Should I Stay or Should I Go? Neighbors' Effects on University Enrollment*

Andrés Barrios Fernández†
London School of Economics

JOB MARKET PAPER

November, 2018
(Latest Version)

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 quasi-random variation generated by a discontinuity in 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. Having a close neighbor going to university the year before, increases applicants' enrollment probability by 10 percentage points. This effect is particularly strong in areas where university attendance is low and among individuals who are more likely to interact; 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 their direct beneficiaries. In the case of student loans, I find that in areas where university attendance is low this indirect effect represents more than 15% of their direