subsidies (PSS) on top of the vouchers. We obtain the list of beneficiaries from the Ministry of Education. This information is used to perform robustness checks described in Appendix section 1.B.

In order to create the master sample of our analysis, we restrict our attention to *incumbent* pupils, i.e. students who started first grade in a given school when the FSD had not been yet introduced (see section 1.4). Furthermore, we discard students that attended non subsidized private schools at some point between grades 1 and 4. This is motivated by the fact that the FSD reform only applies to publicly subsidized schools. Moreover, we do not know whether a given non subsidized private school was already offering a longer school day or started providing it at some point after it became compulsory for other types of schools. Therefore, students attending non subsidized private schools cannot serve as a control group.

The master sample consists of around 600,000 fourth-grade test takers; they started primary school between 2002 and 2010 in schools that had not yet adopted the FSD and took the test between 2005 and 2013. It follows that schools attended by pupils in the master sample had not switched to the longer school day by 2002. Given that the first transitions to the single-shift occurred in 1997, our sample of schools consists of mid to late adopters.

As discussed in section 1.4, a threat to identification could arise if schools adopted the FSD based on the trend or transitory component of test scores. Figure 1.2 plots

[30]The database of the 2015 Time-Use Survey can be downloaded from `https://www.ine.cl/estadisticas/menu-sociales/enut`.

[31]The last tender took place in 2004.

coefficients from an event study exercise where the specification reads:

$$Y_{ist} = \eta_s + \theta_t + \sum_{\rho=-5}^{-2} \beta^p \mathbb{1}(p_{st} = \rho) + \sum_{\rho=0}^{4} \beta^\rho \mathbb{1}(p_{st} = \rho) + \varepsilon_{ist} \tag{1.4}$$

Y_{ist} is the reading or mathematics score of student i who takes the SIMCE test in school s in year t. $p_{st} = t - E_s$ is the distance (in years) from the event, which is the introduction of the FSD in at least one grade in school s. η_s and θ_t are school and year fixed effects, respectively. The FSD is implemented in event-year 0 and coefficients β^ρ show how different scores in event-year ρ are relative to event-year -1, which is taken as the reference year. Schools are observed up to 5 years before and after the introduction of the longer school day and the sample consists of all schools where students in the master sample enrolled in first grade.[32] For both subjects, there appear not to be evident trends in the pre-adoption period, suggesting that test scores were not trending either downward or upward before schools decided to implement longer daily schedules. Furthermore, there are no evident spikes or dips in test scores just before the introduction of the FSD. On the other hand, from event-year 1 scores start increasing, suggesting a positive effect of the FSD on achievement. We will then provide a formal estimation based on our identification strategy in section 1.6.

Table 1.4 reports summary statistics for pupils in the master sample. Column (1) pools all students together, whereas columns (2) to (4) split schoolchildren according to the type of school (public, charter without tuition fees or charter with tuition fees) they attended in first grade. In the vast majority of households (87 per cent) parents do not have university education.[33] Only 15 per cent of students have more than 50 books at home; 55 per cent of the households have a computer at home and slightly less than one third also have a connection to the Internet.[34] The first grade attendance rate is

[32]The sample is unbalanced, meaning that not all schools are observed in every event-year. Given the calendar of SIMCE tests, using a balanced sample would significantly reduce the number of event-years that we can observe. For this exercise, we also use the 1999 and 2002 waves of the SIMCE test.

[33]We build a variable that measures parental education by setting $ParentalEd = max(MotherEd, FatherEd)$, where $MotherEd$ and $FatherEd$ are the highest mother and father's academic attainment, respectively; if the information for either one of the two parents is missing, we rely on the level of education achieved by the other parent.

[34]Information about students' backgrounds, i.e. parental education and resources at home, is drawn from parent surveys. Since these variables are observed at the end of grade 4, they could be affected by a student's exposure to the FSD (for example, if longer school days have

very high (94 per cent) and 3 per cent of pupils repeat first grade. On average, there are 35 students in a first grade class.

When splitting students according to the type of establishment they started primary school in, it emerges that public schools and charter schools without tuition fees cater to relatively similar students. On the other hand, schoolchildren attending charter schools that charge tuition fees live in more affluent households. Test scores are lowest in public schools and highest in charter schools with tuition fees.

1.6 Results

In this section we first discuss the average effect of the FSD on all pupils in the master sample (sub-section 1.6.1). We then explore whether and how effects are heterogeneous depending on the resources available in the household where the student lives (sub-section 1.6.2) and on the type of school she attends (sub-section 1.6.3). Heterogeneous effects of additional instruction time are interesting to study because increasing the amount of time that pupils spend at school reduces the amount of time they can devote to other activities outside of school. The return to longer school schedules therefore could depend on the absolute quality of time use at school, which can vary across schools, and on its relative quality with respect to time use outside of school, which can vary among students.

When investigating heterogeneous effects, we estimate a richer version of specification (1.2), whereby we also interact the treatment and all controls with a dummy variable D that captures a given dimension of heterogeneity.[35] We report estimates coming from the preferred linear IV specification that includes the full set of controls (which we denote as FE-IV2).[36]

an effect on parents' labour supply); for this reason, they are not included in the regression specifications, which only feature pre-determined controls. Furthermore this information is missing for around 15 per cent of schoolchildren in the sample.

[35] A fully interacted specification yields estimates that are equivalent to those obtained from separately estimating two regressions on the sub-sample where $D = 0$ and on the sub-sample where $D = 1$.

[36] $ExpFSD4$ and $ExpFSD4 \times D$ are instrumented with $PotExpFSD4$ and $PotExpFSD4 \times D$, respectively.

1.6.1 The Effect of the FSD on Achievement

Table 1.5 reports results from regression specification (1.2). We start by discussing coefficients when we estimate the most parsimonious specification, which only includes school and year fixed effects, and we do not instrument actual years of exposure to the FSD (specification FE1, column 1). These estimates point to a virtually null effect on reading and a negative impact on mathematics. Including pre-determined controls listed in section 1.4, however, changes the picture significantly (specification FE2, column 2): the effect of an additional year of exposure to the FSD is positive for both subjects, although it is only statistically significant for reading (0.011σ). This indicates that controlling for first-grade status (pass or repeat) is important because repeaters, who are low performers, spend more years at school and are therefore more likely to be exposed to the FSD at some point.

As mentioned in section 1.4, a non negligible fraction of students transfer from one school to another between grades 1 and 4. Furthermore, the availability of longer daily schedules appears to influence mobility across schools. Appendix Figure 1.C2 shows the evolution of transfers of pupils attending grades 1 to 4 at the school-year level, in a 5-year window around the implementation of the FSD. The estimated event study specification reads:

$$Y_{st} = \eta_s + \theta_t + \sum_{\rho=-5}^{-2} \beta^p \mathbb{1}(p_{st} = \rho) + \sum_{\rho=0}^{4} \beta^\rho \mathbb{1}(p_{st} = \rho) + \varepsilon_{st} \tag{1.5}$$

Following the introduction of longer schedules schools experience a decline in the outflow of pupils; at the same time, although the pre-adoption pattern is more scattered, inflows of students appear to increase, with a spike in the year of adoption. As a result, net transfers (i.e. the difference between transfers into and transfers out of a given school) grow, by up to 5 pupils per year. Appendix Table 1.C1 further shows that, among schoolchildren who belong to the master sample, those who transfer are negatively selected, as they have a slightly lower attendance rate in grade 1 (93 per cent versus 95 per cent) and are more than twice as likely to repeat first grade.[37] More-

[37]First-grade attendance rates have a very low dispersion, so that a 1 percentage point difference amounts to almost one fifth of a standard deviation.

over, it emerges that pupils tend to transfer towards schools that offer the FSD; while transferring and non transferring students have a very similar "potential" exposure (i.e. the exposure they would have experienced had they remained in their first-grade schools), the former end up with a much higher actual exposure.[38] Partly because of fewer transfers out and more transfers in, the number of students per class in grades 1 to 4 increases after the adoption of the FSD (Appendix Figure 1.C3) by an amount that is however modest (at most around 1.50 more pupils) when compared to the average class size in primary schools.

These patterns motivate the decision to instrument actual exposure ($ExpFSD4$) with the exposure a student would have experienced had she never transferred ($PotExpFSD4$). When adopting the IV approach, estimates are remarkably stable across the most parsimonious specification (FE-IV1, column 3) and the specification featuring all controls (FE-IV2, column 4). An additional year of exposure to the FSD significantly raises reading test scores by 0.024σ. The effect on mathematics test scores lies in the narrow range 0.007-0.008σ, but is not statistically significant. As shown in the same table, the instrumental variable displays a positive and strong relationship with the treatment, as the first stage coefficient is statistically significant and equal to 0.72, implying that for slightly less than 30 per cent of pupils the real exposure and the "potential" exposure do not coincide.

In columns (5) and (6) of Table 1.5, we relax the assumption that every additional year of exposure has the same effect on achievement. We estimate the preferred IV specification in a fully non parametric way, by introducing a set of dummies for every possible level of exposure to the FSD and setting 0 years of exposure as the reference category.[39] The non parametric specification reveals that the effect of longer schedules

[38]This also holds true also when restricting the comparison to students who never repeat between grade 1 and grade 4.

[39]The non-parametric specification therefore reads:

$$Y_{ist} = \eta_s + \theta_t + \sum_{k=1}^{2} \beta_k \mathbb{1}(ExpFSD4_{ist} = k) + \beta_3 \mathbb{1}(ExpFSD4_{ist} >= 3) + \gamma X_{ist} + \delta Z_{st} + \varepsilon_{ist} \quad (1.6)$$

Specification (1.6) highlights that we collapse 3 and 4 years of actual exposure into a unique category, as only very few pupils (i.e. students who repeat first grade in the year when the school adopts the longer schedules) attend all 4 grades under the FSD scheme. In the IV specification, the set of dummies that capture every possible level of actual exposure to the FSD are instrumented by a set of dummies that capture every possible level of exposure a increases more than linearly with exposure. Three years of exposure are associated with a 0.114-0.116σ increase of reading test scores, significant at the 1 per cent level, and a 0.057-0.058σ increase of mathematics test scores, significant at the 5 per cent level.

The IV estimates therefore show that the FSD has a positive effect on learning outcomes, which increases more than linearly with exposure and is stronger for reading than for mathematics. The stronger impact on reading may depend on the fact that a larger fraction of additional instruction time is devoted to Spanish than to mathematics (Table 1.3). The pattern of coefficients in the fully non parametric specification is consistent with added instruction time in earlier grades having a positive effect on achievement in later grades. Moreover, as the passage from a two-shift to a one-shift scheme implies a re-organization of daily routines, it may also be explained by the presence of adaptation costs that eventually fade away over time.

A possible remaining concern is that other events may happen in a school around the adoption of the FSD and affect learning outcomes in the following years. In Appendix section 1.B we show that our estimates are robust to: *i)* restricting the attention to pupils who started primary schools in establishments that most likely did not expand their infrastructure at the same time when the FSD was adopted; *ii)* controlling for a policy granting further subsidies to disadvantaged schoolchildren since 2008.

1.6.2 Heterogeneity by Students' Backgrounds

In this sub-section we explore whether the effect of the FSD varies depending on the characteristics of the environment students are exposed to when they are not in school. We focus our analysis on the role of household resources, as reflected by parental education and the availability of books and ICT technologies at home.[40] We rely on this information to distinguish schoolchildren from a more privileged background (for whom $D = 0$) from others (for whom $D = 1$).

Table 1.6 shows that longer schedules appear not to benefit in a significant way pupils

student would have experienced had she never transferred out of her first grade school.

[40]This information is drawn from parent surveys administered alongside the test. The non-response rate is similar across the variables and is around 15 per cent. This explains the smaller sample size.

from more advantaged backgrounds. An additional year of exposure to the FSD does not raise by a statistically significant amount reading and mathematics scores for children living in households where at least one parent has some university education (columns 1 and 4), there are more than 50 books (columns 2 and 5), or both a computer and a connection to the Internet are available (columns 3 and 6). On the other hand, reading scores increase by a significant amount for pupils living in households where neither parent has any university education (0.022σ), there are at most 50 books at home (0.022σ), and either a computer or a connection to the Internet is not available (0.024σ).[41] Also mathematics scores increase by a larger amount, which however never becomes significantly different from 0.

It has to be noted, however, that the documented difference, as captured by the interaction term $ExpFSD4 \times D$, although large in size, is not statistically significant. With this caveat concerning the precision of the estimates in mind, the analysis provides suggestive evidence that returns to an additional hour of instruction time tend to be larger for students who have fewer resources and opportunities available at home and for whom, therefore, the relative quality of time spent at school is higher.

Drawing on information coming from the 2015 Chilean Time-Use Survey, Appendix Table 1.C2 shows that pupils from privileged backgrounds indeed receive more support outside of school. We restrict our attention to households where there is at least one child aged 5-18 and we divide them into two groups, depending on whether either the head of the household or the head's spouse has any university education ($U_{hh} = 1$) or not ($U_{hh} = 0$). In households where $U_{hh} = 1$, the percentage of heads of household and heads' spouses who declare that they help their children with their homework is 48 per cent, whereas this percentage drops to 33 per cent in households where $U_{hh} = 0$ (column 1). Summing up the minutes that they dedicate to helping with homework on a given day of the working week and on a given day at the weekend, there is a 14-minute difference in favor of households where $U_{hh} = 1$ (column 2). Assuming a uniform distribution of help across the days of the week, this would translate into a difference of around 50 minutes per week. It is also interesting to look at support by

[41] These figures are the sum of coefficients related to the main term $ExpFSD4$ and the interaction term $ExpFSD4 \times D$. They are significant at the 1 per cent level.

other providers, in the form of tutoring outside of school. 5 per cent of pupils aged 12-18 and living in households where $U_{hh} = 0$ receive some tutoring, as opposed to 12 per cent of students living in households where $U_{hh} = 1$ (column 3). In terms of minutes per day, the former receives tutoring for less than half the time than the latter (column 4).

Appendix Table 1.C3 draws information about the frequency of mathematics homework from the teacher surveys administered alongside the SIMCE test in 2011, 2012 and 2013. The limited period for which this information is available does not allow us to study the evolution of homework's frequency around the adoption of the FSD in an event study framework such as the one in (1.5). Panel A considers all schools and shows that the frequency of homework is lower in schools with the FSD than in establishments without it. For example, the percentage of teachers assigning homework after every class is roughly 20 per cent in schools where the FSD is not in place, while it drops to about 12 per cent in schools that feature longer schedules. In panel B we restrict our attention to schools that had not adopted longer schedules by 2011 and we compare the frequency of homework between the years 2011 and 2013. In establishments that did not adopted the FSD in 2012 or 2013 (column 2), the frequency is very similar in the the two years. On the other hand, homework is assigned much less frequently in 2013 than in 2011 in establishments that switched to the FSD by that year (column 3).[42] Overall, there is therefore suggestive evidence that longer school schedules are associated with less homework. If the productivity of homework is higher for schoolchildren from advantaged backgrounds, because they have more support at home, the reduction of its frequency that seems to be associated with longer school schedules could be one of the mechanisms that explains the documented heterogeneity.

1.6.3 Heterogeneity by School Type

The absolute quality of time use is likely to be the primary driver of additional instruction time's effectiveness. It is therefore important to study the contribution

[42] As an example, in 2011 around 52 per cent of teachers working in schools that had not adopted the policy declared that they had assigned homework after almost every class. This figure remained the same in schools that had not adopted the FSD by 2013, while it fell to 31.82 per cent in schools that adopted it in 2012 or 2013.

of school characteristics in shaping returns to longer schedules. The Chilean school system provides an attractive setting because public and charter schools, whilst being both publicly subsidized, differ in terms of ownership and in the degree of autonomy. As explained in sub-section 1.3.1, charter schools have more autonomy than public schools over the management of school resources and the design of the curricula. Appendix Table 1.C4 reports answers to school surveys administered alongside the 2006 and 2009 waves of PISA tests, which ask about the tasks over which the principal or the governing body of the school have considerable responsibility. The sample consists of all public and charter schools that offer primary education.[43] It confirms that principals as well as the governing bodies of charter schools have greater autonomy in designing the curricula, as they can decide the course offer and the course content more frequently. Moreover, they are more likely to be responsible for the budget formulation and allocation. They are also in charge of taking personnel decisions, in terms of recruitment, promotions and dismissals.

We therefore study whether the FSD has a differential effect in public and charter schools. To this end, we compare pupils who attended a public school in grade 1 (for whom $D = 0$) to students who enrolled in a charter school that does not charge tuition fees in grade 1 (for whom $D = 1$). The choice of focusing on these two types of schools is motivated by the fact that, as shown in Table 1.4, charter schools that charge tuition fees cater to more affluent pupils, whereas public schools and charter schools without tuition fees serve pupils of similarly lower socio-economic status. As we aim to uncover the role of school institutions, we do not want to capture differences related to students' characteristics.

Table 1.7 shows that returns to additional instruction time are higher for students starting primary school in no-fee charter schools than for those who enrolled in public schools. The difference, as captured by the interaction term $ExpFSD4 \times D$, is sizable for both subjects and statistically significant with regards to reading (columns 1 and 5). The effect of an additional year of exposure to the FSD on reading test scores is more than three times larger for students starting primary school in no-fee charter schools

[43] PISA tests are administered to pupils aged 15. We therefore restrict our attention to secondary schools that also offer primary education, which explains the very small sample size.

(0.061σ) as opposed to in public schools (0.019σ). Mathematics scores are raised by a statistically insignificant 0.012σ for pupils attending grade 1 in public schools; the coefficient more than doubles to 0.029σ for enrollees of no-fee charter schools, although the associated p-value is slightly above 0.1. Pupils attending public schools and no-fee charter schools are similar. If anything, those attending the latter type of school are slightly more affluent. In sub-section 1.6.2 we provided suggestive evidence that longer schedules are less beneficial to students from advantaged backgrounds. Hence, small differences in student characteristics across the two types of institutions should not be responsible for the large observed differences in returns to longer schedules. In columns (2) to (8) we further show that the documented differential effect remains remarkably stable when we add the interaction between exposure to the FSD and a dummy capturing parental education (columns 2 and 6), the availability of books at home (columns 3 and 7), and the availability of ICT technologies at home (columns 4 and 8).[44]

As discussed earlier, charter schools have more autonomy over the design of the curriculum and personnel decisions. According to survey evidence provided in sub-section 1.3.2, public and charter schools allocate the additional instruction time across subjects in a similar way. We therefore focus on personnel decisions. Figures 1.3 and 1.4 plot coefficients from the event study specification outlined in (1.5), where the outcomes are various measures of teaching inputs at the school-year level. Consistent with the need to provide more instruction hours, Figure 1.3 shows that total teachers' contract hours and teaching hours increase after the adoption of the FSD (top panel). Moreover, the pattern of coefficients is similar across public and no-fee charter schools and confidence intervals overlap. Indeed, Appendix Table 1.C5 shows that the difference between public and no-fee charter schools is significant only in the final event-years. When total contract and teaching hours are divided by the number of classes (bottom panel), the differences between the two types of schools are never significant.[45]

[44]In columns (2) to (8) all controls are also interacted with the dummy capturing a given characteristic of the household in which the student lives. Furthermore, specifications are estimated on the sub-sample of pupils for whom all background characteristics (parental education, number of books at home, availability of ICT technologies at home) are non missing. The number of observations is nonetheless slightly different across specifications because of differences in the number of singletons that are dropped.

[45]The coefficients presented in Table 1.C5 come from a richer version of specification (1.5), where event year dummy and calendar year fixed effects are also interacted with a dummy (D_s) taking value 1 if the school is a no-fee charter school. This specification allows to test whether

An increase in the number of total contract and teaching hours can be achieved by adjusting both the number of teachers and the number of contract/teaching hours per teacher. Figure 1.4 shows that the number of teachers grow both in public and no-fee charter schools, but the increase is significantly higher in charter schools that do not charge tuition fees. On the contrary, contract hours per teacher increase significantly more in public schools. Appendix Table 1.C5 confirms that these differences are statistically significant. Also teaching hours per teacher grow more in no-fee charter schools than in public establishments, with the difference been slightly smaller. Therefore no-fee charter schools rely more than public schools on expanding the number of teachers, whereas they resort less to increasing teachers' workload. It appears that autonomy over personnel decisions allows charter schools to adjust the teaching input in a different way.

Appendix Table 1.C6 reports teachers' opinions in 2005 about the FSD, dividing them according to the type of school (public or charter) in which they teach. Public school teachers display a lower degree of satisfaction with longer daily schedules. Only 45 per cent of them judge the FSD as "good or very good", compared to 54 per cent of charter school teachers. This may signal that the workload of teachers in public schools increases excessively following the introduction of the FSD. This could in turn negatively affect the absolute quality of additional time use in public schools, contributing to explain the lower returns to longer schedules.[46]

differences between public and charter schools are statistically significant.

[46]Bellei [2009] and Berthelon et al. [2016] find that the effect of the FSD on achievement is larger in public schools. Bellei [2009] focuses on pupils attending grade 10 in 2001 and in 2003, while Berthelon et al. [2016] examine the effect of the FSD at grade 2 in a single year (2012). In both cases, the sample of schools and the population studied are different from the ones we investigate, and the sample periods mostly do not overlap. Furthermore, to avoid that our estimates are confounded by differences in students' characteristics, we compare public schools only to charter schools that do not charge tuition fees, because we provide evidence that they serve pupils from similar backgrounds. Additionally, we show that the larger effect of the FSD found for students starting primary school in no-fee charter schools persists once we also allow longer schedules to affect schoolchildren of higher and lower socio-economic status in a different way.

1.7 Conclusions

Academics and policymakers have long been interested in understanding the determinants of school effectiveness. With the goal of improving pupils' academic achievement, many countries undertake costly educational reforms in order to expand the resources available to schools. Some other countries, including US and UK, have also implemented reforms that grant schools more autonomy.

In this paper we study whether the effects of a large-scale expansion of instruction time in Chilean primary education are different across public and no-fee charter schools. While catering for similar students, no-fee charter schools enjoy greater levels of autonomy. We find that school institutions matter in shaping returns to longer schedules and that pupils who start primary school in no-fee charter schools benefit more than those who start in public schools. In addition, as the returns to additional instruction hours depend not only on the absolute quality of time-use at school, but also on the relative quality of time spent outside of school, we further examine whether the effects of the instruction time expansion are heterogeneous across schoolchildren from different backgrounds. We provide evidence suggesting that returns could be larger for students of lower socio-economic status who have less support at home.

To address these questions, we take advantage of the introduction of the Full School Day (FSD) in publicly subsidized (i.e. public and charter) schools. We exploit within-school variation in years of exposure to longer schedules by the end of grade 4 across several cohorts of pupils starting primary education in a given school between 2002 and 2010 and taking a national level standardized test at the end of grade 4. To limit the confounding effect of changes in the characteristics of pupil intake, we restrict our attention to cohorts of *incumbent* pupils, i.e. students who started primary education in schools that had not yet adopted the FSD and who possibly became exposed to it at some stage of their first four years of education. Furthermore, we account for potentially endogenous mobility across schools after first grade by instrumenting actual exposure with the exposure a student would have experienced, had she never transferred out of the school where she attended first grade.

We first document a positive average effect of longer schedules on reading and mathematics test scores, although estimates for the latter are not statistically significant. It is important to highlight that relying on legacy enrollment cohorts for the sake of identification implies that we are able to study only the first years after the implementation of the FSD. Insofar as it takes time to adjust, short-term effects —which are interesting to study because most large-scale input expansion programs would entail some initial adaptation challenges— may be lower than long-term ones.

Exploiting the features of the Chilean setting, we explore whether school institutions affect returns to additional instruction time. To answer this question, we compare pupils who started primary education in no-fee charter schools with pupils who started in public schools. We focus on these two types of establishments because, while differing in terms of ownership and levels of autonomy, they serve similar students. Charter schools that charge tuition fees, on the other hand, cater to children of higher socioeconomic status. This choice therefore attenuates concerns that the estimates reflect differences in the types of students who attend different types of schools.

Charter schools enjoy more autonomy over the design of the curriculum and personnel decisions. Since survey evidence suggests that public and charter schools allocate additional instruction time across subjects in a similar way, we turn to staff decisions and we uncover a significant difference. No-fee charter schools and public schools increased the total contract and teaching hours required to provide longer schedules in a different way: no-fee charter schools relied more on hiring additional teachers and less on increasing working hours per teacher than public schools. If public school teachers deal more often with a higher workload, which would be consistent with them reporting a lower satisfaction with the FSD, then the absolute quality of time use in public school may be lower, thus contributing to explaining the lower benefits we document for them.

Our findings are in line with the growing literature showing that charter schools are associated with improved learning outcomes (see section 1.2). Lavy [2015] documents that the productivity of instructional time is larger in schools that have more autonomy over staff and budget decisons. We find a similar result and provide suggestive evidence that autonomy over staff decisions seems important when expanding instruction time, because providing longer schedules requires adjusting the teaching input. In general, our results suggest that school institutions and governance matter for the effectiveness of various education policies. Further analysis on complementarities between school inputs and institutions could be a promising avenue for future research.

Because increasing the amount of time pupils spend at school reduces the time for activities outside of school, we also check whether the effects of the FSD vary depending on the resources that students have available at home, as captured by parental education, the number of books, and the availability of a computer and a connection to the Internet. We show that pupils from advantaged backgrounds do not significantly benefit from longer schedules, while the rest of the students in our sample do. Although large in size, the difference in the benefits enjoyed by individuals from different backgrounds are not statistically significant. With this caveat in mind, our findings suggest that for children with fewer resources at home the relative quality of additional time spent at school is higher, thus yielding larger benefits. According to information contained in teachers' surveys, the adoption of the FSD is likely associated with a reduced frequency of homework. In light of the limited support, as shown by time-use surveys, that pupils of low socio-economic status receive outside of school, substituting autonomous study at home with supervised learning at school may be a mechanism underlying the documented heterogeneity.

These findings are in line with the results of Lavy [2015].[47] If also confirmed in other settings, they would suggest that the amount of time spent at school may play a role in reducing inequality in learning opportunities. As pupils from different backgrounds are exposed to the same school inputs for a larger part of the day, the role of household inputs —the quality of which varies greatly— may become less important. This is likely to be especially true if, as in the Chilean setting, the additional instruction time does not entail an expansion of the curriculum. Indeed, in a setting in which increased weekly instruction hours are accompanied by an expansion of the curriculum, Huebener et al. [2017] document a widening gap between high- and low-performing

[47]When restricting the analysis to a sub set of developing countries that include Chile, Lavy [2015] finds a stronger effect among schoolchildren from highly educated families. However, he does not provide country-specific estimates that allow to verify what is the estimated effect in the case of Chile.

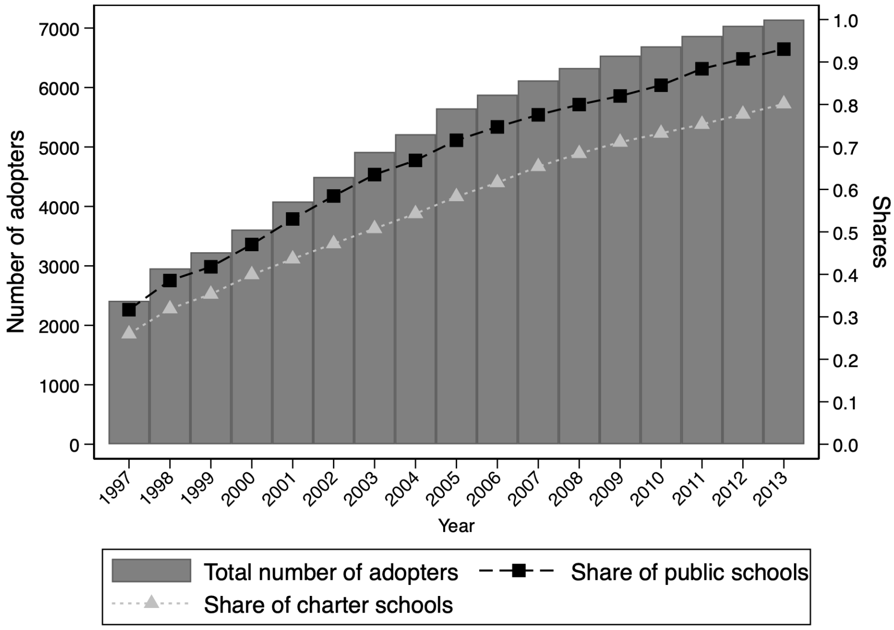
German pupils.

Finally, policies that extend the length of the school day may also affect non-academic outcomes. Berthelon and Kruger [2011], for instance, show that the FSD reduces the incidence of teenage motherhood among girls and of youth crime, with the effects concentrated among poorer families; Contreras and Sepúlveda [2016] report a positive effect of the FSD on labor force participation and employment of single mothers whose youngest child is eligible to attend longer schedules. Studying these outcomes goes beyond the scope of this paper, but it is important to bear them in mind when evaluating this type of policies.

Acknowledgments

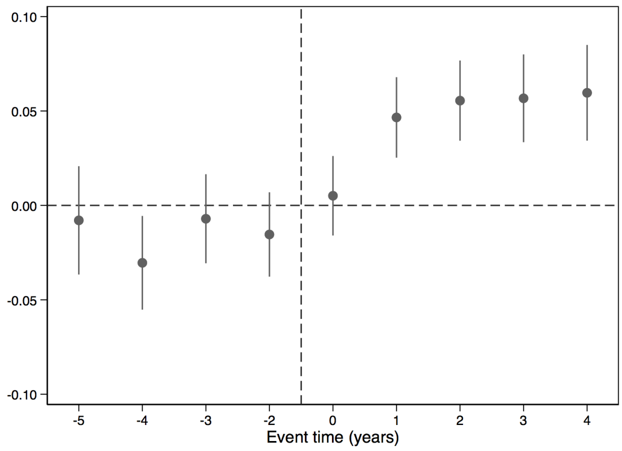
We thank Steve Pischke, Sandra McNally, Steve Machin, Esteban Aucejo, Daniele Checchi, Andy Eyles, Henrik Kleven, Guy Michaels, Alan Manning, Camille Landais and Johannes Spinnewijn for many useful comments. We are grateful for the comments received at the LEER Workshop on Education Economics 2017, the International Workshop on Applied Economics of Education 2017, the Annual Conference of the European Association of Labour Economists 2017, the Bank of Italy Human Capital Workshop 2018 and several LSE internal seminars. We are also grateful to the Chilean Ministry of Education and to the Chilean Education Quality Agency for giving us access to the administrative data we use in this project. We also thank the Microdata Center of the University of Chile, and the DESUC of the Catholic University of Chile for giving us access to survey data. Finally we are also grateful for the comments made by researchers of the Center of Public Studies and of the Research Center of the Chilean Ministry of Education. The views expressed herein are those of the authors and do not necessarily reflect the views of the London School of Economics or of the Bank of Italy.

Figure 1.1: FSD Adoption over the Period 1997-2013

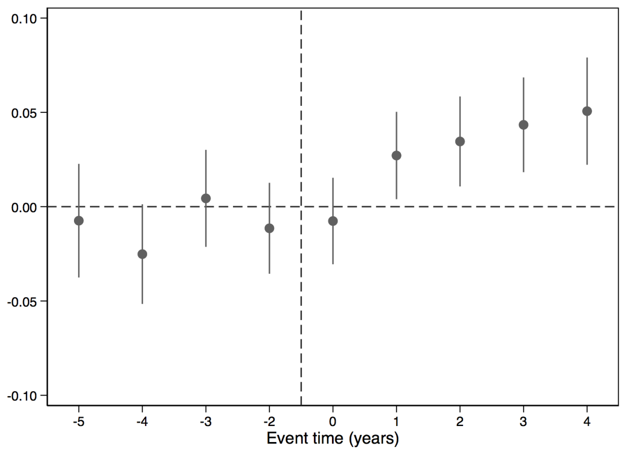


Notes: The figure illustrates the pattern of adoption of the FSD in primary publicly subsidized schools over the period 1997-2013. On the left axis it plots the number of schools that had adopted the policy by a given year; on the right axis it displays the share of public and charter schools that had implemented the FSD by a given year.

Figure 1.2: Evolution of Test Scores relative to 1 Year before the FSD Adoption



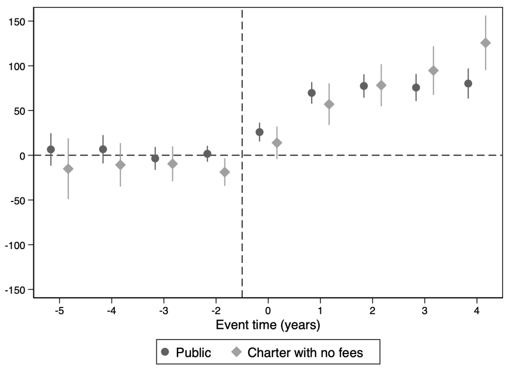
(a) Reading



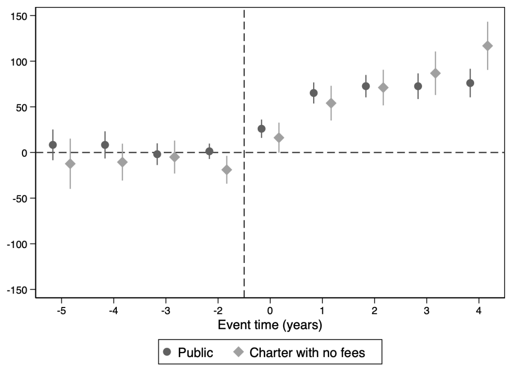
(b) Mathematics

Notes: Panels (a) and (b) plot coefficients, alongside 95 per cent confidence intervals, from the event study specification (1.4). The FSD is adopted in event-year 0 and coefficients show how different reading and mathematics test scores are in event-year ρ relative to event-year -1, which is taken as the reference year. The sample consists of all schools where students in the master sample enrolled in first grade. All specifications include school and year fixed effects. Standard errors are clustered at the school level.

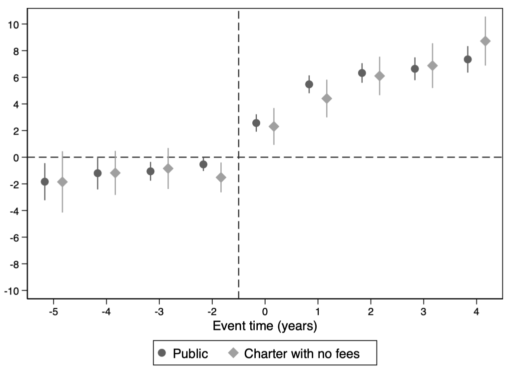
Figure 1.3: Evolution of Contract and Teaching Hours relative to 1 Year before the FSD Adoption



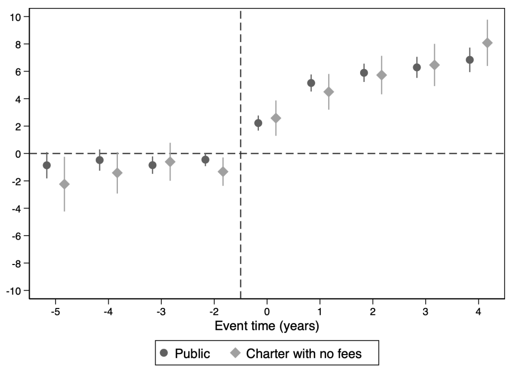
(a) Contract hours



(b) Teaching hours



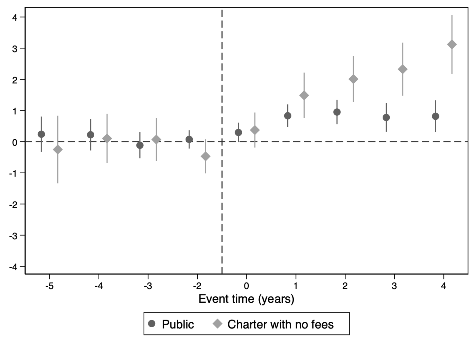
(c) Contract hours/n. of classes



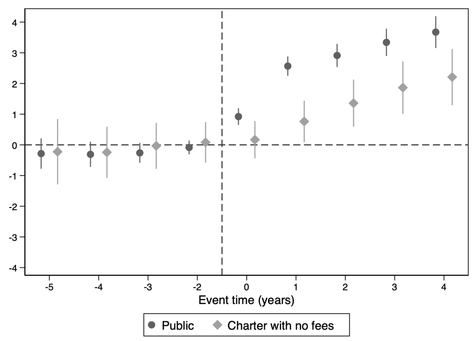
(d) Teaching hours/n. of classes

Notes: Panels (a) to (d) plot coefficients, alongside 95 per cent confidence intervals, from the event study specification outlined in (1.5). The FSD is adopted in event-year 0 and coefficients show how different contract and teaching hours are in event-year ρ relative to event-year -1, which is taken as the reference year. The sample consists of all schools where students in the master sample enrolled in first grade. All specifications include school and year fixed effects. Standard errors are clustered at the school level.

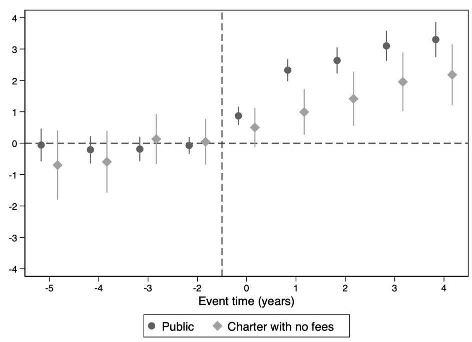
Figure 1.4: Evolution of Number of Teachers, Contract Hours per Teacher and Teaching Hours per Teacher relative to 1 Year before the FSD Adoption



(a) Number of teachers



(b) Contract hours per teacher



(c) Teaching hours per teacher

Notes: Panels (a) to (c) plot coefficients, alongside 95 per cent confidence intervals, from the event study specification outlined in (1.5). The FSD is adopted in event-year 0 and coefficients show how different the number of teachers, contract hours per teacher and teaching hours per teacher are in event-year ρ relative to event-year -1, which is taken as the reference year. The sample consists of all schools where students in the master sample enrolled in first grade. All specifications include school and year fixed effects. Standard errors are clustered at the school level.

Table 1.1: Daily Schedules with and without the FSD

	FSD (1)	No FSD (2)
Minimum number of hours per day	7.08	5.00
Example of daily schedule	5 days under FSD: 08:00-15.05 4 days under FSD: 08.00 - 15.45	Morning shift: 08:00-13:00 Afternoon shift: 14:00-19:00

Notes: The table reports the minimum number of hours students spend at school every day and the daily schedule with and without the FSD in place, inclusive of time devoted to breaks. The minimum number of hours is prescribed in the law. Schools can freely choose the time at which the school day starts. The daily schedules are built assuming that the full school day and the morning shift start at 8.00, while the afternoon shift starts at 14.00.

Table 1.2: Hours of Instruction per Week and Student Voucher with and without the FSD

Subject/Grades	1st - 4th	
	FSD	No FSD
	(1)	(2)
Mathematics	6	6
Spanish	6	6
Natural and Social Sciences	6	6
Physical Education	3	3
Arts and Music	4	4
Technology	3	3
Others	2	2
School Free Choice	8	0
Total	38	30
Student Voucher (U.S.E.)	1.99	1.45

Notes: The table reports weekly subject-specific and total instruction time with and without the FSD, for grades 1 to 4. The information comes from the Decree 625 of the Ministry of Education published in 2003 (`http://bcn.cl/253tx`). It also reports the amount of the student voucher with and without the FSD, expressed in educational subsidy units (U.S.E). This information comes from the version of the DFL2/1996 of the Ministry of Education published in May, 2003 (`http://bcn.cl/1uy40`). These units underwent some modifications since the implementation of the FSD reform.

Table 1.3: Use of Time under the FSD in Primary Schools (Hours per Week)

Subject	All Schools		Public Schools		Charter Schools	
	Core Time	Free Choice Time	Core Time	Free Choice Time	Core Time	Free Choice Time
	(1)	(2)	(3)	(4)	(5)	(6)
Spanish	5.47	2.39	5.39*	2.49	5.59*	2.24
	(0.98)	(1.64)	(0.81)	(1.59)	(1.18)	(1.71)
Mathematics	5.19	1.48	5.14	1.55	5.25	1.37
	(0.94)	(1.31)	(0.78)	(1.34)	(1.13)	(1.26)
Social Sciences	3.83	0.17	3.84	0.15	3.81	0.19
	(0.81)	(0.56)	(0.74)	(0.56)	(0.91)	(0.55)
Natural Sciences	3.89	0.49	3.91	0.47	3.85	0.51
	(0.73)	(0.94)	(0.70)	(0.93)	(0.77)	(0.96)
Foreign Languages	2.03	0.27	1.90***	0.16***	2.22***	0.43***
	(0.70)	(0.75)	(0.59)	(0.57)	(0.80)	(0.93)
Technology	2.03	0.01	2.00	0.004	2.05	0.02
	(0.53)	(0.12)	(0.52)	(0.07)	(0.54)	(0.18)
Art	3.12	0.06	3.09	0.07	3.17	0.06
	(0.81)	(0.35)	(0.77)	(0.33)	(0.86)	(0.38)
Sports	2.10	0.04	2.04**	0.028	2.19**	0.06
	(0.61)	(0.27)	(0.50)	(0.21)	(0.74)	(0.34)
Religion	1.92	0.04	1.89	0.00***	1.97	0.10***
	(0.47)	(0.28)	(0.51)	(0.00)	(0.38)	(0.43)
Number of Schools	387		229		158	

Notes: The table reports hours per week allocated to different subjects in fifth grade for a representative sample of urban schools that adopted the FSD by 2005 and were surveyed by the Studies Directorate of the Sociology Faculty at the Catholic University of Chile (*DESUC*). "Core Time" excludes "Free Choice" time. *, **, *** indicate that the number of hours allocated to a given subject is significantly different between public and charter schools at the 10, 5 and 1 per cent level, respectively. Standard deviations are in parentheses.

Table 1.4: Summary Statistics

	All	Public	Charter	
			No Tuition Fees	Tuition Fees
	(1)	(2)	(3)	(4)
Students demographics				
Female	0.49	0.50	0.49	0.49
Age at school entry	6.60	6.60	6.60	6.59
Parental education				
Less than university	0.87	0.92	0.91	0.78
Books at home				
At most 50	0.85	0.89	0.88	0.79
Other resources at home				
Computer	0.55	0.42	0.50	0.72
Internet	0.31	0.20	0.27	0.46
Schools Characteristics				
First grade average class size	34.68	33.69	35.37	35.53
First grade enrollment	82.33	83.09	72.39	86.56
Academic performance				
Reading test score	-0.04	-0.17	-0.12	0.18
Mathematics test score	-0.04	-0.17	-0.15	0.19
First-grade attendance rate	94.23	93.88	94.83	94.33
End of first-grade status (1=repeat)	0.03	0.03	0.04	0.02
N. of students	604532	270417	114074	218495

Notes: The table reports summary statistics for the sample of fourth graders who start primary school between 2002 and 2010 in publicly subsidized schools that had not yet adopted the FSD. Parental education refers to the highest educational attainment among the mother and the father; in case the information is missing for one parent, it refers to the education level of the other parent. All figures are expressed as fractions, except from averages referring to the age of pupils, class size, enrollment, test scores and the attendance rate. Test scores are standardized by year and subject (including also pupils who are not in the master sample) to have mean equal to 0 and standard deviation equal to 1. The number of observations in columns (2) to (4) does not sum to the number of observations in column (1) because for 2.39 per cent of charter schools we could not find information about the tuition fees.

Table 1.5: Effect of the FSD on Test Scores

	Linear specification				Non parametric specification	
	FE1	FE2	FE-IV1	FE-IV2	FE-IV1	FE-IV2
	(1)	(2)	(3)	(4)	(5)	(6)
				A. Reading		
Years under FSD	0.002	0.011***	0.024***	0.024***		
	(0.003)	(0.003)	(0.006)	(0.006)		
Years under FSD = 1					0.023	0.022
					(0.015)	(0.015)
Years under FSD = 2					0.029**	0.030**
					(0.015)	(0.014)
Years under FSD = 3					0.116***	0.114***
					(0.021)	(0.021)
First stage coefficient			0.720***	0.720***		
			(0.005)	(0.005)		
Kleibergen-Paap rk Wald F statistic			23614.36	24416.41	5136.66	5265.46
N. of students	596108	596108	596108	596108	596108	596108
				B. Mathematics		
Years under FSD	-0.007**	0.005	0.007	0.008		
	(0.003)	(0.003)	(0.007)	(0.007)		
Years under FSD = 1					-0.014	-0.015
					(0.016)	(0.016)
Years under FSD = 2					-0.003	-0.002
					(0.017)	(0.016)
Years under FSD = 3					0.057**	0.058**
					(0.023)	(0.023)
First stage coefficient			0.719***	0.720***		
			(0.005)	(0.005)		
Kleibergen-Paap rk Wald F statistic			23460.87	24294.13	5140.07	5278.70
N. of students	596281	596281	596281	596281	596281	596281
Student-level controls	No	Yes	No	Yes	No	Yes
School-level controls	No	Yes	No	Yes	No	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table reports the effect of the FSD on reading and mathematics test scores. Estimates in columns (1), (3) and (5) are based on a parsimonious specification that only includes as controls school fixed effects and year fixed effects. Estimates in other columns are based on a richer specification that features an additional set of controls. Specifically, student-level controls include: gender, age at school entry, as well as the attendance rate and the status (pass or repeat) at the end of grade 1. School-level controls include averages of the students' characteristics at the school level, as well as enrollment and average class size in first grade. The effect of the FSD is assumed to be linear in exposure in columns (1) to (4), whereas it is allowed to vary in a fully non-parametric way in columns (5) and (6). In specifications FE-IV1 and FE-IV2 the treatment (i.e. actual years of exposure to the FSD by the end of grade 4) is instrumented with the exposure a student would experience had she never transferred out of her first-grade school. Standard errors are clustered at the school level and are reported in parenthesis.
*, ** and *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 1.6: Heterogeneous Effects of the FSD on Test Scores by Students' Socio-economic Background

	Reading			Mathematics		
	$D = \mathbb{1}(NoUni.)$	$D = \mathbb{1}(Books \leq 50)$	$D = \mathbb{1}(NoICT)$	$D = \mathbb{1}(NoUni.)$	$D = \mathbb{1}(Books \leq 50)$	$D = \mathbb{1}(NoICT)$
	(1)	(2)	(3)	(4)	(5)	(6)
Years under FSD	0.007	0.012	0.012	-0.003	0.011	0.004
	(0.012)	(0.011)	(0.009)	(0.012)	(0.012)	(0.010)
Years under FSD $\times$ D	0.015	0.010	0.012	0.011	0.005	0.004
	(0.013)	(0.011)	(0.010)	(0.012)	(0.011)	(0.011)
Kleibergen-Paap rk Wald F statistic	2781.84	11502.69	5644.29	2797.02	11528.94	5643.12
N. of students	532970	529879	517249	534473	531369	518676
Student-level controls	Yes	Yes	Yes	Yes	Yes	Yes
School-level controls	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table shows the effect of the FSD on reading and mathematics test scores by different measures of students' socio-economic background. Every column shows an enriched version of specification (1.2) where the treatment and all controls listed in the notes to Table 1.5, including school and year fixed effects, are also interacted with a dummy D, capturing a relevant dimension of heterogeneity. In columns (1) and (4) D takes value 1 if no parent in the household has some university education, and 0 otherwise. In columns (2) and (5) D takes value 1 if there are at most 50 books at home, and 0 otherwise. In columns (3) and (6) D takes value 1 if there is not a computer or an Internet connection at home, and 0 otherwise. Actual years of exposure to the FSD are instrumented with years of exposure a student would accumulate had she never transferred from the school where she attended first grade. Standard errors are clustered at the school level and are reported in parenthesis.
*, ** and *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 1.7: Heterogeneous Effects of the FSD on Test Scores by School Type

	Reading				Mathematics			
		$D = \mathbb{1}(NoUni.)$	$D = \mathbb{1}(Books \leq 50)$	$D = \mathbb{1}(NoICT)$		$D = \mathbb{1}(NoUni.)$	$D = \mathbb{1}(Books \leq 50)$	$D = \mathbb{1}(NoICT)$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years under FSD	0.019**	-0.001	0.023	0.011	0.012	-0.007	0.008	0.003
	(0.009)	(0.021)	(0.017)	(0.015)	(0.009)	(0.021)	(0.017)	(0.016)
Years under FSD × no-fee charter	0.042**	0.047**	0.046**	0.045**	0.017	0.017	0.015	0.018
	(0.019)	(0.019)	(0.019)	(0.020)	(0.021)	(0.021)	(0.021)	(0.020)
Years under FSD × SES char. D		0.022	-0.003	0.009		0.024	0.010	0.015
		(0.021)	(0.017)	(0.015)		(0.020)	(0.017)	(0.015)
Kleibergen-Paap rk Wald F statistic	828.30	605.18	609.12	632.25	839.28	602.71	606.90	630.61
N. of students	377856	319187	319234	319254	377719	320061	320108	320129
Student-level controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School-level controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table shows the effect of the FSD on reading and mathematics test scores by the type of school (public or charter without tuition fees) that the student attended in grade 1. Estimates in columns (1) and (5) are based on an enriched version of specification (1.2), where the treatment and all controls listed in the notes to Table 1.5, including year fixed effects, are also interacted with a dummy taking value 1 if the school is a charter establishment without tuition fees, and 0 if it is public. In columns (2) and (6) the treatment and all controls are further interacted with a dummy that takes value 1 if none of the parents has some university education, and 0 otherwise. In columns (3) and (7) the treatment and all controls are further interacted with a dummy that takes value 1 if there are at most 50 books at home, and 0 otherwise. In columns (4) and (8) the treatment and all controls are further interacted with a dummy that takes value 1 if either the computer or a connection to the Internet are not available at home, and 0 otherwise. Actual years of exposure to the FSD are instrumented with years of exposure a student would accumulate had she never transferred from the school where she attended first grade. Standard errors are clustered at the school level and are reported in parenthesis.
*, ** and *** denote significance at the 10, 5 and 1 per cent level, respectively.

Appendices

1.A Public and Charter Schools

Public and charter schools are subject to different regulations. This translates into charter schools enjoying more autonomy and flexibility over curriculum, budgetary and personnel decisions.

Public schools are either managed by the Municipal Department of Education (*DAEM*) or by private non-profit education corporations. While the director of the *DAEM* is a teacher, corporations are led by a board of directors who do not need to be teachers and whose president is the major of the municipality. Under both management schemes decisions related to the allocation of resources and to hiring/firing school staff are taken at the municipality level and school principals are not necessarily involved. Charter schools are instead private organizations and all relevant decisions are taken by the school authorities.

The working conditions of the employees of public schools are regulated by the *Estatuto de los Profesionales de la Educación*. The relevant regulation for charter schools is the *Código del Trabajo*, the labor code that applies to all firms in Chile. Appointments of public school teachers are decided by a commission that is formed by the Major, the director of the *DAEM* or the education corporations, as well as one randomly selected teacher from the schools in the municipality. Priority is given to spouses of teachers already working in the municipality. The salary of public school teachers is fixed according to a national scale that takes into account experience, training, specific difficult situations (such as teaching in rural, remote or deprived areas) and responsibilities. Firing is subject to many restrictions. It is possible only if one of the following conditions are met: *i*) school enrollment decreases; *ii)* the national curriculum undergoes changes that justify the decision; *iii)* schools' merges; *iv)* protracted poor performance (see below). Teachers having tenured positions enjoy a greater job security.[48] In any case, firings have to be justified in the Annual Plan of Educational Development that needs the approval of the Provincial Office of the Ministry of Education. Charter schools are instead free to set their own recruitment and dismissal criteria. Wages and the other working conditions are subject to the same regulations that apply to private firms.

There are also differences in the evaluation of teachers. The *Estatuto de los Profesionales de la Educación* originally set some criteria for assessing teachers performance, but they were never fully implemented. In 2003 a new evaluation system was agreed. Nevertheless it is quite lax and in practice very few teachers receive poor evaluations. In principle, teachers could be fired if they fare unsatisfactorily in two or three consecutive evaluations. School principals are not accountable based on the school performance and they can be fired only in case of a grave fault, while poor evaluations can result in assigning them to smaller schools. Charter schools can instead set their own evaluation systems and the consequences in case of poor performance.

[48]The *Estatuto de los Profesionales de la Educación* contemplates two type of contracts, *titular* and *contratado*. The first type of contract affords a greater job security, as it offers a tenured position.

1.B Robustness Checks

As a first robustness check, we show that specification (1.2) delivers similar estimates if we also control for a set of characteristics of teachers in the school s where the student started first grade in year t. Specifically, we include as controls the share of female teachers, the share of teachers with an education degree and teachers' average age.[49] Appendix Table 1.B1 shows that results are virtually unchanged.

A possible concern not addressed by specification (1.2) is that other events may happen in a school around the time of FSD adoption and affect learning outcomes in the following years. Our estimates would then also capture the effects of other changes to the school environment.

The first potential confounder to check is infrastructure investment, as some schools had to expand their infrastructure prior to switching to a single-shift scheme. Funds disbursed for this purpose covered costs related to replicating the existing infrastructure on a larger scale, not to improving it. Nonetheless, to address this issue, we replicate our analysis on the sample of pupils who started first grade in schools that did not receive public funds for expanding infrastructure. These establishments are unlikely to have made substantial changes to their facilities prior to lengthening the school day. Columns (1) and (2) of Appendix Table 1.B2 report estimates that are in a similar range as those coming from the full sample of schools. An additional year of exposure to the FSD raises reading test scores by 0.020σ. The effect on mathematics test scores is virtually 0. According to this exercise, infrastructure investment does not appear to be an important alternative driver of our estimates.

In 2008 Chile introduced a Preferential School Subsidy scheme (*Subvención Escolar Preferencial*, or PSS henceforth) which grants schools an additional subsidy for each disadvantaged student they cater to.[50] To check whether our estimates are also cap-

[49]These controls are not included in the baseline specification because they are not available for the year 2002 and are missing for some schools in other years. In this regression specification, we assume that the teaching staff in 2002 is the same as that observed in 2003, so as not to drop one year of observations.

[50]The receipt of the subsidy is conditional upon schools developing a pedagogical plan that outlines how additional funds are used to improve learning outcomes and upon allowing for an external evaluation of the results achieved. See Santiago et al. [2013] for more info.

turing the roll-out of the subsidy, we implement two exercises. First, we enrich specification (1.2) with controls for the individual exposure to the PSS scheme (i.e. the number of grades during which the student received the subsidy) by grade 4 and the average share of pupils benefiting from the PSS scheme in the schools attended by a student in grades 1 to 4. Second, we estimate specification (1.2) on the sub-sample of cohorts never exposed to the PSS (i.e. those starting primary education before 2005). In both cases, coefficients are similar to those coming from the main specification and, if anything, in the case of reading they are slightly larger (Table 1.B2, columns 3 and 5).

Table 1.B1: Effect of the FSD on Test Scores including Teacher Controls

	Linear specification				Non parametric specification	
	FE2	FE2	FE-IV2	FE-IV2	FE-IV2	FE-IV2
	(1)	(2)	(3)	(4)	(5)	(6)
			A. Reading			
Years under FSD	0.011***	0.011***	0.024***	0.024***		
	(0.003)	(0.003)	(0.006)	(0.006)		
Years under FSD = 1					0.022	0.023
					(0.015)	(0.015)
Years under FSD = 2					0.030**	0.032**
					(0.014)	(0.014)
Years under FSD = 3					0.114***	0.115***
					(0.021)	(0.021)
First stage coefficient			0.720***	0.723***		
			(0.005)	(0.005)		
Kleibergen-Paap rk Wald F statistic			24416.41	24689.20	5265.46	5244.84
N. of students	596108	578112	596108	578112	596108	578112
			B. Mathematics			
Years under FSD	0.005	0.004	0.008	0.009		
	(0.003)	(0.003)	(0.007)	(0.007)		
Years under FSD = 1					-0.015	-0.012
					(0.016)	(0.016)
Years under FSD = 2					-0.002	0.001
					(0.016)	(0.016)
Years under FSD = 3					0.058**	0.060***
					(0.023)	(0.023)
First stage coefficient			0.720***	0.723***		
			(0.005)	(0.005)		
Kleibergen-Paap rk Wald F statistic			24294.13	24506.65	5278.70	5253.62
N. of students	596281	578281	596281	578281	596281	578281
Student-level controls	Yes	Yes	Yes	Yes	Yes	Yes
School-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Teacher-level controls	No	Yes	No	Yes	No	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table reports the effect of the FSD on reading and mathematics test scores. Estimates in columns (1), (3) and (5) are based on the specification with baseline controls. Student-level controls include: gender, age at school entry, as well as the attendance rate and the status (pass or repeat) at the end of grade 1. School-level controls include averages of the students' characteristics at the school level, as well as enrollment and average class size. Estimates in columns (2), (4) and (6) include also controls referring to teachers' characteristics when the students attend grade 1. Specifically, they are the share of female teachers, teachers' average age and the share of teachers with an education degree. The treatment in specifications FE2 is actual years of exposure to the FSD by the end of grade 4, while in specifications FE-IV2 is instrumented with the exposure a student would experience had she never transferred out of her first grade school. Standard errors are clustered at the school level and are reported in parenthesis.
*, ** and *** denote significance at the 10, 5 and 1 per cent level, respectively.

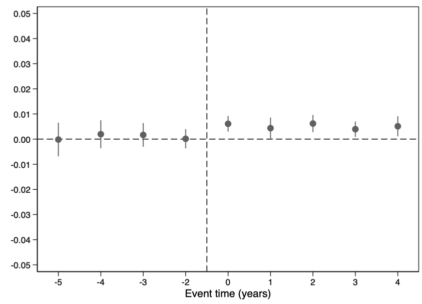
Table 1.B2: Effect of the FSD on Test Scores - Robustness Checks

	No infrastructure funds		PSS			
	Reading	Mathematics	Reading	Mathematics	Reading	Mathematics
	(1)	(2)	(3)	(4)	(5)	(6)
Years under FSD	0.020**	-0.001	0.027***	0.011	0.026**	0.002
	(0.010)	(0.011)	(0.006)	(0.007)	(0.012)	(0.013)
Number of students	379449	379691	596020	596190	291057	291085
Kleibergen-Paap rk Wald F statistic	9202.99	9259.78	21528.75	21474.33	11287.84	11198.08
Student-level controls	Yes	Yes	Yes	Yes	Yes	Yes
School-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

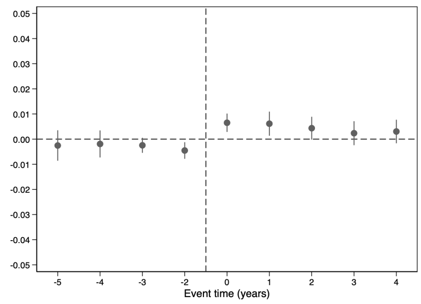
Notes: The table presents the results from a set of specifications that check the robustness of the main estimates of the effects of the FSD on test scores. All specifications include school and year fixed effects, and actual exposure to the FSD is instrumented with the exposure a student would accumulate had she never transferred from the school where she attended first grade. In columns (1) and (2), specification (1.2) is estimated on the sub-sample of pupils in the master sample who start first grade in schools that did not receive public funds for expanding their infrastructure. In columns (3) and (4) specification (1.2) is enriched with two additional controls, on top of those listed in the notes to Table 1.5: individual exposure to the Preferential Subsidy Scheme (PSS) policy by grade 4 and the average share of pupils benefiting from the PSS in the schools attended by the student in grades 1 to 4. In columns (5) and (6) , specification (1.2) is estimated on the sub-sample of cohorts never exposed to the Preferential Subsidy Scheme (i.e. cohorts starting primary education between 2002 and 2004). Standard errors are clustered at the school level and are reported in parenthesis.
*, ** and *** denote significance at the 10, 5 and 1 per cent level, respectively.

1.C Additional Figures and Tables

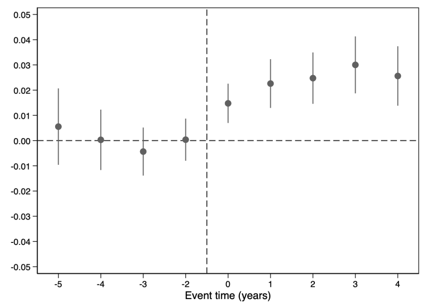
Figure 1.C1: Evolution of the Number of Subjects Relative to 1 Year before the FSD Adoption



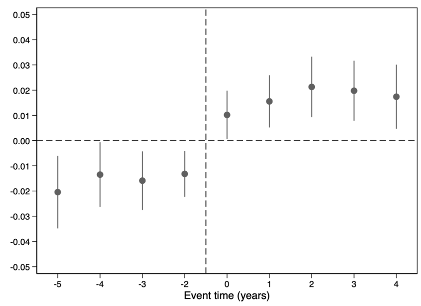
(a) Spanish



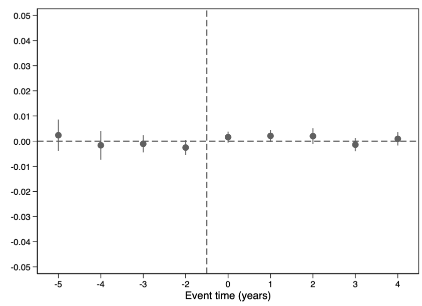
(b) Mathematics



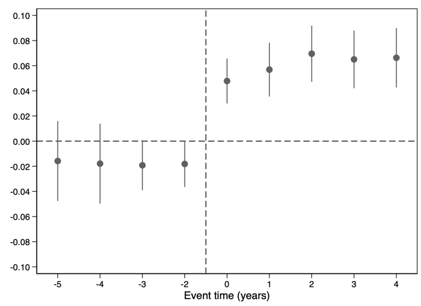
(c) Foreign languages



(d) Tutorials and workshops



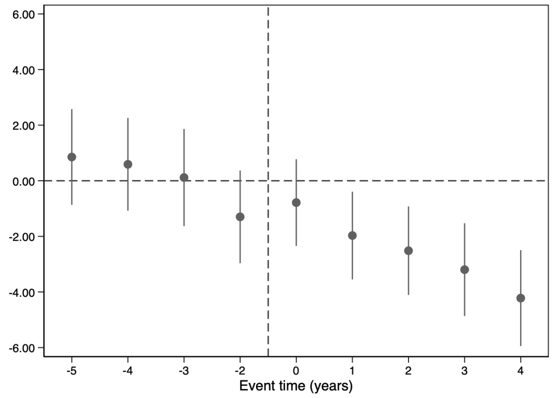
(e) Sports



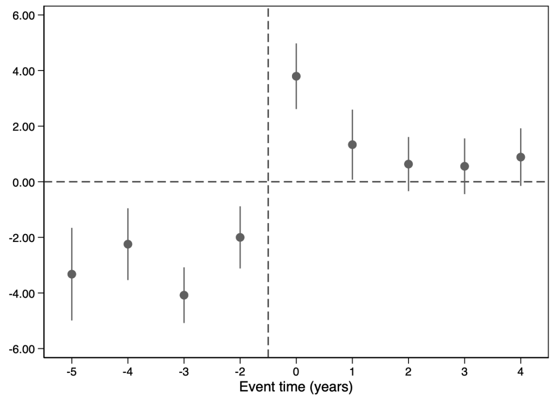
(f) Total n. of subjects

Notes: Panels (a) to (f) plot coefficients, alongside 95 per cent confidence intervals, from the event study specification outlined in (1.1). The FSD is adopted in event-year 0 and coefficients show how different the number of subjects taught is in event-year p relative to event-year -1, which is taken as the reference year. The sample consists of all schools where students in the master sample enrolled in first grade. All specifications include school, grade and year fixed effects, as well as their interactions. Standard errors are clustered at the school-grade level.

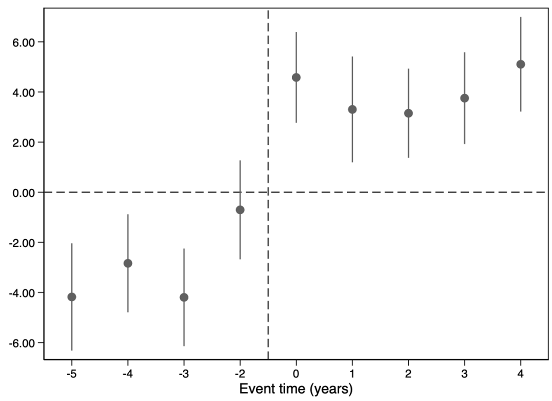
Figure 1.C2: Evolution of Transfers relative to 1 Year before the FSD Adoption



(a) Transfers out of the school



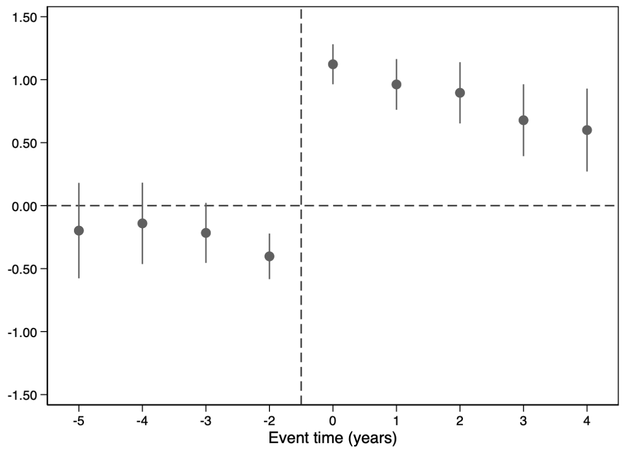
(b) Transfers into the school



(c) Net transfers (In - Out)

Notes: Panels (a), (b) and (c) plot coefficients, alongside 95 per cent confidence intervals, from the event study specification outlined in (1.5). The FSD is adopted in event-year 0 and coefficients show how different the number of transfers in grades 1 to 4 is in event-year p relative to event-year -1, which is taken as the reference year. The sample consists of all schools where students in the master sample enrolled in first grade. All specifications include school and year fixed effects. Standard errors are clustered at the school level.

Figure 1.C3: Evolution of Class Size relative to 1 Year before the FSD Adoption



Notes: The figure plots coefficients, alongside 95 per cent confidence intervals, from the event study specification outlined in (1.5). The FSD is adopted in event-year 0 and coefficients show how different the average class size in grades 1 to 4 is in event-year ρ relative to event-year -1, which is taken as the reference year. The sample consists of all schools where students in the master sample enrolled in first grade. All specifications include school and year fixed effects. Standard errors are clustered at the school level.

Table 1.C1: Characteristics of Students in the Master Sample who do and do not Transfer

	First grade Academic Performance		FSD Exposure	
	Attendance	Repetition	Real	Potential
	(1)	(2)	(3)	(4)
Do not transfer between grades 1 and 4	94.64	0.02	0.52	0.51
Transfer between grades 1 and 4	93.44	0.05	1.35	0.42

Notes: Columns (1) and (2) show the average attendance rate in grade 1 for students in the master sample and the fraction of them who repeat grade 1, distinguishing pupils who never transfer between grades 1 and 4 from those who transfer. Columns (3) and (4) display their average actual exposure to the FSD by grade 4 as well their average "potential" exposure, i.e. the years of exposure a student would experience had she never transferred out of her first-grade school.

Table 1.C2: Support Received by Students Outside of School

	Help with homework from household head and head' spouse		Tutoring	
	1 = Yes	Hours	1 = Yes	Hours
	(1)	(2)	(3)	(4)
No university	0.328	0.557	0.050	0.124
University	0.483	0.791	0.119	0.280
Observations	6214	4564	2677	2666

Notes: The table shows the amount of support that students receive outside of school, depending on whether they live in a household where one among the household head and the head' spouse has some university education (row "University") or not (row "No university"). The units of observation are the households heads and their spouses (in households where these is at least one child aged 5-18) when the question is whether they provide help with homework. The units of observations are pupils aged 12-18 (younger children are not interviewed) when the question is whether they receive tutoring outside of school. Information is drawn from the 2015 Chilean Time-Use Survey (*Encuesta nacional sobre uso del tiempo*).

Table 1.C3: Frequency of Mathematics Homework

	A. All schools		
	No FSD	FSD	
	(1)	(2)	(3)
Every class	20.04%	11.80%	
Almost every class	50.39%	39.35%	
Some classes	28.42%	46.33%	
Never	1.15%	2.51%	
N. of Teachers	3294	18494	
	B. Schools that had not adopted the FSD by 2011		
	2011	2013	
		No FSD	FSD
Every class	22.51%	18.81%	7.95%
Almost every class	51.88%	52.10%	31.82%
Some classes	24.50%	28.34%	59.09%
Never	1.11%	0.74%	1.14%
N. of Teachers	902	808	88

Notes: The table reports information about the frequency of mathematics homework, drawn from the 2011, 2012 and 2013 waves of the teacher surveys administered alongside the SIMCE test. Panel A compares the frequency of homework in schools with and without the FSD. Panel B focuses on schools that had not adopted the FSD by 2011 and compare homework frequency in 2011 and 2013. In 2013, schools are divided according to whether they switched to longer schedules by that year (column 3) or not (column 2).

Table 1.C4: Differences in School Autonomy between Public and Charter Schools

	Public schools (1)	Charter schools (2)
Textbook use	95	98
Courses content	39	57
Courses offer	80	96
Formulate budget	31	97
Allocate budget	65	98
Hire teachers	28	99
Fire teachers	10	97
Set starting salaries	1	87
Increase salaries	1	87
Observations	62	85

Notes: The table reports the percentage of schools in which the principal or the governing body have a considerable responsibility over the listed tasks. Information comes from the 2006 and 2009 school surveys administered alongside PISA tests. The sample consists of all public or charter schools in the Chilean PISA sample that also offer primary education.

Table 1.C5: Evolution of Teacher related Inputs relative to 1 Year before the FSD Adoption

	Contract HH.	Teaching HH.	$\frac{Contract HH.}{N. of Classes}$	$\frac{Teaching HH.}{N. of Classes}$	N. of Teachers	$\frac{Contract HH.}{N. of Teachers}$	$\frac{Teaching HH.}{N. of Classes}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Event-year -5	5.801	7.693	-1.799***	-0.856*	0.219	-0.287	-0.067
	(7.171)	(6.583)	(0.538)	(0.462)	(0.218)	(0.279)	(0.297)
Event-year -4	6.804	8.192	-1.170**	-0.482	0.227	-0.311	-0.223
	(6.436)	(5.908)	(0.483)	(0.414)	(0.195)	(0.250)	(0.267)
Event-year -3	-3.554	-2.088	-1.031**	-0.835**	-0.115	-0.267	-0.186
	(5.992)	(5.500)	(0.450)	(0.386)	(0.182)	(0.233)	(0.248)
Event-year -2	0.725	0.638	-0.508	-0.413	0.050	-0.094	-0.075
	(5.286)	(4.852)	(0.397)	(0.340)	(0.160)	(0.205)	(0.219)
Event-year 0	25.310***	25.521***	2.626***	2.283***	0.291**	0.904***	0.867***
	(4.764)	(4.373)	(0.357)	(0.307)	(0.145)	(0.185)	(0.197)
Event-year 1	68.160***	63.904***	5.487***	5.166***	0.796***	2.550***	2.326***
	(4.870)	(4.471)	(0.365)	(0.314)	(0.148)	(0.189)	(0.202)
Event-year 2	76.502***	72.051***	6.371***	5.946***	0.930***	2.900***	2.641***
	(5.071)	(4.655)	(0.380)	(0.327)	(0.154)	(0.197)	(0.210)
Event-year 3	74.470***	71.325***	6.676***	6.286***	0.745***	3.339***	3.085***
	(5.312)	(4.876)	(0.399)	(0.342)	(0.161)	(0.206)	(0.220)
Event-year 4	78.879***	75.229***	7.339***	6.847***	0.784***	3.661***	3.311***
	(5.581)	(5.123)	(0.419)	(0.360)	(0.169)	(0.217)	(0.231)
Event-year -5 × D_s	-20.288	-19.524	-0.006	-1.341	-0.445	0.073	-0.648
	(15.072)	(13.835)	(1.134)	(0.974)	(0.458)	(0.585)	(0.625)
Event-year -4 × D_s	-17.040	-18.298	0.023	-0.904	-0.104	0.087	-0.371
	(13.860)	(12.722)	(1.039)	(0.892)	(0.421)	(0.538)	(0.574)
Event-year -3 × D_s	-5.540	-2.519	0.181	0.221	0.204	0.251	0.316
	(12.935)	(11.873)	(0.970)	(0.832)	(0.393)	(0.502)	(0.536)
Event-year -2 × D_s	-18.923	-18.884*	-1.002	-0.909	-0.495	0.208	0.143
	(11.976)	(10.993)	(0.897)	(0.770)	(0.364)	(0.465)	(0.496)
Event-year 0 × D_s	-11.246	-9.220	-0.474	0.143	0.096	-0.746*	-0.386
	(10.282)	(9.438)	(0.771)	(0.662)	(0.312)	(0.399)	(0.426)
Event-year 1 × D_s	-11.131	-10.071	-1.099	-0.717	0.700**	-1.717***	-1.264***
	(10.479)	(9.619)	(0.785)	(0.674)	(0.318)	(0.407)	(0.434)
Event-year 2 × D_s	1.279	-1.266	-0.262	-0.206	1.082***	-1.556***	-1.223***
	(10.792)	(9.906)	(0.809)	(0.695)	(0.328)	(0.419)	(0.447)
Event-year 3 × D_s	20.155*	15.900	0.255	0.263	1.599***	-1.440***	-1.057**
	(11.193)	(10.275)	(0.840)	(0.721)	(0.340)	(0.435)	(0.464)
Event-year 4 × D_s	45.760***	40.727***	1.330	1.148	2.317***	-1.382***	-1.057**
	(11.666)	(10.708)	(0.875)	(0.751)	(0.354)	(0.453)	(0.483)
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N. of school-years	19074	19074	18828	18828	19074	19074	19074

Notes: The table reports coefficients from a richer version of the event study specification outlined in (1.5) where calendar year fixed effects and event years are also interacted with a dummy D_s taking value 1 if the school is a no-fee charter school, and 0 otherwise. The FSD is adopted in event-year 0 and coefficients show how different total contract hours, teaching hours, contract hours per class, teaching hour per class, total number of teachers, contract hours per teacher and teaching hours per teacher are in event-year ρ relative to event-year -1, which is taken as the reference year. The sample consists of all schools where students in the master sample enrolled in first grade. All specifications include school and year fixed effects. Standard errors are clustered at the school level. *, ** and *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 1.C6: Teachers opinion about the FSD

	Public Schools (1)	Charter Schools (2)
Good or very good	44.99%	54.44%
Not bad, not good	32.99%	29.89%
Bad or very bad	22.02%	15.67%

Notes: The table reports the opinion of teachers about the FSD, dividing them according to the school (public or charter) in which they teach. Information is drawn from the 2005 wave of the *Encuesta Longitudinal Docente* implemented by the *Centro de Microdatos* of the *Universidad de Chile*.

Chapter 2

Should I Stay or Should I go? Neighbors' Effects on University Attendance

Andrés Barrios Fernández
LSE

Abstract

This paper investigates whether the decision to attend university depends on university enrollment of close neighbors. I create a unique dataset combining detailed geographic information and educational records from different government agencies in Chile, and exploit the quasi-random variation generated by the rules that determine eligibility for student loans. I find that close neighbors have a large and significant impact on university enrollment of younger applicants. Potential applicants are 10 percentage points more likely to attend university if a close neighbor enrolled the year before. This effect is particularly strong in areas where university attendance is low and among neighbors who are more likely to interact with each other; the effect decreases both with physical and social distance and is weaker for individuals who have spent less time in the neighborhood. I also show that the increase in university attendance is mediated by an increase in applications rather than by an improvement on academic performance. These results suggest that policies that expand access to university generate spillovers on the peers of the direct beneficiaries. I show that in areas where university attendance is low the indirect effect of student loans represents more than 15% of the direct effect on enrollment.

Keywords: Neighbors' Effects, University Access, University Enrollment.

JEL classification: I21, I24, R23, R28.

2.1 Introduction

Despite high returns to schooling and governmental efforts to improve educational attainment, university enrollment remains low among disadvantaged individuals in both developing and developed countries. While not all of these individuals would benefit from a university education, enrollment is low even among those with high academic potential.[1] This situation is partially explained by the absence of enough funding opportunities, but there is growing evidence that the lack of information, support, and encouragement also plays an important role in schooling decisions [Hoxby and Avery, 2013, Carrell and Sacerdote, 2017].[2] This evidence also shows that the barriers preventing students to take full advantage of their education opportunities are higher in areas where university attendance is low, suggesting that the neighborhoods where individuals live matter.[3]

This paper builds on these findings and investigates whether potential applicants' decision to attend university is affected by university enrollment of close neighbors. Although the role of peers in education has been widely studied, this paper is among the first looking at how they influence enrollment in higher education.[4] By addressing this question I contribute to understanding if neighborhood effects are driven at least in

[1]Figure 2.E1 in the appendix shows that in the case of Chile —the setting studied in this paper— the gap in university enrollment persists along the ability distribution.

[2]Hoxby and Avery [2013] show that high achieving individuals from areas with low educational attainment in the United States apply to less selective schools than similar students from other areas. This, despite the fact that better schools would admit and provide them with more generous funding. This undermatching phenomenon has also been studied by Black et al. [2015], Griffith and Rothstein [2009] and Smith et al. [2013]. There is also a vast literature looking at the role of information frictions in schooling investment. Attanasio and Kaufmann [2014], Hastings et al. [2015] and Jensen [2010] study these frictions in Mexico, Chile and Dominican Republic respectively. Bettinger et al. [2012] and Hoxby and Turner [2015] look at them in the United States, and Oreopoulos and Dunn [2013] in Canada. Carrell and Sacerdote [2017] on the other hand, argues that interventions that increase university enrollment work not because they provide additional information, but instead because they compensate for lack of support and encouragement. Lavecchia et al. [2016] discusses these frictions and different behavioral barriers that may explain why some individuals do not take full advantage of education opportunities.

[3]This is also consistent with recent studies on neighborhood effects like Chetty et al. [2016] and Chetty and Hendren [2018a] that show that exposure to better neighborhoods increases the probability of college enrollment. Burdick-Will and Ludwig [2010] discusses the literature on neighborhood effects and education attainment.

[4]Bifulco et al. [2014] studies how having classmates with a college-educated mother affects college enrollment. Mendolia et al. [2018] investigates how peers' ability affect performance on high stake exams and on university attendance. Carrell et al. [2018] looks at the effect of having disruptive peers at the elementary school on long-term outcomes, including college enrollment.

part by exposure to better peers. This in contrast to only being driven by exposure to better institutions (i.e. schools, public infrastructure, security). In addition, this question is relevant from a policy perspective because these neighbors' effects would imply that programs that expand access to university generate externalities that should be incorporated into the design and evaluation of this type of policies.

I study these neighbors' effects in Chile, taking advantage of the fact that eligibility for student loans depends on students scoring above a cutoff on the university admission exam and that eligibility for this type of funding increases university enrollment [Solis, 2017]. Exploiting the discontinuity generated by this cutoff rule, I implement a fuzzy RD using potential applicants' enrollment as outcome and instrumenting their neighbors' enrollment with an indicator of their eligibility for student loans.

To conduct this analysis, I create a unique dataset that combines detailed geographic information and educational records from multiple government agencies. This allows me to identify potential applicants and their neighbors, and to follow them throughout high school and in the transition to higher education.

A key challenge for the identification of neighbors' effects is to distinguish between social interactions and correlated effects. In this context, correlated effects arise because individuals are not randomly allocated to neighborhoods and because once in the neighborhood, they are exposed to similar institutions and shocks. Since potential applicants who have a close neighbor near the student loans eligibility cutoff are very similar, the fuzzy RD used in this paper allows me to rule out the estimated effects to be driven by differences in individual or neighborhood characteristics, eliminating in this way concerns about correlated effects.

In addition, if peers' outcomes have an effect on each other, this gives rise to what Manski [1993] described as the "reflection problem". This paper focuses on potential applicants who decide whether or not to enroll in university one year after their neighbors. Thus, neighbors' decision should not be affected by what potential applicants do one year later. This lagged structure and the fact that the variation on neighbors' enrollment only comes from eligibility for funding allows me to abstract from the "reflection problem".

Based on this empirical analysis, I provide three sets of results. Firstly, I find that having a close neighbor going to university has a large and significant impact on potential applicants' university enrollment, increasing it by around 10 percentage points. I also show that this effect is stronger when individuals are more likely to interact. Only the closest neighbors seem to matter and the effect quickly decays with distance, completely disappearing after 200 meters. The effects also seem to be stronger among neighbors who are closer in gender and socioeconomic status, and for individuals who have lived in the neighborhood for longer.

Secondly, I show that having a close neighbor eligible for student-loans increases university enrollment of potential applicants. I study how this indirect effect of eligibility for funding changes depending on the university attendance rates observed at different municipalities and find that it is stronger in low attendance areas, where it represents more than 15% of the direct effect of student loans on enrollment. Neighbors are not the only peers that may affect potential applicants. I also study what happens in the case of siblings and find that a similar indirect effect arises in this context. Potential applicants with an older sibling eligible for student loans are also more likely to enroll in university.

Finally, I show that the increase in university enrollment documented for both neighbors and siblings, is mediated by an increase in the number of potential applicants taking the university admission exam and applying to university and to financial aid. I find no effects on their attendance or on their academic performance during high school.

These results are consistent with two broad classes of mechanisms. First, neighbors may increase university enrollment of potential applicants by providing them with relevant information about applications, returns and the overall university experience. Second, they could also affect the costs of going to university. Although with the data that I have available I cannot perfectly distinguish between them, I present suggestive evidence that information is the mechanism behind the documented responses.

This paper contributes to existing research in several ways. Firstly, it contributes to the literature on peers' effects. Since the publication of the Coleman Report [Coleman, 1966], peers' effects in education have been widely studied.[5] Although the majority of these studies have focused in the classroom, others have looked at neighborhoods [Goux and Maurin, 2007, Gibbons et al., 2013, 2017], and at the family [Goodman et al., 2015, Dustan, 2018, Joensen and Nielsen, 2018].[6] Others have studied peer effects in higher education [Sacerdote, 2001, Zimmerman, 2003, Stinebrickner and Stinebrickner, 2006, Foster, 2006, Lyle, 2007, Carrell et al., 2009, Feld and Zölitz, 2017]. Few of them find sizeable effects on academic performance. However, as pointed out by Hoxby and Weingarth [2005] and further discussed by Lavy et al. [2012], Burke and Sass [2013] and Imberman et al. [2012] this could be a consequence of assuming linear-in-means average effects; indeed, when relaxing this assumption, they find large peers' effects for some groups of individuals. On the other hand, studies looking at "social" outcomes like program participation, group membership, church attendance, alcohol consumption, drug use, teenage pregnancy and criminal behavior find much bigger effects [Case and Katz, 1991, Gaviria and Raphael, 2001, Sacerdote, 2001, Duflo and Saez, 2003, Boisjoly et al., 2006, Maurin and Moschion, 2009, Mora and Oreopoulos, 2011, Dahl et al., 2014].

This paper is novel not only because of the outcome and type of peers analyzed, but also because in the Chilean setting I can overcome a challenge commonly faced by previous studies, which is how to define the relevant peer group. Defining it as the whole class or looking at neighbors' effects using a too large definition of neighborhood may dilute the effects of the actual peers. The detailed information I have on neighbors allows me to study how the effects evolve with physical and social distance.

Secondly, this paper contributes to the literature on underinvestment in higher education. This literature has shown that especially in disadvantaged contexts, individuals face constraints that prevent them of taking full advantage of the education opportunities that they have available. The hypotheses most commonly studied for explaining

[5] See for instance Hoxby [2000], Boozer and Cacciola [2001], Hoxby [2002], Hanushek et al. [2003], Angrist and Lang [2004], Burke and Sass [2013], Ammermueller and Pischke [2009], Lavy and Schlosser [2011], Mora and Oreopoulos [2011], Imberman et al. [2012], Lavy et al. [2012], Sojourner [2013], Bursztyn and Jensen [2015]

[6] These papers study the relation between education investment decisions of siblings. Goodman et al. [2015] look at correlations between siblings' college-major choices, Dustan [2018] at the choice of high school and Joensen and Nielsen [2018] at the subjects taken during high school. Although related to my work, the focus of these studies is on different margins.

this phenomenon are liquidity and information constraints, but there is also evidence that other behavioral constraints play a role.[7] In this paper, I add a new element to the analysis by investigating the role of neighbors and siblings on the university enrollment decision. These peers could contribute to reduce some of the frictions previously discussed. In addition, by exploiting variation that comes from a funding program, I can study their indirect effects (i.e. effects of funding on the peers of the direct beneficiaries).

Finally, it adds to the literature on neighborhood effects. This literature has shown that exposure to a better neighborhood as a child reduces teenage pregnancy, improves future earnings and increases the probability of college enrollment [Chetty et al., 2014, 2016, Chetty and Hendren, 2018a,b].[8] However, from these results we cannot tell to what extent the observed effects are driven by exposure to better peers or to better institutions (i.e. schools, infrastructure, security).[9] This paper focuses on the role of peers by exploiting a source of variation that allows the identification of neighbors' effects keeping the neighborhood where individuals live fixed.

The rest of the paper is organized in seven sections. The second section describes the Chilean higher education system, while the third describes the data. The fourth section discusses the identification strategy, and the fifth the main results of the paper. The sixth section looks at siblings and investigates responses of potential applicants in other educational outcomes. The seventh section discusses mechanisms and relate the main results of the paper to previous findings. Finally, section eight concludes.

[7]Examples of papers studying liquidity constraints include Dynarski [2000], Seftor and Turner [2002], Dynarski [2003], Long [2004], van der Klaauw [2002], Solis [2017]; on the other hand examples of papers investigating information frictions include Bettinger et al. [2012], Busso et al. [2017], Dinkelman and Martínez A. [2014], Hastings et al. [2015, 2016], Hoxby and Turner [2015], Oreopoulos and Dunn [2013], Wiswall and Zafar [2013], Booij et al. [2012], Nguyen [2008], Castleman and Page [2015]. Carrell and Sacerdote [2017] on the other hand, argues that differences in support and encouragement are key to explain this. Lavecchia et al. [2016] discusses the literature on behavioral constraints that may explain why some individuals do not take full advantage of their education opportunities.

[8]This has been an active area of research in the last decade. Damm and Dustmann [2014], Fryer and Katz [2013], Kling et al. [2005, 2007], Ludwig et al. [2012] are examples of papers exploiting experimental or quasi experimental variation to study neighborhood effects on mental health, wellbeing, criminal behavior, among others.

[9] The policy implications of these two alternative explanations are very different. As Burdick-Will and Ludwig [2010] point out, if neighborhood effects are mainly driven by the quality of local institutions, then educational attainment could be improved by investing in these institutions without having to move disadvantaged individuals to different areas.

2.2 Higher Education in Chile

This section describes the higher education system in Chile. It begins characterizing the institutions that offer this level of education; continues explaining the university admission system and finishes discussing the main financial aid programs available in the country.

2.2.1 Institutions and Inequality in the System

In Chile, higher education is offered by three types of institutions: vocational centers, professional institutes, and universities. Only universities can grant academic degrees, and in 2017 they attracted 48.1% of the students entering higher education.

Despite the expansion experienced by the higher education system in the last decades, inequality in access to university remains high.[10] According to the national household survey (CASEN), in 2015 individuals in the top decile of the income distribution were 3.5 times more likely to attend university than students in the bottom decile.[11]

Although part of this inequality can be explained by differences in academic potential measured by students' performance in standardized tests in grade 10, Figure 2.1 shows that the gap in university enrollment persists along the ability distribution. This figure also shows that while on average, low-income students are less likely to attend university, in some municipalities their enrollment is higher in comparison to wealthier students from other locations.

2.2.2 University Admission System

In Chile, there are public and private universities. All the public universities and 9 out of 43 private universities are part of the Council of Chilean Universities (CRUCH),

[10]According to figures of the Ministry of Education, the number of students going to university was five times bigger in 2017, than in 1990. The number of students going to professional institutes increased by 10 times over the same period. In the case of vocational centers, it doubled.

[11]Figure 2.E2 in the appendix illustrates university attendance rates for the whole income distribution. Also according to the CASEN, the main reasons for not attending higher education among individuals between 18 and 24 years old are personal (49.7%) and economic (47.5%). Academic performance is mentioned in less than 1.5% of the cases as a reason for not going to higher education.

an organization that was created to improve coordination and to provide advice to the Ministry of Education in matters related to higher education. For-profit universities are forbidden under the Chilean law.

The CRUCH universities and since 2012 other eight private universities select their students using a centralized deferred acceptance admission system that only considers students' performance in high school and in a national level university admission exam (PSU).[12] The PSU is taken in December, at the end of the Chilean academic year, but students typically need to register before mid-August.[13] Since 2006 all the students graduating from public and voucher schools are eligible for a fee waiver that in practice makes the PSU free for them.[14]

The universities that do not participate in the centralized system have their own admission systems.[15] Although they could use their own entrance exams, the PSU still plays an important role in the selection of their students, mostly due to strong financial incentives for both the students and the institutions.[16] For instance, the largest financial aid programs for university require students to score above a cutoff in the PSU.

[12] The PSU has four sections: Language, mathematics, social sciences and natural sciences. The raw scores obtained by students in each of these sections are adjusted to obtain a normal distribution of scores with mean 500 and standard deviation 110. The extremes of the distribution are truncated to obtain a minimum score of 150 and a maximum score of 850 in each section. In order to apply, students need to take language, mathematics and at least one of the other sections. Universities are free to set the weights allocated to these instruments for selecting students. Students apply to their programs of interest using an online platform. They are asked to rank up to 10 programs according to their preferences. Places are then allocated using an algorithm of the Gale-Shapley family that matches students to programs using their preferences and scores as inputs. Once a student is admitted in one of her preferences, all the others are dropped.

[13] In 2017, the registration fee for the PSU was CLP 30,960 (USD 47).

[14] Around 93% of the high school students in Chile attend public or voucher schools. The entire registration process operates through an online platform that automatically detects the students' eligibility for the scholarship.

[15] I observe enrollment in all the universities of the country, independently of the admission system they use.

[16] Firstly creating a new test generate costs for both the institutions and the applicants. Secondly part of the public resources received by higher education institutions depends on the performance of their first-year students in the PSU. This mechanism was a way of rewarding institutions that attracted the best students of each cohort. It was eliminated in 2016, but it was in place during the period analyzed in this study.

2.2.3 Financial Aid

In Chile, the majority of the financial aid comes from the government. There are two student loans and multiple scholarship programs designed to fund studies in different types of higher education institutions. The allocation of these benefits is in charge of the Ministry of Education. This section, briefly describes the programs that fund university degrees, emphasizing the rules that generate the discontinuities that are later exploited in this paper.

Students that need financial aid have to apply to it between October and November using an online platform (this is before taking the PSU). After verifying if the information provided by applicants is correct, the Ministry of Education informs them to which benefits they are eligible. Something similar occurs once the PSU scores are published; the Ministry of Education incorporates this new information to the system and updates the list of benefits that students could receive based on their performance. This allows them to consider the funding they have available before applying and enrolling in higher education.

There are two student loans programs: solidarity fund credit (FSCU) and state guaranteed credit (CAE). The former can be used solely in CRUCH universities, while the latter can be used in any higher education institution.[17] In order to be eligible for these loans, students need to obtain an average PSU score (language and mathematics) above 475 and come from households in the bottom 90% of the income distribution.[18]

Eligibility for student loan creates a discrete jump in the probability of university enrollment amongst financial aid applicants [Solis, 2017] . This is the discontinuity that I exploit to study the indirect effect of university attendance on the neighbors and siblings of the loans' beneficiaries.

The majority of the scholarship programs are allocated following a similar logic; the

[17]Although both programs are currently very similar, during the period under study they had several differences; for instance, while the annual interest rate of the FSCU was 2%, in the case of the CAE it varied between 5% and 6%. On top of that, while the repayment of the FSCU has always been income contingent, the CAE used to have fixed installments.

[18]In the case of the FSCU, they need to come from households in the bottom 80% of the income distribution; the CAE, on the other hand, used to be focused on students in the bottom 90% of the income distribution, but since 2014 the loan is available to anyone that satisfies the academic requirements.

main difference is that the academic requirements are higher (i.e. PSU average above 550) and that they are focused on students from more disadvantaged backgrounds. I do not use this discontinuity because it does not affect enrollment; once students have access to subsidized credits such as CAE and FSCU, offering them a scholarship does not make a difference in their decision to attend university. There are also a few programs that instead of requiring a minimum score in the PSU, allocate funding based on high school performance. These programs are relatively small, both in terms of beneficiaries and in terms of the support they offer.

Given that in Chile universities have complete freedom to define their tuition fees, the government sets a reference tuition fee for each program and institution as a way to control public expenditure. These reference tuition fees define the maximum amount of funding that a student can receive from the government in a specific program.[19] At the university level, the reference tuition fee is around 80% of the actual fee. This means that students need to cover the additional 20% using their own resources, taking a private loan or if available, applying to scholarships offered by their institutions.

2.3 Data

This section describes the sources of the data and the sample used to study the effects of neighbors on potential applicants' probability of enrolling in university.

2.3.1 Data Sources

This paper combines administrative data of different government agencies, including the Chilean Ministry of Education and the Department of Evaluation, Assessment and Educational Records (DEMRE) of the University of Chile, which is the agency in charge of the PSU. In addition, it uses data from the Ministry of Social Development, from the Education Quality Agency and from the Census.

This data makes it possible to follow students throughout high school. It contains information on demographic characteristics, attendance and academic performance (GPA)

[19]The only exception to this rule is given by the CAE. In this case, students still cannot receive more than the reference tuition fee through the CAE, but they can use it to complement scholarships or the FSCU, up to the real tuition fee.

for each individual in every grade. In addition, it registers the educational track chosen by students and also schools characteristics such as their administrative dependence (i.e. public, voucher, private) and the municipality where they are located. All this information is available from 2002 onwards, meaning that the first cohort that I can follow between grades 9 and 12 is the one completing high school in 2005.

I also observe all the students who register for taking the PSU. As discussed in Section 2.2, the PSU is free for students graduating from public and voucher high schools, so the majority signs up for the test even if they do not plan to apply to university.[20] Apart from observing the scores that students obtain in each one of the sections of this exam, the data contains information on applications to the universities that participate of the centralized admission system (see Section 2.2 for more details). This includes the list of all the programs to which students apply and their admission status. The PSU registers also contain demographic and socioeconomic variables of the students and their families, including household income, parental education, parents' occupations and family size. These variables are later used to study if the identifying assumptions of the RD are satisfied, and to perform heterogeneity analyses. These registers also include students' addresses and a unique identifier of parents. This information is used to identify neighbors and siblings.[21]

The Ministry of Education keeps records of all the applications and the allocation of financial aid. The type and amount of benefits are only observed for individuals who enroll in higher education, what means that it is not possible to know if students not going to higher education were actually offered funding. However, the eligibility rules are clear and all the applicants satisfying the academic and socioeconomic requirements should be offered a student-loan or a scholarship.

Finally, I also observe enrollment in higher education. These records contain individual-level data of students attending any higher education institution in the country, and

[20] In the period that I study, more than 85% of the high school graduates appear in the registers of the PSU.

[21] The information on demographic and socioeconomic variables, addresses and parents is not available for all the students in the registers. Some of it can be recovered from the secondary and higher education registers. The baseline specifications do not use controls. Observations with missing values are not used when performing heterogeneity analyses.

report the programs and institutions in which students are enrolled.[22] This data, as the data on financial aid, is available from 2006 onwards.

The neighbors' and siblings' samples combines information from all these datasets.[23] The former includes students that appear in the PSU registers between 2006 and 2012, while the latter students that appear in the PSU registers between 2006 and 2015. The difference in the years included in each sample is just driven by data availability.

2.3.2 Sample Definition

This section describes the steps and restrictions imposed on the data to build the estimation sample. The first step in this process is to match potential applicants observed in time t with their neighbors observed in $t-1$.

To make this possible, I geocoded students' addresses. Since these addresses do not include postcodes, the geocoding process was very challenging, especially in regions with high levels of rural population where the street names are not well defined. Therefore, this study focuses on three regions where the identification of neighbors was easier and that together represent more than 60% of the total population of the country: *Metropolitana of Santiago*, *Bio-bío* and *Valparaíso*.[24]

After geocoding the addresses, potential applicants of year t were matched to their 60 closest neighbors registered for taking the PSU in $t-1$. Then, the demographic, socioeconomic and academic variables from other datasets were added to potential applicants and their neighbors. Finally, each individual was linked to their respective census block and neighborhood unit. Census blocks are the smallest geographic units used in the census, and in urban areas they usually coincide with an actual block; the neighborhood units correspond to subareas within a municipality.[25] They were de-

[22] This dataset includes students enrolled in university, professional institutes and vocational centers.

[23] Although the focus of this paper is on neighbors, I also investigate what happens with potential applicants when an older sibling goes to university T years before her. The sample used for this purpose is described in Section 2.A in the appendix.

[24] Even in these regions, it was not possible to geocode 100% of applicants' addresses. I identified addresses for near 85% of the sample. This implies that for some applicants, only a subset of close neighbor was identified. Unless the missing neighbors are selected in a systematic way, this should work against finding effects.

[25] Standard errors are clustered at this level.

fined by the Ministry of Social Development to decentralize certain local matters and to foster citizen participation and community-based management. After this process, I end with a sample of more than 550,000 potential applicants and their respective neighbors.

To build the estimation sample, I apply two additional restrictions. I only keep individuals graduating from regular education programs no more than 3 years before registering for the PSU (i.e. no remedial programs), and individuals who were between 17 and 22 years old when taking the test. In addition, I drop applicant-neighbor pairs in which the applicant completes high school before the neighbor. Finally, I also drop the pairs in which applicants and neighbors are siblings. These restrictions made me lose around one third of the observations.[26]

The main analyses focus on potential applicants and their closest neighbor, but I also study how they are affected by other individuals that live close to them (i.e. n-th closest neighbor, best neighbor among n and best neighbor within d meters). In all these cases, I work only with potential applicants whose neighbors apply to financial aid; these are the only neighbors that could change their university enrollment decision based on eligibility for student-loans. As a consequence of this last restriction, another third of the original sample is lost. Note that this restriction is only imposed on neighbors, and does not affect potential applicants.[27]

The first two columns of table 2.1 present summary statistics for the sample of potential applicants and their closest neighbors. The third column characterizes all the students in the PSU registers between 2007 and 2012.

Potential applicants and their closest neighbors are very similar. The only relevant differences are in academic variables. Neighbors, who by definition apply to financial aid, are more likely to have chosen the academic track during high school. They also obtain better scores in the PSU, a result that is in part driven by the fact that more of them actually take the test. Despite the restrictions imposed to build this sample,

[26]Note that these restrictions do not affect the internal validity of the identification strategy.

[27]Once more this restriction does not affect the internal validity of the identification strategy. The restriction is only imposed on the neighbors. Potential applicants are in my sample even if they do not apply for funding.

potential applicants look very similar to the rest of the individuals that appear in the PSU registers.

2.4 Identification Strategy

The identification of neighbors' effects is challenging. Families are not randomly allocated to neighborhoods and once in the neighborhoods they face similar circumstances, which makes it difficult to distinguish between social interactions and correlated effects. In addition, if peers' outcomes have an effect on each other, this gives rise to what Manski [1993] described as the "reflection problem".

This paper studies how close neighbors going to university in year $t-1$ affect individuals that could apply to university in year t. Since neighbors decide whether to enroll or not into university before the potential applicants, their decision should not be affected by the decision of potential applicants. If the decision of the younger applicant does not affect the decision of the older neighbor the "reflection problem"disappears.

To identify these neighbors' effects I exploit the fact that eligibility for student loans depends on the score obtained in the PSU. This allows me to implement a fuzzy RD instrumenting neighbors' university enrollment (U_n) with an indicator variable that takes value 1 if the student's PSU score is above the student loans eligibility cutoff (L_n). This means that the variation on neighbors' university enrollment comes only from eligibility for funding. Thus, even if the decision of the younger applicant would affect the decision of the older neighbor, by using this instrument I am able to abstract from the "reflection problem".

In addition, since neighbors around the student loans eligibility threshold are very similar, this approach also eliminates concerns related to correlated effects. [28]

Using this strategy, I estimate the following specification:

$$U_{at} = \alpha + \beta_n U_{nt-1} + \mu_t + \varepsilon_{at} \tag{2.1}$$

[28] Apart from neighbors, also the neighborhoods and the potential applicants who live near them are similar.

Where U_{at} is the university enrollment status of potential applicant a on year t and U_{nt-1} is the university enrollment status of neighbor n on year $t-1$.

Note, this specification only includes neighbor n. In order to interpret β_n as local average treatment effect (LATE) of neighbor n on potential applicant a, in addition to the IV assumptions discussed by Imbens and Angrist [1994], we need to assume that university enrollment of contemporaneous peers does not affects applicants' own university enrollment (Section 2.B in the appendix discuss this in detail).[29]

If this assumption is not satisfied, β_n can be interpreted as a reduced form parameter capturing not only the effect of neighbor n on potential applicant a, but also the effects that other neighbors affected by n generate on a. This is still a relevant parameter from a policy perspective.

For the RD estimation, I use optimal bandwidths computed according to Calonico et al. [2014b] and provide parametric and non-parametric estimates. 2SLS estimates come from specifications that assume a flexible functional form for the running variable and instrument U_{nt-1} with a dummy variable that indicates if neighbor n was eligible for student loans on $t-1$, L_{nt-1}. Non-parametric estimates come from local polynomials regressions that use a triangular kernel to give more weight to observations closer to the cutoff. The implementation of this approach follows Calonico et al. [2014a] and Calonico et al. [2017].

Section 2.D in the appendix presents a series of analyses that investigates if the assumptions required for the validity of the RD estimates are satisfied. First, it shows that there are no discontinuities at the cutoff in a rich set of demographic, socioeconomic and academic characteristics of potential applicants and their neighbors.

Second, it provides evidence that there is no manipulation of the running variable around the cutoff. In order to study this, I implement the density discontinuity test suggested by Cattaneo et al. [2018a].[30]

[29]Considering the timing of the application and enrollment process, individuals have limited scope to respond to university enrollment of their contemporaneous peers. Figure 2.E4 shows that this is the case and that contemporaneous peers do not seem to affect potential applicants' enrollment

[30]In this setting, it is not easy to think of a way in which applicants could manipulate the running variable. All the PSU process, from the creation to the correction of the tests, is carried out under strict measures of security. In addition, final scores are the result of a transformation

In addition to the robustness checks just mentioned, I also study if potential applicants' decision of going to university has an effect on their older neighbors. As discussed earlier, there should be no effect in this case, something that is corroborated by the results of this exercise.

Finally, section 2.D also shows that the results are robust to different bandwidths choices and that there are no jumps like the ones observed at the student loans eligibility cutoff in other points where there should not be.

2.5 Results

This section discusses the main findings of the paper. It uses the definitions introduced in Section 2.4 according to which potential applicants are individuals that could go to university on year t, while the neighbors are individuals that applied to university on year $t-1$. This section begins by looking at what happens with potential applicants' enrollment probability when their closest neighbor is eligible for student loans and goes to university.[31] Then, it incorporates other close neighbors to the analysis and studies how the effect evolves with physical distance. It concludes by investigating heterogeneous effects by social distance and by the university enrollment rates observed in potential applicants' municipalities.

2.5.1 Effect of the closest neighbor on potential applicants' enrollment

In order to study how potential applicants' enrollment probability changes when their closest neighbor goes to university, I estimate a specification like the one presented in equation 2.1, instrumenting neighbors' university enrollment with their eligibility for student loans.

Panel (a) of Figure 2.2 illustrates the first stage of this exercise. It shows that neighbors' probabilities of going to university increase by around 18 percentage points when

that adjusts raw scores so that they follow a normal distribution. This makes it difficult to know ex ante the exact number of correct answers needed to be just above the cutoff. Considering this, it seems very unlikely that potential applicants could manipulate their neighbors' scores.

[31]Section 2.C in the appendix studies how potential applicants respond to what happens to other neighbors, including the best among n and the best within d meters.

they become eligible for a loan. This figure, significantly different from zero, captures the direct effect of student loans on university enrollment. According to it, this type of funding roughly doubles the probability of going to university for students with PSU scores near the eligibility threshold.

Panel (b), on the other hand, illustrates the reduced form. It shows that potential applicants whose closest neighbor is eligible for a student loan in year $t-1$ are around 2 percentage points more likely to enroll in university on year t. This figure is statistically different from zero and measures part of the indirect effect of offering funding for university. According to this result, student loans not only have an effect on their direct beneficiaries, but also on the close neighbors of these beneficiaries. This indirect effect represents more than a 10% of the direct effect of student loans on university enrollment.[32]

If this reduced form effect only works through neighbors taking-up the student loans and going to university, the first stage and reduced form estimates can be combined to estimate the effect of exposure to a close neighbor going to university on potential applicant's university enrollment. Table 2.2 presents estimates obtained using a parametric and non-parametric approach. The first two columns show 2SLS estimates, while the third and fourth show estimates obtained using linear and quadratic local polynomials instead. According to these results, potential applicants' probability of going to university increases by more than 10 percentage points when their closest neighbor enrolls in university. This figure is statistically different from zero, and represents around one third of the enrollment probability of individuals at the cutoff.

This estimate would be an upper bound of the effect of neighbors' enrollment on applicants' enrollment if having a close neighbor eligible for funding makes potential applicants more aware of funding opportunities, independently if the neighbor goes or

[32]Results in Table 2.E2 complement these analyses by investigating if the difference in enrollment persist one year after the shock. The estimates reported on this table are very similar to the ones discussed in this section, suggesting that the compliers of the IV do not drop out at higher rate than always takers. To study this I define two outcomes. The first indicates if applicants enroll and remain enrolled in any university one year after they decides to enroll. The second indicates if applicants enroll and remain in the same university one year after enrollment. Both outcomes are 0 for applicants who do not enroll and for applicants who enroll, but dropout.

not to university.[33] However, the information intervention implemented by Busso et al. [2017] among grade 12 students in Chile, shows that in this setting learning about funding opportunities alone does not generate responses like the ones I find,[34], alleviating concerns related to this type of violations to the exclusion restriction.

2.5.2 How do neighbors' effects evolve with distance?

This section investigates how neighbors' effects evolve with physical and social distance. Both types of distance can be relevant if they affect the likelihood of interactions between individuals.

All the results discussed so far have focused on the closest neighbor. However, there could be other neighbors that are relevant for potential applicants. In order to study this, I estimate the same baseline specification presented in Section 2.4, but replacing university enrollment of the closest neighbor by university enrollment of the n-th closest neighbor.

Thus, I estimate independent specifications to study how each one of the eight closest neighbors affect potential applicants' university enrollment. As discussed in Section 2.4 to interpret the results of this specification as the effect of neighbor n on potential applicant a, we need to assume that university enrollment of contemporaneous peers does not affects individuals' university enrollment.[35] If this assumption is not satisfied, then the estimated coefficient can be interpreted as a reduce form parameter that captures not only the effect of neighbor n on applicant a, but also the effect that other individuals affected by n have on a.[36]

[33]If the applicant becomes aware of the funding opportunities only when the neighbor uses it and goes to university, then this would be a mechanism through which exposure works and not a violation to the exclusion restriction.

[34]This intervention provided students with tailored information about funding opportunities and labor market outcomes of graduates from different programs. They find no extensive margin responses. They find no increase in enrollment to non-selective or selective institutions.

[35]I show that this indeed seems to be the case in Figure 2.E4.

[36]A more detailed discussion on this is presented in Section 2.B. An alternative approach to study this would be to include the enrollment status of multiple neighbors simultaneously in the same specification. In case of counting with instruments for the enrollment of each neighbor it would be possible to proceed in a similar way as I do now. In my setting, this is not possible. The instrument I have is valid only locally. In addition, it is relevant only for neighbors that apply for financial aid. To estimate a specification like this one, I would need to find applicants with many neighbors applying for funding and with PSU scores close enough to the eligibility threshold. Unfortunately, this type of potential applicants are scarce in my sample.

Panel A on Figure 2.3 reports OLS and RD estimates for this analysis. Each dot corresponds to the estimates obtained from the eight independent regressions mentioned in the previous paragraph. The horizontal axis, apart from reporting the relative distance to the applicant, presents in parenthesis the average distance between the n-th closest neighbor and applicant a. According to the figure, on average potential applicants live at 40 meters from their closest neighbor registered for the PSU the previous year, and at about 60 meters from the second closest one. The RD estimates, represented by blue circles, quickly decay. The coefficient associated to the second closest neighbor is around 5 percentage points, and in the case of the third closest neighbor it is below 3 percentage points. In addition, only the coefficient associated to the closest neighbor is significantly different from zero. The pattern observed in the case of OLS is substantially different. Although there is a small drop on the size of the coefficient, they are very persistent.

In order to study how the effects evolve with physical distance, I estimate an additional specification in which potential applicants and their ten closest neighbors are pooled together. I present two set of results. The first one comes from splitting the sample in three equal parts depending on the distance between potential applicants and their neighbors. The second one comes from a specification that uses the whole sample and adds an interaction between neighbors' university enrollment and distance.

As illustrated in Panel B of Figure 2.3 the pattern of the RD estimates presented in blue are consistent with the results on Panel A. The effect of neighbors on potential applicants decays with distance, becoming non-significant at 100 meters and reaching 0 at 200 meters. As before, the OLS estimates are persistent, and in this case they even seem to increase a little bit.

The difference between OLS and RD estimates illustrate the relevance of correlated effects in this context. As discussed earlier, the composition of neighborhoods is not random, which means that individuals who live relatively close to each other are similar in many dimensions (i.g. household income, parental education). In addition, these individuals live under similar circumstances, and are exposed to similar institutions and shocks. Thus, it is not surprising to find a persistent correlation in outcomes of neighbors, even if they do not interact with each other.

In the context of peers' effects, these results also highlight the importance of using an appropriate reference group. The results discussed in this section suggest that interactions between neighbors occur at a very local level. Therefore, using a too broad definition of neighborhood could dilute the effect of the relevant peers (i.e. what happens with individuals living 200 meters apart does not seem to be relevant for potential applicants).

The extent to which individuals interact with each other is not only determined by physical distance. In the rest of this section I study how the effects evolve depending on social distance and depending on time spent at the neighborhood. Given the results just discussed, I focus my attention only on the closest neighbor and to study heterogeneity I split the sample in different sub groups.

The results in Table 2.3 suggest that the effects are bigger when potential applicants are closer to their neighbors in socioeconomic status and gender. In the case of age, a similar pattern emerges, but the differences is smaller. This could be due to the fact that age differences between individuals registered for taking the PSU in consecutive years are not huge.[37] Although the precision of these estimates does not allow me to rule out that they are equal, finding that the coefficients are larger when individuals are closer in social terms is consistent with the idea that interactions between neighbors are important for these effects to arise.

In line with these results, Table 2.4 shows that the effect seems to be stronger for potential applicants who have lived for longer in the neighborhood and for the cases in

[37] Socioeconomic status is measured by an index that combines information on household income, parental education, health insurance and high school administrative dependence. This index is build by extracting the first component from a principal component analysis that included household income, parental education, health insurance and high school administrative dependence. Using this index, potential applicants and neighbors are classified in three socioeconomic groups; they are defined as similar if the absolute difference between the indexes of the applicant and her closest neighbor are within $1 - \sigma$ of the index. Table 2.E1 in the appendix present additional heterogeneity analyses. According to these results, students coming from very disadvantaged backgrounds or who follow the vocational track during high school are less responsive. This suggest that the effects are driven by potential applicants who are better prepared for the PSU and for whom it is easier to score above the student loans eligibility threshold and to be admitted in some university if they decide to apply. The estimated effect is also bigger for females, although the difference with the effect estimated for males is not statistically significant.

which neighbors plan to continue living with their parents in case of going to university (i.e. plan to remain in the neighborhood). The effect is also stronger for potential applicants whose mothers do not work outside the household. The time spent by the applicants and other members of their families at the neighborhood may strengthen the relations between neighbors, increasing in this way the likelihood of exposure and interactions.

2.5.3 Urban Segregation and Inequality in University Enrollment

As discussed in Section 2.2, access to university is very unequal in Chile. Given the high levels of urban segregation that exist in the country, this also translates into spatial inequality. The map in Figure 2.4 illustrates this for Santiago, Chile's capital city. The red areas in the map correspond to municipalities where on average 20% of potential potential applicants go to university, while the green areas represent municipalities where this figure is above 50%.

According to the results discussed in previous section, programs that expand access to university generate indirect effects on the close peers of the direct beneficiaries. The estimates obtained when looking at potential applicants and their closest neighbor indicate that the indirect effects of student loans represent a little bit more than 10% of their direct effect. In order to estimate the full extent of these indirect effects, we would need to investigate if they also emerge between other peers[38] In addition, we would need to consider that potential applicants who enroll in university as a consequence of these indirect effect could also affect university enrollment of other individuals in the future, making the indirect effect to grow over time.

So far, the analyses have assumed that direct and indirect effects are constant across different regions. However, they may change depending on the number of individuals that usually goes to university in these areas. I study this by estimating the direct and indirect effect of student loans independently for low, mid and high attendance municipalities.[39]

[38]According to the results discussed in Section 2.5.2, in the context of neighbors these spillover effects seem to be very local. Section 2.6.1 studies indirects effects between siblings.

[39]The map in Figure 2.4 illustrates the geographic distribution of these three groups for Santiago, Chile's capital city.

Figure 2.5 presents the results of this exercise. The top panel shows first stage estimates, the panel in the middle reduced form estimates, and the panel at the bottom 2SLS estimates. These last estimates capture the effects of neighbors' enrollment on potential applicants' enrollment.

The pattern illustrated in this figure shows that the direct effect (i.e. the share of individuals who take up student loans and go to university) is bigger in areas where university attendance rates are higher. The reduced form results and the exposure effects on the other hand seem stronger in low and mid attendance areas. Indeed, in high attendance areas these coefficient are non-significant and are considerably smaller.[40]

Although the standard errors of these estimates do not allow me to conclude that they are statistically different, these results shows that indirect effects are relevant in low and mid attendance areas. They represent roughly a 15% of the direct effects, indicating that exposure to neighbors who are eligible for funding and go to university affects the enrollment of potential applicants. This suggests that policies that increase exposure to these type of neighbors would also increase enrollment in areas where university attendance is relatively low.

A back of the envelope calculation shows that in case of increasing exposure to university-going neighbors in low attendance municipalities to the levels observed in those with high attendance, the gap in university enrollment would drop by around 5 percentage points (i.e. enrollment in low attendance areas would rise from 20% to 25%).[41]

The previous exercise does not say anything about how to increase exposure. An alternative would be to relax the criteria defining eligibility for funding in areas where attendance is low. However, not everyone who is offered funding goes to university. According to the first stage results, in these areas eligibility for student loans increases

[40] To enter my estimation sample, potential applicants need to have neighbor with a PSU score close enough to the eligibility threshold. In areas with very high attendance, this is not very common. In order to obtain three samples of similar size, the high attendance areas include places where university attendance varies between 40% and 75%. The potential applicants of high attendance municipalities that appear in my estimation sample come from places where attendance is closer to the lower bound of this range.

[41] This exercise assumes that local treatment effects are a good representation of average treatment effects. In addition, it ignores general equilibrium responses. This figure comes from multiplying the difference in exposure between high and low attendance municipalities, and the 2SLS estimate of the effect of exposure for low attendance municipalities.

the probability of enrollment by about 15 percentage points. Assuming that this number is a good approximation of how individuals below the current eligibility cutoff would respond in case of being offered funding, the direct effect of a policy that lowers the cutoff by 50 points can be computed multiplying the share of people with scores in the new eligibility range and the first stage coefficient. In municipalities with low university attendance, a policy like this this would increase enrollment by 3 percentage points. Given that the indirect effect is proportional to the direct effect, this would make the indirect effect small as well. One year after the student loans expansion, the increase in enrollment generated by the indirect effect would be equal to 0.5 percentage points. Assuming that these additional individuals going to university also affect the enrollment of other applicants in the future, the increase in enrollment that the indirect effect would be generating after five years would be equal to 0.6 percentage points, representing a 20% of the direct effects.[42] Note that policies with larger direct effects generating a more significant increase in exposure would also have more relevant indirect effects in absolute terms.

A final consideration to think about the design of policies to expand access to university is that depending on the mechanisms behind these neighbors' effects, there could be more efficient ways of providing potential applicants with what university going neighbors give them. I discuss mechanisms in Section 2.7.

2.6 Siblings and Other Educational Outcomes

This section starts by investigating if indirect effects as the ones discussed in previous sections also arise among siblings. Then it studies how university enrollment of neighbors and siblings affect other educational outcomes of potential applicants, to understand what are the margins that they adjust that result on the increase I document in university enrollment.

[42]To obtain this last figure, I assume that each individual induced to enroll in university as a consequence of exposure also affect other potential applicants. Thus, the indirect effects after 5 years can be computed as $IE = 0.03 \cdot \frac{1-\beta^5}{1-\beta}$

2.6.1 Siblings Effects

Neighbors are not the only peers that may affect university enrollment of potential applicants. If indirect effects as the ones described in previous sections also arise in other settings, this is something that we would like to incorporate to the evaluation and design of policies that seek to expand access to university.

As discussed in Section 2.3, apart from identifying neighbors, my data allows me to identify siblings. I use this data to study how having an older sibling going to university affects potential applicants' university enrollment. To do this, I estimate the same specification used in the case of neighbors, but replacing neighbors by siblings.

Although the siblings' sample is similar to the neighbors' sample, it is worth mentioning that it covers a longer period of time —2006 to 2015— and that the potential applicants in this sample (i.e. the younger siblings) obtain higher PSU scores than potential applicants in the neighbors sample.

The top panel of Figure 2.6 shows that siblings who are eligible for student loans are around 16 percentage points more likely to enroll in university than to those who are not eligible. This figure, statistically different from zero, represents the direct effect of student loans in this group.

The second panel presents the reduced form. It shows that potential applicants with an older sibling eligible for student loans are around 2.5 percentage points more likely to go to university than those whose older sibling are not eligible. This indirect effect is slightly bigger than in the case of neighbors. This result is consistent with the idea that exposure is relevant. If interactions between siblings are more intense than between neighbors, this could explain the difference in the coefficient.[43]

As in the case of neighbors, if these effects are purely going through older siblings taking up the loans and going to university, the first stage and reduced form results can be combined to obtain an estimate of the effect of exposure to siblings going to university. Table 2.5 presents these results. The first two columns show 2SLS estimates, while the third and fourth columns show estimates obtained using linear and quadratic local

[43] The samples used to estimate neighbors and siblings are different. This could also be behind the differences in the effects documented for neighbors and siblings.

polynomials. According to these figures, having an older sibling going to university increases potential applicants' probability of going by 15 percentage points. As in the case of the reduced form, this coefficient is also bigger than in the case of neighbors.

In this setting however, satisfying the exclusion restriction is more challenging. Apart from the transmission of information about funding opportunities, having an older sibling going to university with a student loans could affect the household budget constraint. This could explain at least some part of the response observed in younger siblings.

However, it is important to consider that student loans only cover a share of the tuition fees.[44] This means that students and their families still need to pay part of them, in addition to other costs involved in going to university, including the foregone earnings of labor. If families have limited resources, then sending one child to university, even with a student-loan, should reduce the chances of going for the younger ones.[45]

There could also be scenarios where the student loans could relax the household budget constraint in a more significant way.[46] Thus, it cannot be ruled out that at least some part of the effects found for siblings are driven by changes in household resources.

2.6.2 Other Educational Outcomes

This section looks at changes on the academic performance and on the application decisions of potential applicants. This allows us to identify the margins that potential applicants adjust and that mediate the increase in enrollment documented in previous sections. To study this, I employ once more the fuzzy RD used in previous sections, but this time to investigate how these other outcomes change when a close neighbor or an older sibling goes to university.

[44] student loans cover up to the reference tuition fee set by the government. See Section 2.2 for more details.

[45] An exception to this could be given by siblings whose age difference is big enough to allow the older one to graduate before the younger one applies to university. In this case, the older sibling could help to fund the younger sibling studies. However, I do not find differences depending on the age gap between siblings. These results are not presented in the paper, but are available upon request.

[46] This would be the case if for instance parents were able to save or borrow to pay for exactly one university degree; or if having one child in university would change their willingness to borrow.

According to the results presented in table 2.6, potential applicants with a close peer (i.e. closest neighbor or sibling) going to university are more likely to take the PSU, and to apply for financial aid and to university; [47] they are also more likely to be eligible for student loans and to take them up. I find no effects on attendance or academic performance during high school, [48] and the documented improvement on PSU scores is driven by the extensive margin response mentioned earlier.[49]

Although the coefficients on the application responses are not always precisely estimated, they represent an important fraction of the changes in potential applicants' enrollment. This suggests that the increase in enrollment is driven by a change in the decision to apply. This is consistent with the results on undermatching discussed by Hoxby and Avery [2013] and Black et al. [2015], suggesting that there are students who despite having the potential to be admitted into university and to receive funding for it do not apply.

2.7 Discussion

The results presented in this paper show that exposure to close peers that enroll in university increases the probability of potential applicants' enrollment and that the effects are stronger when potential applicants and their peers are more likely to interact.

These results are consistent with two broad classes of mechanisms. Firstly, neighbors and siblings could affect potential applicants' university enrollment by expanding their access to relevant information. Alternatively, they could affect the costs of going to university.

There is vast evidence that information frictions affect individual schooling decisions in both developing and developed countries. Jensen [2010] for instance shows that pro-

[47] I only observe applications to universities that use the centralized admission system described in section 2.2. These are the applications I use as outcome.

[48] In Chile, the GPA scale goes from 1.0 to 7.0. The minimum GPA to pass to the next grade or to finish high school is 4.0.

[49] I replaced missing scores in the PSU by 0 (or -475 after centering the PSU scores around the student loans eligibility threshold). Thus, if potential applicants with neighbors or siblings going to university are more likely to take the admission test, this automatically creates an increase on average performance (i.e. they are less likely to have -475 points in the PSU).

viding information on returns to education to grade 8 students in Dominican Republic increased the years of high school completed.

Students seem to face similar information frictions in the case of higher education. Hoxby and Turner [2015] study this in the US and shows that in low-income areas, even high achieving students know little about costs, quality and the overall college experience. The situation in Chile is similar. Hastings et al. [2016] document that students from disadvantaged groups have limited and imprecise information on returns to education. These results suggest that university enrollment could be increased by tackling these information constraints.[50]

An alternative way in which neighbors and siblings could affect potential applicants' enrollment is by affecting the costs of going to university. This would be the case for instance if they face a social sanction for going to university. Austen-Smith and Fryer [2005] formalizes this idea and shows that individuals may choose to underinvest in education to gain acceptance in their social group. Along this line, Bursztyn and Jensen [2015] finds that students respond to peers pressure and that when effort is observable they adjust it according to the prevalent social norm (i.e. reduce effort when peers view it as something bad, increase effort when peers value it). This is not the only way in which peers could affect the costs and benefits of going to university. They could also be affected if for instance individuals enjoy spending time with their peers or if they are competitive and want to surpass their peers achievements.[51]

Although I cannot perfectly distinguish between these two classes of mechanisms, not finding responses on high school attendance or an improvement on academic performance, suggests that individuals are not experiencing relevant changes on the costs and benefits of going to university. If this were the case, we would expect them to increase the effort they make to be admitted in college, something that does not seem to be occurring. The results seem more consistent with the transmission of information.

[50]Providing relevant information on returns to higher education though is challenging. As shown by Hastings et al. [2015] these returns can be very different depending on the institution and program attended. n addition, if higher education institutions charge fees, then information about funding opportunities may also be relevant.

[51]Costs and benefits nest multiple ways in which these peers' effects may arise. Changes in aspirations, models of competition, the existence of social norms or models of interdependent preferences can be accommodated to this framework.

However, interventions providing Chilean students with tailored information about funding opportunities and returns to higher education —like Busso et al. [2017] and Hastings et al. [2015] — do not find extensive margin responses and even when looking at the type of institution and program that students attend, they do not find large effects. Similarly, Bettinger et al. [2012] finds no relevant responses in college enrollment to a pure information intervention in the United States. However, when complementing the information with personalized support to fill the application for financial aid, they find that college enrollment increases by a similar magnitude to the one documented in this paper. Also in the United States, Carrell and Sacerdote [2017] designed an intervention to investigate what makes programs that foster college attendance effective. They argue that what makes these programs effective is not the information that they provide, but rather their ability to compensate for the lack of encouragement and support that students receive at home or at the school.

According to these results, expanding information on funding and returns to education has not been very effective in increasing university enrollment. Nevertheless, the information transmitted by peers could be different to the one traditionally provided in information interventions. It may be different on its content,[52] but it might also be more relevant because it comes from someone closer.[53]

With the data that I have available, I cannot tell exactly what potential applicants learn from their peers. This is a potential avenue for future research that would also contribute to gain a better understanding of the mechanisms behind my results.

[52] Apart from learning about funding and returns to education, potential applicants may receive information about the application process, the likelihood of being successful and other elements related to the whole university experience. Hoxby and Turner [2013] shows that providing high achieving applicants from disadvantaged backgrounds with this type of information and an application fee waiver changes the set of colleges to which they apply. This reduces the gap on the type of college to which high achieving students from different backgrounds attend by 5 percentage points.

[53] Nguyen [2008] finds that individuals are able to process information on returns to education in a sophisticated way, and that they respond differently depending on who provides the information.

2.8 Conclusions

Recent studies have shown that especially in disadvantaged contexts individuals face constraints that prevent them from taking full advantage of the education opportunities that they have available. In the context of university enrollment, financial constraints are relevant, but there is growing evidence that the lack of information, support, and encouragement also plays an important role in this context. In both developing and developed countries, these constraints seem to be more relevant in areas where exposure to university is lower.

This paper investigates whether potential applicants' decision to attend university is affected by university enrollment of close neighbors. To address this question, I use rich administrative data from Chile and take advantage of the variation generated by the rules that define eligibility for student loans. Exploiting this quasi-random variation, I implement a fuzzy RD that allows me to eliminate concerns about correlated effects and to abstract from the 'reflection problem'.

I find that neighbors have a large and significant impact on the university enrollment of potential applicants. Having a close neighbor going to university increases their enrollment probability by about 10 percentage points. I also show that this effect is stronger when the interactions between neighbors are more likely to occur. Indeed, only the closest neighbors seems to matter, and the effects decline quickly with distance, disappearing after 200 meters. The effect also seems to be stronger when potential applicants and their neighbors are closer in terms of gender and socioeconomic status and when they have spent more time in the neighborhood. The fact that neighbors' effects are very local highlights the relevance of using an appropriate reference group when studying peers' effects.

In addition, I show that student loans generate indirect effects on close peers of their direct beneficiaries. In the case of neighbors, this indirect effect seems to be stronger in municipalities with low university attendance rates, where it represents a 15% of the direct effect of student loans on enrollment. I find that a similar indirect effect arises in the context of siblings. These externalities should be incorporated to the evaluation and design of funding programs, and could also be relevant in the context of other policies that seek to expand access to university.

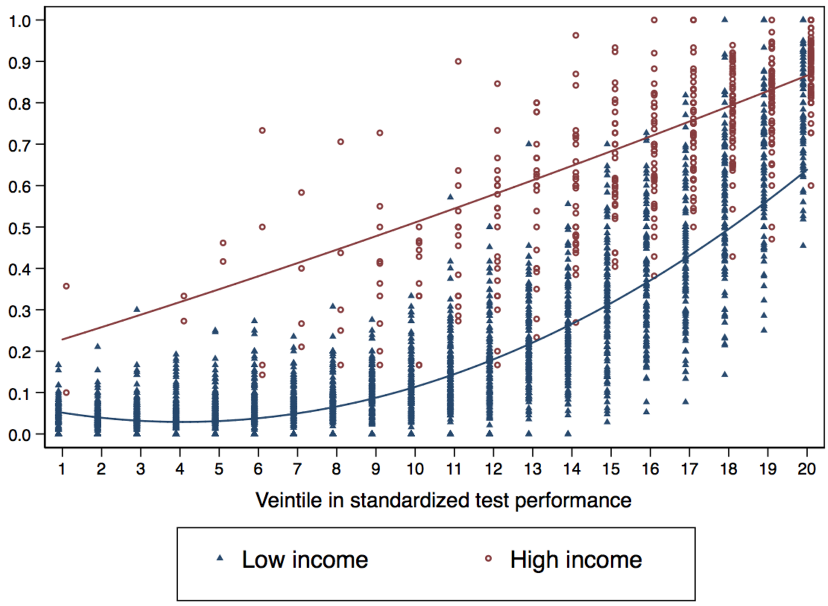
My main results are consistent with two broad classes of mechanisms, both related to some of the constraints that may affect individuals' schooling decisions. First, neighbors may increase university enrollment of potential applicants by providing them with relevant information about applications, returns and the overall university experience. Second, they could also affect the costs and benefits of going to university. Although with the data that I have available I cannot perfectly distinguish between them, finding no increase on potential applicants' effort or academic performance suggests that information is the mechanism behind my results.

Note however that interventions providing students with information on funding opportunities and returns to education have not been very effective at increasing college enrollment. This suggests that the information that potential applicants receive from their peers is different. It may be different on its content, but it could also be more relevant because it comes from someone closer. Investigating what potential applicants learn from their university-going neighbors and siblings seems a promising avenue for future research. Addressing this question would also contribute to gaining a better understanding of the mechanisms behind peers effects in this and other settings.

Acknowledgments

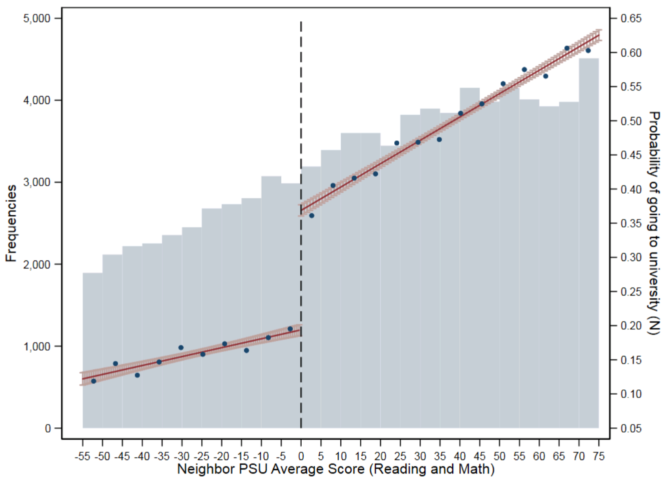
I thank Steve Pischke and Johannes Spinnewijn for their guidance and advice. I also thank Esteban Aucejo, Christopher Neilson, Sandra McNally, Philippe Aghion, Christopher Avery, Tim Besley, Peter Blair, Taryn Dinkelman, Joshua Goodman, Xavier Jaravel, Camille Landais, Steve Machin, Alan Manning, Guy Michaels, Daniel Reck, Amanda Pallais, Bruce Sacerdote and Olmo Silva for many useful comments, as well as seminar participants at LSE, at the IZA Workshop: The Economics of Education and at the EDP Jamboree 2018. Finally, I thank the Chilean Ministries of Education and Social Development, the Education Quality Agency, the Department of Assessment, Evaluation and Academic Registers (DEMRE) of the University of Chile for giving me access to the administrative data I use in this project.

Figure 2.1: University Enrollment by Household Income, Ability Level and Municipality

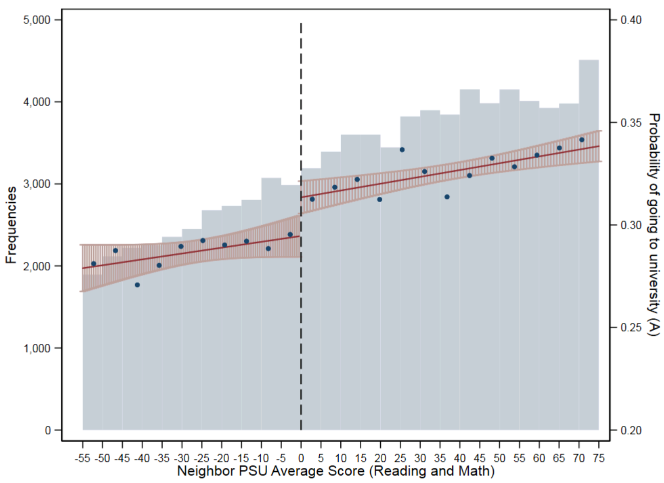


Notes: This figure illustrates the share of low and high income students enrolling in the university by ability level and municipality. Blue triangles represent the shares of low-income students, while red circles represent the shares of high-income students. The figure also presents quadratic fits of university enrollment on ability. The red line comes from a quadratic fit of high-income students attendance shares, while the blue from a similar exercise for low-income students. Ability is measured by students performance in grade 10 mathematics standardized test. University enrollment is measured 3 years later; if students do not repeat or dropout, this is one year after they complete high school. The sample includes students taking the standardized test in 2006, 2008, 2010 and 2012. Shares are computed only for municipalities for which at least 10 students were observed in each income-ability group.

Figure 2.2: First Stage and Reduced Form of Neighbors' RD



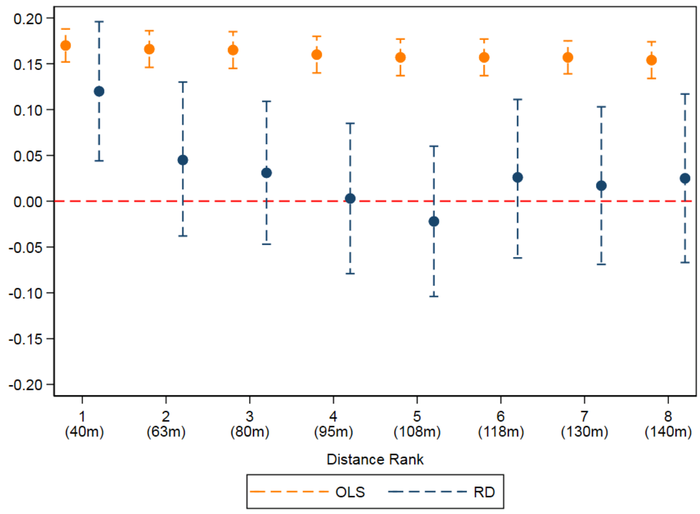
(a) First Stage: Neighbors' Probability of going to University



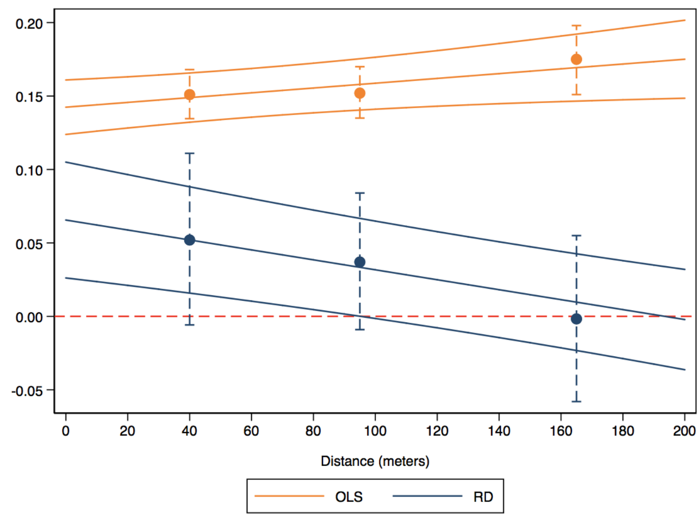
(b) Reduced Form: Potential Applicants' Probability of going to University

Notes: This figure illustrates the first stage and reduced form of the neighbors' RD. The first panel shows how neighbors' probability of going to university evolves with the score they obtain in the PSU. The second panel shows how potential applicants' probability of going to university evolves with the PSU score of their closest neighbor. The PSU score is centered around the student-loans eligibility threshold. Each dot represents the share of neighbors (panel 1) or potential applicants (panel 2) going to university at different ranges of neighbors' PSU scores. The red lines come from linear regressions of the outcome on the running variable on each side of the eligibility threshold, and the shadow around them to 95% confidence intervals. The blue bars in the background illustrate the distribution of the neighbors' scores in the PSU. The range used for these plots corresponds to optimal bandwidths computed following Calonico et al. [2014b].

Figure 2.3: Effect of Neighbors on Potential Applicants by Distance



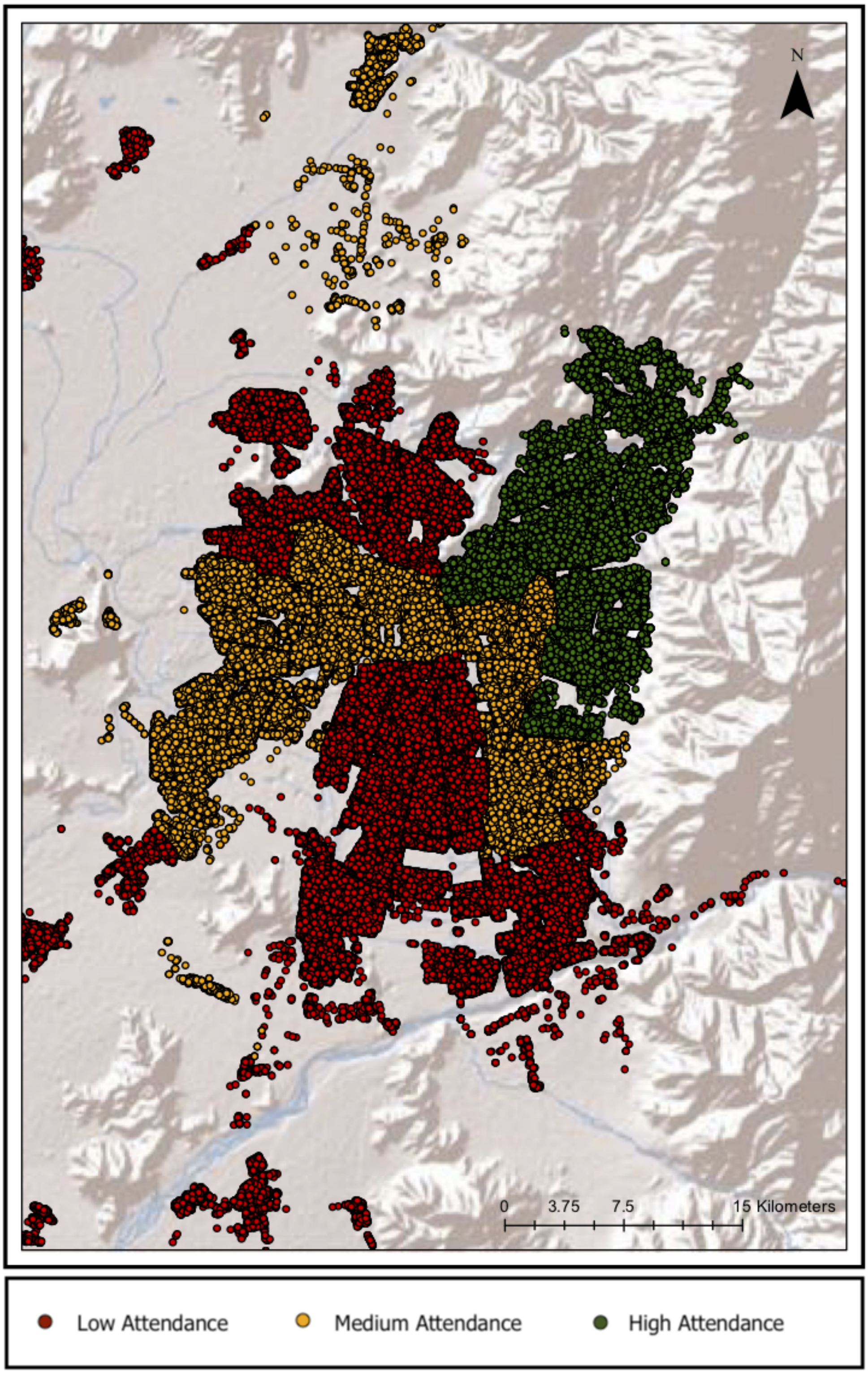
(a) Evolution of Effects by Distance Rank



(b) Evolution of Effects by Physical Distance

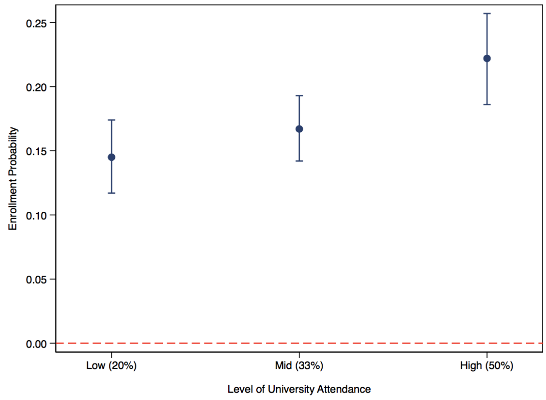
Notes: This figure illustrates how the effects of different neighbors on potential applicants evolve with distance. These coefficients come from specifications that include a potentially different linear function of the running variable on each side of the cutoff. The estimation uses optimal bandwidths computed following Calonico et al. [2014b] for the main specification.

Figure 2.4: University Attendance across Municipalities in Santiago

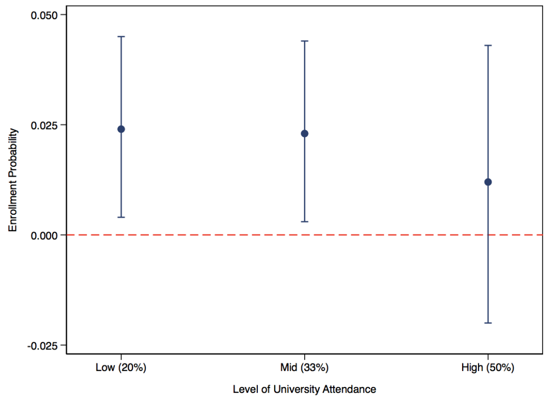


Notes: The figure illustrates the share of potential applicants going to university in different municipalities of Santiago between 2006 and 2012. In red areas the average attendance is 20%, in yellow areas 33% and in green areas 50%.

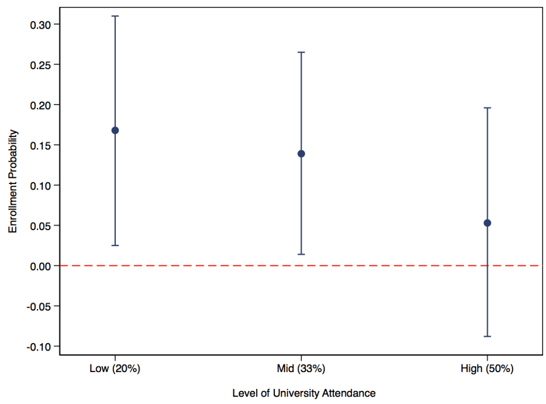
Figure 2.5: Neighbors' Effects by Municipality Level of Attendance



(a) First Stage



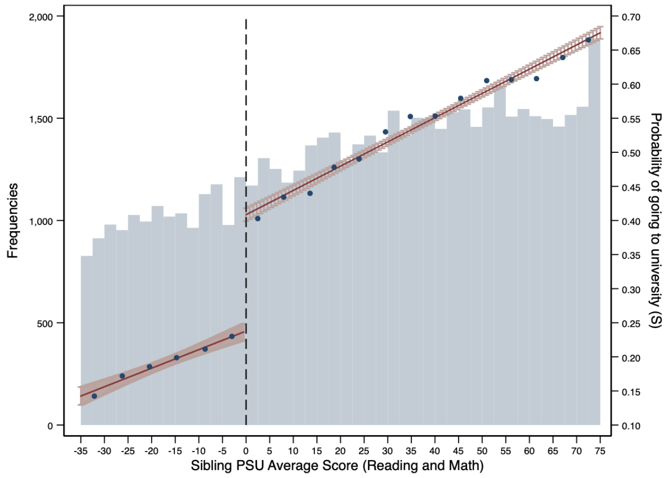
(b) Reduced Form



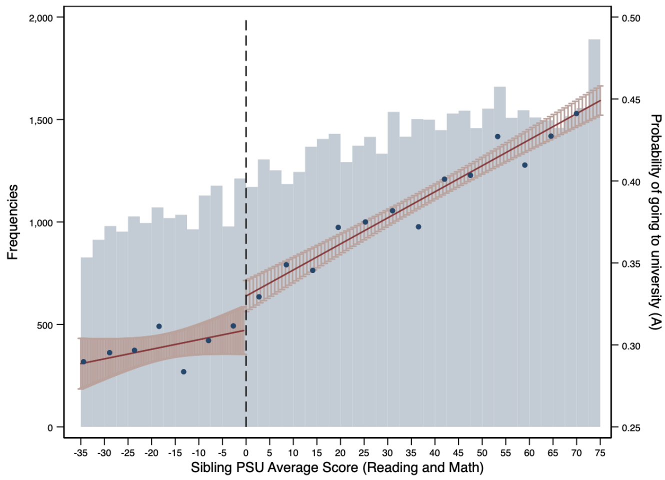
(c) 2SLS Estimates

Notes: The figure illustrates how neighbors' effects evolve depending on the level of attendance of the municipality of potential applicants. The dots represent coefficients from three different samples: low, mid and high attendance municipalities. The specification used controls by a linear polynomial of the running variable. The bandwidth used correspond to the optimal bandwidths computed for the whole sample. The lines represent 95% confidence intervals. Standard errors are clustered at the neighborhood unit level.

Figure 2.6: First Stage and Reduced Form of Siblings' RD



(a) First Stage: Siblings' Probability of going to Univeristy



(b) Reduced Form: Potential Applicants' Probability of going to University

Notes: This figure illustrates the first stage and reduced form of the siblings RD. The first panel shows how siblings' probability of going to university evolves with the score they obtain in the PSU. The second panel shows how potential applicants' probability of going to university evolves with the PSU score of their older sibling. The PSU score is centered around the student-loans eligibility threshold. Each dot represents the share of siblings (panel 1) or potential applicants (panel 2) going to university at different ranges of PSU scores. The red lines correspond come from linear regression of the outcome on the running variable on both sides of the eligibility threshold. The shadow around them to 95% confidence intervals. The blue bars in the background illustrate the distribution of the siblings' scores in the PSU. The range used for these plots corresponds to optimal bandwidths that were computed following Calonico et al. [2014b].

Table 2.1: Summary Statistics

	Neighbors (1)	Potential Applicants (2)	Whole country (3)
1. Demographic characteristics			
Female	0.56	0.53	0.54
Age when taking the PSU	18.30	17.72	18.08
2. Socioeconomic characteristics			
Low Income ($\leq$ 288K CLP)	0.57	0.54	0.57
Mid Income ($\leq$ 864K CLP)	0.36	0.36	0.30
High Income ($>$ 864K CLP)	0.07	0.10	0.13
Parental ed. = primary ed.	0.07	0.08	0.13
Parental ed. = secondary ed.	0.55	0.55	0.52
Parental ed. = other	0.01	0.01	0.01
Parental ed. = vocational he	0.09	0.07	0.06
Parental ed. = professional he	0.09	0.08	0.05
Parental ed. = university	0.20	0.21	0.23
3. Academic characteristics			
Public high school	0.37	0.36	0.41
Charter high school	0.58	0.57	0.49
Private high school	0.05	0.07	0.10
Education track = academic	0.74	0.65	0.66
Education track = vocational	0.26	0.35	0.34
High school GPA	5.72	5.57	5.48
Score in the PSU (centered at the cutoff)	47.76	-9.90	-20.40
4. Family structure			
Family size	4.46	4.62	4.48
Household head = father	0.61	0.61	0.59
Household head = mother	0.31	0.31	0.28
Household head = other	0.08	0.08	0.13
Distance to closest neighbor (km)	0.05	0.05	
Age difference	1.56	1.56	
Observations	193,101	193,101	1,316,117

Notes: Columns (1) and (2) present summary statistics for potential applicants and their closest neighbors. Column (3) for all potential applicants in the country.

Table 2.2: Effect of Neighbors on Potential Applicants' University Enrollment

	2SLS-1 (1)	2SLS-2 (2)	CCT-1 (3)	CCT-2 (4)
Neighbor goes to university (t-1)	0.104*** (0.037)	0.134*** (0.047)	0.118** (0.053)	0.104* (0.062)
First stage	0.178*** (0.009)	0.168*** (0.010)	0.171*** (0.011)	0.172*** (0.013)
Reduced form	0.019*** (0.007)	0.023*** (0.008)		
Year fixed effects	Yes	Yes	Yes	Yes
N. of students	83,894	133,911	83,894	133,911
PSU Polynomial	1	2	1	2
Bandwidth	(55-73.5)	(75.5-133.5)	(55-73.5)	(75.5-133.5)
Kleibergen-Paap F statistic	423.32	271.34		

Notes: The table presents the estimated effects of neighbors on potential applicants' university enrollment. Columns 1 and 2 present two stages least squares estimates using a linear and quadratic polynomial of PSU respectively. Columns 3 and 4 use instead local polynomials following Calonico et al. [2014b]. Optimal bandwidths are used in all the specifications. In parenthesis, standard errors clustered at neighborhood unit level. $*p-value < 0.1 **p-value < 0.05 ***p-value < 0.01$

Table 2.3: Effect of Neighbors on Potential Applicants by Social Distance

	Socioeconomic Status		Gender		Age	
	Same	Different	Same	Different	1 year	> 1 year
	(1)	(2)	(3)	(4)	(5)	(6)
Neighbor goes to university (t-1)	0.153***	0.049	0.130**	0.084	0.113**	0.092
	(0.048)	(0.060)	(0.051)	(0.054)	(0.054)	(0.067)
First stage	0.188***	0.168***	0.183***	0.173***	0.168***	0.226***
	(0.011)	(0.011)	(0.010)	(0.011)	(0.010)	(0.018)
Reduced form	0.029***	0.008	0.024**	0.014	0.019**	0.021
	(0.009)	(0.010)	(0.009)	(0.009)	(0.009)	(0.015)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N. of potential applicants	42680	42404	43378	41706	50979	18283
Kleibergen-Paap Wald F Statistic	270.55	251.63	310.81	262.93	281.18	155.32

Notes: The table presents the estimated effects of neighbors on potential applicants' university enrollment by different measures of social distance. Columns 1 and 2 study how the effects change with differences in socioeconomic status, columns 3 and 4 with gender and finally columns 5 and 6 with age. All specifications include a linear polynomial of the closest neighbor or sibling PSU score; it is allowed to be different on both sides of the student-loans eligibility threshold. Optimal bandwidths are used in all the specifications and were computed following Calonico et al. [2014b]. In parenthesis, standard errors clustered at neighborhood unit level. *p-value<0.1 **p-value<0.05 ***p-value<0.01

Table 2.4: Effect of Neighbors on Potential Applicants by Time at the Neighborhood

	Time at the neighborhood		Neighbors remain-leave		Mother works outside the hh.	
	$\geq$ 4 years	$<$ 4 years	Remain	Leave	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)
Neighbor goes to university (t-1)	0.134***	0.041	0.118***	0.068	0.182***	0.066
	(0.051)	(0.100)	(0.044)	(0.080)	(0.055)	(0.054)
First stage	0.191***	0.218***	0.177***	0.202***	0.174***	0.179***
	(0.012)	(0.024)	(0.010)	(0.018)	(0.011)	(0.010)
Reduced form	0.026***	0.009	0.021***	0.014	0.032***	0.012
	(0.010)	(0.022)	(0.008)	(0.016)	(0.009)	(0.010)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N. of potential applicants	42509	8283	61327	16705	38696	41220
Kleibergen-Paap F statistic	264.90	80.98	306.68	119.53	240.93	290.35

Notes: The table presents the estimated effects of neighbors on potential applicants' university enrollment by different characteristics of potential applicants and their neighbors. Columns 1 and 2 show how the effects change depending on the time potential applicants have lived in the neighborhood. Columns 3 and 4 compare potential applicant whose neighbors say that they will remain or leave the neighborhood in case of going to university. Columns 5 and 6 compare potential applicants depending on mothers' occupation. All specifications include a linear polynomial of the closest neighbor or sibling PSU score; it is allowed to be different on both sides of the student-loans eligibility threshold. Optimal bandwidths are used in all the specifications and were computed following Calonico et al. [2014b]. In parenthesis, standard errors clustered at neighborhood unit level. *p-value<0.1 **p-value<0.05 ***p-value<0.01

Table 2.5: Effect of Siblings on Potential Applicants' University Enrollment

	2SLS-1 (1)	2SLS-2 (2)	CCT-1 (3)	CCT-2 (4)
Sibling goes to university (t-T)	0.137** (0.053)	0.172** (0.068)	0.145** (0.064)	0.168** (0.079)
First stage	0.170*** (0.009)	0.156*** (0.010)	0.157*** (0.011)	0.160*** (0.013)
Reduced form	0.023*** (0.009)	0.027** (0.011)		
Year fixed effects	Yes	Yes	Yes	Yes
N. of students	56,767	94,205	56,767	94,205
PSU Polynomial	1	2	1	2
Bandwidth	(37.5-73.0)	(59.5 - 128.0)	(37.5-73.0)	(59.5 - 138.0)
Kleibergen-Paap F statistic	359.64	222.09		

Notes: The table presents the estimated effects of siblings on potential applicants' university enrollment. Columns 1 and 2 present two stages least squares estimates using a linear and quadratic polynomial of PSU respectively. Columns 3 and 4 use instead local polynomials following Calonico et al. [2014b]. Optimal bandwidths are used in all the specifications. In parenthesis, standard errors clustered at family level. $*p-value < 0.1$ $**p-value < 0.05$ $***p-value < 0.01$

Table 2.6: Effect of Neighbors and Siblings on Potential Applicants' Academic Performance and Application Behavior

	Neighbors (1)	Siblings (2)
	Panel A - Academic Performance	
High school GPA	0.000	0.119*
	(0.001)	(0.066)
High school attendance	0.009	0.018
	(0.009)	(0.016)
PSU Performance	34.970**	32.137*
	(15.405)	(17.373)
	Panel B - Application Behavior	
Take PSU	0.065**	0.068**
	(0.025)	(0.032)
Apply to financial aid	0.053	0.125**
	(0.042)	(0.052)
Eligible for financial aid	0.051**	0.071
	(0.022)	(0.064)
Take up financial aid	0.085**	0.121**
	(0.034)	(0.059)
Apply to CRUCH universities	0.054	0.228***
	(0.041)	(0.063)
Active application to CRUCH universities	0.059	0.103*
	(0.039)	(0.061)

Notes: The table presents the estimated effects of neighbors and siblings on potential applicants' academic performance and application behavior. Column 1 presents the results for neighbors ($n = 81,587$) and column 2 for siblings ($n = 45,372$). All specifications include a linear polynomial of the closest neighbor or sibling PSU score; it is allowed to be different on both sides of the student-loans eligibility threshold. Optimal bandwidths are used in all the specifications and were computed following Calonico et al. [2014b]. In parenthesis, standard errors clustered at neighborhood unit level. *p-value<0.1 **p-value<0.05 ***p-value<0.01

Appendices

2.A Siblings Sample

Although the focus of this paper is on neighbors, I also investigate what happens with potential applicants when an older sibling goes to university T years before her. The sample that I use for this purpose is similar to the one used to study neighbors effects, but it includes students that appear in the PSU registers between 2006 and 2015.

When registering for the PSU, potential applicants report their parents national id number. Using this information, I managed to identify 273,806 pairs of siblings. Proceeding in the same way as with neighbors, I restrict the sample to 17-22 years old students completing high school in regular educational programs no more than 3 years before registering for the PSU. These restrictions reduce the sample size by 13.8%. I further restrict the sample to potential applicants whose siblings apply to financial aid; they are the only ones that could change their decisions based on students-loan eligibility. As before, this restriction is not imposed on potential applicants, but it reduces the sample size and I end up working with roughly half of the original observations. Table 2.A1 presents the summary statistics for this sample.

Table 2.A1: Summary Statistics - Siblings' Sample

	Siblings (1)	Potential Applicants (2)
1. Demographic characteristics		
Female	0.55	0.54
Age	18.06	17.75
2. Socioeconomic characteristics		
Low Income	0.52	0.51
Mid Income	0.38	0.38
High Income	0.09	0.11
Parental ed. = primary ed.	0.07	0.07
Parental ed. = secondary ed.	0.51	0.51
Parental ed. = other	0.01	0.01
Parental ed. = vocational he	0.09	0.08
Parental ed. = professional he	0.09	0.12
Parental ed. = university	0.23	0.21
3. Academic characteristics		
Public high school	0.40	0.34
Charter high school	0.55	0.60
Private high school	0.05	0.05
Education track = academic	0.77	0.76
Education track = vocational	0.23	0.24
High school GPA	5.84	5.75
Score in the PSU (centered at the cutoff)	52.89	20.90
4. Family structure		
Family size	5.03	4.77
Household head = father	0.73	0.70
Household head = mother	0.23	0.26
Household head = other	0.04	0.04
Age difference	3.89	3.89
Observations	135,658	135,658

Notes: ColumnS (1) and (2) present summary statistics for potential applicants and their siblings.

2.B Identification Strategy: Further Discussion

Traditionally, peers' effects have been modeled using a linear-in-means function. This implicitly assumes that all peers are equally important. Since in this case, there is available a measure of proximity between peers, it is possible to assume a more flexible functional form:

$$U_{at} = \alpha + \sum_{n \in N_a} \beta_{n\tau} U_{n\tau} + \varepsilon_{it} \tag{2.2}$$

Where, N_a is the set of relevant neighbors for potential applicant a and U_{nt} is a dummy variable indicating if the $n-th$ neighbor goes to university in t.

As discussed in section 2.4 neighbors decide whether to enroll or not into university before potential applicants. Thus, their decision should not be affected by what potential applicants do after them. This implies that N_a does not include younger neighbors (i.e. neighbors that could potentially apply to university in the future).[54]

This paper focuses on the effects of neighbors going to university one year before potential applicants. To highlight this, equation 2.2 can be rearranged as follows:

$$U_{at} = \alpha + \beta_{mt-1} U_{mt-1} + \sum_{n \in N_a \backslash U_{mt-T}} \beta_{n\tau} U_{n\tau} + \varepsilon_{it} \tag{2.3}$$

The coefficient β_{mt-1} can be consistently identified if $Cov(U_{mt-1}, \varepsilon_{it}) = 0$. This implies that there are no correlated effects, and that potential applicant a_t does not affect the decision of neighbor $mt-1$.

There are many reasons why we could want to estimate a more parsimonious function. For instance, if we do not observe all the relevant neighbors, or if the type of variation used to identify these effects imposes some restrictions that prevent us from using all the information available.

Consider the following simplified specification:

$$U_{at} = \alpha + \beta_{mt-1} U_{mt-1} + v_{it} \tag{2.4}$$

[54]If younger applicants' decision enter equation 2.2, instrumenting enrollment of the older neighbor with student-loan eligibility would be enough to solve the reflection problem.

In this case, to consistently estimate β_{mt-1} we need $Cov(U_{mt-1}, v_{it}) = 0$. This means that in addition to the conditions discussed for equation 2.3, we need $(Cov(U_{at}, U_{n\tau}) \cdot (Cov(U_{mt-1}, U_{n\tau})) = 0 \ \forall \ \{n, \tau\} \neq \{m, t-1\}$. To discuss the implications of this additional condition we can analyze three cases:

- Contemporaneous applicants: $\tau = t$
- Neighbors in t-1: $\tau = t - 1$
- Neighbors in t-T: $\tau = t - T$ (with $T > 1$).

Note that for the first two cases, the absence of contemporaneous peers' effects is sufficient.[55] To satisfy the assumption in the third case we would need to assume that neighbors applying two or more years before potential applicants do not directly affect them (i.e. they are not part of the structural equation).

This last assumption can be relaxed if as in this case we have an instrument for university enrollment. Instead of assuming that neighbors two or more years apart do not enter the structural equation, we would need to assume that $(Cov(Z_{mt-1}, U_{n\tau-T})) = 0$.

If the decisions of contemporaneous and younger peers enter equation 2.2, β_n can still be interpreted as a reduce form parameter capturing not only the effect of the $n-th$ closest neighbor on a, but also the effects that other neighbors affected by n could have generate on a. This is still a relevant parameter from a policy perspective.

A fuzzy RD can be thought as a particular case of IV. This means that my estimates will be consistent under the following assumptions:

A1. Independence:

The instrument L_n needs to be independent of the enrollment decision of both, the potential applicant and her neighbor. In my setting, this will only be true around the student loans eligibility treshold and after conditioning on neighbors' performance in the PSU.

A2. Relevance:

The instrument L_n needs to change the enrollment decision of neighbors U_n. First-

[55] We are already assuming that younger applicants' decision are not part of equation 2.2.

stage regressions in section 2.5 show that this is indeed the case.[56]

A3. Exclusion:

The instrument only affects potential applicants enrollment U_i through the change it induces in neighbors' university attendance. This implies that neighbors eligibility for student loans does not have a direct effect on the enrollment decision of potential applicants.

A4. Monotonicity:

Finally, the monotonicity assumption requires eligibility for student loans to weakly increase neighbors enrollment. In this setting, it is difficult to think in any reasons that would make individuals to decide not to enroll in university because they are eligible for financial aid.[57]

According to Imbens and Angrist [1994], under this set of assumptions the IV estimates are consistent and can be interpreted as local average treatment effects (LATEs). In this setting, this means that my estimates will capture the effect of having a neighbor near the student loans eligibility threshold going to university.

[56] In line with the results of Solis [2017] I find that being eligible for student loans roughly doubles the probabilities of going to university at the eligibility cutoff.

[57] Note that if for some reason individuals dislike student loans or other types of funding, they could reject them and pay the tuition fees with their own resources.

2.C Other Neighbors Definitions

The results discussed on section 2.5 focus on the closest neighbor. However, there could be other neighbors that are relevant for potential applicants. To investigate this, I identify the best neighbor among the closest 3 and 5, and the best living within 75 and 100 meters from potential applicants.

When implementing these exercises, the sample size decreases with the radius being analyzed. The student-loans cutoff is relatively low (percentile 40 in the PSU distribution); this makes it more difficult to find individuals that being the best of a group are at the same time close enough to the cutoff. This not only affects the precision of the estimates, but also the composition of the sample used to estimate the effects of interest.

The characteristics of areas where the best neighbor in 100 meters is close enough to the cutoff may be very different to those where the best in 200 meters is close. Thus, these results do not tell us much about how neighbors effect evolve with distance. Each estimate comes from a different sample, what means that apart from distance to the relevant neighbor many other things may be changing.

Table 2.C1 presents the results of these analysis. When looking at the effect of the best neighbor among 3 or the best neighbor within 75 meters the coefficient obtained is in the same range as the one discussed in the main section. In this case they are only significant at a 90% level what in part reflects the fact that sample sizes are smaller in this case. When looking at the best neighbor among the closest 5, the coefficient is bigger and significant at a 95% level. This result is consistent with the idea that the effects of exposure are stronger when there are fewer people going to university. Finally, when looking at the effect of the best neighbor within 100 meters, the coefficient drops and becomes not statistically different from 0.

Table 2.C1: Effects of other Close Neighbors on Potential Applicants' University Enrollment

	Best among "n" closest		Best within "d" meters	
	3 closest	5 closest	75 meters	100 meters
	(1)	(2)	(3)	(4)
Probability of attending U(2).	0.091*	0.140**	0.098*	0.057
	(0.048)	(0.071)	(0.055)	(0.059)
Score in the PSU	0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.207***	0.145***	0.212***	0.221***
	(0.024)	(0.036)	(0.028)	(0.030)
Year fixed effects	Yes	Yes	Yes	Yes
N. of students	58594	24504	50439	32361
Kleibergen-Paap F statistic	219.41	87.24	200.82	148.32
Dep. var mean	0.254	0.220	0.260	0.247

Notes: The table presents results for specifications that study the effect of other close neighbors on potential applicants' university enrollment. Columns 1 and 2 look at the effect of the best neighbor among the closest 3 and 5, while columns 3 and 4 look at the effect of the best neighbor within 75 and 50 meters. All specifications include a linear polynomial of PSU which slope is allowed to differ at both sides of the cutoff. All specifications use optimal bandwidths computed according to Calonico et al. [2014b]. In parenthesis, standard errors clustered at neighborhood unit level. *p-value<0.1 **p-value<0.05 ***p-value<0.01

2.D Robustness Checks

In this section, I study if the identification assumptions of my empirical strategy are satisfied. I start by investigating if there is evidence of manipulation in the running variable, and then I check if other variables that could be related to the decision of going to university present jumps around the student loans eligibility threshold. I continue showing the results of placebo exercises and the robustness of my estimates to different bandwidths choices. I finish this section discussing some issues that could emerge due to missing observations.

2.D.1 Manipulation of the running variable

A common concern in the context of a regression discontinuity is if individuals can strategically manipulate the running variable affecting in this way their treatment status.

In this case, it would mean that potential applicants have the ability of affecting the average PSU score of their older neighbors and siblings. As discussed in section 2.2, the PSU is a national level test which application and marking processes are completely centralized. In addition, given that the scores of students in each section of the test are normalized, students do not know ex ante the exact number of correct answers they need to be above the eligibility cutoff.

All this makes it very difficult, even for the students taking the test to manipulate their score around the threshold. Considering this, it seems very unlikely that potential applicants can strategically affect it.

In the context of neighbors, a way in which potential applicants could change the score they obtain in the PSU would be to move to a different neighborhood. However, the results on movers and no-movers presented in section 2.5 do not support this hypothesis. In addition, in the next section I show that there are no jumps in neighbors' characteristics around the cutoff; so, if potential applicants are moving to areas where neighbors are more likely to be eligible for student loans, they are not using any of the socioeconomic and academic variables I study to select them.

I further investigate manipulation by looking at the density of the PSU scores around the eligibility threshold implementing the test suggested by Cattaneo et al. [2018a]. Figures 2.D1 and 2.D7 show that there is no evidence to reject the null hypothesis of a continuous density of neighbors PSU scores around the eligibility threshold. In the case of neighbors, the p-value of the test is 0.7791, whereas in the case of siblings it is 0.5968. Therefore, the results I find do not seem to be driven by manipulation of the running variable.

2.D.2 Discontinuities in potential confounders

A second concern in the context of an RD is the existence of other discontinuities around the cutoff that may explain the differences we observe in the outcome of interest.

Taking advantage of a rich vector of demographic, socioeconomic and academic variables, I study if there are discontinuities in any of them around the threshold.

Figure 2.D2 summarizes these results for neighbors, and figure 2.D8 for siblings. They illustrate the estimated discontinuities at the cutoff and their 95% confidence intervals. To estimate these discontinuities I use optimal bandwidths following Calonico et al. [2014a]. In both figures, the left panel looks at characteristics of the older peer, and the right panel at characteristics of potential applicants.

I do not find any significant difference in older peers and potential applicants characteristics around the threshold. In the case of neighbors, there is a close-to-significant difference in parental education. Neighbors to the right of the cutoff seem to come from households where the parents are more likely to have attended higher education; in the case of potential applicants, this difference is clearly not significant. In addition, the magnitudes of these coefficients are quite small and the differences in university enrollment documented in section 2.5 are robust to the inclusion of neighbors' parental education as control. Indeed, they are robust to the inclusion of all the variables in these figures. [58]

[58] This specification is not presented here, but is available upon request.

2.D.3 Placebo exercises

This setting allows me to perform placebo exercise to study if the potential applicants' enrollment decision has any effect on the decision of their older neighbors or siblings. Given the timing of both decisions, we should not find any effect; what happens with potential applicants in t, should not change the probabilities of going to university of their older peers in t-T.

Figures 2.D3 and 2.D9 illustrate the results when performing this exercise in the same sample I use in when estimating the main results. Table 2.D1 presents the estimated coefficients of this exercise; table 2.D2 presents the results of a similar exercise but using a different sample. This time, I include neighbors and siblings who do not apply to financial aid and keep in the sample only potential applicants who apply to financial aid. It is reassuring not finding discontinuities around the eligibility threshold; both, the levels and slopes seem to be continuous around it. As in section 2.5, tables 2.D1 and 2.D2 present the estimated coefficient using two stages least squares and local polynomials. The coefficients are small and never significant.

In addition to this robustness check, I also study if there are significant discontinuities in points different to the student loans eligibility threshold. Since in these points there is no first stage, we should not find jumps like the ones we observe around the threshold. Figure 2.D4 presents these results for neighbors and siblings. As can be appreciated, none of these jumps is significant.

2.D.4 Different bandwidths

In this section, I study how sensible are my results to the bandwidth choice. Optimal bandwidths try to balance the loss of precision suffered when narrowing the window of data points used to estimate the effect of interest, with the bias generated by using points that are too far from the relevant cutoff.

Figures 2.D5 and 2.D10 presents the estimated coefficients when using bandwidths that go from 0.4 to 1 times the optimal bandwidth. These results correspond to specifications that use polynomial of degree 1 on both sides of the eligibility threshold. Changing the bandwidths does not make an important difference on the estimated coefficients.

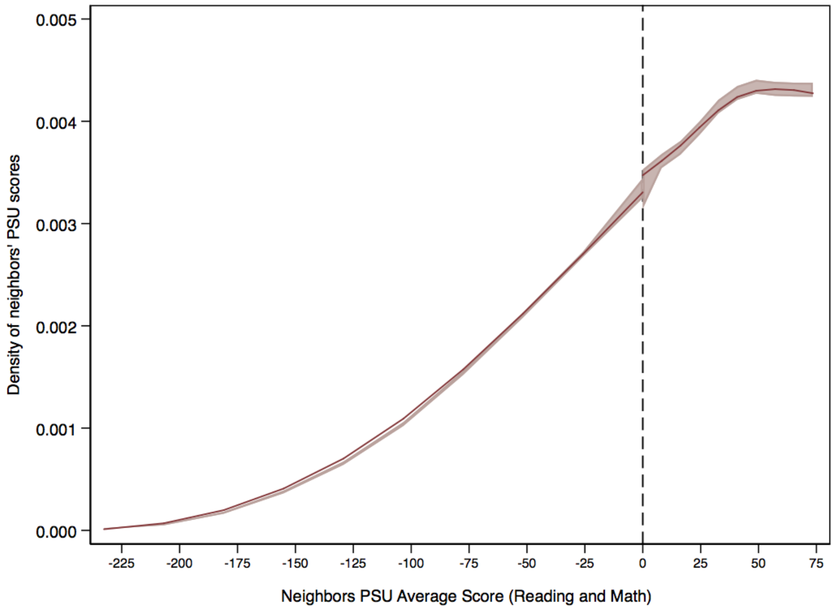
2.D.5 Missing students

In this section, I discuss how missing information about applicants and their older peers could affect my results. As mentioned in Section 2.3, to identify neighbors I rely on the geocoding process of addresses; since the addresses I use do not include post-code, finding them was not always possible. This is especially the case in rural areas, where there is no precise information on the names of all the roads and locations. In this geocoding process, I loss around 15% of my sample.

To analyze how serious this threat could be, I present an additional exercise just focusing in the Metropolitan Region of Santiago; in this area the geocoding rate of success was higher. Table 2.D3 presents the results to this exercise.

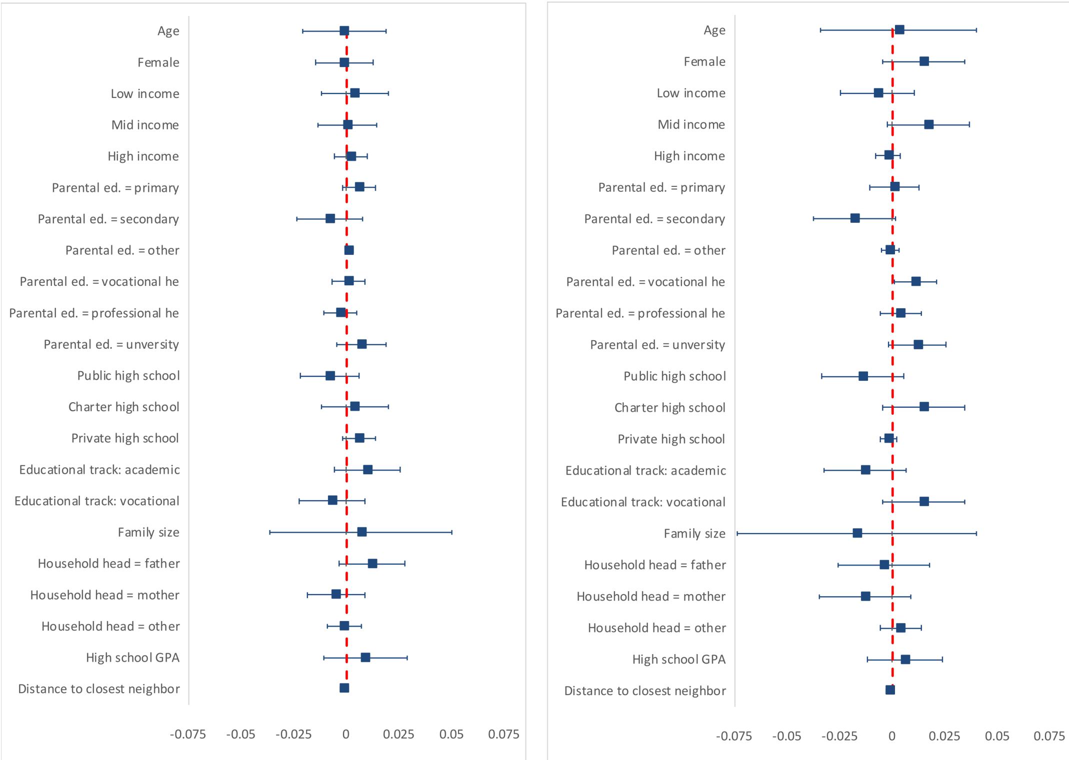
The coefficients obtained in both cases are slightly bigger than the ones I discuss in the rest of the paper. However, they are not significantly different from them. In part, this can reflect differences between students in the academic and vocational track of high school and between students from urban and rural areas.

Figure 2.D1: Density of Neighbors' PSU Scores around the Student Loans Eligibility Threshold)



Notes: This figure illustrates the density of neighbors PSU scores around the student loans eligibility thresholds. The density and its confidence intervals on each side of the cutoff were estimated following Cattaneo et al. [2018a]. This chart complements the formal test they suggest to study discontinuities in the distribution of the running variable around the relevant threshold. In this case its $p - value$ is 0.7791. This means there is no statistical evidence to reject the null hypothesis of a smooth density around the threshold.

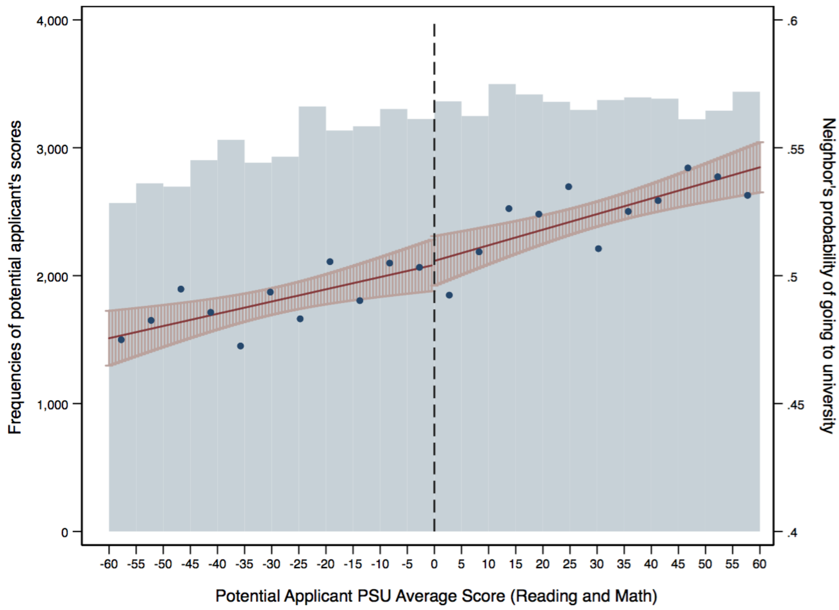
Figure 2.D2: Discontinuities in other Covariates at the Cutoff



(a) Potential Applicants (b) Neighbors

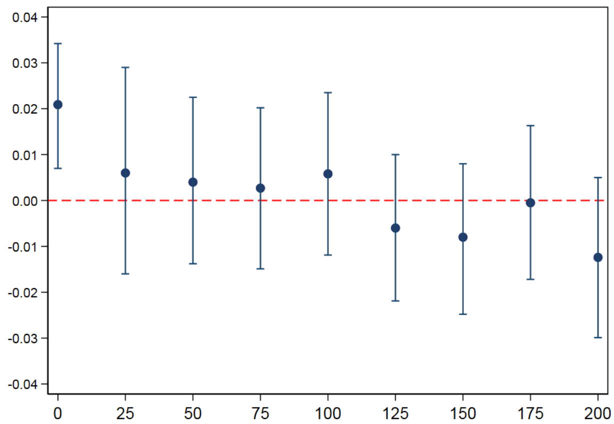
Notes: This figure illustrates the coefficients obtained when studying discontinuities in other variables that could potentially affect the outcome of interest. The left panel presents the results for potential applicants, while the right panel for neighbors. Apart from the coefficients, the figures illustrate 95% confidence intervals. The dashed red line correspond to 0. The coefficients were obtained using optimal bandwidths that were computed following Calonico et al. [2014b].

Figure 2.D3: Placebo Exercise: Effect of Potential Applicants (t) on Neighbors (t-1)

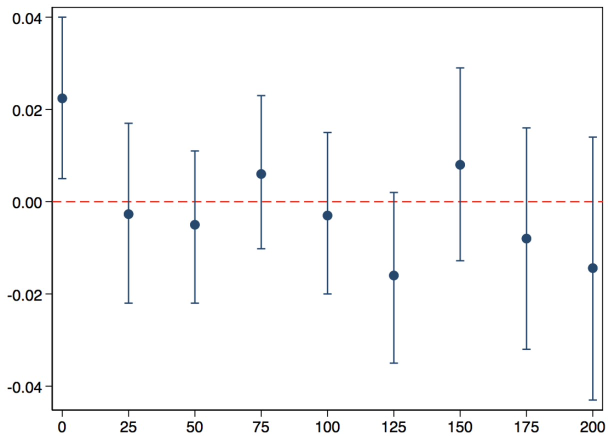


Notes: This figure illustrates the reduced form of a placebo exercise. It shows how neighbors' probability of going to university evolves with the PSU score of potential applicants. The PSU score is centered around the student-loans eligibility threshold. Each dot represents the share of neighbors going to university at different ranges of potential applicants PSU scores. The red lines correspond to linear approximations of these shares, and the shadow around them to 95% confidence intervals. The blue bars in the background illustrate the distribution of the potential applicants' scores in the PSU. The range used for these plots corresponds to optimal bandwidths that were computed following Calonico et al. [2014b].

Figure 2.D4: Neighbors and Siblings Placebo Cutoffs



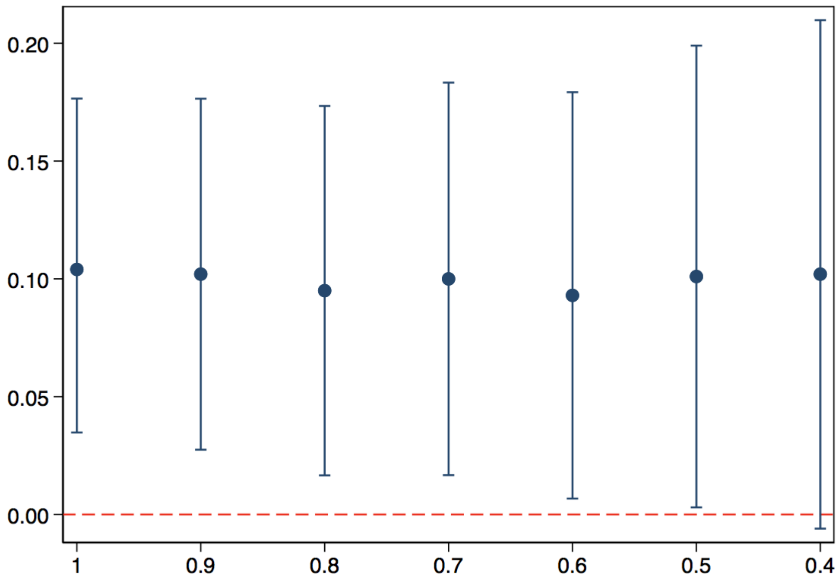
(a) Neighbors



(b) Siblings

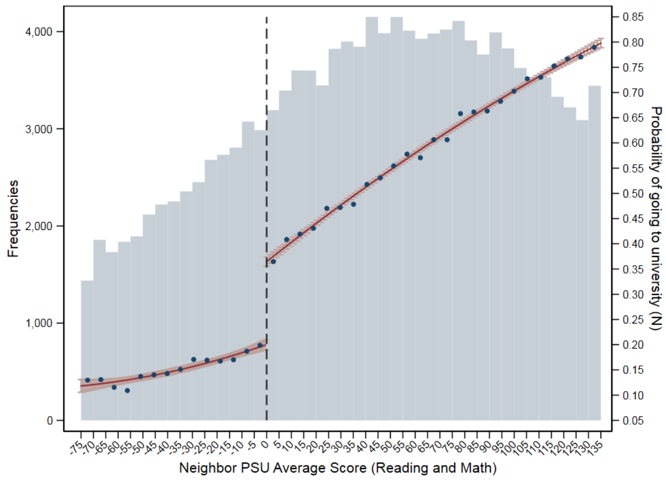
Notes: This figure illustrates the reduced form coefficients for the different cutoffs. The top panel illustrates the results for neighbors, and the panel at the bottom for siblings. Apart from the coefficients, the figures illustrate 95% confidence intervals. Standard errors are clustered at the neighborhood unit level.

Figure 2.D5: Neighbors' Effects with Different Bandwidths

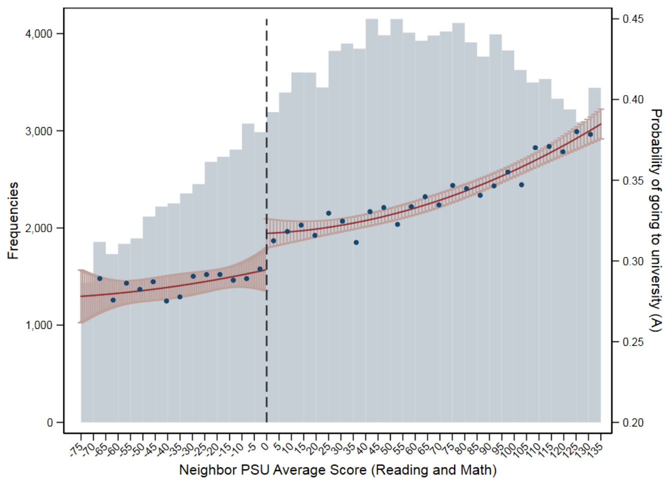


Notes: This figure illustrates the coefficients obtained when studying neighbors' effects using different bandwidths. The dots represent the coefficients, and the lines illustrate 95% confidence intervals.

Figure 2.D6: First Stage and Reduced Form of Neighbors' RD (P2)



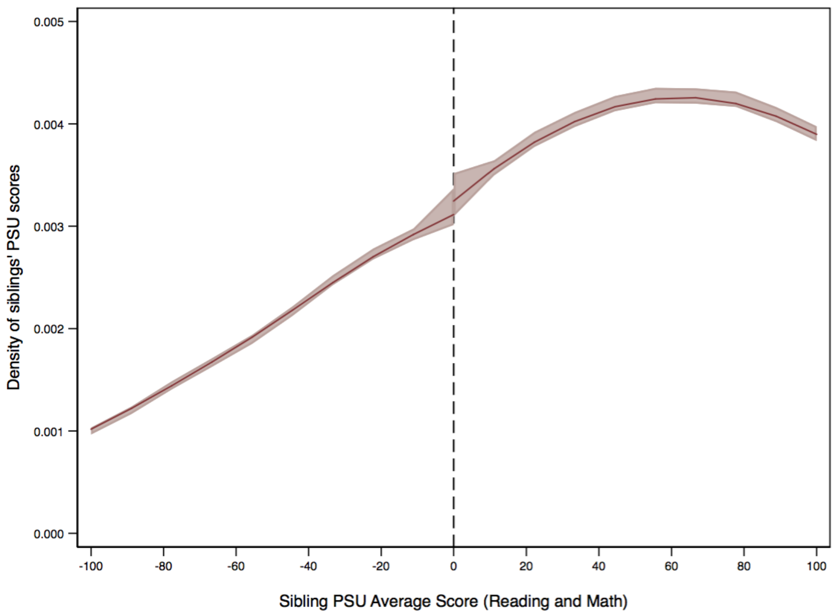
(a) First Stage: Neighbors' Probability of going to Univeristy



(b) Reduced Form: Potential Applicants' Probability of going to University

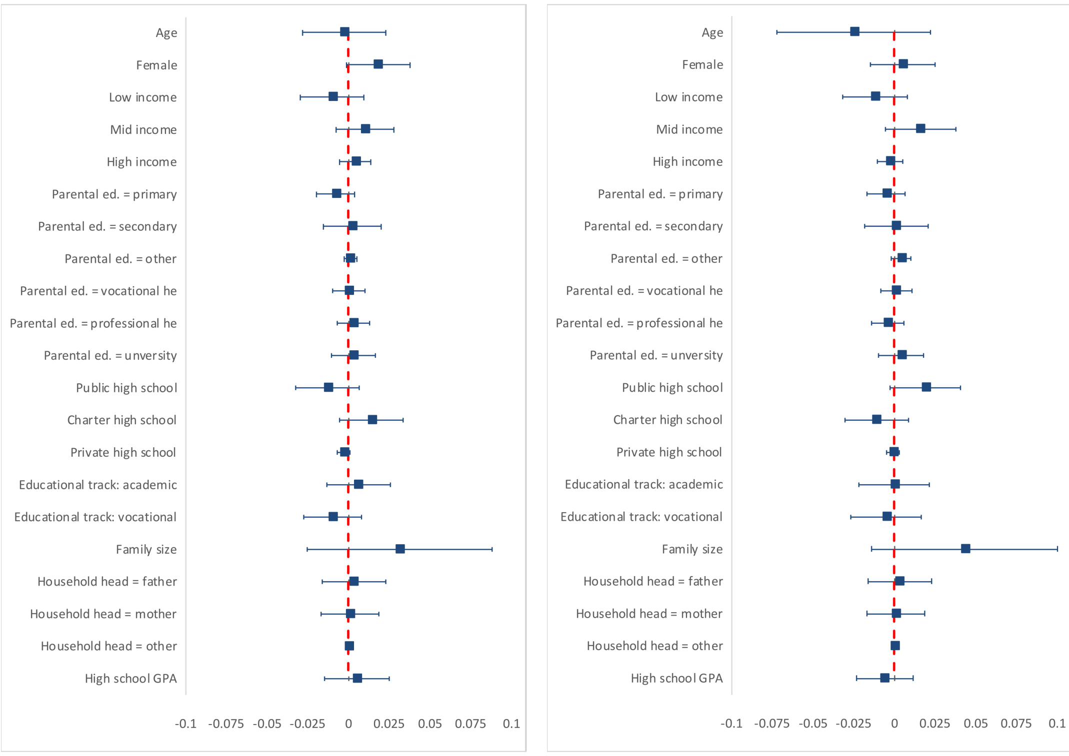
Notes: This figure illustrates the first stage and reduced form of the neighbors rd. The first panel shows how neighbors' probability of going to university evolves with the score they obtain in the PSU. The second panel shows how potential applicants' probability of going to university evolves with the PSU score of their closest neighbor. The PSU score is centered around the student-loans eligibility threshold. Each dot represents the share of neighbors (panel 1) or potential applicants (panel 2) going to university at different ranges of PSU scores. The red lines correspond to quadratic approximations of these shares, and the shadow around them to 95% confidence intervals. The blue bars in the background illustrate the distribution of the neighbors' scores in the PSU. The range used for these plots corresponds to optimal bandwidths that were computed following Calonico et al. [2014b].

Figure 2.D7: Density of Siblings' PSU Scores around the Student Loans Eligibility Threshold)



Notes: This figure illustrates the density of siblings PSU scores around the student loans eligibility thresholds. The density and its confidence intervals on each side of the cutoff were estimated following [Cattaneo et al., 2018a]. This chart complements the formal test they suggest to study discontinuities in the distribution of the running variable around the relevant threshold. In this case the test statistic is 0.4479 and the $p-value$ is 0.5968. This means there is no statistical evidence to reject the null hypothesis of a smooth density around the threshold.

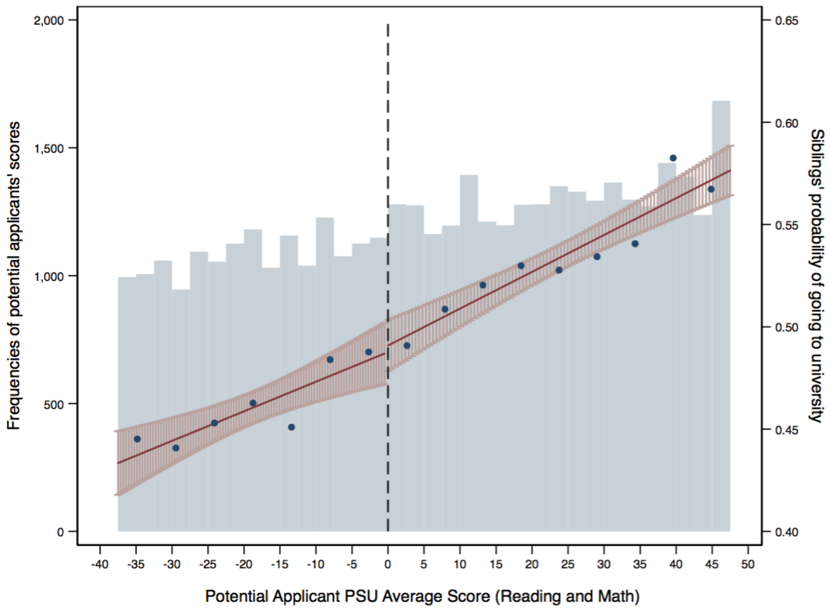
Figure 2.D8: Discontinuities in other Covariates at the Cutoff - Siblings



(a) Potential Applicants (b) Siblings

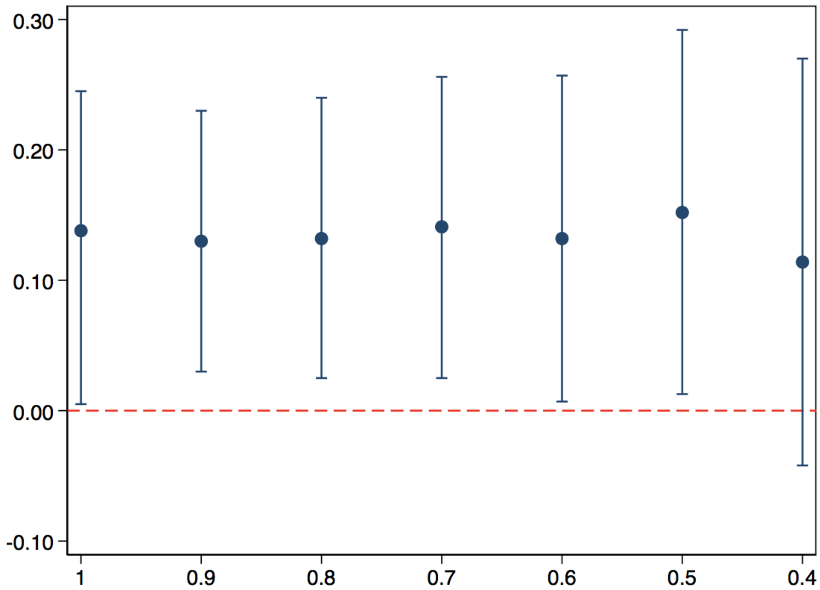
Notes: This figure illustrates the coefficients obtained when studying discontinuities in other variables that could potentially affect the outcome of interest. The left panel presents the results for potential applicants, while the right panel for siblings. Apart from the coefficients, the figures illustrate 95% confidence intervals. The dashed red line correspond to 0. The coefficients were obtained using optimal bandwidths that were computed following Calonico et al. [2014b].

Figure 2.D9: Placebo Exercise: Effect of Potential Applicants (t) on Siblings (t-T)



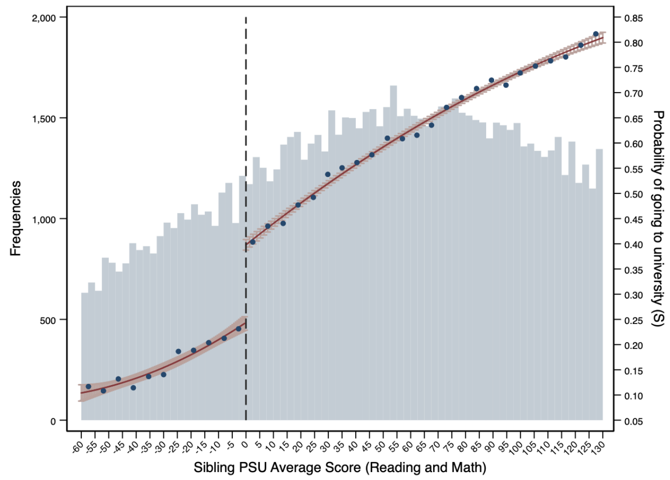
Notes: This figure illustrates the reduced form of a placebo exercise. It shows how siblings' probability of going to university evolves with the PSU score of potential applicants. The PSU score is centered around the student-loans eligibility threshold. Each dot represents the share of siblings going to university at different ranges of potential applicants PSU scores. The red lines correspond to linear approximations of these shares, and the shadow around them to 95% confidence intervals. The blue bars in the background illustrate the distribution of the potential applicants' scores in the PSU. The range used for these plots corresponds to optimal bandwidths that were computed following Calonico et al. [2014b].

Figure 2.D10: Siblings' Effects with Different Bandwidths

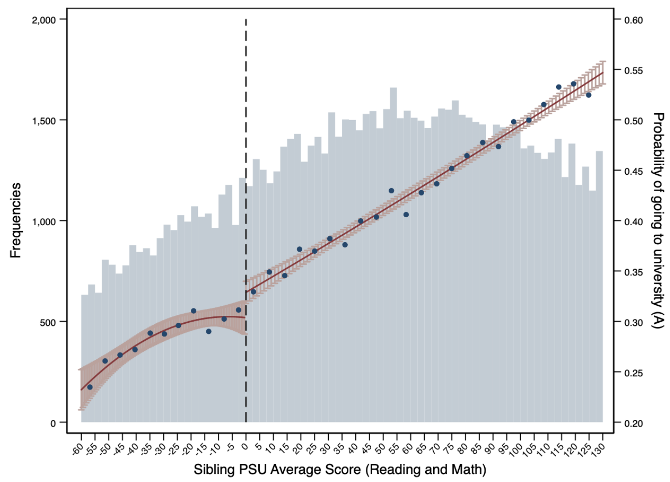


Notes: This figure illustrates the coefficients obtained when studying siblings' effects using different bandwidths. The dots represent the coefficients, and the lines illustrate 95% confidence intervals.

Figure 2.D11: First Stage and Reduced Form of Siblings RD (P2)



(a) First Stage: Siblings' Probability of going to Univeristy



(b) Reduced Form: Potential Applicants' Probability of going to University

Notes: This figure illustrates the first stage and reduced form of the siblings rd. The first panel shows how siblings' probability of going to university evolves with the score they obtain in the PSU. The second panel shows how potential applicants' probability of going to university evolves with the PSU score of their sibling. The PSU score is centered around the student-loans eligibility threshold. Each dot represents the share of siblings (panel 1) or potential applicants (panel 2) going to university at different ranges of PSU scores. The red lines correspond to quadratic approximations of these shares, and the shadow around them to 95% confidence intervals. The blue bars in the background illustrate the distribution of the siblings' scores in the PSU. The range used for these plots corresponds to optimal bandwidths that were computed following Calonico et al. [2014b].

Table 2.D1: Effect of Potential Applicants on Neighbors and Siblings University Enrollment

	2SLS-1 (1)	2SLS-2 (2)	CCT-1 (3)	CCT-2 (4)
		Panel A - Neighbors		
Potential applicant goes to university (t+1)	0.027	-0.006	-0.029	-0.029
	(0.072)	(0.091)	(0.115)	(0.116)
Constant	0.754***	0.751***		
	(0.015)	(0.019)		
First stage coefficient	0.098***	0.099***	0.098***	0.104***
	(0.006)	(0.008)	(0.008)	(0.009)
Year fixed effects	Yes	Yes	Yes	Yes
N. of students	76,349	101,222	76,349	101,222
PSU Polynomial	1	2	1	2
Bandwidth	(60.5-60.5)	(71.5-92.5)	(60.5-60.5)	(71.5-92.5)
Kleibergen-Paap F statistic	267.64	169.22		
		Panel B - Siblings		
Potential applicant goes to university (t+T)	0.011	-0.040	-0.002	0.001
	(0.072)	(0.078)	(0.086)	(0.093)
Constant	0.741***	0.730***		
	(0.017)	(0.018)		
First stage coefficient	0.136***	0.133***	0.128***	0.128***
	(0.008)	(0.009)	(0.008)	(0.010)
Year fixed effects	Yes	Yes	Yes	Yes
N. of students	41,185	85,787	41,185	85,787
PSU Polynomial	1	2	1	2
Bandwidth	(38-47.5)	(80-110.5)	(38-47.5)	(80-110.5)
Kleibergen-Paap F statistic	294.94	264.19		

Notes: The table presents the results of a placebo exercise in which I estimate the effects of potential applicants (t) on neighbors university enrollment (t-1). Columns 1 and 2 present two stages least squares estimates using a linear and quadratic polynomial of PSU respectively. Columns 3 and 4 use instead local polynomials following Calonico et al. [2014b]. Optimal bandwidths are used in all the specifications. In parenthesis, standard errors clustered at neighborhood unit level. *p-value<0.1 **p-value<0.05 ***p-value<0.01

Table 2.D2: Placebo - Effect of Potential Applicants on Neighbors and Siblings University Enrollment (II)

	2SLS-1 (1)	2SLS-2 (2)	CCT-1 (3)	CCT-2 (4)
	Panel A - Neighbors			
Potential applicant goes to university (t+1)	0.019	0.017	0.017	-0.033
	(0.048)	(0.064)	(0.069)	(0.083)
Constant	0.394***	0.395***		
	(0.013)	(0.016)		
First stage coefficient	0.140***	0.138***	0.137***	0.138***
	(0.006)	(0.008)	(0.007)	(0.009)
Year fixed effects	Yes	Yes	Yes	Yes
N. of students	97,987	112,389	97,987	112,389
PSU Polynomial	1	2	1	2
Bandwidth	(48.0-84.5)	(74.0-87.5)	(48.0-84.5)	(74.0-87.5)
Kleibergen-Paap F statistic	571.14	322.57		
	Panel B - Siblings			
Potential applicant goes to university (t+T)	0.043	0.005	0.016	-0.018
	(0.057)	(0.079)	(0.081)	(0.094)
Constant	0.449***	0.444***		
	(0.016)	(0.021)		
First stage coefficient	0.166***	0.151***	0.153***	0.152***
	(0.009)	(0.010)	(0.011)	(0.012)
Year fixed effects	Yes	Yes	Yes	Yes
N. of students	35,394	64,136	35,394	64,136
PSU Polynomial	1	2	1	2
Bandwidth	(45.0-50.0)	(68.0-108.5)	(45.0-50.0)	(68.0-108.0)
Kleibergen-Paap F statistic	347.16	234.42		

Notes: The table presents the results of a placebo exercise in which I estimate the effects of potential applicants (t) on neighbors university enrollment (t-1). Columns 1 and 2 present two stages least squares estimates using a linear and quadratic polynomial of PSU respectively. Columns 3 and 4 use instead local polynomials following Calonico et al. [2014b]. Optimal bandwidths are used in all the specifications. In parenthesis, standard errors clustered at neighborhood unit level. *p-value<0.1 **p-value<0.05 ***p-value<0.01

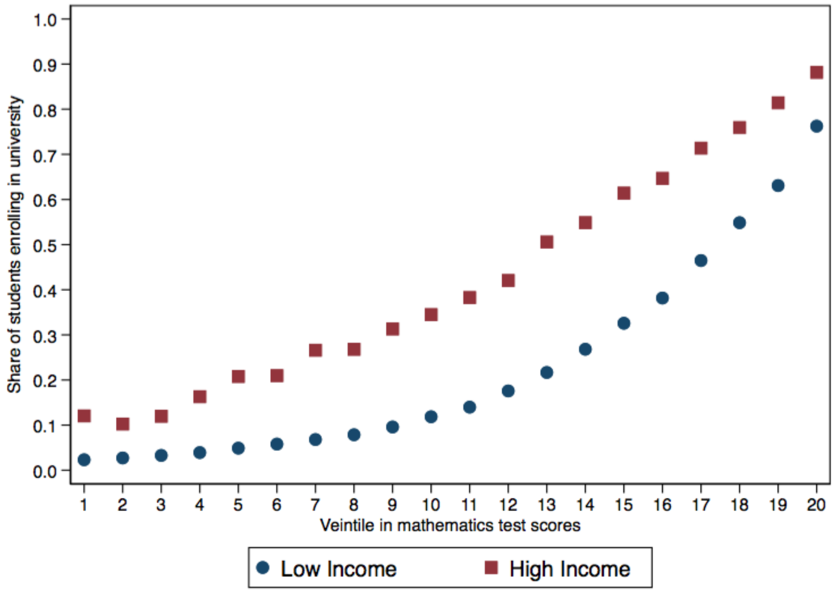
Table 2.D3: Effect of Neighbors on Potential Applicants University Enrollment (Metropolitan Region of Santiago)

	2SLS-1 (1)	2SLS-2 (2)	CCT-1 (3)	CCT-2 (4)
Neighbor goes to university (t-1)	0.129** (0.060)	0.168** (0.075)	0.151** (0.075)	0.128 (0.092)
First stage coefficient	0.132*** (0.011)	0.130*** (0.014)	0.137*** (0.009)	0.140*** (0.012)
Year fixed effects	Yes	Yes	Yes	Yes
N. of students	56,225	75,120	56,225	75,120
PSU Polynomial	1	2	1	2
Bandwidth	(55.0-84.0)	(73.5-113.0)	(55.0-84.0)	(73.5-113.0)
Kleibergen-Paap F statistic	151.92	88.51		

Notes: The table presents the results of analysis similar to those presented in table 2.2 for focusing in RM. Columns 1 and 2 present two stages least squares estimates using a linear and quadratic polynomial of PSU respectively. Columns 3 and 4 use instead local polynomials following Calonico et al. [2014b]. Optimal bandwidths are used in all the specifications. In parenthesis, standard errors clustered at neighborhood unit level. *p-value<0.1 **p-value<0.05 ***p-value<0.01

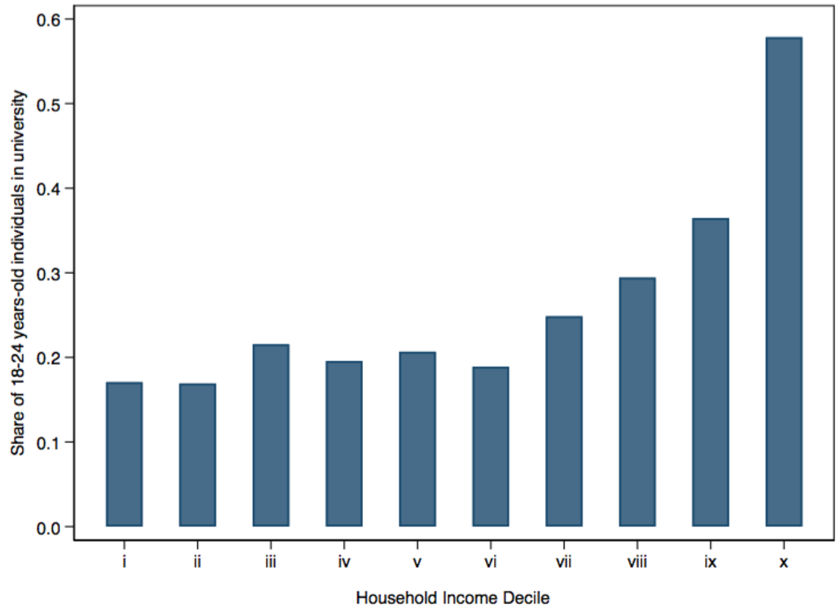
2.E Additional Figures and Tables

Figure 2.E1: Share of Students going to University vs Performance in Mathematics Standardized Test



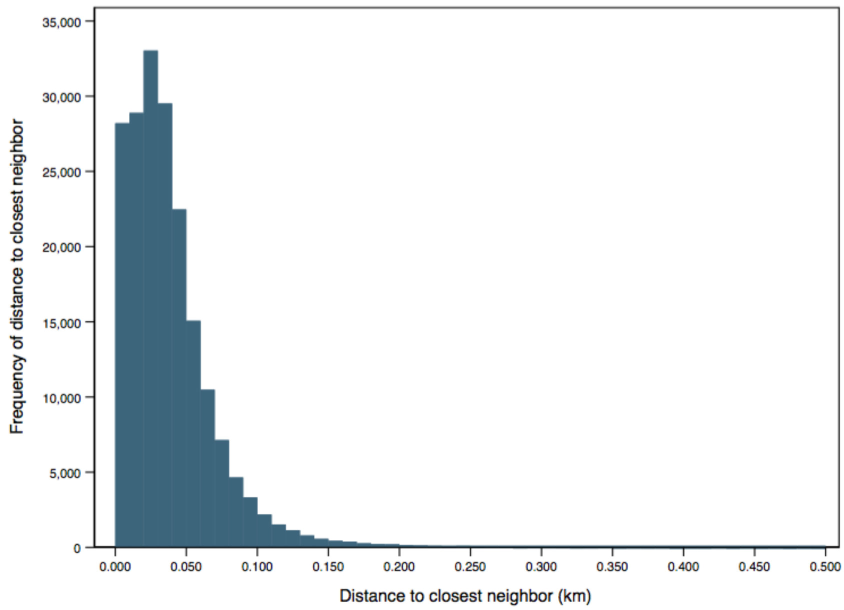
Notes: This figure illustrates how the gap in university enrollment observed across income groups evolves with ability. Ability is measured by students performance in grade 10 mathematics standardized test. University enrollment is measured 3 years later; if students do not repeat or dropout, this is one year after they complete high school. The blue dots correspond to low-income students, while the red squares correspond to high-income students. Low-income students come roughly from households in the bottom 20% of the income distribution, while high-income students from households in the top 20%. The statistics in this table are based on the sample of students in grade 10 in 2006, 2008, 2010 and 2012.

Figure 2.E2: Share of Students going to University vs Household Income (2015)



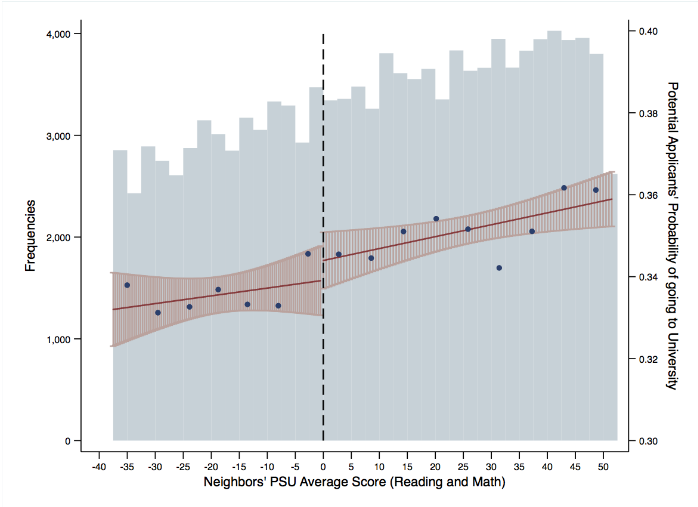
Notes: This figure illustrates the relationship between the share of 18 to 24 years old individuals going to university in 2015 and their household income. It was build using data from the Chilean national household survey, CASEN (http://observatorio.ministeriodesarrollosocial.gob.cl/casen-multidimensional/casen/basedatos.php).

Figure 2.E3: Distribution of Distance between Potential Applicants and their Closest Neighbor

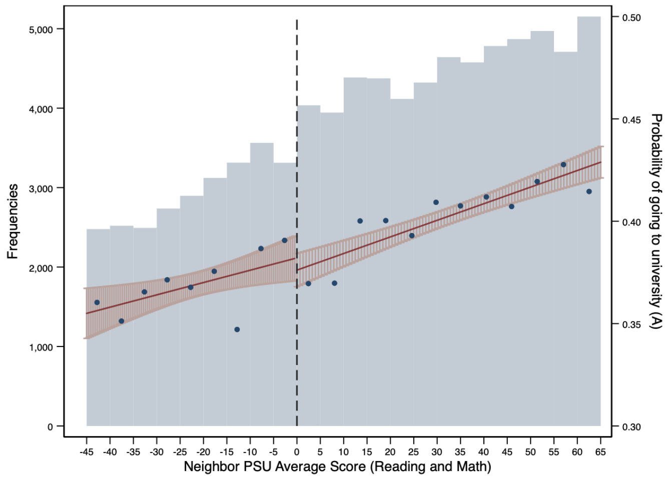


Notes: This figure illustrates the distribution of distance between potential applicants' household and their closest neighbor. Potential applicants are individuals that appear in the PSU registers between 2007 and 2012. Their neighbors are individuals that appear in the PSU registers one year before them.

Figure 2.E4: Reduced Form of Neighbors' RD: Neighbors applying the Same year and Two years before Applicants



(a) Neighbors apply on the same year



(b) Neighbors apply two years before

Notes: This figure illustrates reduced forms similar to the one presented in Figure 2.2, but this time focusing on the closest neighbor applying to university the same year as the applicant (panel 1) and two years before the applicant (panel 2). Both panels show how potential applicants' probability of going to university evolves with the PSU score of their neighbors. The PSU score is centered around the student-loans eligibility threshold. Each dot represents the share of potential applicants going to university at different ranges of neighbors' PSU scores. The red lines come from linear regressions of the outcome on the running variable on each side of the eligibility threshold, and the shadow around them to 95% confidence intervals. The blue bars in the background illustrate the distribution of the neighbors' scores in the PSU. The range used for these plots corresponds to optimal bandwidths computed following Calonico et al. [2014b].

Table 2.E1: Heterogeneity in the Effects of Closest Neighbor on Potential Applicants' Enrollment

	Socioeconomic Status			High School Track		Gender	
	Bottom 30%	Between 30%-70%	Top 30%	Academic track	Vocational track	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Neighbor goes to university (t-1)	0.033	0.150**	0.113	0.130**	0.046	0.083*	0.131**
	(0.039)	(0.065)	(0.095)	(0.053)	(0.036)	(0.050)	(0.054)
First stage	0.192***	0.174***	0.160***	0.175***	0.182***	0.189***	0.168***
	(0.012)	(0.012)	(0.015)	(0.010)	(0.012)	(0.011)	(0.010)
Reduced form	0.006	0.026**	0.018	0.023**	0.008	0.016*	0.022**
	(0.007)	(0.011)	(0.015)	(0.009)	(0.007)	(0.009)	(0.009)
N. of potential applicants	31279	30899	22906	52943	32141	39684	45400
Kleibergen-Paap F statistic	259.58	211.11	116.53	282.26	227.30	299.66	276.96

Notes: The table presents the estimated effects of neighbors on potential applicants' university enrollment depending on socioeconomic, academic and demographic variables. Columns 1 to 3 study how the effect of neighbors and siblings on potential applicants change depending on the socioeconomic status of potential applicants. Socioeconomic status is measured through an index that incorporate income level, parental education, health insurance and the high school administrative dependence. Columns 4 and 5 do the same, but distinguishing by the high school track followed by potential applicants. Finally, columns 6 and 7 look at heterogeneous effects by gender. All specifications include years fixed effects and a linear polynomial of the closest neighbor or sibling PSU score; it is allowed to be different on both sides of the student-loans eligibility threshold. Optimal bandwidths are used in all the specifications and were computed following Calonico et al. [2014b]. In parenthesis, standard errors clustered at neighborhood unit or household level. *p-value<0.1 **p-value<0.05 ***p-value<0.01

Table 2.E2: Persistence of the Effect of Neighbors on Potential Applicants' University Enrollment

	Pr. of remaining in the:			
	System		University	
	2SLS-1	2SLS-2	2SLS-1	2SLS-2
	(1)	(2)	(3)	(4)
Neighbor goes to university (t-1)	0.106***	0.084**	0.116**	0.088*
	(0.037)	(0.036)	(0.048)	(0.047)
First stage	0.178***	0.168***	0.178***	0.168***
	(0.009)	(0.010)	(0.009)	(0.010)
Reduced form	0.019***	0.015**	0.019***	0.015*
	(0.007)	(0.006)	(0.008)	(0.008)
Years fixed effects	Yes	Yes	Yes	Yes
N. of students	85084	134796	85084	134796
PSU Polynomial	1	2	1	2
Kleibergen-Paap F statistic	425.27	269.38	425.27	269.38

Notes: The table presents the estimated effects of neighbors on potential applicants' permanence in the system 1 year after enrollment. Columns 1 and 2 look at permanence in any university, while columns 3 and 4 in the same university in which potential applicants enrolled in their first year. In both cases, the outcome is 1 for applicants who enroll and remain enrolled one year later; it is 0 for applicants who do not enroll at all or who enroll but dropout after their first year. 2SLS estimates come from specifications that control for linear or quadratic polynomial of PSU which slopes are allowed to change at the cutoff. Bandwidths are the same used in the specifications presented in Table 2.2. In parenthesis, standard errors clustered at neighborhood unit level. *p-value<0.1 **p-value<0.05 ***p-value<0.01

Chapter 3

Siblings' Effects on Human Capital Investment Decisions: Evidence from Chile and Croatia

Andrés Barrios Fernández
LSE

Marin Drlje
CERGE-EI, Charles University

Dejan Kovac
Princeton University

Christopher Neilson
Princeton University

Abstract

This paper investigates how the probability of applying and enrolling in a particular university (or program) changes if an older sibling enrolls in it. We study this in Chile and Croatia taking advantage of the deferred acceptance admission systems (DA) that universities use in both countries to select their students. Exploiting the quasi-random variation generated by thousands of admission cutoffs, we find that individuals are more likely to apply and to enroll in a university (or program) if an older sibling is admitted and enrolls in it before. These siblings' spillovers persist even when, due to age differences, siblings are unlikely to attend university at the same time. We discuss five classes of mechanisms that could drive these results and present evidence consistent with the transmission of information being a relevant driver. Individuals are more likely to follow their older siblings when they enroll in programs with higher retention rates and where graduates perform better in the labor market. Older siblings' experience in university also seems to be important, suggesting that individuals learn through siblings if specific programs or universities are a good match for them.

Keywords: Siblings' Effects, University Choice, Program Choice.

JEL classification: I21, I24.

3.1 Introduction

The growth experienced by higher education in the last decades has generated great interest in understanding how individuals choose if and where to continue their education after high school. Although individual returns to schooling remain high, there is vast evidence that these returns mask substantial heterogeneity[1] and that some individuals face constraints that prevent them from making optimal choices when applying to university.[2]

The role of family —especially of parents— and peers in education have been widely studied. However, we know far less about how siblings affect human capital investment decisions. In this paper, we investigate how individuals' application and enrollment decisions are affected by the program and university attended by their older siblings. To the best of our knowledge, we are the first studying the causal relationship between siblings' university and program choices.

We study this in Chile and Croatia, exploiting the fact that in both countries universities select their students using centralized deferred acceptance admission systems (DA) that allocate applicants to programs based on their declared preferences and on a score that combines academic performance in high school and in a national level admission exam. These selection systems give rise to admission cutoffs in all oversubscribed programs (i.e. programs with more applicants than available seats). Taking advantage of the variation generated by these cutoffs, we instrument enrollment with an indicator of admission and use a fuzzy RD to investigate how having an older sibling enrolling

[1] Oreopoulos and Petronijevic [2013] presents a review of works studying returns to higher education. Hastings et al. [2013] and Reyes et al. [2013] document heterogeneous returns to postsecondary education in Chile, one of the countries we study in this paper.

[2] Hoxby and Avery [2013] show that high achieving individuals from areas with low educational attainment in the United States apply to less selective schools than similar students from other areas. This, despite the fact that better schools would admit them and provide more generous funding. This undermatching phenomenon has also been studied by Black et al. [2015], Griffith and Rothstein [2009] and Smith et al. [2013]. There is also a vast literature looking at the role of information frictions in schooling investment. Attanasio and Kaufmann [2014], Hastings et al. [2015] and Jensen [2010] study these frictions in Mexico, Chile and Dominican Republic respectively. Bettinger et al. [2012] and Hoxby and Turner [2015] look at them in the United States, and Oreopoulos and Dunn [2013] in Canada. Carrell and Sacerdote [2017] on the other hand, argues that interventions that increase university enrollment work not because they provide additional information, but because they compensate for lack of support and encouragement. Lavecchia et al. [2016] discusses these frictions and different behavioral barriers that may explain why some individuals do not take full advantage of education opportunities.

in a specific university (or program) affects individuals' probabilities of applying and enrolling in the same university (or program).

A key challenge for the identification of siblings' effects is to distinguish between social interactions and correlated effects. In our setting, correlated effects arise because siblings share genetic characteristics and grow up under very similar circumstances. Thus, it is not surprising to find a high correlation between their outcomes. Our empirical strategy compares individuals whose older siblings are marginally admitted or rejected from a specific program. Since these individuals are very similar both in observable and unobservable characteristics, we can rule out concerns about correlated effects.

In addition, if siblings' outcomes simultaneously affect each other, this gives rise to what Manski [1993] described as the reflection problem. In our setting individuals apply and enroll in university after their older siblings. The lagged structure of their decisions and the fact that the variation that we exploit in older siblings' enrollment comes only from the admission cutoffs, allow us to abstract from the reflection problem.

Studying this is not only relevant to understand how family networks and in particular siblings affect human capital investment decisions, but also because these siblings' effects may have important policy implications. First, in the context of policies that change the pool of students admitted to specific programs and institutions (i.e. affirmative action), these spillovers would imply that these type of policies, in addition to having an effect on their direct beneficiaries, would also have an indirect effect on other members of their network. Second, if the reason why individuals respond to their older siblings' choices is incomplete information, this would mean that there is space to improve the match between students and institutions by providing better information.

Our findings show that despite the differences that exist between Chile and Croatia, in both countries siblings' spillovers are very similar. Individuals whose older siblings "marginally enroll"in their target university are 10 percentage points more likely to apply to the same university.[3] While in Chile this translates into an increase of 5 percentage points (50%) in the probability of enrolling in the older sibling's target uni-

[3]We use the term "marginal enrollment"to highlight the fact that these results come from a fuzzy RD that compares individuals whose older siblings were marginally admitted or rejected from specific programs (or universities).

versity, in Croatia this figure is close to 10 percentage points (33%). We also find that in both countries individuals are more likely to apply and enroll in the target program of their older siblings when they "marginally enroll" in it. In Chile, we document an increase of around 2.5 percentage points (50%) in applications and 0.3 percentage points (33%) in enrollment; the same figures for Croatia are 4.5 percentage points (35%) and 1.5 percentage points (75%) respectively. In addition, we show that the effects are stronger for individuals who resemble their older siblings in terms of gender and academic potential.

Our main results are consistent with five broad classes of mechanisms. First, the effects could be driven by the convenience of attending university together. Second, they could be driven by competition between siblings or by changes in parental expectations. Third, having an older siblings enrolling in a specific pogram or institution could affect aspirations. Fourth, the results could be driven by path dependence or by an increase in the salience of the older sibling's program. Finally, the effects could be the result of the transmission of information between siblings. We investigate all of these alternatives, and present suggestive evidence that information is an important mechanism behind the effects we find.

This paper contributes to existing research in several ways. First, it adds to the literature studying how family and peers influence educational attainment; Björklund and Salvanes [2011] reviews the literature studying the role of family, while Sacerdote [2011] and Sacerdote [2014] review the literature studying the role of peers. Despite all the attention that these topics have received in the economic literature, little is known about how siblings affect human capital investment decisions. Dustan [2018] and Joensen and Nielsen [2018] are among the first looking at siblings' effects, but they focus on high school. The former uses a similar approach to the one used in this paper to study siblings' spillovers on the choice of high school in Mexico, while the latter exploits quasi-random variation induced by a policy change in Denmark to investigate how the decision to take advanced mathematics and science courses during high school affect younger siblings' course choice.

Goodman et al. [2015] investigates the relationship between siblings' university choice;

they study this in the United States and find that the correlation between siblings' applications are much stronger than among similar classmates. However, the lack of an exogenous source of variation on older siblings' choices prevents them from obtaining causal estimates. Barrios-Fernandez [2018] studies neighbors' and siblings' spillovers in the access to university in Chile, and finds that having a close neighbor or sibling going to university increases the probability of reaching this level of education, especially in areas where university attendance has traditionally been low. Our paper complements this work by exploiting a different source of variation and by focusing on the choice of university and program, rather than in the decision to attend university.

Finally, this paper contributes to the literature that studies how individuals choose universities and programs to apply and enroll. This has been an active area of research in the last decades, and has looked at the role of costs, information and more recently at some behavioral responses.[4] This paper adds a new element by analyzing the role of family networks on these choices.

The rest of the paper is organized in seven sections. Section 3.2 describes the university systems of Chile and Croatia, Section 3.3 the data, and Section 3.4 the empirical strategy and the samples we use. Section 3.5 presents the main results and Section 3.6 places them in the context of previous findings and discusses potential mechanisms. Finally, Section 3.7 concludes.

3.2 Institutions

This section describes the university admission systems of Chile and Croatia, emphasizing the rules that generate the discontinuities that we later exploit to identify spillovers among siblings.

[4] Papers investigating costs and liquidity constraints include Dynarski [2000], Seftor and Turner [2002], Dynarski [2003], Long [2004], van der Klaauw [2002], and Solis [2017]; papers studying the role of information include Bettinger et al. [2012], Busso et al. [2017], Dinkelman and Martínez A. [2014], Hastings et al. [2015, 2016], Hoxby and Turner [2015], Oreopoulos and Dunn [2013], Wiswall and Zafar [2013], Booij et al. [2012], Nguyen [2008], Castleman and Page [2015]. Carrell and Sacerdote [2017] argues that differences in support and encouragement are key to explain differences in college enrollment. Lavecchia et al. [2016] discusses the literature on behavioral constraints that may explain why some individuals do not take full advantage of their education opportunities.

3.2.1 University Admission System in Chile

In Chile, there are public and private universities. All the public universities and 9 out of 43 private universities are part of the Council of Chilean Universities (CRUCH), an organization that was created to improve coordination and to provide advice to the Ministry of Education in matters related to higher education.

The CRUCH universities, and since 2012 a group of eight private universities, select their students using a deferred acceptance admission system that only considers students' performance in high school and in a national level university admission exam (PSU).[5] The PSU is taken in December, at the end of the Chilean academic year, but students typically need to register before mid-August.[6] Since 2006 all students graduating from public and voucher schools are eligible for a fee waiver that makes the PSU free for them.[7]

Universities publish the list of programs and vacancies offered for the next academic year well before the PSU examination date. Concurrently, they inform the weights allocated to high school performance and to each section of the PSU to compute the application score for each program.

With this information available and after receiving their PSU scores, students apply to their programs of interest using an online platform. They are asked to rank up to 10 programs according to their preferences. Places are then allocated using an algorithm of the Gale-Shapley family that matches students to programs using their preferences and scores as inputs. Once a student is admitted to one of her preferences, the rest of her applications are dropped. As shown in Figure 3.1, this system generates a sharp discontinuity in admission probabilities in each program with more applicants than vacancies.

[5] The PSU has four sections: language, mathematics, social sciences and natural sciences. The raw scores obtained by students in each of these sections are adjusted to obtain a normal distribution of scores with mean 500 and standard deviation 110. The extremes of the distribution are truncated to obtain a minimum score of 150 and a maximum score of 850 in each section. In order to apply to university, students need to take language, mathematics and at least one of the other sections. Universities are free to set the weights allocated to these instruments for selecting students.

[6] In 2017, the registration fee for the PSU was CLP 30,960 (USD 47).

[7] Around 93% of high school students in Chile attend public or voucher schools. The entire registration process operates through an online platform that automatically detects the students' eligibility for the fee waiver.

Universities that do not use the centralized system have their own admission processes.[8] Although they could use their own entrance exams, the PSU still plays an important role in the selection of their students, mostly due to the existence of strong financial incentives for both students and institutions.[9] For instance, the largest financial aid programs available for university studies require students to score above a certain threshold in the PSU.

The coexistence of these two selection systems means that being admitted to a university that uses the centralized platform does not necessarily translate into enrollment. Once students receive an offer from a university they are free to accept or reject it without any major consequence. This also makes it possible for some students originally rejected from a program to receive a later offer. Figure 3.1 illustrates how the admission to a program translates into enrollment.

3.2.2 University Admission System in Croatia

In Croatia, there are 49 universities. Since 2010, all of them select their students using a centralized admission system managed by the National Informational System for University Application (NISpVU).

As in the case of Chile, NISpVU uses a deferred acceptance admission system that focuses primarily on students' performance in high school and in a national level university exam.[10] The national exam is taken in late June, approximately one month after the end of the Croatian academic year. Students however, are required to submit a free of charge online registration form by mid-February.

[8]From 2007, we observe enrollment in all the universities of the country independently of the admission system they use.

[9]Firstly, creating a new test would generate costs for both the institutions and the applicants. Secondly, for the period studied in this paper, part of the public resources received by higher education institutions depended on the PSU performance of their first-year students. This mechanism, eliminated in 2016, was a way of rewarding institutions that attracted the best students of each cohort.

[10]In rare cases, certain universities are allowed to consider additional criteria for student assessment. For example, Academy of Music assigns 80% of admission points based on an in-house exam. These criteria are known well in advance, and are clearly communicated to students through NISpVU. Students are required to take the obligatory part of the national exam, comprising mathematics, Croatian and foreign language. In addition, students can choose to take up to 6 voluntary subjects. Students' performance is measured as a percentage of the maximum attainable score in a particular subject.

Universities disclose the list of programs and vacancies, together with program specific weights allocated to high school performance and performance on each section of the national exam roughly half a year before the application deadline. This information is transparently organized and easily accessible through an interactive online platform hosted by NISpVU.

Students are free to submit a rank of up to 10 programs from the moment they register in the online platform. The system allows them to update these preference until mid-July, moment in which students are allocated to programs based on their last list of preferences. As in Chile, vacancies are allocated using a Gale-Shapley algorithm, giving rise to similar discontinuities in admission probabilities (Figure 3.1).

Before the final deadline, the system allows students to learn their position on the queue for each one of the programs to which they apply. This information is regularly updated to take into account the changes that applicants make in their list of preferences. In this paper we focus on the first applications submitted by students after receiving their scores in the national admission test. Since some of them change their applications before the deadline, admission based on these applications does not translate one-to-one into enrollment (Figure 3.1).[11]

Two important differences with the Chilean system, are that in Croatia all universities use the centralized admission system and that rejecting an offer is costly for students.[12]

3.3 Data

In this paper we exploit administrative data provided by different public agencies from Chile and Croatia. In both countries, the main data sources are the agencies in charge of the centralized university admission system: DEMRE in Chile, and NISpVU and ASHE in Croatia.

From DEMRE we got access to individual level data of all the students registered to

[11]We focus on the first applications students submit after learning their exam performance to avoid endogeneity issues in admission results that may arise from some students learning about the system and being more active in modifying their applications before the deadline.

[12]Enrollment tuition fees are covered by the Ministry of Science and Education. If a student rejects an offer, eligibility for tuition waiver is lost.

take the PSU between 2004 and 2015. This data contains information on students' performance in high school and in the different sections of the university admission exam. It also contains information on demographic and socioeconomic characteristics of individuals and on their applications and enrollment to the universities that select students through the centralized admission system. To identify siblings, we exploit the fact that when registering for the exam, students provide the national id number of their parents. Using this unique identifier we can match all siblings that correctly reported this number for at least one of their parents.[13]

In the case of Chile, we complement this information with registers from the Ministry of Education and from the National Council of Education. In this data we observe enrollment for all the institutions offering higher education in the country between 2007 and 2015, information that allows us to build program-year specific measures of retention for the cohorts entering the system in 2006 or later. In these registers, we also observe some program and institution characteristics, including past students' performance in the labor market (i.e. employment and annual earnings). Finally, using the registers of the Ministry of Education we are also able to match students to their high schools and observe their academic performance before they start higher education.

NISpVU and ASHE provided us with similar data for Croatia. These individual registers contain information on students' performance in high school and in the different sections of the university admission exam, and on applications and enrollment to all Croatian universities between 2012 and 2018. These registers include the home address of students, information that we exploit to identify siblings. We define as siblings two individuals living in exactly the same address at the moment of registration for the university admission exam.

Using this data, we were able to identify around $190,000$ pairs of siblings in Chile and $13,000$ in Croatia in which the older sibling had at least one active application to an oversubscribed program (i.e. admitted or in the waiting list). Their characteristics, as well as the characteristics of the rest of the students registered for taking the university admission exams in the same period, are presented in Table 3.1.

[13] For the period that we study 79.2% of the students in the registers report a valid national id number for at least one of their parents. 77.0% report the national id number of their mother.

In both countries, the sample of siblings is very similar to the rest of the population in terms of demographic characteristics. The bigger differences arise when looking at socioeconomic and academic variables. In Chile, individuals whose older siblings had at least one active application come from wealthier and more educated households. In both countries, they are more likely to have followed the academic track in high school and to perform better in high school and in the university admission test.

These differences are not surprising. After all, the siblings samples contain individuals from families in which at least one child had an active application to a selective program (i.e. oversubscribed programs) in the past. On top of this, the institutions that use the centralized admission system in Chile are on average more selective than the rest. Thus, individuals with active applications to these universities are usually better candidates than the average student in the whole population of PSU takers.

3.4 Empirical Strategy

The identification of siblings' effects is challenging. In the first place, since siblings share genetic characteristics and grow up under very similar circumstances, it is not surprising to find that their outcomes —including the program and university that they attend— are highly correlated. Thus, a first identification challenge consists in distinguishing these correlated effects from the effects generated by interactions among siblings. In addition, if siblings' outcomes simultaneously affect each other, this gives rise to what Manski [1993] described as the reflection problem. In our setting, given that older siblings decide to apply and enroll in university before their younger siblings, this is less of a concern (i.e. decisions that have not yet taken place should not affect current decisions). However, there could still be cases in which siblings decide together the university and program that they want to attend and therefore we need an empirical strategy to address this potential threat.

To overcome the identification challenges described in the previous paragraph, we exploit thousands of cutoffs generated by the DA systems that Chilean and Croatian universities use to select their students. Taking advantage of the discontinuities created by these cutoffs on admission probabilities, we use a Regression Discontinuity

(RD) design to investigate how older siblings' admission to their target university (or program) affects the probabilities that their younger siblings apply and enroll in the same target university (or program).

Since individuals whose older siblings are marginally admitted or rejected from a specific program are very similar, the RD allows us to rule out the estimated effects to be driven by differences on individual or family characteristics, eliminating in this way concerns about correlated effects. Moreover, considering that the variation that we exploit in the program in which older siblings enroll comes only from their admission status and cannot be affected by the choices that their younger siblings will make in the future, we can abstract from the reflection problem.[14]

As discussed in Section 3.2, rejecting an offer does not have any major consequence for Chilean students. As a result, there is a non-negligible share of applicants that despite being admitted to a particular university or program, decide not to enroll. Thus, when studying how older siblings' actual enrollment affects their younger siblings, we use a fuzzy RD in which older siblings' enrollment in a specific program is instrumented with an indicator of admission.

In the case of Croatia, we follow a similar approach. Although in this setting rejecting an offer is costly, we use a fuzzy and not a sharp RD because as explained in Section 3.2, we focus our attention on the first application students submit after receiving their results in the university admission exam. Since some individuals modify their applications in the weeks following the exam results, admission to the first set of preferences does not translate one-to-one into enrollment.[15]

This paper investigates how individuals' probabilities of applying and enrolling in specific programs and universities change when their older sibling is marginally admitted and enrolls in them. The basic idea behind our empirical design consists in defining for each institution and program, the sample of older siblings who are marginally admitted and marginally rejected from them, and then compare the outcomes of their younger

[14]We show that this is indeed the case in a series of placebo exercises that we present in Figures 3.B6 and 3.B7 in the Appendix.

[15]We focus on the first applications submitted after learning the exam scores to avoid endogeneity issues in admission results that may arise by some type of students being more active in modifying their applications in the weeks following the exam.

siblings. Next, we discuss the restrictions used to identify the groups of marginal older siblings in each case.

3.4.1 University Sample

This section describes the restrictions applied to the data in order to build the sample used to study how individuals' probabilities of applying and enrolling in a specific university change when their older sibling is marginally admitted and enrolls in it.

As discussed earlier, the assignment mechanism used in Chile and Croatia results in cutoff scores for each program with more applicants than available seats; these cutoffs correspond to the lowest score among the admitted students. Let c_{jut} be the cutoff for program j in institution u in year t. If the program j in institution u is ranked before the program j' offered by institution u' in student i's preference list, we write $(j, u) \succ (j', u')$.[16] Denoting the application score of individual i as a_{ijut}, we can define marginal students in the university sample as those whose older siblings:

1. listed program j in institution u as a choice, such that all programs preferred to j had a higher cutoff score than j (otherwise assignment to j is impossible):
   $c_{jut} < c_{j'u't} \ \forall \ (j', u') \succ (j, u)$.

2. listed program j in institution u as a choice, such that programs not preferred to j are dictated by an institution different from u (otherwise being above or below the cutoff would not generate variation on the institution attended).

3. had a score sufficiently close to j's cutoff score to be within a given bandwidth bw around the cutoff:
   $|a_{ijut} - c_{jut}| \leq bw$.

This sample includes individuals whose older siblings were rejected from (j, u) ($a_{ijut} < c_{jut}$) and those whose older siblings scored above the admission cutoff ($a_{ijut} \geq c_{jut}$). Since the application list in general contains more than one preference, this means that the same individual may belong to more than one program-institution marginal group.

[16] This notation does not say anything about the optimality of the declared preferences. It only reflects the order stated by individual i.

Figure 3.1 illustrates the probability of admission and enrollment in a given institution around the admission cutoff in Chile and Croatia.

3.4.2 Program Sample

In addition to studying the effect of the institution attended by older siblings, we investigate how individuals' probabilities of applying and enrolling in a specific program change when their older sibling is marginally admitted and enrolls in it. The sample used in this case is very similar to the one described in Section 3.4.1. The only difference is that in this case only the first and third restrictions are applied. This means that in the Program Sample, the institution attended by older siblings does not necessarily change by being above or below the admission cutoff. As far as the program in which they are admitted changes, they will be in the sample.

3.4.3 Identifying Assumptions

As in any other RD setting, the validity of our estimates relies on two key assumptions. First, individuals should not be able to manipulate their application scores around the admission cutoff. The structures of the university admission systems in Chile and Croatia make the violation of this assumption unlikely. However, to confirm this we study if the distribution of the running variable (i.e. older sibling's application score centered around relevant cutoff) is continuous at the cutoff. We do this by implementing the test suggested by Cattaneo et al. [2018b] whose results are presented in Figure 3.2. As expected, we do not detect discontinuities on the distribution of the running variable at the cutoff. Strictly speaking, the density of the running variable needs to be continuous around each admission cutoff. In our analysis, we pool them together because there are thousands of cutoffs in our samples and studying them independently would be impractical.

Second, in order to interpret changes in individuals' outcomes as a result of the admission status of their older siblings, there cannot be discontinuities in other potential confounders at the cutoff (i.e. the only relevant difference at the cutoff must be older siblings' admission). We study this by taking advantage of the availability of a rich set of socioeconomic and demographic characteristics. Figure 3.B1, in Appendix 3.B shows that this is indeed the case. Once more, this analysis is done by pooling together all the admission cutoffs in our samples.

As mentioned at the beginning of this section, to study the effect of older siblings' enrollment (instead of admission) on younger siblings' outcomes we use a fuzzy RD. This approach can be thought as an IV strategy, meaning that in order to interpret our estimates as a local average treatment effects (LATE) we need to satisfy the assumptions discussed by Imbens and Angrist [1994].[17] In this setting, in addition to the usual IV assumptions, we also need to assume that receiving an offer for a specific program does not make the probability of enrolling in a different program bigger than in the absence of the offer (in Appendix 3.A we discuss this in detail). Given the structure of the admission systems that we study, this additional assumption does not seem very demanding.[18]

An additional issue related to the interpretation of our estimates is that as noted by Cattaneo et al. [2016], by pooling together different cutoffs, our estimates correspond to a weighted average of LATEs across programs. This weighted average gives more importance to programs with more applicants in the vicinity of the admission cutoff. Since there could be heterogeneity on the characteristics of individuals around each admission cutoff, and also on the effect of admission and enrollment at each admission cutoff, we need to be careful with the interpretation of this weighted average.

In order to understand what is driving our results we perform a detailed heterogeneity analysis along multiple dimensions including both individual and program characteristics. In addition, Tables 3.B3 and 3.B4 in Appendix 3.B, study how different our

[17]Independence, relevance, exclusion and monotonicity. In this setting, independence is satisfied around the cutoff. The existence of a first stage is shown in Figure 3.1. The exclusion restriction implies that the only way through which older siblings' admission to a program affects younger siblings' outcomes, is by the increase it generates in older siblings' enrollment in that program. Finally, the monotonicity assumption means that admission to a program weakly increases the probability of enrollment in that program (i.e. being admitted into a program does not reduce the enrollment probability in that program).

[18]In the case of Chile, where not all universities use the centralized admission system and where rejecting an offer is not costly for students, this assumption could be violated if for instance universities that do not use the centralized admission system offer scholarships or other types of incentives to attract students marginally admitted to universities that do use it. Although it does not seem very likely for the universities out of the centralized system to define students' incentives based on marginal offers to other institutions, we cannot completely rule out this possibility. In the case of Croatia, where students lose their funding in case of rejecting an offer, violations to this assumption are much less likely.

results are when we re-weight observations around each cutoff by the inverse of the total number of applicants around it. Although the estimates are slightly smaller in this case, the main conclusions still hold.

A final consideration for the interpretation of our results relates to the findings of Barrios-Fernandez [2018]. According to it, the probability of attending university increases with close peers' enrollment. If marginal admission to the programs that we study translates into an increase in total university enrollment, then our estimated results could simply reflect that individuals whose older siblings attend university are more likely to enroll. We address these concerns in Table 3.B5 of Appendix 3.B where we study if older siblings' marginal admission increases their own and their younger siblings' total enrollment. In the case of Chile, we only find a small increase on total enrollment of older siblings. This result is not surprising. As discussed in Section 3.2, the universities that use the centralized admission system in Chile are on average more selective than the rest. This means that individuals rejected from these institutions still have many other alternatives available. In the case of Croatia, we find that marginal admission translates into a more significant increase in older siblings total enrollment.[19] However, we do not find an extensive margin response among younger siblings in any of the two settings.[20]

In Appendix 3.B, we also present multiple additional robustness checks. We show that as expected, changes in the admission status of younger siblings do not have an effect on older siblings; that our estimates are robust to different bandwidth choices and that when replacing the actual cutoffs by placebo ones, there are no significant effects on any of the outcomes that we study.

[19]In Section 3.5 we perform heterogeneity analyses along multiple dimensions. We find that the effects in both countries are driven by individuals whose older siblings are admitted to more selective institutions. Individuals on the margin of admission to selective institutions are less likely to suffer a change on their total enrollment since in case of rejection they still have many alternatives available.

[20]Our results for Chile are still consistent with the findings of Barrios-Fernandez [2018]. The variation that he exploits generates a much bigger difference on older siblings enrollment than the one we document here.

3.5 Results

This section begins by providing additional details about the empirical approach used to estimate the effects of interest. It then discusses how individuals' probabilities of applying and enrolling in a specific university or program change when their older siblings are marginally admitted and enroll in it. It continues by looking at how these effects vary with individual and program characteristics. The section concludes by investigating how individuals' academic performance is affected by the admission and enrollment results of their older siblings.

3.5.1 Method

In all the specifications used in this paper, we pool together observations from all over-subscribed programs and center older siblings' application scores around the relevant admission cutoff. The following expression describes our baseline specification:

$$y_{ijut\tau} = \beta admitted_{iju\tau} + f(a_{iju\tau};\gamma) + \mu_t + \mu_{ju\tau} + \varepsilon_{ijut\tau} \tag{3.1}$$

where,

$y_{ijut\tau}$ is the outcome of interest of the younger sibling of the siblings-pair i applying to university in year t and whose older sibling was near the admission cutoff of program j in university u in year τ.

$admitted_{iju\tau}$ is a dummy variable that takes value 1 if the older sibling of the siblings-pair i was admitted to program j offered by university u on year τ ($a_{iuj\tau} \geq c_{uj\tau}$)

$f(a_{iuj\tau};\gamma)$ is a function of the application score of the older sibling of the siblings-pair i for program j offered by university u on year τ.

μ_t and $\mu_{ju\tau}$ are younger sibling's application year and cutoff-older sibling's application year fixed effects; and ε_{ijut} is an error term.

We estimate parametric and non-parametric versions of this specification. For the parametric approach, $f(a_{iujt\tau};\gamma)$ corresponds to linear or quadratic polynomials of $a_{iuj\tau}$ whose slopes are allowed to change at the cutoff. For the non-parametric approach we follow Calonico et al. [2014b, 2018] and use local polynomials of degree 1 and 2.

In all cases, we use optimal bandwidths computed according to Calonico et al. [2014b] (Figures 3.B2 and 3.B3 in Appendix 3.B illustrates how sensible are our estimates to the choice of bandwidth). In Tables 3.B1 and 3.B2 of Appendix 3.B we also present a parametric specification in which we allow the slope of the running variable to be different for each admission cutoff.[21] The estimates obtained with this specification are very similar to the ones we discuss in this section.

Since all the specifications that we use focus on individuals whose older siblings are near an admissions cutoff, our estimates represents the average effect of older siblings' marginal admission compared to the counterfactual of marginal rejection from a target program.[22]

To study the effect of enrollment —instead of the effect of admission— we instrument older siblings' enrollment ($enrolls_{iju\tau}$) with the indicator of admission ($admitted_{iju\tau}$).

Standard errors must account for the fact that each older sibling may appear several times in our estimation sample if she is near two or more cutoffs. To deal with this situation we use cluster standard errors at family level.

To study heterogeneous effects, we add to the baseline specification an interaction between older siblings' admission and the characteristic along which heterogeneous effects are being investigated (i.e. $admitted_{iju\tau} \times x_{ijut\tau}$). This interaction is also used as an instrument for the interaction between older sibling's enrollment and $x_{ijut\tau}$. In both cases, $x_{ijut\tau}$ is also included as a control.

3.5.2 Mean Effects

This section discusses how older siblings' admission and enrollment in specific universities (or programs) affect their younger siblings' university (or program) choices. When investigating effects on university choice, we use the University Sample defined in Section 3.4.1. For studying the effects on program choice, we use the Program Sample.

[21]The estimation of these specifications is costly in computing time. In addition to the fixed effects included in the baseline specification, we include interactions between the running variable $a_{iju\tau}$ and $\mu_{ju\tau}$, and also between $a_{iju\tau}$, $\mu_{ju\tau}$ and $admitted_{ijut\tau}$.

[22]Strictly speaking, our estimates represent a weighted average of multiple LATEs. See Section 3.4.3 for additional details.

The RD estimates illustrated in Figures 3.3 and 3.4 give consistent causal evidence that students are more likely to apply and to enroll in a university and program if an older sibling was admitted before. The precise reduced form coefficients represented in these figures are presented in Tables 3.2 and 3.3. These tables summarize the reduced form and IV-estimates obtained from the parametric and non-parametric specifications discussed in Section 3.5.1.

As shown in Figure 3.1, although admission does not translates one-to-one into enrollment, there is a strong relation between these two variables. Thus, under the assumptions discussed in section 3.4.3, we can combine the reduced forms and first stages to obtain fuzzy RD estimates for the effect of older siblings' enrollment in their target university and program on the probabilities that their younger siblings apply and enroll in them.

In the case of Chile, having an older sibling "marginally enrolling"[23] in a specific university increases the likelihood of applying to it in the first preference by around 8 percentage points (50%) and in any preference by around 10 percentage points (30%). These changes in applications also translate into an increase of around 5 percentage points (50%) in the probability of enrolling in that university. The results for Croatia are very similar. Younger siblings are 8 percentage points (25%) more likely to apply to a specific university in the first preference and 10 percentage points (18%) more likely to apply to it in any preference if their older sibling "marginally enrolls"in it. This increase in applications translates into a 9 percentage points (30%) increase in the probability of enrollment.

When focusing instead on the effect of older siblings' marginal enrollment on a specific program, the results follow a similar pattern. In the case of Chile, individuals are 0.8 percentage points (40%) more likely to apply to that program in the first preference, 2.8 percentage points (55%) more likely to apply to it in any preference, and 0.3 percentage points (30%) more likely to enroll in it (this last figure is not statistically significant). For Croatia, the same figures are 1.4 percentage points (45%), 4.3 percentage points

[23] "Marginally enrolling"means that the individual was marginally admitted to the program in which he enrolled. We emphasize this to remind the reader that the estimates come from comparing individuals whose older siblings were marginally admitted and marginally rejected from specific programs.

(33%) and 1.6 percentage points (80%) respectively and they are all statistically significant.

These results show that despite the differences that exist between Chile and Croatia in size, location, culture and history, individuals respond to their older siblings' human capital investment decisions in a very similar way.

3.5.3 Effects by Sibling Similarity: Demographic and Academic Characteristics

This section studies if the effects discussed in Section 3.5.2 change depending on how close siblings are in terms of gender, age and academic potential. To measure similarity in academic potential, we use the absolute difference in the high school GPA of siblings. In Croatia, we observe high school GPA only for students completing their secondary education before 2015; this explains the smaller sample used in this part of the analysis for Croatia.

Tables 3.4 and 3.5 summarize the results of this section. The probability of applying to the target university of older siblings after they "marginally enroll"in it does not seem to be differentially affected by the gender of siblings (column 1 of Table 3.4). The coefficient capturing the differential effect for siblings of the same gender is not statistically significant neither in Chile nor in Croatia. When looking at changes in actual enrollment, we find a positive, but only marginally significant difference between same gender and opposite gender siblings (column 4 of Table 3.4). In Chile, the estimated difference represents a 30% of the main effect, while in Croatia it represents more than 50% of the main effect.

The differences by siblings' gender are clearer when looking at changes in the probabilities of applying and enrolling in the older siblings' target programs. Both effects —on applications and on enrollment— are bigger for same gender individuals (columns 1 and 4 of Table 3.5). The difference in the effect on the application probability represents 50% of the main effect in Chile, and 90% of the main effect in Croatia. The differences on the effects on enrollment are even bigger. In Chile, the effect on enrollment is only significant for siblings of the same gender; for them the effect is eight times bigger than for siblings of opposite gender. In the case of Croatia, the effect we find for siblings of the same gender is almost three times the effect found for siblings of the opposite gender.

We find no evidence of heterogeneous effects by age difference (columns 2 and 4 in Tables 3.4 and 3.5). The only statistically significant difference arises when studying changes on the probability of applying to the older siblings' target university in Chile (column 2 in Table 3.4). However, the coefficient capturing the differential effect is very small. According to it, the effect decreases by 0.6 percentage points with age difference. Since the main effect is 12.5 percentage points, this means that the age difference between siblings would need to be bigger than 20 years to make the mean effect disappear.

The difference in academic potential between siblings seems to make a bigger difference on the size of the effects (columns 3 and 6 of Tables 3.4 and 3.5). All the coefficients measuring heterogeneous effects are significant, except for the ones estimated when studying the effects on the probability of applying and enrolling in older siblings' target university in Croatia. As discussed at the beginning of this section, in Croatia we observe high school GPA only for a subsample of individuals. This considerably reduces the sample size affecting our ability to detect significant effects.

In the case of Chile, a difference of 1-σ (128.26) in siblings' high school GPA score makes the effect on the probability of enrolling in the target university and in the target program of the older sibling to practically disappear. In the case of Croatia, a difference of 1-σ (0.57) in siblings' high school GPA reduces the total effect in university and program enrollment by around 30% and 50% respectively (only the second difference is statistically significant).

Table 3.6 provides additional details on how the effects differ by the gender of siblings. For these analyses we split the sample in two groups depending on the gender of the older sibling and estimate a specification that includes an interaction between the treatment and a dummy variable that indicates if the younger sibling is female.

Although splitting the sample results in a loss of precision, we end with a pretty consistent general picture. In both countries, males respond more to what happens with their older siblings when they are also males. Females seem to be less responsive than males to what happens with male older siblings. Interestingly, in the case of Chile, the response of females does not seem to dramatically change with the gender of the older sibling (i.e. they respond in a similar way to what happens with older brothers and sisters).

The results discussed in this section suggest that the effects are stronger when younger and older siblings are of the same gender, especially in the case of males. In addition, our results show that the effects are bigger when siblings are similar in academic potential.

3.5.4 Effects by Program and University Quality

This section studies how the effects documented in Section 3.5.2 change depending on the quality of the target university or target program of the older sibling. We measure quality in terms of students' academic potential, first year dropout rates, graduates' employability and graduates' wages.[24]

Student quality is the only variable studied in this section that we are able to build for both countries. We define the quality of the students in a program in a given year using the average performance of admitted students on the university admission exams. The rest of the variables are available only for Chile. We compute dropout rates for each university and program using individual level data provided by the Ministry of Education. This data allows us to compute dropout rates for all cohorts entering university since 2006.[25] Variables measuring labor market performance of former students are available at program level and are computed by the Ministry of Education with the support of the Chilean Tax Authority.[26]

[24]Employability is measured one year after graduation, whereas wages are measured four years after graduation. We observe them only once for each program-university. This means that in our analysis these variables do not change over time.

[25]The cohorts of older siblings applying to university in 2004 and 2005 are assigned the dropout rates observed for their target programs and universities in 2006. Since some programs disappear from one year to the next, this means that we are not able to complete information for all programs offered in 2004 and 2005.

[26] These figures are only available for programs that were being offered in 2018 and that had more than 4 cohorts of graduates. In addition, the Tax Authority only reports employment and earnings statistics for programs in which they observe at least 10 graduates.

The main results of this section are summarized in Table 3.7. The estimates show that the effects on the probabilities of applying and enrolling in the target university and program of older siblings decrease with dropout rates (column 1), and increase with employment rates (column 2) and earnings (column 3) of former graduates.

To put these differences in context, an increase of 1-σ (0.138) in the share of individuals that dropout from a university after the first year (i.e. enroll in a different university or do not enroll at all in the second year), reduces the effect on the probability of applying to the target university of the older sibling by almost 30% and the effect on the probability of enrolling in it by a little less than 40%. Although the coefficients capturing these differences when focusing on the target program of older siblings are not statistically significant, they are still relevant in size. An increase of 1-σ (0.173) in the share of individuals that drop out from a program after the first year reduces the effect on the probabilities of applying and enrolling in the older siblings' target program by 20% and 40% respectively. In the case of employment rates, an increase of 1-σ (0.094) translates into an increase of 0.24 and 0.28 percentage points on the effects on application and enrollment in older sibling's target university. The same figures when focusing on older siblings' target programs are 0.017 and 0.05 percentage points. Finally, when looking at earnings, a similar picture emerges. Our results indicate that an increase of 1-σ (0.384 MM CLP) in earnings translate into an increase of 5.26 and 3.57 percentage points on the effect of applying and enrolling in older siblings' target university and of 1.61 and 0.34 percentage points on the effect of applying and enrolling in older siblings' target program.

The results obtained when looking at heterogeneity by student quality are very similar. As shown in Figure 3.5, younger siblings are more likely to apply to older siblings' target universities and programs when the admitted students perform better in the university admission exam. The coefficients presented in this figure come from a specification like 3.1 estimated in three independent samples. The levels of selectivity correspond to the bottom, middle and top third of programs according to the average score obtained by their students in the admission exam.[27]

[27]Since our sample only includes programs with positive waiting lists, none of the estimates really includes non-selective programs. This is particularly relevant in the case of Chile, where the less selective institutions are not part of the sample at all.

3.5.5 Effects on Human Capital

We end this section by looking at a different set of outcomes. Until this point the analyses have focused on changes in the applications and enrollment of younger siblings. Here, we study if having an older sibling marginally admitted into a specific program affects students' academic performance during high school or in the university admission exam. Since not all students take the university admission exam, we replace missing values by zero. This means that our estimates capture differences in the actual performance in the exam, but also differences in the probability of taking it. The bandwidths used in this section are the same used in section 3.5.2.

Tables 3.8 and 3.9 show that older siblings' "marginal enrollment"in their target program or university does not generate significant changes on younger siblings' high school performance, or in their performance in the university admission exam.

These results hold for both countries and in the university and in the program sample, indicating that the differences documented in the previous sections in terms of applications and enrollment are not driven by an improvement in younger siblings' academic performance.

3.6 Discussion

The results presented in Section 3.5 show that the path followed by older siblings in higher education affects the choice of university and program of their younger siblings. These results are consistent with the findings of Dustan [2018] and Joensen and Nielsen [2018], that show that similar spillovers exist in the choice of high school in Mexico and in the decision to take advanced mathematics and science courses in secondary education in Denmark. Our findings are also in line with Goodman et al. [2015], who documents that the correlation between university choices of siblings is stronger than the ones observed among similar classmates in the U.S., and with Barrios-Fernandez [2018] who studies externalities in university enrollment in Chile. This last work finds that especially in areas where university attendance is low, having an older neighbor or sibling going to university increases the enrollment probability of younger applicants.[28]

Although documenting the existence of sibling spillovers in university and program choices in two settings as different as Chile and Croatia is interesting in itself, from a policy perspective it is important to also investigate the mechanisms behind these responses. In the rest of this section, we discuss five classes of mechanisms that could be relevant in this context.

A first possibility to explain why individuals are more likely to apply and enroll in the university and program in which their older siblings enroll, is that it may be convenient to go to university together. Attending the same university may reduce commuting and living costs for siblings and could make their overall university experience more enjoyable.[29] Figure 3.6 shows that the effect persists even among siblings whose age difference implies that they would spend very little or no time at all together in university. Along the same line, Figure 3.7 shows that the effects are significant, both when older siblings continue living with their parents and when they move and live by themselves. These results suggest that in our setting, convenience is not the only driver of our results.[30]

A second possibility highlighted by previous research is competition and changes in parental expectations. Indeed, Joensen and Nielsen [2018] argues that the fact that their results are driven by siblings who are close in age and in academic performance is evidence in favor of competition being the main driver of their results. As discussed in the previous paragraph, in our case the results persist even among siblings whose age difference is five years or more. In addition, Table 3.6 shows that although the effects between brothers seem to be stronger, effects between sisters and different-gender siblings are also significant. If competition only arises between same-gender siblings

[28]Note that while Barrios-Fernandez [2018] focuses on the extensive margin (i.e. enroll or not in university), here we analyze the specific choice of university and program. In Appendix 3.B we show that in the case of Chile, admission of older siblings to selective universities does not generate a change in overall enrollment of younger siblings. This is not surprising, because older siblings rejected from the programs we study, still have many other options available.

[29]They may enjoy each other's company, and for younger siblings it may also be useful to have someone close who understands how everything works.

[30]In some settings, the admission systems give some advantage to siblings of current or former students. This however is not a concern in our case; Chile and Croatia use a centralized admission system that select students based only on their academic performance in high school and in a national level admission exam.

or only between siblings close in age, then competition cannot be the only driver of our results. Dustan [2018] on the other hand, claims that finding no heterogeneous effects depending on the difference between selectivity in the target high school and in the high school to which students would go in case of rejection, is evidence against a parental expectations channel. The intuition behind this argument is that if counterfactual high schools are similarly selective, then having a child admitted in one or the other should affect parental expectations in a similar way. Thus, if as shown in table 3.10, the effects persist when there is no difference in selectivity between older siblings' counterfactual universities or programs, we can rule out the effects being driven by a change in parental expectations.[31]

A third alternative that we consider is path dependence or changes in the salience of the alternative chosen by the older sibling. As shown in section 3.5.4, we find that the effects are stronger when the quality of the university or program of the older sibling is higher. This suggests that younger siblings do not follow their older siblings everywhere (i.e. they care about the quality of the program). This last result is consistent with an information channel in which individuals learn about the quality of specific programs or universities through their older siblings.

The results discussed in the previous paragraph are also consistent with an aspirations channel. Seeing an older sibling going to a high-quality institution or program may motivate younger siblings to work harder to achieve something similar. However, not finding heterogeneous effects depending on the selectivity of the counterfactual programs of older siblings (Table 3.10) and as in Barrios-Fernandez [2018], finding no effect on younger siblings' academic performance, goes against this story. If students' aspirations are affected, we would expect to see them exerting more effort preparing for university, something that in our case is not reflected in their high school or university

[31] To study heterogeneous effects by differences in the selectivity of counterfactual universities and programs, we first identified the counterfactual program of each older sibling. For admitted students, this corresponds to the next preference in their application list to which they would have been admitted. To rejected students, this corresponds to the program in which they enrolled. In both countries, Chile and Croatia, selectivity of a program is measured by the average performance of their students in the admission exam. The specifications presented in Table 3.10 augment the baseline specification including an interaction between the treatment and the difference in selectivity between the counterfactual programs, and by controlling by this difference. In the case of Chile the standard deviation of the selectivity measure is close to 100, in the case of Croatia it is equal to 0.129.

admission exam performance.

Finally, older siblings may facilitate the access to relevant information for their younger siblings. As discussed by Barrios-Fernandez [2018], older siblings can transmit different types of information. They may provide information about the application process, about costs and funding opportunities, about universities and programs quality or about the experience of going to university.

Table 3.11 provides evidence consistent with the idea that individuals learn from their older siblings' experience if a specific program or university would be a good match for them. Siblings are similar in many dimensions, and therefore if older siblings have a bad experience in a specific program or university, their younger siblings may infer that applying and enrolling in that program is not a good choice for them. In our data, the best available proxy for older siblings' experience at university is dropout. We use this information to implement two types of analysis.

First, taking advantage of the fact that we observe retention for all the cohorts enrolling in university since 2006, for each individual in our sample we predict program-specific dropout risks. To predict these risks we use a cross-validation Logistic Lasso regression that includes a rich vector of demographic, socioeconomic and academic characteristics.[32]. Then, using this predicted risk we estimate the specifications presented in columns 1 and 3 of Table 3.11.

Second, following Dustan [2018], we estimate similar specifications in which we include an interaction between the treatment and a dummy variable that indicates if the older sibling dropouts from the program or university in which she enrolls,[33] and with a variable that controls by the main effect of dropout (columns 2 and 5 of Table 3.11).[34]. The results of this exercise should be interpreted with caution. Dropping out from a program or university is not random, and although by adding the indicator of dropout

[32]The variables included in these regressions are the ones described in Table 3.1. For the estimation of these models we use 5 folds and we choose the models that minimize the mean-squared prediction error.

[33]Note that the program or university in which older siblings enroll are not necessarily the ones to which they are applying to.

[34]In this case, we study dropout in the 4 years following enrollment. To be able to do this, we restrict the sample to sibling pairs in which the older sibling applies to university before 2011.

in the specification we capture some of the differences that may exist between individuals who remain and leave a program, there could still be differences that we are not able to control for.[35] In addition, the dropout variable can only be built for older siblings who actually enroll in the system. Table 3.B5 in Appendix 3.B shows that marginal admission does not translate into relevant increases in older siblings' total enrollment. However, this change affects the composition of the sample used for this analysis. Columns 3 and 6 of Table 3.11 present the results of a similar specification that incorporates additional terms and allows us to study if the timing of older siblings' dropout affects the responses of their younger siblings.[36]

Although these two exercises are quite different, and having in mind the caveats discussed for the second exercise in the previous paragraph, we find pretty consistent results. The predicted risk of dropout and actual dropout reduce the probability of individuals applying and enrolling in the programs and universities attended by their older siblings. As expected, if individuals apply to university after their older siblings drop out, the likelihood of applying and enrolling in the same university and program is lower than if they apply before the dropout occurs. However, even in this last case there is a decrease in the estimated effects. This suggests that even before the dropout takes place, younger siblings learn that the experience their older siblings are having in a specific program or university has not been great.

Even though the evidence discussed in this section does not allow us to perfectly distinguish the mechanisms that drive our results, it suggests that information and in particular information about the university experience of someone close, might play a relevant role in the choice of university and program.

[35]In addition, note that with this specification we are studying if the effects found when comparing admitted and rejected individuals who remain in the programs in which they enroll, are bigger than the ones found when comparing admitted and rejected individuals who dropout from the program in which they enroll. In general, admitted and rejected individuals enroll in different programs.

[36]This specification also includes controls for the age difference between siblings.

3.7 Conclusions

This paper studies how the probability of applying and enrolling in a specific university or program is affected by the admission and enrollment outcomes of older siblings. By investigating this, we contribute to understanding the role played by family networks on human capital investment decisions after high school.

We study this in Chile and Croatia, taking advantage that in both countries universities select their students using a centralized deferred acceptance admission system that allocates students to programs only taking into account their declared preferences and their performance in high school and in national level admission exams. These admission systems create thousands of discontinuities that we exploit in a fuzzy RD framework to study the causal effect of the university and program attended by older siblings on individuals' application and enrollment decisions. This variation allows us to address the main identification challenges that arise in the context of peers' effects (i.e. correlated effects and the reflection problem).

Our findings show that despite the differences that exist between Chile and Croatia, in the two countries siblings' effects are statistically and economically significant. In both settings, older siblings enrollment in a particular university increases the probability that their younger siblings apply to it by around 10 percentage points. These figures translate into an increase of 5 percentage points (50%) in the probability of enrolling in the same university in Chile, and 10 percentage points (33%) in Croatia. We also find that in both countries younger siblings are more likely to apply and enroll in the program in which their older siblings enroll. In Chile, we document an increase of around 2.5 percentage points (50%) in applications and 0.3 percentage points (33%) in enrollment; the same figures for Croatia are 4.5 percentage points (35%) and 1.5 percentage points (75%).

These results are consistent with five classes of mechanisms. We discuss mechanisms related to convenience, competition and parental expectations, aspirations, path dependence and salience, and information. We provide suggestive evidence that information about the quality of universities and about the quality of the student-program match is an important driver of our results.

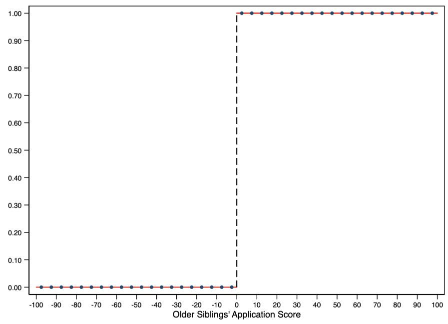
Our findings have important policy implications. They suggest that especially in contexts of incomplete information, policies that change the composition of the admitted students to a given university or program have an indirect effect on the siblings and potentially on other members of the networks of the direct beneficiaries. They also suggest that providing information about the experience that individuals would have in specific universities or programs, could improve their application and enrollment decisions.

Still further research is required to identify the type and accuracy of the information transmitted by siblings, and to find effective ways of closing the information gaps between applicants with different levels of exposure to university.

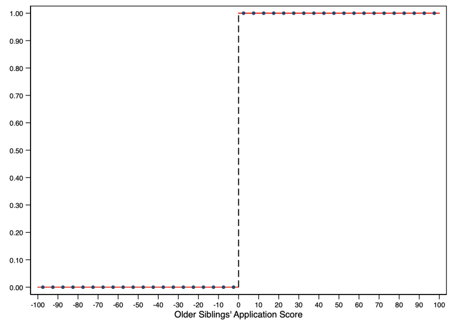
Acknowledgments

We thank Karun Adusumilli, Josh Angrist, Esteban Aucejo, Christopher Avery, Michal Bauer, Randall K. Filer, Sebastian Gallegos, Alan B. Krueger, Jacob N. Shapiro, Peter Blair, Taryn Dinkelman, Joshua Goodman, Jan Hanousek, Xavier Jaravel, Stepan Jurajda, Vasily Korovkin, Camille Landais, Alexandre Mas, Alan Manning, Sandra McNally, Guy Michaels, Daniel Munich, Andreas Menzel, Christian Ochsner, Steve Pischke, Mariola Pytliková, Johannes Spinnewijn, Daniel Reck, Steven Rivkin, Jan Zápal and Kresimir Zigic for many useful comments. We are also grateful to participants at LSE, Princeton University and CERGE-EI internal seminars, and at Umag Conference 2017 "Economics in a Changing World". Finally, we thank the Ministries of Education of Chile and Croatia, the DEMRE, AZVO and ASHE for giving us access to administrative data.

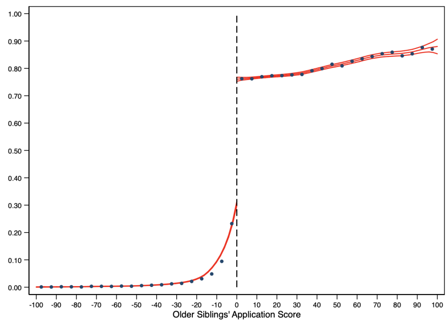
Figure 3.1: Older Siblings' Admission and Enrollment Probability in Target Program at the Admission Cutoff (First Stage)



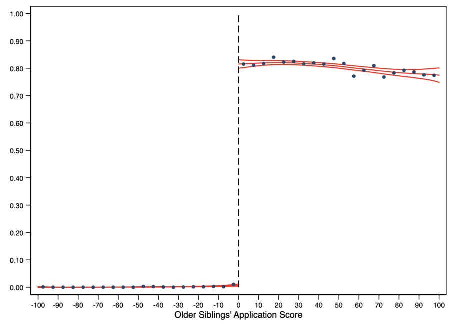
(a) Admission - Chile



(b) Admission - Croatia



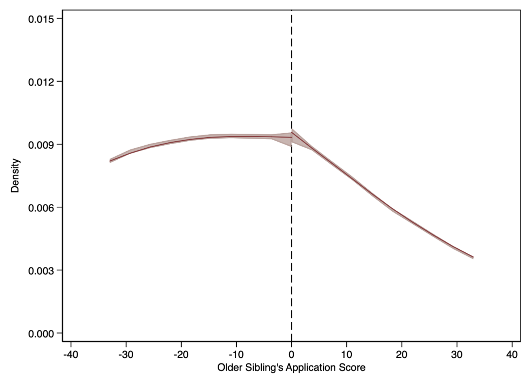
(c) Enrollment - Chile



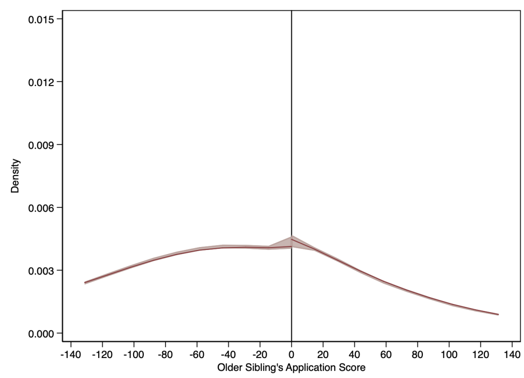
(d) Enrollment - Croatia

Notes: This figure illustrates older siblings' admission and enrollment probabilities around the admission cutoffs of their target programs in Chile and Croatia. Figures (a) and (c) illustrate these probabilities for the case of Chile, while figures (b) and (d) for Croatia. Red lines represent local linear polynomials and 95% confidence intervals. In all cases triangular kernels are used. The bandwidths used for the local polynomials correspond to optimal bandwidths computed according to Calonico et al. [2014b] for the estimation of discontinuities at the cutoff. Blue dots represent sample means of the dependent variable for bins of width 5.

Figure 3.2: Density of Older Siblings' Application Scores at the Target Program Admission Cutoff



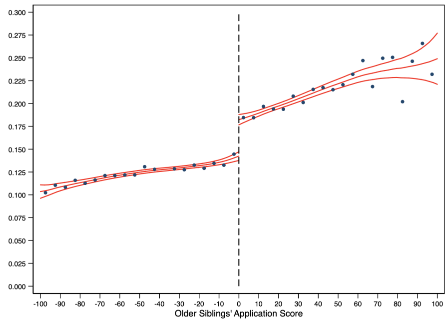
(a) Chile



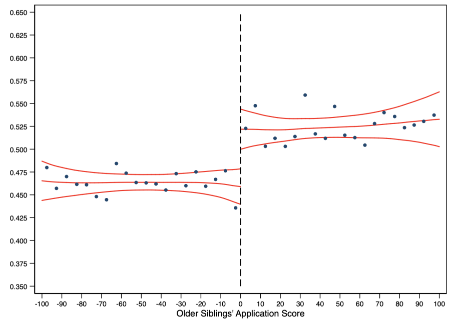
(b) Croatia - First Set of Applications

Notes: This figure illustrates the density of older siblings' application scores around the cutoff. Figure (a) illustrates this density for Chile, while figure (b) for Croatia. Red lines represent local quadratic polynomials and the shadows in the back 95% confidence intervals. In all cases, triangular kernels are used. Bandwidths are estimated according to Cattaneo et al. [2018b]. The p-values associated to the null hypothesis of no jumps at the cutoff are 0.379 and 0.393 respectively.

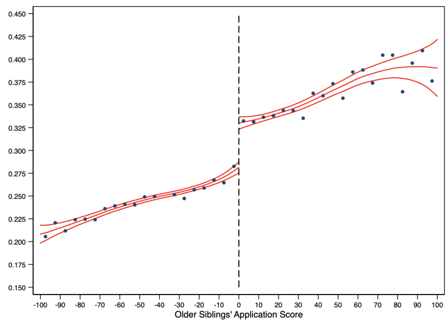
Figure 3.3: Probability of Applying to and Enrolling in Older Sibling's Target University (Reduced Form - P1)



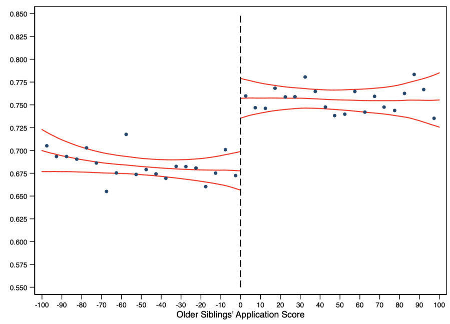
(a) Applies (1st Preference) - Chile



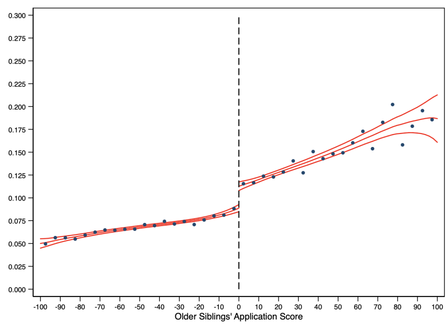
(b) Applies (1st Preference) - Croatia



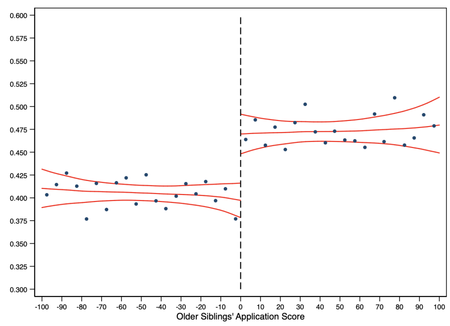
(c) Applies - Chile



(d) Applies - Croatia



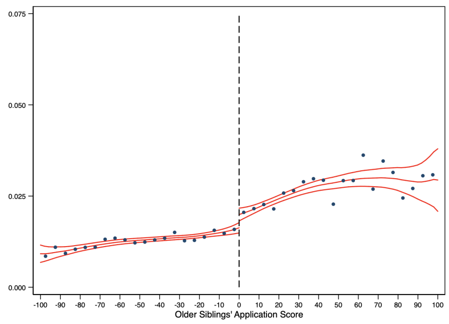
(e) Enrolls - Chile



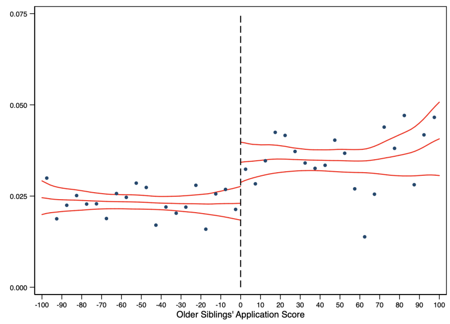
(f) Enrolls - Croatia

Notes: This figure illustrates the probability that younger siblings apply to and enroll in the target university of their older siblings in Chile and Croatia. Figures (a), (c) and (e) illustrate the case of Chile, while figures (b), (d) and (f) the case of Croatia. Red lines correspond to local polynomials of degree 1 and 95% confidence intervals. In all cases triangular kernels are used. The bandwiths used to build these figures correspond to optimal bandwidths computed following Calonico et al. [2014b]. Blue dots represent sample means of the dependent variable for bins of width 5.

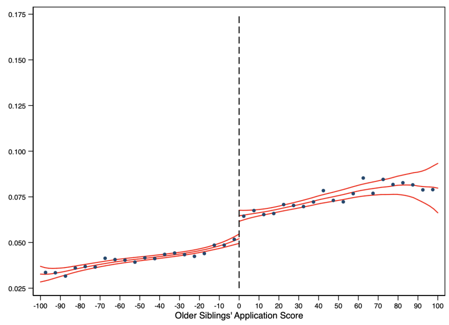
Figure 3.4: Probability of Applying to and Enrolling in Older Sibling's Target Program (Reduced Form - P1)



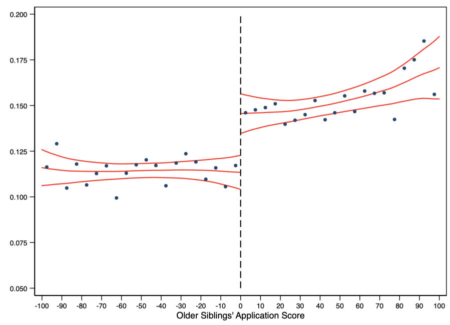
(a) Applies (1st Preference) - Chile



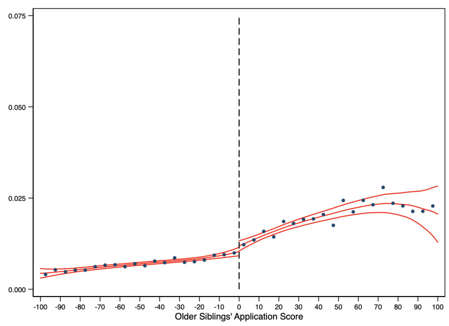
(b) Applies (1st Preference) - Croatia



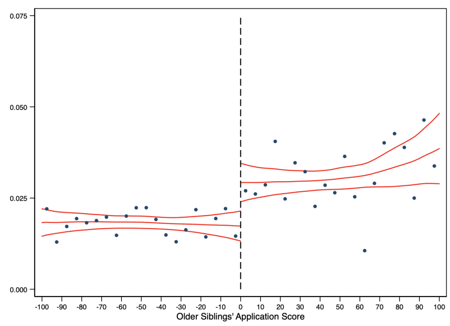
(c) Applies - Chile



(d) Applies - Croatia



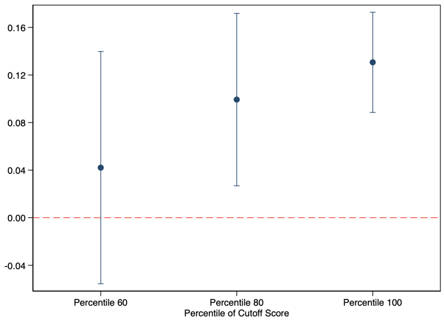
(e) Enrolls - Chile



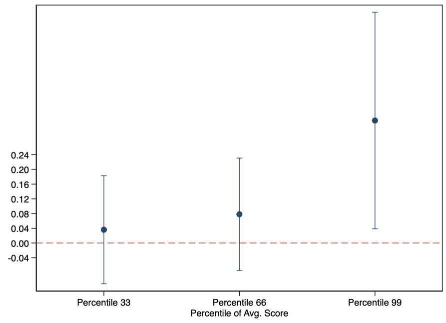
(f) Enrolls - Croatia

Notes: This figure illustrates the probability that younger siblings apply to and enroll in the target program of their older siblings in Chile and Croatia. Figures (a), (c) and (e) illustrate the case of Chile, while figures (b), (d) and (f) the case of Croatia. Red lines correspond to local polynomials of degree 1 and 95% confidence intervals. In all cases triangular kernels are used. The bandwiths used to build these figures correspond to optimal bandwidths computed following Calonico et al. [2014b] for estimating the discontinuities at the cutoff. Blue dots represent sample means of the dependent variable for bins of width 5

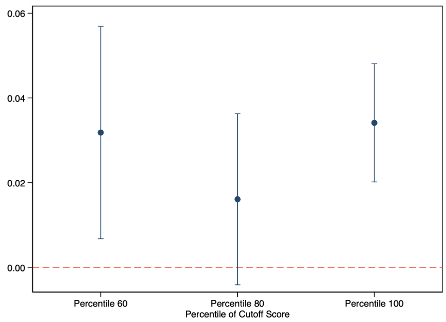
Figure 3.5: Probability of Younger Siblings Applying to Older Siblings' Target Program or University depending Program Selectivity (P1)



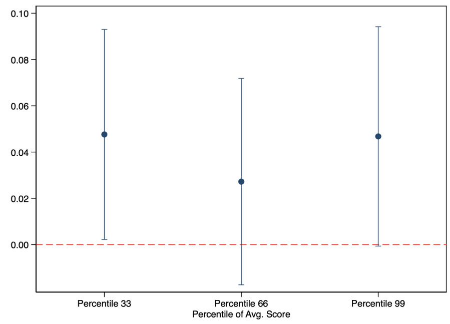
(a) University - Chile



(b) University - Croatia



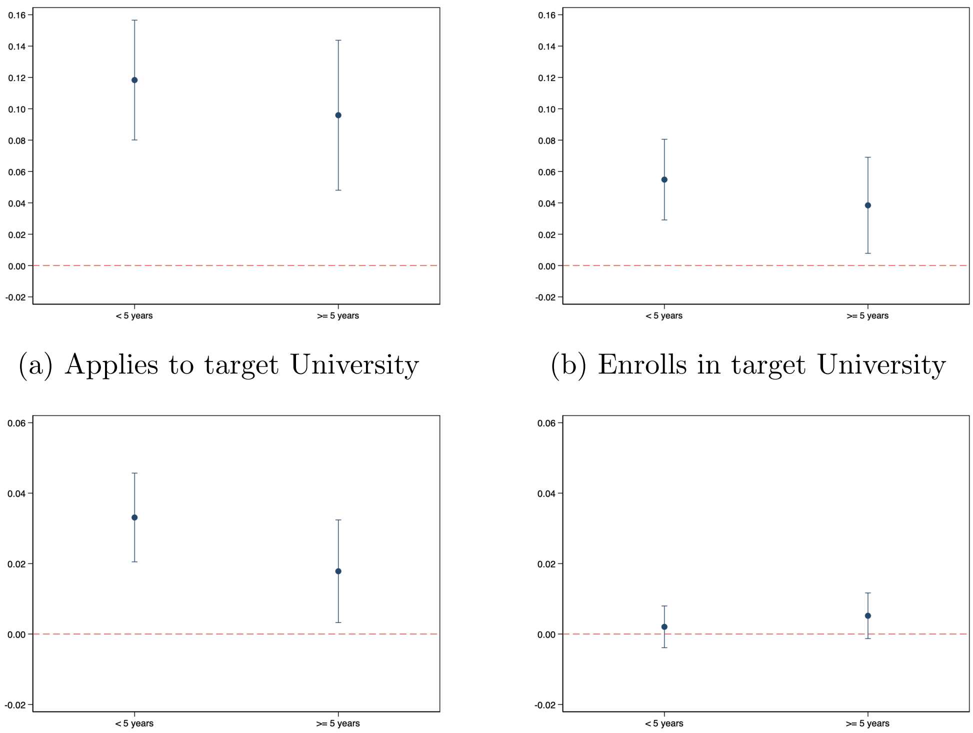
(c) Program - Chile



(d) Program - Croatia

Notes: This figure illustrates how the selectivity of the programs in which older siblings marginally enroll, affects younger siblings' probability of applying to the target university or program of older siblings. Figures (a) and (c) illustrate these probabilities for the case of Chile, while figures (b) and (d) for Croatia. The dots represent estimates from a parametric specification like the one discussed in Tables 3.2 and 3.3. Controls and bandwiths are the same used in these tables. Each coefficient comes from a different sample; the samples were created using the average admission score as criteria. The lines represent 95% confidence intervals. Standard errors are clustered at family level.

Figure 3.6: Probability of Younger Siblings Applying to and Enrolling in Older Siblings' Target University or Program by Age Difference (P1)



(a) Applies to target University

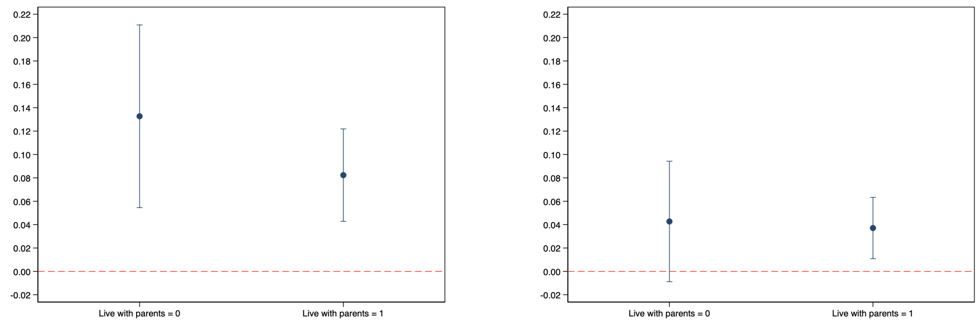
(b) Enrolls in target University

(c) Applies to target Program

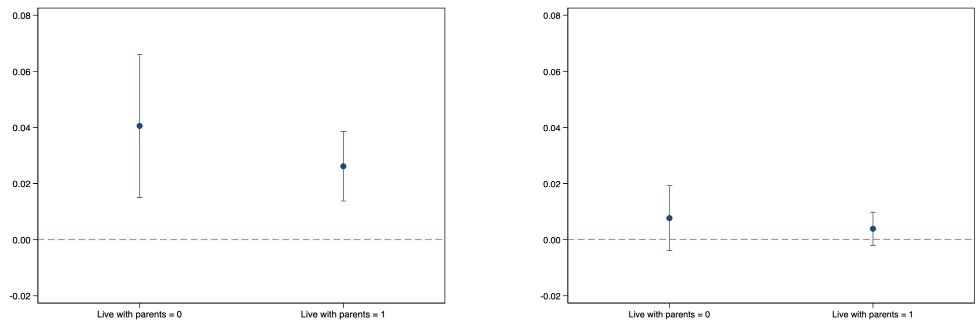
(d) Enrolls in target Program

Notes: This figure illustrates the effect of older siblings' marginal enrollment in their target university or program on the probability that younger siblings apply to and enroll in the same target university or program by age difference in Chile. Figures (a) and (b) focus on the university choice, while figures (c) and (d) on the program choice. The dots represent estimates from a parametric specification like the one discussed in Tables 3.2 and 3.3. Controls and bandwiths are the same used in these tables. Each coefficient comes from a different sample; the samples were created using age difference as criteria. The lines represent 95% confidence intervals. Standard errors are clustered at family level.

Figure 3.7: Probability of Younger Siblings Applying to and Enrolling in Older Siblings' Target University or Program by Older Siblings' Place of Residence (P1)



(a) Applies to the same University (b) Enrolls in the same University



(c) Applies to the same Program (d) Enrolls in the same Program

Notes: This figure illustrates the effect of older siblings' marginal enrollment in their target university or program on the probability that younger siblings apply to and enroll in the same target university or program by older siblings' place of residence in Chile. Figures (a) and (b) focus on the university choice, while figures (c) and (d) on the program choice. The dots represent estimates from a parametric specification like the one discussed in Tables 3.2 and 3.3. Controls and bandwiths are the same used in these tables. Each coefficient comes from a different sample; the samples were created using older siblings' place of residence as criteria. The lines represent 95% confidence intervals. Standard errors are clustered at family level.

Table 3.1: Summary Statistics

	Chile		Croatia	
	Siblings Sample	Whole Sample	Siblings Sample	Whole Sample
	(1)	(2)	(3)	(4)
A. Demographic characteristics				
Female	0.522	0.520	0.572	0.567
	(0.499)	(0.499)	(0.494)	(0.495)
Age when applying	18.787	19.829	18.878	19.158
	(0.607)	(2.484)	(0.621)	(0.963)
Family group	4.800	4.625	2.784	1.925[1]
	(1.507)	(1.607)	(1.287)	(1.198)
Siblings in HE	1.065	0.364	1.053	0.251
	(0.695)	(0.635)	(0.233)	(0.391)
Will live with parents	0.651	0.626		
	(0.477)	(0.484)		
Will live with relatives	0.109	0.127		
	(0.311)	(0.333)		
Will live independently	0.240	0.247		
	(0.427)	(0.431)		
B. Socioeconomic characteristics				
High income ($\geq$ CLP 850M)	0.279	0.128		
	(0.449)	(0.334)		
Mid income (CLP 270M - 850M)	0.400	0.325		
	(0.490)	(0.469)		
Low income ($\leq$ CLP 270)	0.321	0.546		
	(0.467)	(0.498)		
Parental ed: $\leq$ high school	0.100	0.254		
	(0.300)	(0.435)		
Parental ed: high school	0.334	0.386		
	(0.472)	(0.487)		
Parental ed: vocational HE	0.146	0.115		
	(0.353)	(0.319)		
Parental ed: university	0.411	0.234		
	(0.492)	(0.423)		
Health insurance: private	0.354	0.211		
	(0.478)	(0.408)		
Health insurance: public	0.553	0.684		
	(0.497)	(0.465)		
Health insurance: other	0.093	0.105		
	(0.290)	(0.307)		
C. Academic characteristics				
High school: private	0.198	0.104		
	(0.398)	(0.306)		
High school: voucher	0.508	0.485		
	(0.500)	(0.500)		
High school: public	0.287	0.401		
	(0.452)	(0.490)		
High school: gymnasium	0.846	0.673	0.439	0.416[2]
	(0.361)	(0.469)	(0.496)	(0.496)
High school: technical	0.154	0.327	0.561	0.584[2]
	(0.361)	(0.469)	(0.496)	(0.496)
Takes admission test	0.953	0.868	0.865	0.835[2]
	(0.211)	(0.338)	(0.342)	(0.372)
High school GPA score	556.773	519.997	268.373	265.298
	(128.255)	(139.417)	(65.766)	(66.600)
Admission test avg. score	523.252	443.032	312.800	286.247
	(142.840)	(187.849)	(102.568)	(112.787)
D. Institutional characteristics				
Universities	33	33	49	49
Programs $\times$ year	10,994	12,137	2,631	5,146
Observations	187,677	2,823,897	12,947	199,475

Notes: The table present summary statistics for Chile and Croatia. Columns (1) and (3) describe the siblings sample used in this paper, while columns (2) and (4) describe all the students registered to take the university admission test in the same period of time covered by our sample.
[1] In Croatia, *Family group* counts number of siblings within a family.
[2] In Croatia, high school and state exam information is available from 2011 to 2015. This sample has 155,587 observations (corresponding siblings sample has 8,398 observations).

Table 3.2: Probability of Applying to and Enrolling in Older Siblings' Target University

	Applies in 1st Preference		Applies in any Preference		Enrolls	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
Local Polynomial	0.093***	0.089***	0.120***	0.116***	0.057***	0.054***
	(0.013)	(0.012)	(0.019)	(0.018)	(0.010)	(0.009)
2SLS	0.072***	0.081***	0.101***	0.095***	0.045***	0.044***
	(0.012)	(0.011)	(0.015)	(0.014)	(0.010)	(0.009)
Reduced Form	0.033***	0.038***	0.047***	0.045***	0.021***	0.021***
	(0.006)	(0.005)	(0.007)	(0.007)	(0.005)	(0.004)
Observations	73331	152301	73331	152301	73331	152301
Outcome mean	0.16	0.16	0.30	0.29	0.10	0.10
Optimal bandwidth	15.75	42.30	13.18	31.40	18.62	47.40
Kleibergen-Paap Wald F Statistic	5441.60	5905.71	5441.60	5905.71	5441.60	5905.71
			Panel B - Croatia			
Local Polynomial	0.080***	0.090**	0.097***	0.092***	0.090***	0.100***
	(0.026)	(0.030)	(0.028)	(0.031)	(0.024)	(0.029)
2SLS	0.075***	0.070***	0.109***	0.102***	0.084***	0.090***
	(0.019)	(0.023)	(0.019)	(0.024)	(0.018)	(0.023)
Reduced Form	0.063***	0.058***	0.091***	0.085***	0.070***	0.075***
	(0.016)	(0.019)	(0.016)	(0.020)	(0.016)	(0.019)
Observations	12950	17312	12950	17312	12950	17312
Outcome mean	0.32	0.32	0.55	0.56	0.29	0.29
Optimal bandwidth	76.24	115.75	79.65	130.98	87.92	114.18
Kleibergen-Paap Wald F Statistic	6459.56	4214.09	6459.56	4214.09	6459.56	4214.09
Cutoff-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Running Variable Polynomial	1	2	1	2	1	2

Notes: The table presents non-parametric, parametric (2SLS) and reduced form estimates for the effect of older siblings' marginal enrollment in a target university on younger siblings' probability of applying to and enrolling in the same university. The non-parametric specification controls for a linear and quadratic local polynomial of older siblings' application score centered around the target program admission cutoff and for older and younger siblings application years fixed effects. A triangular kernel is used to give more weight to observations around the cutoff. The parametric specification controls for a linear and quadratic polynomial of older siblings' application score centered around the target program admission cutoff. The slope of the running variable is allowed to change at the cutoff. In addition, the parametric specification controls for target program-year, older siblings application year, and younger siblings application year fixed effects. Optimal bandwidths computed according to Calonico et al. [2014b] are used in non-parametric specifications. Parametric specifications use bandwiths of 15 and 35 for linear and quadratic specifications in the case of Chile; the same figures for Croatia are 80 and 120. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 3.3: Probability of Applying to and Enrolling in Older Sibling's Target Program

	Applies in 1st Preference		Applies in any Preference		Enrolls	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
Local Polynomial	0.009***	0.010***	0.027***	0.027***	0.003	0.003
	(0.003)	(0.004)	(0.006)	(0.007)	(0.002)	(0.003)
2SLS	0.008***	0.007**	0.028***	0.025***	0.003	0.002
	(0.003)	(0.003)	(0.005)	(0.006)	(0.002)	(0.003)
Reduced Form	0.004***	0.003**	0.015***	0.012***	0.002	0.001
	0.001)	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)
Observations	136364	214840	136364	214840	136364	2148409
Outcome mean	0.02	0.02	0.06	0.05	0.01	0.01
Optimal bandwidth	28.08	38.51	19.11	35.56	32.10	42.68
Kleibergen-Paap Wald F Statistic	13867.40	9520.72	13867.40	9520.72	13867.40	9520.72
			Panel B - Croatia			
Local Polynomial	0.014***	0.013**	0.043***	0.048***	0.016***	0.014***
	(0.005)	(0.006)	(0.010)	(0.014)	(0.005)	(0.006)
2SLS	0.015***	0.015***	0.034***	0.040***	0.014***	0.016***
	(0.005)	(0.005)	(0.009)	(0.011)	(0.004)	(0.005)
Reduced Form	0.013***	0.013***	0.028***	0.032***	0.012***	0.013***
	(0.004)	(0.004)	(0.007)	(0.009)	(0.003)	(0.004)
Observations	34882	47366	34882	47366	34882	47366
Outcome mean	0.03	0.03	0.13	0.13	0.02	0.02
Optimal bandwidth	72.66	118.67	69.04	105.58	86.87	123.04
Kleibergen-Paap Wald F Statistic	13693.44	10147.08	13693.44	10147.08	13693.44	10147.08
Cutoff-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Running Variable Polynomial	1	2	1	2	1	2

Notes: The table presents non-parametric, parametric (2SLS) and reduced form estimates for the effect of older siblings' marginal enrollment in a target program on younger siblings' probability of applying to and enrolling in the same program. The non-parametric specification controls for a linear or quadratic local polynomial of older siblings' application score centered around the target program admission cutoff and for older and younger siblings application years fixed effects. A triangular kernel is used to give more weight to observations around the cutoff. The parametric specification controls for a linear or quadratic polynomial of older siblings' application score centered around the target program admission cutoff. The slope of the running variable is also allowed to change at the cutoff. In addition, the parametric specification controls for target program-year, older siblings application year, and younger siblings application year fixed effects. Optimal bandwidths computed according to Calonico et al. [2014b] are used in non-parametric specifications. Parametric specifications use bandwiths of 20 and 35 for linear and quadratic specifications in the case of Chile; the same figures for Croatia are 80 and 120. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01

Table 3.4: Probability of Applying and Enrolling in Older Sibling's Target University by Siblings' Similarity

	Applies			Enrolls		
	Same Gender	Δ Age	$\|\Delta$ HS GPA$\|$	Same Gender	Δ Age	$\|\Delta$ HS GPA$\|$
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A - Chile						
Older sibling enrolls	0.094***	0.125***	0.170***	0.037***	0.051***	0.091***
	(0.016)	(0.018)	(0.017)	(0.010)	(0.012)	(0.012)
Interaction	0.014	-0.006**	-0.001***	0.013*	-0.002	-0.001***
	(0.012)	(0.002)	(0.0001)	(0.008)	(0.002)	(0.0001)
Observations	73331	73030	71865	73331	73030	71865
Bandwidth	15	15	15	15	15	15
Kleibergen-Paap Wald F Statistic	2719.59	2709.63	2664.69	2719.59	2709.63	2664.69
Panel B - Croatia						
Older sibling's enrollment	0.114***	0.142**	0.195***	0.065***	0.116***	0.117***
	(0.022)	(0.028)	(0.052)	(0.021)	(0.027)	(0.033)
Interaction	-0.007	-0.011	-0.053	0.037*	-0.011	-0.068
	(0.020)	(0.007)	(0.055)	(0.019)	(0.007)	(0.055)
Observations	12950	12950	2588	12950	12950	2588
Optimal bandwidth	80	80	80	80	80	80
Kleibergen-Paap Wald F Statistic	3229.53	3225.45	648.63	3229.53	3225.45	648.63
Cutoff-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Running Variable Polynomial	1	1	1	1	1	1

Notes: The table presents parametric (2SLS) estimates for the effect of older siblings' marginal enrollment in a target university on younger siblings' probability of applying to and enrolling in the same target university. These specifications use the same set of controls and bandwiths used in the parametric specifications described in Table 3.2. Controls also include the main effect of the variable studied in each column. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 3.5: Probability of Applying and Enrolling in Older Sibling's Target Program by Siblings' Similarity

	Applies			Enrolls		
	Same Gender	Δ Age	$\|\Delta$ HS GPA$\|$	Same Gender	Δ Age	$\|\Delta$ HS GPA$\|$
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
Older sibling's enrollment	0.023***	0.032***	0.056***	0.001	0.002	0.012***
	(0.005)	(0.006)	(0.006)	(0.002)	(0.003)	(0.003)
Interaction	0.010***	-0.001	-0.0003***	0.005***	0.0003	-0.0001***
	(0.004)	(0.001)	(0.00002)	(0.002)	(0.0004)	(0.00001)
Observations	136364	135777	133703	136364	135777	133703
Optimal bandwidth	20	20	20	20	20	20
Kleibergen-Paap Wald F Statistic	6933.23	6906.31	6789.42	6933.23	6906.31	6789.42
			Panel B - Croatia			
Older sibling's enrollment	0.026***	0.045***	0.075***	0.007	0.013**	0.053***
	(0.009)	(0.013)	(0.025)	(0.004)	(0.006)	(0.012)
Interaction	0.023**	-0.003	-0.057**	0.013***	-0.0002	-0.048***
	(0.009)	(0.003)	(0.025)	(0.004)	(0.002)	(0.011)
Observations	36757	36757	8567	36757	36757	8567
Optimal bandwidth	80	80	80	80	80	80
Kleibergen-Paap Wald F Statistic	7220.18	7249.86	1567.76	7220.18	7249.86	1567.76
Cutoff-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Running Variable Polynomial	1	1	1	1	1	1

Notes: The table presents parametric (2SLS) estimates for the effect of older siblings' marginal enrollment in a target program on younger siblings' probability of applying to and enrolling in the same target program. These specifications use the same set of controls and bandwiths used in the parametric specifications described in Table 3.3. Controls also include the main effect of the variable studied in each column. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 3.6: Probability of Applying to and Enrolling in Older Sibling's Target University and Program by Older Sibling's Gender

	University Sample				Program Sample			
	Applies		Enrolls		Applies		Enrolls	
	Male	Female	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A - Chile								
Older sibling's enrollment	0.123***	0.061***	0.062***	0.027*	0.042***	0.023***	0.012***	0.001
	(0.023)	(0.023)	(0.016)	(0.015)	(0.008)	(0.007)	(0.004)	(0.003)
Older sibling's enrollment x Female	0.001	0.032*	-0.020*	0.015	-0.019***	0.001	-0.011***	0.000
	(0.017)	(0.017)	(0.012)	(0.011)	(0.006)	(0.005)	(0.003)	(0.002)
Observations	32302	39129	32302	39129	61982	73014	61982	73014
Bandwidth	15	15	15	15	20	20	20	20
Kleibergen-Paap Wald F Statistic	1337.94	1278.86	1337.94	1278.86	3530.69	3310.96	3530.69	3310.96
Outcome mean	0.30	0.31	0.10	0.10	0.06	0.05	0.01	0.01
Panel B - Croatia								
Older sibling's enrollment	0.126***	0.098***	0.080**	0.044	0.069***	0.031**	0.038***	0.006
	(0.037)	(0.031)	(0.035)	(0.029)	(0.018)	(0.013)	(0.009)	(0.006)
Older sibling's enrollment x Female	-0.001	-0.027	-0.014	0.046*	-0.044***	0.007	-0.031***	0.004
	(0.032)	(0.027)	(0.031)	(0.026)	(0.016)	(0.012)	(0.008)	(0.005)
Observations	5008	7545	5008	7545	14203	22239	14203	22239
Bandwidth	80	80	80	80	80	80	80	80
Kleibergen-Paap Wald F Statistic	1405.97	1651.53	1405.97	1651.53	4025.07	3662.68	4025.07	3662.68
Outcome mean	0.56	0.55	0.29	0.28	0.14	0.12	0.03	0.02
Cutoff-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Running Variable Polynomial	1	1	1	1	1	1	1	1

Notes: The table presents parametric (2SLS) estimates for the effect of older siblings' marginal enrollment in a target university or program on the probability that their younger siblings apply to and enroll in the same target university or program. These specifications use the same set of controls and bandwiths used in the parametric specifications described in Tables 3.2 and 3.3. Controls also include a dummy variable indicating if younger sibling is female. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 3.7: Probability of Applying and Enrolling in Older Siblings' Target University and Program by University and Program Quality (Chile)

	Applies			Enrolls		
	1*st* Year Dropout	Employment	Earnings (CLP MM)	1*st* Year Dropout	Employment	Earnings (CLP MM)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A - University						
Older sibling enrolls	0.138***	0.069**	-0.009	0.068***	0.006	-0.031
	(0.017)	(0.029)	(0.057)	(0.011)	(0.019)	(0.038)
Interaction	-0.288**	0.026*	0.137**	-0.194**	0.030***	0.093**
	(0.085)	(0.016)	(0.064)	(0.058)	(0.011)	(0.043)
Observations	68042	69927	70791	68042	69927	70791
Bandwidth	15	15	15	15	15	15
Kleibergen-Paap Wald F Statistic	2388.57	2183.69	2552.83	2388.57	2183.69	2552.83
Panel B - Program						
Older sibling enrolls	0.030***	0.004	-0.006	0.005*	-0.004	-0.004
	(0.006)	(0.009)	(0.018)	(0.003)	(0.004)	(0.009)
Interaction	-0.035	0.019***	0.042**	-0.014	0.006**	0.009
	(0.027)	(0.005)	(0.021)	(0.013)	(0.003)	(0.010)
Observations	120987	129847	131534	120987	129847	131534
Bandwidth	20	20	20	20	20	20
Kleibergen-Paap Wald F Statistic	6016.47	5732.57	6535.15	6016.47	5732.57	6535.15
Cutoff-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Running Variable Polynomial	1	1	1	1	1	1

Notes: The table presents parametric (2SLS) estimates for the effect of older siblings' marginal enrollment in a target university or program on their younger siblings' probability of applying to and enrolling in the same target university or program. These specifications use the same set of controls and bandwiths used in the parametric specifications described in Tables 3.2 and 3.3. Controls also include the main effect of the variable studied in each column. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 3.8: Effect of Older Siblings' Admission to Target University on Younger Siblings' Academic Performance - University Sample

	Admission Exam (1)	Applies (2)	HS GPA (3)	Language (4)	Mathematics (5)	Social Sciences (6)	Natural Sciences (7)	Enrolls (8)
				Panel A - Chile				
Older sibling's enrollment	0.0001	0.0281*	2.6730	0.4852	4.3694	-2.7605	7.2810	0.0127
	(0.0063)	(0.0158)	(4.0182)	(4.2362)	(4.4129)	(8.9849)	(8.7522)	(0.0152)
Observations	73741	73741	73741	73741	73741	73741	73741	73741
Optimal bandwidth	15	15	15	15	15	15	15	15
Kleibergen-Paap Wald F Statistic	5446.00	5446.00	5446.00	5446.00	5446.00	5446.00	5446.00	5446.00
Outcome mean	0.96	0.58	557.13	526.32	534.91	316.71	341.28	0.38
				Panel B - Croatia				
Older sibling's enrollment	-0.0233		-0.1864	-5.2831	-2.3687**			
	(0.0311)		(0.1291)	(3.5691)	(1.1262)			
Observations	4170		4170	4170	4170			
Optimal bandwidth	80		80	80	80			
Kleibergen-Paap Wald F Statistic	2008.20		2008.20	2008.20	2008.20			2008.20
Outcome mean	0.82		3.22	87.68	22.74			
Cutoff-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Running Variable Polynomial	1	1	1	1	1	1	1	1

Notes: The table presents parametric (2SLS) estimates for the effect of older siblings' marginal enrollment in a target university on the younger siblings' probability of taking the admission exam and applying to university (columns 1 and 2). Columns 3 to 7 present the effects on high school GPA and on the different sections of the admission exam. Finally, column 8 presents the effects on university enrollment. These specifications use the same set of controls and bandwiths used in the parametric specifications described in Table 3.2. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 3.9: Effect of Older Siblings' Admission to Target Program on Younger Siblings' Academic Performance - Program Sample

	Admission Exam (1)	Applies (2)	HS GPA (3)	Language (4)	Mathematics (5)	Social Sciences (6)	Natural Sciences (7)
				Panel A - Chile			
Older sibling's enrollment	0.0024	0.0140	1.4624	3.2740	4.9477*	5.4777	0.1609
	(0.0041)	(0.0100)	(2.5570)	(2.7056)	(2.8092)	(5.6988)	(5.5170)
Observations	136364	136364	136364	136364	136364	136364	136364
Optimal bandwidth	20	20	20	20	20	20	20
Kleibergen-Paap Wald F Statistic	13867.40	13867.40	13867.40	13867.40	13867.40	13867.40	13867.40
Outcome mean	0.96	0.58	556.91	524.23	533.33	314.52	344.03
				Panel B - Croatia			
Older sibling's enrollment	-0.0128		-0.0681	-1.9475	-0.9588		
	(0.0174)		(0.0719)	(2.0214)	(0.6419)		
Observations	12443		12443	12443	12443		
Optimal bandwidth			80	80	80		
Kleibergen-Paap Wald F Statistic	4498.48		4498.48	4498.48	4498.48		
Outcome mean	0.83		3.22	89.12	23.45		
Cutoff-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Running Variable Polynomial	1	1	1	1	1	1	1

Notes: The table presents parametric (2SLS) estimates for the effect of older siblings' marginal enrollment in a target program on the younger siblings' probability of taking the admission exam and applying to university (columns 1 and 2). Columns 3 to 7 present the effects on high school GPA and on the different sections of the admission exam. These specifications use the same set of controls and bandwiths used in the parametric specifications described in Table 3.2. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 3.10: Probability of Applying and Enrolling in Older Sibling's Target Program and University by Difference in Counterfactual Programs' Selectivity

	Same University		Same Program	
	Applies	Enrolls	Applies	Enrolls
	(1)	(2)	(3)	(4)
	Panel A - Chile			
Older sibling's enrollment	0.1081***	0.0438***	0.0283***	0.0049*
	(0.0169)	(0.0112)	(0.0055)	(0.0026)
Interaction	-0.0001	-0.0000	-0.0000	-0.0000
	(0.0003)	(0.0002)	(0.0001)	(0.0000)
Observations	45082	45082	99652	99652
Optimal bandwidth	15	15	20	20
Kleibergen-Paap Wald F Statistic	3153.69	3153.69	7674.01	7674.01
	Panel B - Croatia			
Older sibling's enrollment	0.107***	0.101***	0.034***	0.013***
	(0.021)	(0.020)	(0.009)	(0.004)
Interaction	0.052	0.052	-0.019	0.016
	(0.075)	(0.074)	(0.035)	(0.017)
Observations	10693	10693	34510	34510
Optimal bandwidth	80	80	80	80
Kleibergen-Paap Wald F Statistic	2607.33	2607.33	6854.73	6854.73
Cutoff-Year fixed effects	Yes	Yes	Yes	Yes
Running Variable Polynomial	1	1	1	1

Notes: The table presents parametric (2SLS) estimates for the effect of older siblings' marginal enrollment in their target university and program on younger siblings' probability of applying to and enrolling in the same university and program. The specifications include the same controls and use the same bandwidths described in Tables 3.2 and 3.3. In addition, these specifications control by the difference in the selectivity of older siblings' target program and the chosen program in case of rejection. The effect is allowed to change with this difference in selectivity. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01

Table 3.11: Probability of Applying and Enrolling in Older Siblings' Target University and Program by Older Siblings' Dropout and Dropout Risk

	Applies			Enrolls		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A - Same University						
Older sibling enrolls (OE)	0.123***	0.121***	0.122***	0.064***	0.074***	0.045***
	(0.021)	(0.011)	(0.024)	(0.014)	(0.008)	(0.017)
OE × Older Sibling Drops-out (OD)	-0.111*	-0.069***	-0.047**	-0.090**	-0.044***	-0.038**
	(0.064)	(0.021)	(0.0227)	(0.044)	(0.014)	(0.016)
OE × OD × Applies after OD			-0.044**			-0.011
			(0.017)			(0.012)
Observations	70212	24753	24753	70212	24753	24753
Bandwidth	15	15	15	15	15	15
Kleibergen-Paap Wald F Statistic	2616.67	8968.55	1014.73	2616.67	8968.55	1014.73
Panel B - Same Program						
Older sibling enrolls (OE)	0.021***	0.028***	0.024***	0.004	0.011***	0.007*
	(0.007)	(0.004)	(0.008)	(0.003)	(0.002)	(0.004)
OE × Older Sibling Drops-out (OD)	0.028	-0.015**	-0.007	-0.006	-0.005*	-0.003
	(0.022)	(0.006)	(0.007)	(0.011)	(0.003)	(0.004)
OE × OD × Applies after OD			-0.015**			-0.005
			(0.006)			(0.003)
Observations	130613	49823	49823	130613	49823	49823
Bandwidth	20	20	20	20	20	20
Kleibergen-Paap Wald F Statistic	6645.64	19800.34	2851.33	6645.64	19800.34	2851.33
Cutoff-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Running Variable Polynomial	1	1	1	1	1	1

Notes: The table presents parametric (2SLS) estimates for the effect of older siblings' marginal enrollment in their target university and program on younger siblings' probabilities of applying to and enrolling in the same target program. The controls and bandwidths used in these specifications are the same described in Tables 3.2 and 3.3. In addition, columns (1) and (4) control by older siblings' predicted probability of dropping out from their target program. Columns (2), (3), (5) and (6) control for a dummy variable that indicates if the older sibling actually drops out from the program in which she enrolls. The samples used in these last columns only include individuals whose older siblings enroll in a program. All specifications allow for a differential effect depending on the dropping out probability or actual drop out. Columns (3) and (6) also include an interaction to study how the timing of older siblings' drop out affects younger siblings' application and enrollment decisions. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01

Appendices

3.A Identification Strategy: Further Discussion

This section discusses the assumptions under which our identification strategy provides us with a consistent estimator of the effects of interest. As discussed in Section 3.4.3, a fuzzy RD can be thought as an IV. In what follows, and for ease of notation, we drop time and individual indices t, i, τ and focus our analysis on a specific university-program u. Following this notation, the treatment in which we are interested is:

$$ATE = E[Y_u|O_u = 1] - E[Y_u|O_u = 0],$$

where Y_u is the probability of younger sibling applying to program u, and O_u takes value 1 if the older sibling enrolls in program u and 0 otherwise. In an RD setting, in order to overcome omitted variable bias, we focus only on older siblings who are in a bandwidth bw neighborhood of the university-program u cutoff. For this purpose, denote with adm_u the dummy variable indicating whether older siblings with an application score equal to a_u, were admitted to university-program u with cutoff c_u, and define the following operator:

$$\hat{E}[Y_u] = E[Y_u| \ |a_u - c_u| \leq bw, adm_u \equiv 1_{a_u \geq c_u}].$$

In other words, $\hat{E}$ is an expectation that restricts the sample to older siblings who are around the cutoff c_u and whose risk of assignment is solely determined by the indicator function $1_{a_u \geq c_u}$. Finally, to eliminate concerns related to selection into enrollment, we use adm_u as an instrument for O_u. Denote with I_{jk} a dummy variable that takes value 1 if the younger sibling enrolls in program j when his older sibling enrolls in k, and let's introduce the following notational simplification:

$$R(z) := R|_{Z=z},$$

where $R \in [Y_u, O_u, I_{jk}]$. Introduce now the usual LATE assumptions discussed by Imbens and Angrist [1994], adapted to our setting:

1. Independence of the instrument:

$$\{O_u(1),\ O_u(0),\ I_{jk}(1),\ I_{jk}(0)\} \perp adm_u, \quad \forall j, k$$

2. Exclusion restriction:

$$I_{jk}(1) = I_{jk}(0) = I_{jk}, \quad \forall j, k$$

3. First stage:

$$\hat{E}[O_u(1) - O_u(0)] \neq 0$$

4. Monotonicity:

   (a) Admission weakly increases the likelihood of attending program u

   $$O_u(1) - O_u(0) \geq 0$$

   (b) Admission weakly reduces the likelihood of attending non-offered program $j \neq u$

   $$O_j(1) - O_j(0) \leq 0, \quad \forall j \neq u$$

   In addition to the usual monotonicity assumption that requires that admission to program u cannot discourage students from enrolling in program u, we need to assume an analogous statement affecting other programs $j \neq u$. In particular, we assume that receiving an offer for program u does not encourage enrollment in other program $j \neq u$.

Proposition 1. *Under assumptions* $1-4$*:*

$$\frac{\hat{E}[Y_u|adm_u = 1] - \hat{E}[Y_u|adm_u = 0]}{\hat{E}[O_u|adm_u = 1] - \hat{E}[O_u|adm_u = 0]} = \frac{\sum_{k\neq u} \hat{E}[I_{uu} - I_{uk}|O_u(1) = 1,\ O_k(0) = 1] \times P(O_u(1) = 1,\ O_k(0) = 1)}{P(O_u(1) = 1, O_u(0) = 0)}.$$

Proof. Start with simplifying the first term of the Wald estimator:

$$\begin{aligned}\hat{E}[Y_u|adm_u = 1] &= \hat{E}[Y_u(1) \times adm_u + Y_u(0) \times (1 - adm_u)|adm_u = 1] \quad \text{by assumption 2} \\ &= \hat{E}[Y_u(1)] \quad \text{by assumption 1.}\end{aligned}$$

Applying analogous transformation to all four Wald estimator terms, we obtain:

$$\frac{\hat{E}[Y_u|adm_u = 1] - \hat{E}[Y_u|adm_u = 0]}{\hat{E}[O_u|adm_u = 1] - \hat{E}[O_u|adm_u = 0]} = \frac{\hat{E}[Y_u(1) - Y_u(0)]}{\hat{E}[O_u(1) - O_u(0)]}. \tag{3.2}$$

The numerator of equation 3.2, after applying law of iterated expectations, becomes:

$$\hat{E}[Y_u(1) - Y_u(0)] = \tag{3.3}$$

$$\begin{aligned}&\sum_{k\neq u} \hat{E}[I_{uu} - I_{uk}|O_u(1) = 1,\ O_k(0) = 1] \times P(O_u(1) = 1,\ O_k(0) = 1) \\ -&\sum_{k\neq u} \hat{E}[I_{uu} - I_{uk}|O_u(1) = 0,\ O_u(0) = 1,\ O_k(1) = 1] \\ &\qquad\qquad \times P(O_u(1) = 0,\ O_u(0) = 1,\ O_k(1) = 1) \\ +&\sum_{k\neq u, j\neq u} \hat{E}[I_{uk} - I_{uj}|O_k(1) = 1,\ O_j(0) = 1] \times P(O_k(1) = 1,\ O_j(0) = 1).\end{aligned}$$

Assumption 4.1. implies that there are no defiers, cancelling the second term in the above equation. In addition, assumption 4.2. implies that instrument does not encourage enrollment into program $j \neq u$, cancelling the third term.

Similarly, by the virtue of the assumption 4.1., the denominator of equation 3.2 becomes:

$$\hat{E}[O_u(1) - O_u(0)] = P(O_u(1) = 1, O_u(0) = 0). \tag{3.4}$$

Taken together, 3.3 and 3.4 imply:

$$\frac{\hat{E}[Y_u|adm_u = 1] - \hat{E}[Y_u|adm_u = 0]}{\hat{E}[O_u|Z_u = 1] - \hat{E}[O_u|adm_u = 0]} =$$
$$\frac{\sum_{k \neq u} \hat{E}[I_{uu} - I_{uk}|O_u(1) = 1,\ O_k(0) = 1] \times P(O_u(1) = 1,\ O_k(0) = 1)}{P(O_u(1) = 1, O_u(0) = 0)}.$$

$\square$

As asymptotic 2SLS estimator converges to Wald ratio, we interpret the β_{2SLS} as the local average treatment effect identified through compliers (students enrolled to cutoff program when offered admission).

3.B Robustness Checks

This section investigates if the identification assumptions of our empirical strategy are satisfied. In Section 3.4 we already show that there is no evidence of manipulation of the running variable. Thus, in what follows we check if other variables that could affect individuals' application and enrollment decisions present jumps at the cutoff and if the results are robust to different bandwidths. We continue by performing two types of placebo exercises. In the first we study if similar effects arise when looking at placebo cutoffs (i.e. cutoffs that do not affect older siblings' admission). In the second we analyze if similar effects arise when looking at the effect of the younger sibling enrollment on older siblings decisions.

3.B.1 Discontinuities in Potential Confounders

A first concern in the context of an RD is the existence of other discontinuities around the cutoff that could explain the differences we observe in our outcomes of interest.

Taking advantage of a rich vector of demographic, socioeconomic and academic variables, we study if there is evidence of discontinuities in any of them around the threshold.

Figure 3.B1 summarizes this result. It plots the estimated discontinuities at the cutoff and their 95% confidence intervals. To estimate these discontinuities we controll for a linear polynomial of the running variable which slope is allowed to change at the cutoff. Using the same bandwidths reported for linear specifications in section 3.5 we find no statistically significan jump at the cutoff.

3.B.2 Different Bandwidths

In this section, we study how sensible are our main results to the bandwidth used. Optimal bandwidths try to balance the loss of precision suffered when narrowing the window of data points used to estimate the effect of interest, with the bias generated by using points that are too far from the relevant cutoff.

Figures 3.B2 and 3.B3 show how the estimated coefficients change when reducing the bandwidth used in the estimations. Although as expected the standard errors increase while we reduced the sameple, the coefficients that we obtain are very stabel.

3.B.3 Placebo Exercises

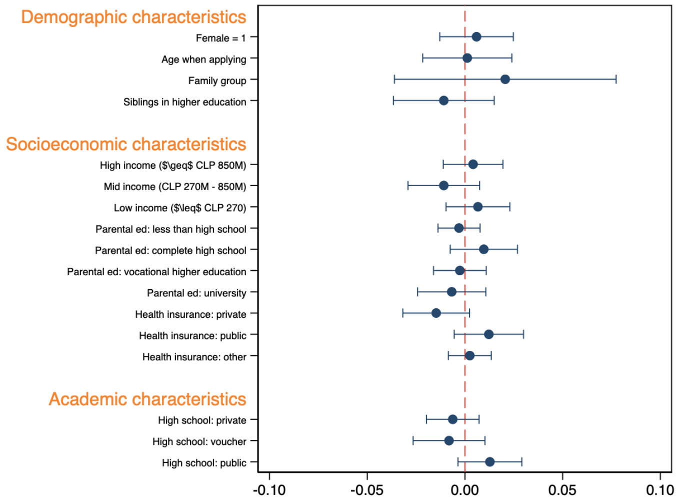
This setting allows us to perform two types of placebo exercises. First, in Figures 3.B4 and 3.B5 we show that only at the real cutoff we observe a discontinuity on younger siblings outcomes This is not surprising since these fake cutoffs do not generate any increase in older siblings' admission. In addition, in Figures 3.B6 and 3.B7 we study if younger siblings' enrollment affect the application decisions of their older siblings. Since younger siblings apply to university after their older siblings, being marginally admitted or rejected from a program or university should not affect what happens with older siblings. These figures show that this is indeed the case.

3.B.4 Alternative Specifications and Total Enrollment

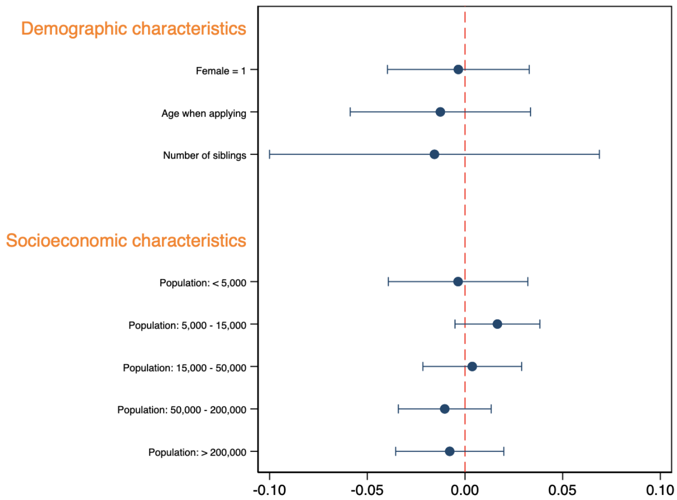
Figures 3.B8 and 3.B9 and in Tables 3.B1, 3.B3, 3.B2 and 3.B4 we study how robust are our estimates to the degree of the local polynomial used, to allow the running variable to have different slopes for each cutoff-program and to re-weighting the observations by the inverse of the total number of applicants in the proximity of each cutoff. The results are quite robust to these changes, and although when reweighing the observations the coefficients are slightly smaller, the general picture remains unchanged.

Finally, Table 3.B5 investigates if marginal admission of older siblings translates into an increase in total enrollment (i.e. enrollment in any university of the system) for them or for their younger siblings. We did not find evidence of extensive margin responses in neither of the countries we study. Thus, according to these results our findings are not driven by this a general increase on younger siblings enrollment. In the case of older siblings, while in Chile we observe a relatively small increase in total enrollment, in Croatia we find a bigger change. This is not surprising because the group of universities studied in Chile is more selective than the ones we study in Croatia. This means that in the case of Chile, older siblings' have still many available universities in case of rejection.

Figure 3.B1: Discontinuities in other Covariates at the Cutoff



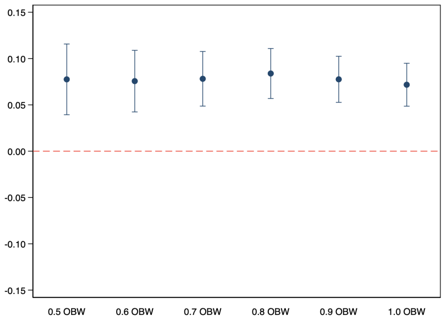
(a) Chile



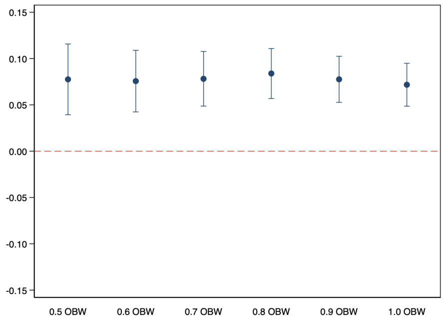
(b) Croatia

Notes: This figure illustrates the estimated jumps at the cutoff for a vector of socioeconomic and demographic characteristics. These estimates come from parametric specifications that control for a linear polynomial of the running variable. As the main specifications, these also include program-year fixed effects. Figure (a) illustrates this for Chile, while figure (b) for Croatia. The points represent the estimated coefficient, while the lines 95% confidence intervals.

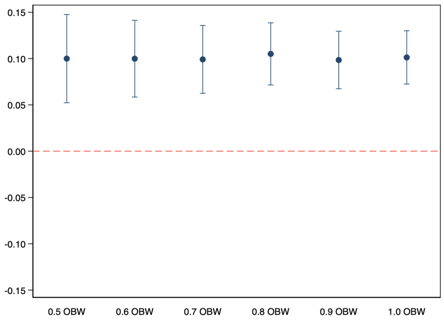
Figure 3.B2: Application and Enrollment Probabilities of Younger Siblings Depending on their Older Siblings' University - Different Bandwidths (P1)



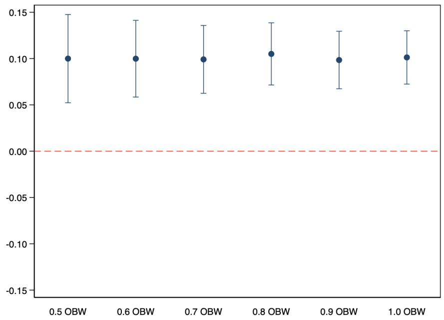
(a) Applies (1st Preference) - Chile



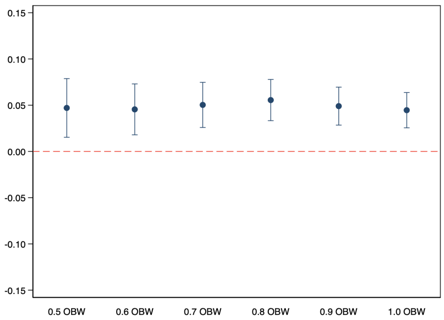
(b) Applies (1st Preference) - Croatia



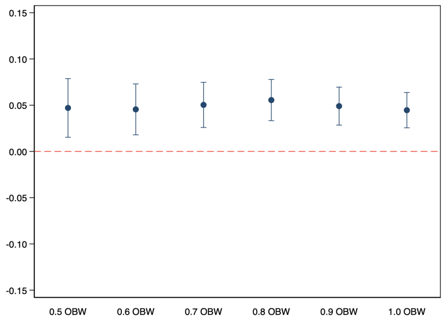
(c) Applies - Chile



(d) Applies - Croatia



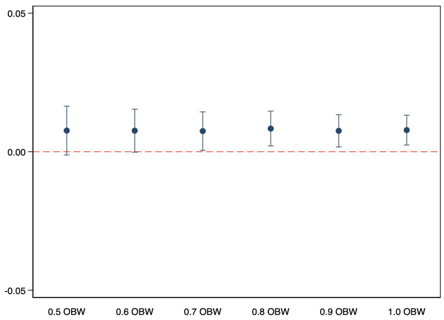
(e) Enrolls - Chile



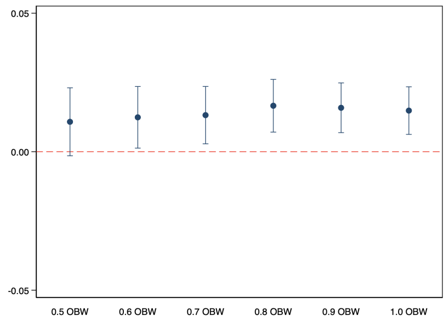
(f) Enrolls - Croatia

Notes: This figure illustrates how being admitted to a specific program changes the probability that younger siblings apply to and enroll in the same university in Chile and Croatia. Figures (a), (c) and (e) illustrate the case of Chile, while figures (b), (d) and (f) the case of Croatia. The coefficients and their confidence intervals come from parametric specifications that control for a linear polynomial of the running variable. Standard errors are clustered at the family level. The bandwiths used to build these figures correspond to multiples of the optimal bandwidths computed following Calonico et al. [2014b].

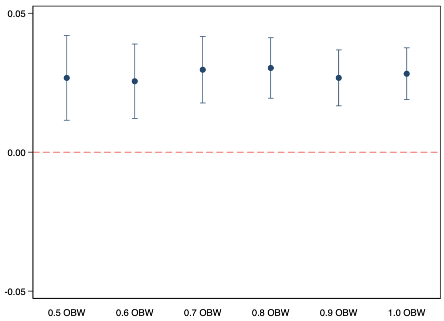
Figure 3.B3: Application and Enrollment Probabilities of Younger Siblings Depending on their Older Siblings' Program - Different Bandwidths (P1)



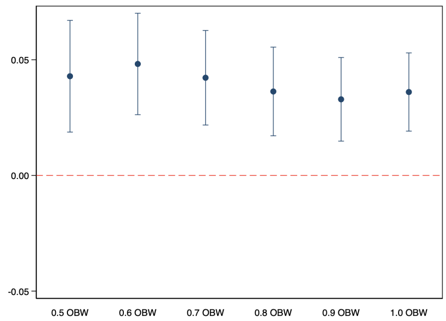
(a) Applies (1st Preference) - Chile



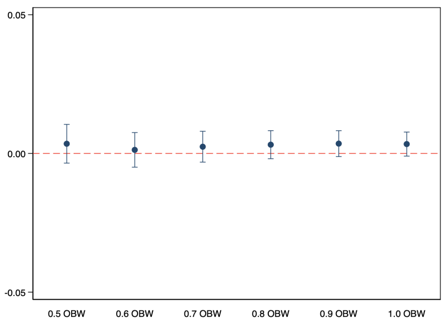
(b) Applies (1st Preference) - Croatia



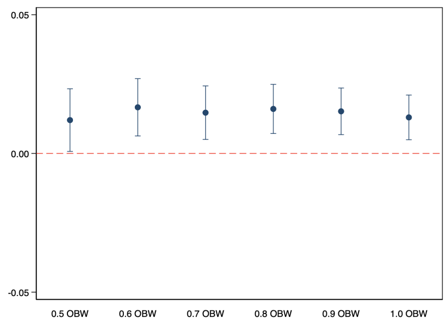
(c) Applies - Chile



(d) Applies - Croatia



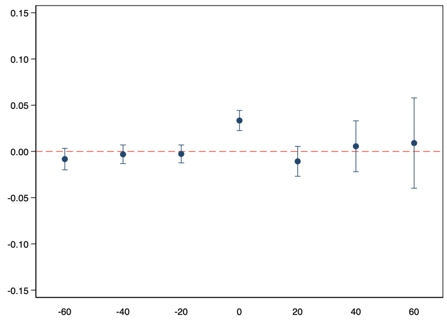
(e) Enrolls - Chile



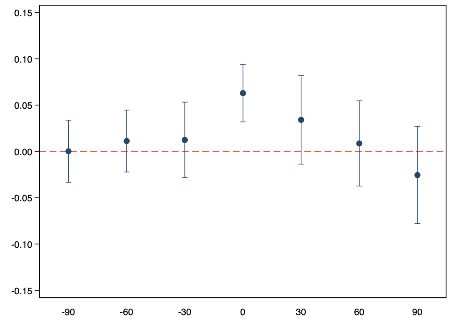
(f) Enrolls - Croatia

Notes: This figure illustrates how being admitted in a specific program changes the probability that younger siblings apply to and enroll in the same program in Chile and Croatia. Figures (a), (c) and (e) illustrate the case of Chile, while figures (b), (d) and (f) the case of Croatia. The coefficients and their confidence intervals come from parametric specifications that control for a linear polynomial of the running variable. Standard errors are clustered at the family level. The bandwiths used to build these figures correspond to multiples of the optimal bandwidths computed following Calonico et al. [2014b].

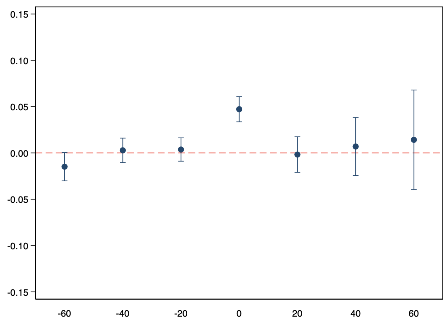
Figure 3.B4: Placebo - Application and Enrollment Probabilities of Younger Siblings Depending on their Older Siblings' University (P1)



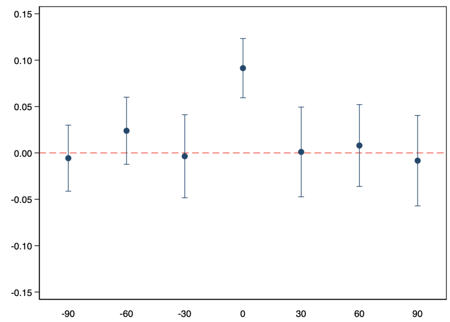
(a) Applies (1st Preference) - Chile



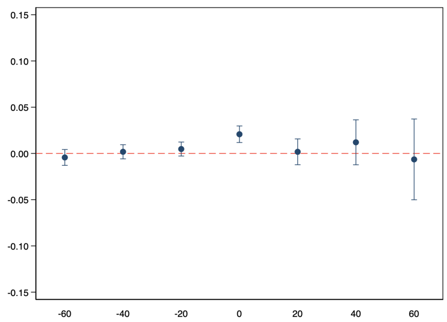
(b) Applies (1st Preference) - Croatia



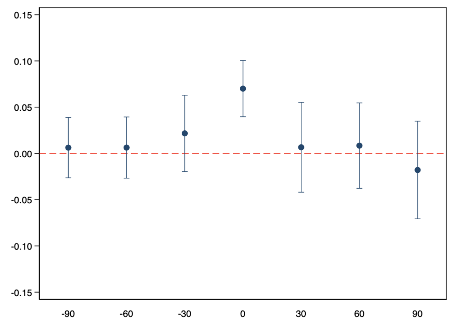
(c) Applies - Chile



(d) Applies - Croatia



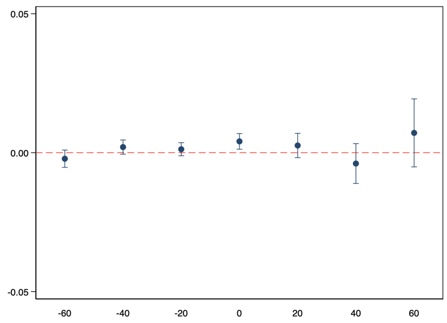
(e) Enrolls - Chile



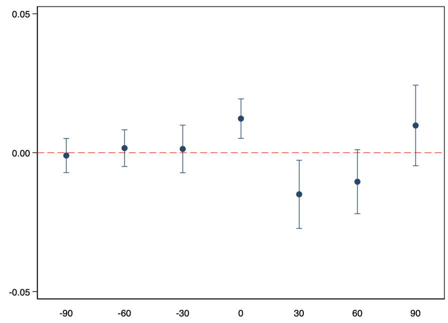
(f) Enrolls - Croatia

Notes: This figure illustrates how being admitted in a specific program changes the probability that younger siblings apply to and enroll in the same university in Chile and Croatia. Figures (a), (c) and (e) illustrate the case of Chile, while figures (b), (d) and (f) the case of Croatia. The coefficients and their confidence intervals come from parametric specifications that control for a linear polynomial of the running variable. Standard errors are clustered at the family level. The bandwiths used to build these figures were computed following Calonico et al. [2014b]. Fake cutoffs are used to illustrate that the discontinuity only arises when using the real cutoff.

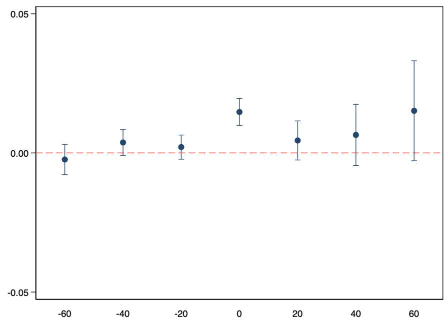
Figure 3.B5: Placebo Cutoffs - Application and Enrollment Probabilities of Younger Siblings depending on their Older Siblings' Program (P1)



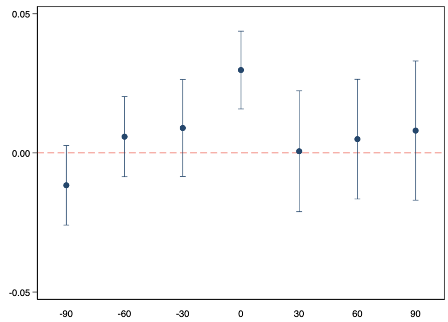
(a) Applies (1st Preference) - Chile



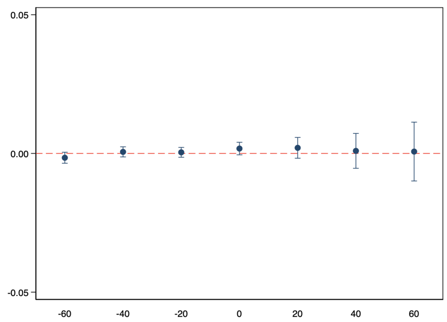
(b) Applies (1st Preference) - Croatia



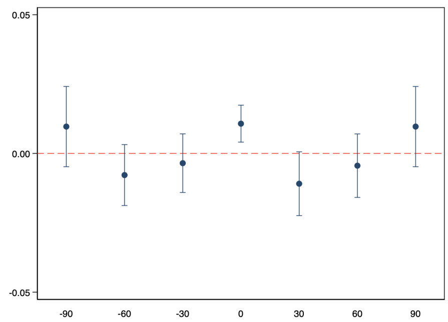
(c) Applies - Chile



(d) Applies - Croatia



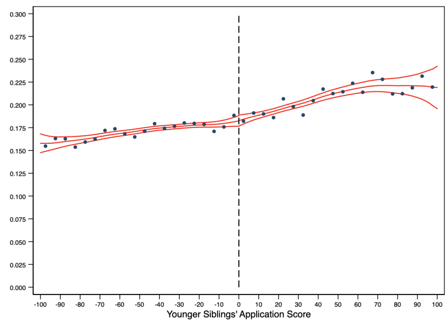
(e) Enrolls - Chile



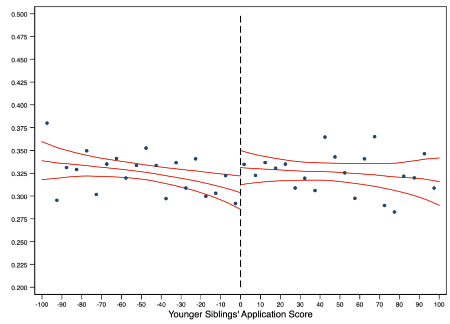
(f) Enrolls - Croatia

Notes: This figure illustrates how being admitted in a specific program changes the probability that younger siblings apply to and enroll in the same program in Chile and Croatia. Figures (a), (c) and (e) illustrate the case of Chile, while figures (b), (d) and (f) the case of Croatia. The coefficients and their confidence intervals come from parametric specifications that control for a linear polynomial of the running variable. Standard errors are clustered at the family level. The bandwiths used to build these figures were computed following Calonico et al. [2014b]. Fake cutoffs are used to illustrate that the discontinuity only arises when using the real cutoff.

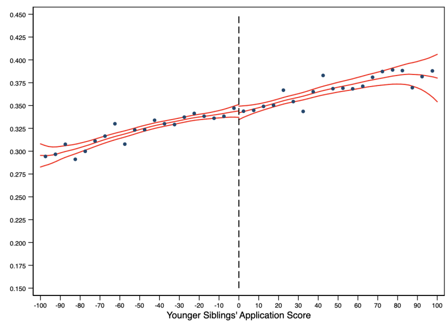
Figure 3.B6: Placebo - Application and Enrollment Probabilities of Older Siblings depending on their Younger Sibling's University (P1)



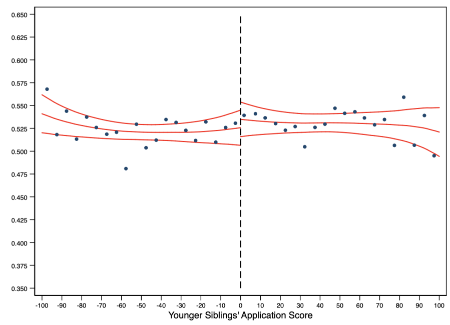
(a) Applies (1st Preference) - Chile



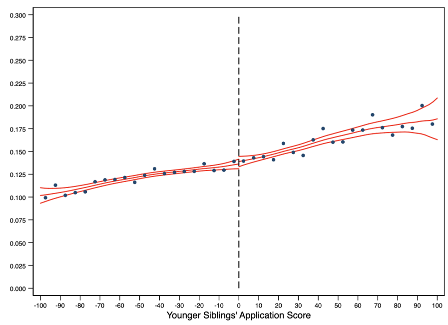
(b) Applies (1st Preference) - Croatia



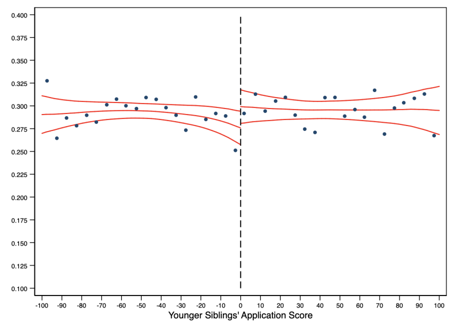
(c) Applies - Chile



(d) Applies - Croatia



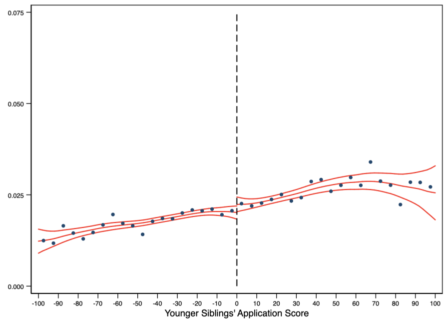
(e) Enrolls - Chile



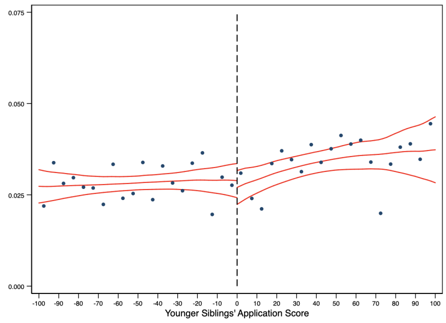
(f) Enrolls - Croatia

Notes: This figure illustrates how younger siblings affect the university to which older siblings apply to and enroll in in Chile and Croatia. Figures (a), (c) and (e) illustrate the case of Chile, while figures (b), (d) and (f) the case of Croatia. Red lines correspond to local polynomials of degree 1 and 95% confidence intervals. In all cases triangular kernels are used. The bandwidths used to build these figures correspond to optimal bandwidths computed following Calonico et al. [2014b] for estimating the discontinuities at the cutoff. Blue dots represent sample means of the dependent variable for bins of width 5.

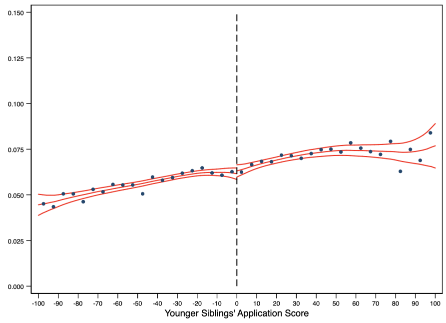
Figure 3.B7: Placebo - Application and Enrollment Probabilities of Older Siblings Depending on their Younger Siblings' Program (P1)



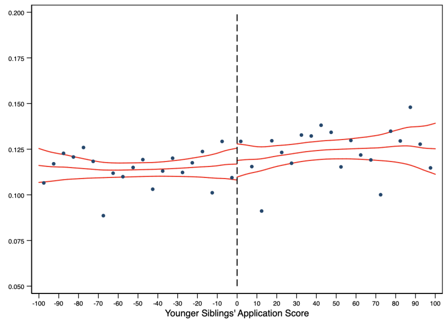
(a) Applies (1st Preference) - Chile



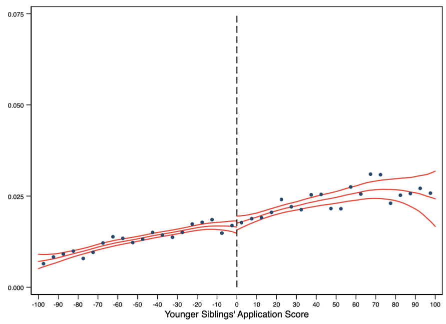
(b) Applies (1st Preference) - Croatia



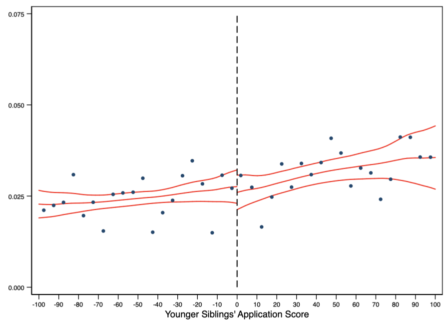
(c) Applies - Chile



(d) Applies - Croatia



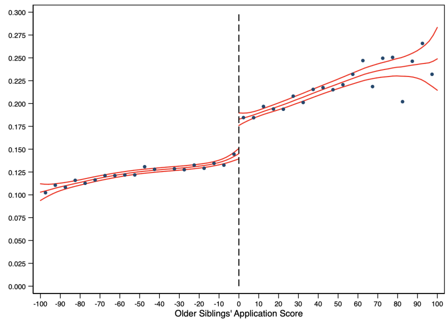
(e) Enrolls - Chile



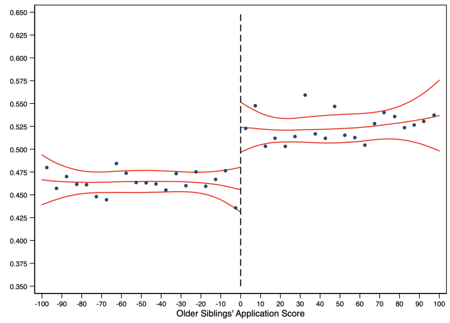
(f) Enrolls - Croatia

Notes: This figure illustrates how younger siblings affect the program to which older siblings apply to and enroll in in Chile and Croatia. Figures (a), (c) and (e) illustrate the case of Chile, while figures (b), (d) and (f) the case of Croatia. Red lines correspond to local polynomials of degree 1 and 95% confidence intervals. In all cases triangular kernels are used. The bandwiths used to build these figures correspond to optimal bandwidths computed following Calonico et al. [2014b] for estimating the discontinuities at the cutoff. Blue dots represent sample means of the dependent variable for bins of width 5.

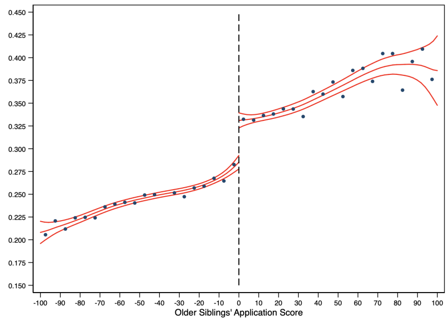
Figure 3.B8: Probability of Applying and Enrolling in Older Sibling's Target University (Reduced Form - P2)



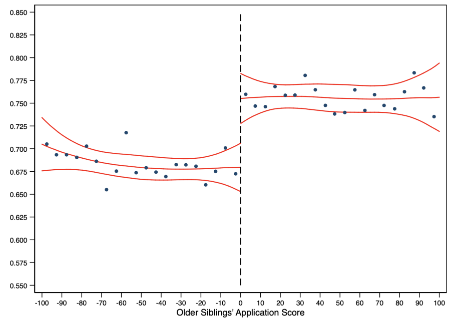
(a) Applies (1st Preference) - Chile



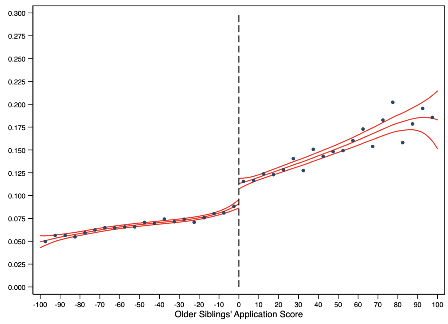
(b) Applies (1st Preference) - Croatia



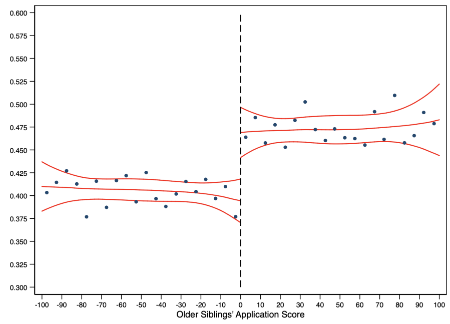
(c) Applies - Chile



(d) Applies - Croatia



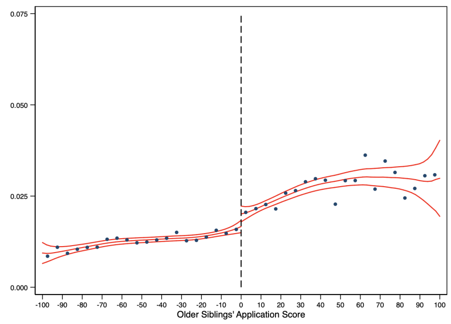
(e) Enrolls - Chile



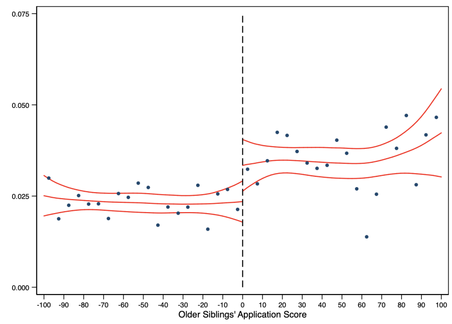
(f) Enrolls - Croatia

Notes: This figure illustrates the probability that younger siblings apply to and enroll in the target university of their older siblings in Chile and Croatia. Figures (a), (c) and (e) illustrate the case of Chile, while figures (b), (d) and (f) the case of Croatia. Red lines correspond to local polynomials of degree 2 and 95% confidence intervals. In all cases triangular kernels are used. The bandwiths used to build these figures correspond to optimal bandwidths computed following Calonico et al. [2014b] for estimating the discontinuities at the cutoff. Blue dots represent sample means of the dependent variable for bins of width 5.

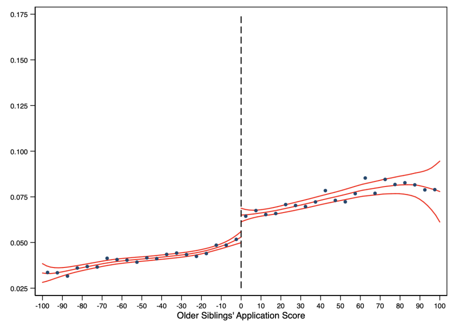
Figure 3.B9: Probability of Applying to and Enrolling in Older Sibling's Target Program (Reduced Form - P2)



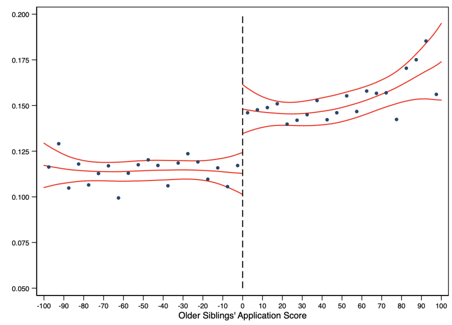
(a) Applies (1st Preference) - Chile



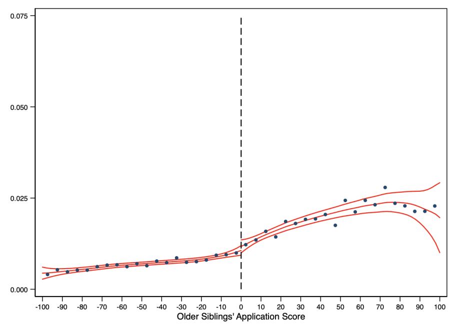
(b) Applies (1st Preference) - Croatia



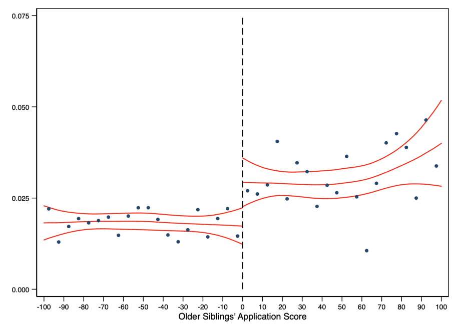
(c) Applies - Chile



(d) Applies - Croatia



(e) Enrolls - Chile



(f) Enrolls - Croatia

Notes: This figure illustrates the probability that younger siblings apply to and enroll in the target program of their older siblings in Chile and Croatia. Figures (a), (c) and (e) illustrate the case of Chile, while figures (b), (d) and (f) the case of Croatia. Red lines correspond to local polynomials of degree 2 and 95% confidence intervals. In all cases triangular kernels are used. The bandwiths used to build these figures correspond to optimal bandwidths computed following Calonico et al. [2014b] for estimating the discontinuities at the cutoff. Blue dots represent sample means of the dependent variable for bins of width 5

Table 3.B1: Probability of Applying to and Enrolling in Older Sibling's Target University (Cutoff-specific Slopes for Running Variable)

	Applies in 1st Preference		Applies in any Preference		Enrolls	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
Local Polynomial	0.093***	0.089***	0.120***	0.116***	0.057***	0.054***
	(0.013)	(0.012)	(0.019)	(0.018)	(0.010)	(0.009)
2SLS	0.076***	0.075***	0.106***	0.092***	0.048***	0.041***
	(0.014)	(0.014)	(0.018)	(0.017)	(0.012)	(0.011)
Reduced Form	0.037***	0.037***	0.052***	0.045***	0.024***	0.020***
	(0.007)	(0.007)	(0.009)	(0.009)	(0.006)	(0.006)
Observations	71447	152301	71447	152301	71447	152301
Outcome mean	0.16	0.16	0.30	0.29	0.10	0.10
Optimal bandwidth	15.75	42.30	13.08	31.40	18.62	47.43
Kleibergen-Paap Wald F Statistic	3858.03	4390.98	3858.03	4390.99	3858.03	4390.99
			Panel B - Croatia			
Local Polynomial	0.080***	0.090**	0.097***	0.092***	0.090***	0.100***
	(0.026)	(0.030)	(0.028)	(0.031)	(0.024)	(0.029)
2SLS	0.080***	0.081***	0.107***	0.115***	0.085***	0.096***
	(0.025)	(0.037)	(0.026)	(0.038)	(0.025)	(0.036)
Reduced Form	0.068***	0.067**	0.090***	0.096***	0.072***	0.080***
	(0.021)	(0.031)	(0.022)	(0.031)	(0.021)	(0.030)
Observations	12526	17312	12526	17312	12526	17312
Outcome mean	0.32	0.32	0.55	0.56	0.29	0.29
Optimal bandwidth	76.24	115.75	79.65	130.98	87.92	114.18
Kleibergen-Paap Wald F Statistic	4019.77	1945.21	4019.77	1945.21	4019.77	1945.21
Running Variable Polynomial	1	2	1	2	1	2

Notes: The table presents non-parametric, parametric (2SLS) and reduced form estimates for the effect of older siblings' marginal enrollment in a target university on younger siblings' probability of applying to and enrolling in the same university. The non-parametric specification controls for a linear or quadratic local polynomial of older siblings' application score centered around the target program admission cutoff and for older and younger siblings application years fixed effects. A triangular kernel is used to give more weight to observations around the cutoff. The parametric specification controls for a linear or quadratic polynomial of older siblings' application score centered around the target program admission cutoff. This polynomial is allowed to have different slopes for each target program; the slope of the running variable is also allowed to change at the cutoff. In addition, the parametric specification controls for target program-year, older siblings application year, and younger siblings application year fixed effects. Optimal bandwidths computed according to Calonico et al. [2014b] are used in non-parametric specifications. Parametric specifications use bandwiths of 15 and 35 for linear and quadratic specifications in the case of Chile; the same figures for Croatia are 80 and 120. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 3.B2: Probability of Applying to and Enrolling in Older Sibling's Target Program (Cutoff-specific slopes for running Variable)

	Applies in 1st Preference		Applies in any Preference		Enrolls	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
Local Polynomial	0.009***	0.010***	0.027***	0.027***	0.003	0.003
	(0.003)	(0.004)	(0.006)	(0.007)	(0.002)	(0.003)
2SLS	0.010***	0.009**	0.029***	0.027***	0.003	0.000
	(0.003)	(0.004)	(0.005)	(0.007)	(0.003)	(0.003)
Reduced Form	0.005***	0.004**	0.016***	0.014***	0.001	0.000
	(0.002)	(0.002)	(0.003)	(0.003)	(0.001)	(0.002)
Observations	135229	214840	135229	214840	135229	214840
Outcome mean	0.02	0.02	0.06	0.05	0.01	0.01
Optimal bandwidth	28.08	38.51	19.11	35.56	32.10	42.68
Kleibergen-Paap Wald F Statistic	11573.41	7965.26	11573.41	7965.26	11573.41	7965.26
			Panel B - Croatia			
Local Polynomial	0.014***	0.011*	0.045***	0.048***	0.016***	0.014***
	(0.006)	(0.006)	(0.012)	(0.014)	(0.005)	(0.006)
2SLS	0.016***	0.016**	0.044***	0.051***	0.014***	0.017***
	(0.005)	(0.007)	(0.010)	(0.013)	(0.005)	(0.006)
Reduced Form	0.013***	0.013***	0.036***	0.042***	0.012***	0.014***
	(0.004)	(0.006)	(0.008)	(0.011)	(0.004)	(0.005)
Observations	36529	48611	36529	48611	36529	48611
Outcome mean	0.03	0.03	0.13	0.13	0.02	0.02
Optimal bandwidth	72.66	118.67	69.04	105.58	86.87	123.04
Kleibergen-Paap Wald F Statistic	12089.16	7917.66	12089.16	7917.66	12089.16	7917.66
Cutoff-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Running Variable Polynomial	1	2	1	2	1	2

Notes: The table presents non-parametric, parametric (2SLS) and reduced form estimates for the effect of older siblings' marginal enrollment in a target program on younger siblings' probability of applying to and enrolling in the same program. The non-parametric specification controls for a linear or quadratic local polynomial of older siblings' application score centered around the target program admission cutoff and for older and younger siblings application years fixed effects. A triangular kernel is used to give more weight to observations around the cutoff. The parametric specification controls for a linear or quadratic polynomial of older siblings' application score centered around the target program admission cutoff. This polynomial is allowed to have different slopes for each target program; the slope of the running variable is also allowed to change at the cutoff. In addition, the parametric specification controls for target program-year, older siblings application year, and younger siblings application year fixed effects. Optimal bandwidths computed according to Calonico et al. [2014b] are used in non-parametric specifications. Parametric specifications use bandwiths of 20 and 35 for linear and quadratic specifications in the case of Chile; the same figures for Croatia are 80 and 120. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01

Table 3.B3: Probability of Applying to and Enrolling in Older Siblings' Target University (Re-weighted)

	Applies in 1st Preference			Applies in any Preference			Enrolls	
	(1)	(2)		(3)	(4)		(5)	(6)
				Panel A - Chile				
2SLS	0.061***	0.067***		0.082***	0.067***		0.032**	0.044***
	(0.016)	(0.018)		(0.020)	(0.022)		(0.014)	(0.015)
Reduced Form	0.025***	0.027***		0.033***	0.027***		0.013**	0.017***
	(0.007)	(0.007)		(0.008)	(0.009)		(0.006)	(0.006)
Observations	73331	152301		73331	152301		73331	152301
Outcome mean	0.16	0.15		0.29	0.29		0.10	0.10
Optimal bandwidth	15	35		15	35		15	35
Kleibergen-Paap Wald F Statistic	2576.80	2319.29		2576.80	2319.29		2576.80	2319.29
				Panel B - Croatia				
2SLS	0.090***	0.085***	0.102***	0.095***	0.087***	0.113***		
	(0.024)	(0.030)	(0.024)	(0.030)	(0.024)	(0.030)		
Reduced Form	0.074***	0.070***	0.084***	0.078***	0.071***	0.093***		
	(0.020)	(0.025)	(0.020)	(0.025)	(0.020)	(0.025)		
Observations	12950	17312		12950	17312		12950	17312
Outcome mean	0.32	0.32		0.55	0.56		0.29	0.29
Optimal bandwidth	80	120		80	120		80	120
Kleibergen-Paap Wald F Statistic	3981.46	2474.69	3981.46	2474.69	3981.46	2474.69		
Cutoff-Year fixed effects	Yes	Yes		Yes	Yes		Yes	Yes
Running Variable Polynomial	1	2		1	2		1	2

Notes: The table presents parametric (2SLS) and reduced form estimates for the effect of older siblings' marginal enrollment in a target university on younger siblings' probability of applying to and enrolling in the same university. The parametric specification controls for a linear and quadratic polynomial of older siblings' application score centered around the target program admission cutoff. The slope of the running variable is allowed to change at the cutoff. In addition, specifications include controls for target program-year, older siblings application year, and younger siblings application year fixed effects. These specifications re-weight observations in order to give equal importance to all the programs in the sample. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 3.B4: Probability of Applying to and Enrolling in Older Sibling's Target Program (Re-weighted)

	Applies in 1st Preference			Applies in any Preference			Enrolls	
	(1)	(2)		(3)	(4)		(5)	(6)
				Panel A - Chile				
2SLS	0.003	0.003		0.024***	0.016*		0.001	0.002
	(0.003)	(0.004)		(0.007)	(0.008)		(0.003)	(0.004)
Reduced Form	0.001	0.001		0.011***	0.007*		0.000	0.001
	(0.002)	(0.002)		(0.003)	(0.004)		(0.001)	(0.002)
Observations	135229	214840		135229	214840		135229	214840
Outcome mean	0.01	0.01		0.05	0.05		0.01	0.01
Optimal bandwidth	20	35		20	35		20	35
Kleibergen-Paap Wald F Statistic	5791.85	3479.05		5791.85	3479.05		5791.85	3479.05
				Panel B - Croatia				
2SLS	0.019***	0.020***	0.026***	0.021*	0.012***	0.013**		
	(0.005)	(0.006)	(0.009)	(0.011)	(0.005)	(0.006)		
Reduced Form	0.015***	0.016***	0.021***	0.017*	0.010***	0.011**		
	(0.004)	(0.005)	(0.007)	(0.009)	(0.004)	(0.005)		
Observations	36529	48611		36529	48611		36529	48611
Outcome mean	0.03	0.03		0.13	0.13		0.02	0.02
Kleibergen-Paap Wald F Statistic	8076.13	5369.30	8076.13	5369.30	8076.13	5369.30		
Cutoff-Year fixed effects	Yes	Yes		Yes	Yes		Yes	Yes
Running Variable Polynomial	1	2		1	2		1	2

Notes: The table presents parametric (2SLS) and reduced form estimates for the effect of older siblings' marginal enrollment in a target program on younger siblings' probability of applying to and enrolling in the same program. The specifications controls for a linear or quadratic polynomial of older siblings' application score centered around the target program admission cutoff. The slope of the running variable is also allowed to change at the cutoff. In addition, specifications controls for target program-year, older siblings application year, and younger siblings application year fixed effects. These specifications re-weight observations in order to give equal importance to all the programs in the sample. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 3.B5: Probability of Enrolling in University Depending on Older Siblings' Admission to Target Program

	Younger Sibling		Older Sibling	
	(1)	(2)	(3)	(4)
	Panel A - Chile			
Local Polynomial	0.002	0.002	0.013***	0.013***
	(0.009)	(0.008)	(0.006)	(0.005)
Reduced Form	-0.002	-0.004	0.017***	0.019***
	(0.006)	(0.006)	(0.004)	(0.004)
Observations	101955	206940	101955	206940
Outcome mean	0.53	0.53	0.93	0.92
Bandwidth	15	35	15	35
	Panel B - Croatia			
Local Polynomial	0.001	0.000	0.141***	0.142***
	(0.010)	(0.012)	(0.010)	(0.010)
Reduced Form	-0.003	0.000	0.123***	0.131***
	(0.007)	(0.008)	(0.007)	(0.008)
Observations	36757	48611	36757	48611
Outcome mean	0.90	0.90	0.88	0.85
Bandwidth	80	120	80	120
Cutoff-Year fixed effects	Yes	Yes	Yes	Yes
Running Variable Polynomial	1	2	1	2

Notes: The table presents non-parametric and parametric estimates for the effect of older siblings' marginal admission in a target program on younger and older siblings' probability of enrolling in any university of the system. The non-parametric specification controls for a linear or quadratic local polynomial of older siblings' application score centered around the target program admission cutoff. While older siblings' application year fixed effects are used in all specifications, younger siblings' application year fixed effects are only used in columns (1) and (2). A triangular kernel is used to give more weight to observations around the cutoff. The parametric specification controls for a linear and quadratic polynomial of older siblings' application score centered around the target program admission cutoff. The slope of the running variable is allowed to change at the cutoff. In addition, the parametric specification controls for target program-year, older siblings' application year, and in the specifications of columns (1) and (2) by younger siblings' application year fixed effects. In the case of Chile, we observe enrollment for all the universities of the system from 2007 onwards. Thus, the sample is adjuster accordingly. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

References

Atila Abdulkadiroğlu, Joshua D Angrist, Susan M Dynarski, Thomas J Kane, and Parag A Pathak. Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots. *The Quarterly Journal of Economics*, 126(2):699–748, 2011.

Atila Abdulkadiroğlu, Joshua D Angrist, Peter D Hull, and Parag A Pathak. Charters Without Lotteries: Testing Takeovers in New Orleans and Boston. *The American Economic Review*, 106(7):1878–1920, 2016.

Jorge M Agüero and Trinidad Beleche. Test-mex: Estimating the Effects of School Year Length on Student Performance in Mexico. *Journal of Development Economics*, 103: 353–361, 2013.

Rita Almeida, Antonio Bresolin, Bruna Pugialli Da Silva Borges, Karen Mendes, and Naercio A Menezes-Filho. Assessing the Impacts of Mais Educacao on Educational Outcomes: Evidence between 2007 and 2011. *World Bank Policy Research Working Paper 7644*, 2016.

Andreas Ammermueller and Jörn-Steffen Pischke. Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study. *Journal of Labor Economics*, 27(3):315–348, jul 2009. ISSN 0734-306X. doi: 10.1086/603650.

Joshua D Angrist and Kevin Lang. Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *American Economic Review*, 94(5):1613–1634, nov 2004. ISSN 0002-8282. doi: 10.1257/0002828043052169.

Joshua D Angrist, Parag A Pathak, and Christopher R Walters. Explaining Charter

School Effectiveness. *American Economic Journal: Applied Economics*, 5(4):1–27, 2013.

Orazio Attanasio and Katja Kaufmann. Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender. *Journal of Development Economics*, 109(C):203–216, 2014.

Esteban M Aucejo and Teresa Foy Romano. Assessing the Effect of School Days and Absences on Test Score Performance. *Economics of Education Review*, 55:70–87, 2016.

David Austen-Smith and Roland G. Fryer. An economic analysis of "acting white". *The Quarterly Journal of Economics*, 120(2):551–583, 2005. ISSN 00335533, 15314650.

Andres Barrios-Fernandez. Should I Stay or Should I go? Neighbors' Effects on University Enrollment. *SSRN*, 2018. URL `http://dx.doi.org/10.2139/ssrn.3278251`.

Erich Battistin and Elena Claudia Meroni. Should we Increase Instruction Time in Low Achieving Schools? Evidence from Southern Italy. *Economics of Education Review*, 55:39–56, Dec 2016.

Patrick L. Baude, Marcus Casey, Eric A. Hanushek, Gregory R. Phelan, and Steven G. Rivkin. The evolution of charter school quality. *Economica*, forthcoming. doi: 10.1111/ecca.12299. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/ecca.12299`.

Cristián Bellei. Does Lengthening the School Day Increase Students' Academic Achievement? Results from a Natural Experiment in Chile. *Economics of Education Review*, 28(5):629–640, 2009.

Matias Berthelon, Diana I Kruger, and Veronica Vienne. Longer School Schedules and Early Reading Skills: Effects from a Full-day School Reform in Chile. *IZA DP n. 10282*, 2016.

Matias E Berthelon and Diana I Kruger. Risky Behavior Among Youth: Incapacitation Effects of School on Adolescent Motherhood and Crime in Chile. *Journal of public economics*, 95(1):41–53, 2011.

E. P. Bettinger, B. T. Long, P. Oreopoulos, and L. Sanbonmatsu. The Role of Application Assistance and Information in College Decisions: Results from the H&R Block Fafsa Experiment. *The Quarterly Journal of Economics*, 127(3):1205–1242, aug 2012. ISSN 0033-5533. doi: 10.1093/qje/qjs017.

Julian R Betts and Eric Johnson. A Test of Diminishing Returns to School Spending. *mimeographed, University of California San Diego*, 1998.

Robert Bifulco, Jason M. Fletcher, Sun Jung Oh, and Stephen L. Ross. Do high school peers have persistent effects on college attainment and other life outcomes? *Labour Economics*, 29:83–90, aug 2014. ISSN 09275371. doi: 10.1016/j.labeco.2014.07.001.

Anders Björklund and Kjell G. Salvanes. Chapter 3 - Education and Family Background: Mechanisms and Policies. volume 3 of *Handbook of the Economics of Education*, pages 201 – 247. Elsevier, 2011. doi: https://doi.org/10.1016/B978-0-444-53429-3.00003-X. URL `http://www.sciencedirect.com/science/article/pii/B978044453429300003X`.

Sandra E. Black, Kalena E. Cortes, and Jane Arnold Lincove. Academic undermatching of high-achieving minority students: Evidence from race-neutral and holistic admissions policies. *American Economic Review*, 105(5):604–10, May 2015. doi: 10.1257/aer.p20151114.

Johanne Boisjoly, Greg J Duncan, Michael Kremer, Dan M Levy, and Jacque Eccles. Empathy or antipathy? the impact of diversity. *American Economic Review*, 96(5): 1890–1905, 2006. ISSN 0002-8282.

Adam S. Booij, Edwin Leuven, and Hessel Oosterbeek. The role of information in the take-up of student loans. *Economics of Education Review*, 31(1):33–44, 2012. ISSN 02727757. doi: 10.1016/j.econedurev.2011.08.009.

Michael A. Boozer and Stephen E. Cacciola. Inside the 'Black Box' of Project STAR: Estimation of Peer Effects Using Experimental Data. Working Papers 832, Economic Growth Center, Yale University, June 2001.

Julia Burdick-Will and Jens Ludwig. Neighborhood and Community Initiatives. In *Targeting Investments in Children: Fighting Poverty When Resources are Limited*, NBER Chapters, pages 303–321. National Bureau of Economic Research, Inc, November 2010.

Mary A. Burke and Tim R. Sass. Classroom Peer Effects and Student Achievement. *Journal of Labor Economics*, 31(1):51–82, 2013. ISSN 0734-306X. doi: 10.1086/666653.

Leonardo Bursztyn and Robert Jensen. How Does Peer Pressure Affect Educational Investments? *The Quarterly Journal of Economics*, 130(3):1329–1367, aug 2015. ISSN 0033-5533. doi: 10.1093/qje/qjv021.

Matias Busso, Taryn Dinkelman, A. Claudia Martínez, and Dario Romero. The effects of financial aid and returns information in selective and less selective schools: Experimental evidence from Chile. *Labour Economics*, 2017. ISSN 09275371. doi: 10.1016/j.labeco.2016.11.001.

Sebastian Calonico, Matias D. Cattaneo, and Rocío Titiunik. Robust data-driven inference in the regression-discontinuity design. *Stata Journal*, 2014a. ISSN 1536867X. doi: 10.1257/jel.48.2.281.

Sebastian Calonico, Matias D. Cattaneo, and Rocio Titiunik. Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, 2014b. ISSN 00129682. doi: 10.3982/ECTA11757.

Sebastian Calonico, Matias D Cattaneo, Max H Farrell, Rocío Titiunik, Stephane Bonhomme, Ivan Canay, David Drukker, Kosuke Imai, Michael Jansson, Lutz Kilian, Pat Kline, Xinwei Ma, Andres Santos, and Gonzalo Vazquez. Regression Discontinuity Designs Using Covariates. 2017.

Sebastian Calonico, Matias D Cattaneo, Max H Farrell, and Rocio Titiunik. Regression Discontinuity Designs using Covariates. *Review of Economics and Statistics*, (0), 2018.

David Card and Alan B Krueger. Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. *Journal of Political Economy*, 100(1):1–40, 1992.

Scott Carrell and Bruce Sacerdote. Why do college-going interventions work? *American Economic Journal: Applied Economics*, 9(3):124–51, July 2017. doi: 10.1257/app.20150530. URL `http://www.aeaweb.org/articles?id=10.1257/app.20150530`.

Scott E. Carrell, Mark Hoekstra, and Elira Kuka. The long-run effects of disruptive peers. *American Economic Review*, 108(11):3377–3415, November 2018. doi: 10.1257/aer.20160763.

Scott E. Carrell, Richard L. Fullerton, and James E. West. Does Your Cohort Matter? Measuring Peer Effects in College Achievement. *Journal of Labor Economics*, 27(3):439–464, 2009. ISSN 0734-306X. doi: 10.1086/600143.

Anne Case and Lawrence Katz. The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths. Technical report, National Bureau of Economic Research, Cambridge, MA, may 1991.

Benjamin L. Castleman and Lindsay C. Page. Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates? *Journal of Economic Behavior & Organization*, 115:144–160, jul 2015. ISSN 01672681. doi: 10.1016/j.jebo.2014.12.008.

Maria A Cattaneo, Chantal Oggenfuss, and Stefan C Wolter. The More, the Better? The Impact of Instructional Time on Student Performance. *Education Economics*, pages 1–13, 2017.

Matias D. Cattaneo, Luke Keele, Rocío Titiunik, and Gonzalo Vazquez-Bare. Interpreting regression discontinuity designs with multiple cutoffs. *The Journal of Politics*, 78(4):1229–1248, 2016. doi: 10.1086/686802. URL `https://doi.org/10.1086/686802`.

Matias D Cattaneo, Michael Jansson, and Xinwei Ma. Manipulation Testing based on Density Discontinuity. *Stata Journal*, 18(1):234–261, 2018a.

Matias D Cattaneo, Michael Jansson, and Xinwei Ma. Manipulation Testing based on Density Discontinuity. *The Stata Journal*, 18(1):234–261, 2018b.

Pedro Cerdan-Infantes and Christel Vermeersch. More Time is Better: An Evaluation of the Full Time School Program in Uruguay. *World Bank Policy Research Working Paper 4167*, 2007.

Raj Chetty and Nathaniel Hendren. The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects*. *The Quarterly Journal of Economics*, 133(3):1107–1162, aug 2018a. ISSN 0033-5533. doi: 10.1093/qje/qjy007.

Raj Chetty and Nathaniel Hendren. The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates*. *The Quarterly Journal of Economics*, 133(3):1163–1228, aug 2018b. ISSN 0033-5533. doi: 10.1093/qje/qjy006.

Raj Chetty, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. Where is the Land of Opportunity? The Geopgraphy of Intergenerational Mobility in the United States. *Quarterly Journal of Economics*, 129(4):1553–1623, 2014. ISSN 0033-5533. doi: 10.1093/qje/qju022.Advance.

Raj Chetty, Nathaniel Hendren, and Lawrence F. Katz. The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. *American Economic Review*, 106(4):855–902, apr 2016. ISSN 0002-8282. doi: 10.1257/aer.20150572.

James S. Coleman. Equality of Educational Opportunity Study. sep 1966. ISSN 1066-5684. doi: 10.3886/ICPSR06389.v3.

Dante Contreras and Paulina Sepúlveda. Effect of lengthening the school day on mother's labor supply. *The World Bank Economic Review*, 31(3):747–766, 2016.

Gordon B. Dahl, Katrine V. Løken, and Magne Mogstad. Peer effects in program participation. *American Economic Review*, 2014. ISSN 00028282. doi: 10.1257/aer.104.7.2049.

Anna Piil Damm and Christian Dustmann. Does growing up in a high crime neighborhood affect youth criminal behavior? †. *American Economic Review*, 104(6): 1806–1832, 2014. ISSN 0002-8282.

Taryn Dinkelman and Claudia Martínez A. Investing in Schooling In Chile: The Role of Information about Financial Aid for Higher Education. *Review of Economics and Statistics*, 96(2):244–257, may 2014. ISSN 0034-6535. doi: 10.1162/REST_a_00384.

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

Will Dobbie and Roland G Fryer Jr. The Medium-term Impacts of High-achieving Charter Schools. *Journal of Political Economy*, 123(5):985–1037, 2015.

Will Dobbie, Roland G Fryer, and G Fryer Jr. Are High-quality Schools Enough to Increase Achievement Among the Poor? Evidence from the Harlem Children's Zone. *American Economic Journal: Applied Economics*, 3(3):158–187, 2011.

Esther Duflo and Emmanuel Saez. The Role of Information and Social Interactions in Retirement Plan Decisions : Evidence from a Randomized Experiment. *The Quarterly Journal of Economics*, 118(3):815–842, 2003.

Andrew Dustan. Family networks and school choice. *Journal of Development Economics*, 134(June 2017):372–391, sep 2018. ISSN 03043878. doi: 10.1016/j.jdeveco.2018.06.004.

Susan Dynarski. Hope for Whom? Financial Aid for the Middle Class and its Impact on College Attendance. *NBER(National Bureau of Economic Research)*, (7756), 2000.

Susan M Dynarski. Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion. *American Economic Review*, 93(1):279–288, feb 2003. ISSN 0002-8282. doi: 10.1257/000282803321455287.

Eric Eide and Mark H Showalter. The Effect of School Quality on Student Performance: A Quantile Regression Approach. *Economics letters*, 58(3):345–350, 1998.

Andrew Eyles and Stephen Machin. The Introduction of Academy Schools to England's Education. *Journal of the European Economic Association*, 07 2018. ISSN 1542-4766. doi: 10.1093/jeea/jvy021. URL `https://doi.org/10.1093/jeea/jvy021`.

Andrew Eyles, Stephen Machin, and Sandra McNally. Unexpected School Reform: Academisation of Primary Schools in England. *Journal of Public Economics*, 2017.

Jan Feld and Ulf Zölitz. Understanding peer effects: On the nature, estimation, and channels of peer effects. *Journal of Labor Economics*, 35(2):387–428, 2017. doi: 10.1086/689472.

David Figlio, Kristian L. Holden, and Umut Ozek. Do Students Benefit From Longer School Days? Regression Discontinuity Evidence From Florida's Additional Hour of Literacy Instruction. *Economics of Education Review*, 67:171–183, 2018.

Gigi Foster. It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism. *Journal of Public Economics*, 90(8-9):1455–1475, sep 2006. ISSN 00472727. doi: 10.1016/j.jpubeco.2005.12.001.

Jr. Fryer, Roland G. and Lawrence F. Katz. Achieving escape velocity: Neighborhood and school interventions to reduce persistent inequality. *American Economic Review*, 103(3):232–37, May 2013. doi: 10.1257/aer.103.3.232.

Alejandro Gaviria and Steven Raphael. School-based peer effects and juvenile behavior. *Review of Economics and Statistics*, 83(2):257–268, 2001. ISSN 0034-6535.

Stephen Gibbons, Olmo Silva, and Felix Weinhardt. Everybody Needs Good Neighbours? Evidence from Students' Outcomes in England. *The Economic Journal*, 123 (571):831–874, sep 2013. ISSN 00130133. doi: 10.1111/ecoj.12025.

Stephen Gibbons, Olmo Silva, and Felix Weinhardt. Neighbourhood Turnover and Teenage Attainment. *Journal of the European Economic Association*, 15(4):746–783, aug 2017. ISSN 1542-4766. doi: 10.1093/jeea/jvw018.

Joshua Goodman. Flaking out: Student Absences and Snow Days as Disruptions of Instructional Time. *NBER Working Paper Series*, 2014.

Joshua Goodman, Michael Hurwitz, Jonathan Smith, and Julia Fox. The relationship between siblings' college choices: Evidence from one million sat-taking families. *Economics of Education Review*, 48:75 – 85, 2015. ISSN 0272-7757. doi: https://doi.org/10.1016/j.econedurev.2015.05.006.

Dominique Goux and Eric Maurin. Close Neighbours Matter: Neighbourhood Effects on Early Performance at School. *The Economic Journal*, 117(523):1193–1215, oct 2007. ISSN 0013-0133. doi: 10.1111/j.1468-0297.2007.02079.x.

Amanda L Griffith and Donna S Rothstein. Can't get there from here: The decision to apply to a selective college. *Economics of Education Review*, 28(5):620–628, 2009.

Jeff Grogger. Does School Quality Explain the Recent Black/White Wage Trend? *Journal of Labor Economics*, pages 231–253, 1996.

Benjamin Hansen. School Year Length and Student Performance: Quasi-experimental Evidence. *Available at SSRN 2269846*, 2011.

Eric A. Hanushek, John F. Kain, Jacob M. Markman, and Steven G. Rivkin. Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18(5): 527–544, sep 2003. ISSN 0883-7252. doi: 10.1002/jae.741.

Justine Hastings, Christopher Neilson, and Seth Zimmerman. The Effects of Earnings Disclosure on College Enrollment Decisions. Technical report, National Bureau of Economic Research, Cambridge, MA, jun 2015.

Justine S Hastings, Christopher A Neilson, and Seth D Zimmerman. Are some degrees worth more than others? evidence from college admission cutoffs in chile. Working Paper 19241, National Bureau of Economic Research, July 2013. URL `http://www.nber.org/papers/w19241`.

Justine S. Hastings, Christopher A. Neilson, Anely Ramirez, and Seth D. Zimmerman. (Un)informed college and major choice: Evidence from linked survey and administrative data. *Economics of Education Review*, 51:136–151, apr 2016. ISSN 02727757. doi: 10.1016/j.econedurev.2015.06.005.

Diana Hincapie. Do Longer School Days Improve Student Achievement? Evidence from Colombia. *IDB Working Paper n. 679*, 2016.

Caroline Hoxby. Peer Effects in the Classroom: Learning from Gender and Race Variation. NBER Working Papers 7867, National Bureau of Economic Research, Inc, August 2000.

Caroline Hoxby. The Power of Peers: How does the makeup of a classroom influence achievement? *Education next*, 2(2):57–63, jan 2002. ISSN 00295507. doi: 10.1049/ep.1985.0257.

Caroline Hoxby and Christopher Avery. The Missing "One-Offs": The Hidden Supply of High-Achieving, Low-Income Students. *Brookings Papers on Economic Activity*, 2013(1):1–65, 2013. ISSN 1533-4465. doi: 10.1353/eca.2013.0000.

Caroline M Hoxby and Sarah Turner. Informing Students about Their College Options : A Proposal for Broadening the Expanding College Opportunities Project. *The Hamilton Project*, (June), 2013.

Caroline M. Hoxby and Sarah Turner. What High-Achieving Low-Income Students Know About College. *American Economic Review*, 105(5):514–517, 2015. ISSN 0002-8282. doi: 10.1257/aer.p20151027.

Caroline Minter Hoxby and Gretchen Weingarth. Taking race out of the equation: School reassignment and the structure of peer effects. *Mimeograph*, 18(2005):2007, 2005. doi: 10.1.1.75.4661.

Mathias Huebener, Susanne Kuger, and Jan Marcus. Increased Instruction Hours and the Widening Gap in Student Performance. *Labour Economics*, 2017.

Guido W. Imbens and Joshua D. Angrist. Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475, 1994. ISSN 00129682.

Scott A. Imberman, Adriana D. Kugler, and Bruce I. Sacerdote. Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees. *American Economic Review*, 102(5):2048–2082, aug 2012. ISSN 0002-8282. doi: 10.1257/aer.102.5.2048.

Robert Jensen. The (Perceived) Returns to Education and the Demand for Schooling *. *Quarterly Journal of Economics*, 125(2):515–548, may 2010. ISSN 0033-5533. doi: 10.1162/qjec.2010.125.2.515.

Juanna Schrøter Joensen and Helena Skyt Nielsen. Spillovers in education choice. *Journal of Public Economics*, 157(November 2015):158–183, jan 2018. ISSN 00472727. doi: 10.1016/j.jpubeco.2017.10.006.

Jeffrey R. Kling, Jens Ludwig, and Lawrence F. Katz. Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment*. *The Quarterly Journal of Economics*, 120(1):87–130, 2005. doi: 10.1162/0033553053327470.

Jeffrey R Kling, Jeffrey B Liebman, and Lawrence F Katz. Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119, 2007. doi: 10.1111/j.1468-0262.2007.00733.x.

Adam M Lavecchia, Heidi Liu, and Philip Oreopoulos. Behavioral economics of education: Progress and possibilities. In *Handbook of the Economics of Education*, volume 5, pages 1–74. Elsevier, 2016.

Victor Lavy. Do Differences in Schools' Instruction Time Explain International Achievement Gaps? Evidence from Developed and Developing Countries. *The Economic Journal*, 125(588):F397–F424, 2015.

Victor Lavy. Expanding School Resources and Increasing Time on Task: Effects of a Policy Experiment in Israel on Student Academic Achievement and Behavior. *Journal of the European Economic Association.*, forthcoming.

Victor Lavy and Analía Schlosser. Mechanisms and Impacts of Gender Peer Effects at School. *American Economic Journal: Applied Economics*, 3(2):1–33, apr 2011. ISSN 1945-7782. doi: 10.1257/app.3.2.1.

Victor Lavy, M. Daniele Paserman, and Analia Schlosser. Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom. *Economic Journal*, 122(559):208–237, 2012. ISSN 0013-0133.

Jong-Wha Lee and Robert J Barro. Schooling Quality in a Cross–section of Countries. *Economica*, 68(272):465–488, 2001.

Bridget Terry Long. Does the Format of a Financial Aid Program Matter? The Effect of State In-Kind Tuition Subsidies. *Review of Economics and Statistics*, 86(3):767–782, aug 2004. ISSN 0034-6535. doi: 10.1162/0034653041811653.

Jens Ludwig, Greg J. Duncan, Lisa A. Gennetian, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling, and Lisa Sanbonmatsu. Neighborhood effects on the long-term well-being of low-income adults. *Science*, 337(6101):1505–1510, 2012. ISSN 0036-8075. doi: 10.1126/science.1224648.

David S Lyle. Estimating and Interpreting Peer and Role Model Effects from Randomly Assigned Social Groups at West Point. *Review of Economics and Statistics*, 89(2): 289–299, may 2007. ISSN 0034-6535. doi: 10.1162/rest.89.2.289.

Charles F. Manski. Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, 1993. ISSN 00346527. doi: 10.2307/2298123.

Dave E Marcotte. Schooling and Test Scores: A Mother-natural Experiment. *Economics of Education Review*, 26(5):629–640, 2007.

Dave E Marcotte and Steven W Hemelt. Unscheduled School Closings and Student Performance. *Education Finance and Policy*, 3(3):316–338, 2008.

Eric Maurin and Julie Moschion. The Social Multiplier and Labor Market Participation of Mothers. *American Economic Journal: Applied Economics*, 1(1):251–272, jan 2009. ISSN 1945-7782. doi: 10.1257/app.1.1.251.

Silvia Mendolia, Alfredo R Paloyo, and Ian Walker. Heterogeneous effects of high school peers on educational outcomes. *Oxford Economic Papers*, 70(3):613–634, 2018. doi: 10.1093/oep/gpy008.

Elena Claudia Meroni and Giovanni Abbiati. How do Students React To Longer Instruction Time? evidence from Italy. *Education Economics*, 24(6):592–611, Nov 2016.

Fernando Mönckeberg B. Prevention of Undernutrition in Chile Experience Lived by an Actor and Spectator. *Revista Chilena de Nutrición*, 2003.

Toni Mora and Philip Oreopoulos. Peer effects on high school aspirations: Evidence from a sample of close and not-so-close friends. *Economics of Education Review*, 30 (4):575–581, aug 2011. ISSN 02727757. doi: 10.1016/j.econedurev.2011.01.004.

Trang Nguyen. Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar *. 2008.

OECD. Dare to Share: Germany's Experience Promoting Equal Partnership in Families. 2016a.

OECD. How is Learning Time Organised in Primary and Secondary Education? 2016b. doi: "http://dx.doi.org/10.1787/5jm3tqsm1kq5-en". URL `"/content/workingpaper/5jm3tqsm1kq5-en"`.

Philip Oreopoulos and Ryan Dunn. Information and College Access: Evidence from a Randomized Field Experiment. *Scandinavian Journal of Economics*, 115(1):3–26, 2013. ISSN 03470520. doi: 10.1111/j.1467-9442.2012.01742.x.

Philip Oreopoulos and Uros Petronijevic. Making College Worth It: A Review of the Returns to Higher Education. *The Future of Children*, 23(1):41–65, 2013. ISSN 10548289, 15501558. URL `http://www.jstor.org/stable/23409488`.

Rasyad A Parinduri. Do Children Spend Too Much Time in Schools? Evidence from a Longer School Year in Indonesia. *Economics of Education Review*, 41:89–104, 2014.

Erika A Patall, Harris Cooper, and Ashley Batts Allen. Extending the School Day or School Year: A Systematic Review of Research (1985–2009). *Review of Educational Research*, 80(3):401–436, 2010. ISSN 0034-6543.

Jörn-Steffen Pischke. The Impact of Length of the School Year on Student Performance and Earnings: Evidence From the German Short School Years. *The Economic Journal*, 117(523):1216–1242, 2007.

Loreto Reyes, Jorge Rodríguez, and Sergio S Urzúa. Heterogeneous Economic Returns to Postsecondary Degrees: Evidence from Chile. Working Paper 18817, National Bureau of Economic Research, February 2013. URL `http://www.nber.org/papers/w18817`.

Steven G Rivkin and Jeffrey C Schiman. Instruction Time, Classroom Quality, and Academic Achievement. *The Economic Journal*, 125(588):F425–F448, 2015.

Ronald Rizzuto and Paul Wachtel. Further Evidence on the Returns to School Quality. *Journal of Human Resources*, pages 240–254, 1980.

MA Ruz Pérez and Ángela Madrid Valenzuela. Evaluación Jornada Escolar Completa (Resumen Ejecutivo). *Santiago de Chile: Pontificia Universidad Católica de Chile, Dirección de Estudios Sociológicos*, 2005.

Bruce Sacerdote. Peer Effects with Random Assignment : Results for Dartmouth Roommates. *The Quarterly Journal of Economics*, 116(2):681–704, 2001.

Bruce Sacerdote. Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far? In *Handbook of the Economics of Education*, pages 249–277. 2011. ISBN 1574-0692. doi: 10.1016/B978-0-444-53429-3.00004-1.

Bruce Sacerdote. Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward? *Annual Review of Economics*, 6(1):253–272, aug 2014. ISSN 1941-1383. doi: 10.1146/annurev-economics-071813-104217.

Paulo Santiago, Francisco Benavides, Charlotte Danielson, Laura Goe, and Deborah Nusche. *Teacher Evaluation in Chile 2013*. OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, 2013.

Neil S. Seftor and Sarah E Turner. Back to School: Federal Student Aid Policy and Adult College Enrollment. *The Journal of Human Resources*, 37(2):336, 2002. ISSN 0022166X. doi: 10.2307/3069650.

David P. Sims. Strategic Responses to School Accountability Measures: It's all in the Timing. *Economics of Education Review*, 27(1):58–68, 2008. ISSN 0272-7757.

Jonathan Smith, Matea Pender, and Jessica Howell. The full extent of student-college academic undermatch. *Economics of Education Review*, 32:247–261, 2013.

Aaron Sojourner. Identification of Peer Effects with Missing Peer Data: Evidence from Project STAR. *The Economic Journal*, 123(569):574–605, jun 2013. ISSN 00130133. doi: 10.1111/j.1468-0297.2012.02559.x.

Alex Solis. Credit Access and College Enrollment. *Journal of Political Economy*, 125 (2):562–622, apr 2017. ISSN 0022-3808. doi: 10.1086/690829.

Ralph Stinebrickner and Todd R. Stinebrickner. What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds. *Journal of Public Economics*, 90(8-9):1435–1454, sep 2006. ISSN 00472727. doi: 10.1016/j.jpubeco.2006.03.002.

Wilbert van der Klaauw. Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach*. *International Economic Review*, 43(4):1249–1287, nov 2002. ISSN 0020-6598. doi: 10.1111/1468-2354.t01-1-00055.

Matthew Wiswall and Basit Zafar. Determinants of college major choice: Identification using an information experiment. *Review of Economic Studies*, 82(2):791–824, 2013. ISSN 1467937X. doi: 10.1093/restud/rdu044.

Ludger Woessmann. The importance of school systems: Evidence from international differences in student achievement. *Journal of Economic Perspectives*, 30(3):3–32, September 2016. doi: 10.1257/jep.30.3.3. URL `http://www.aeaweb.org/articles?id=10.1257/jep.30.3.3`.

Ludger Wößmann. Schooling Resources, Educational Institutions and Student Performance: The International Evidence. *Oxford bulletin of economics and statistics*, 65 (2):117–170, 2003.

David J. Zimmerman. Peer Effects in Academic Outcomes: Evidence from a Natural Experiment. *Review of Economics and Statistics*, 85(1):9–23, feb 2003. ISSN 0034-6535. doi: 10.1162/003465303762687677.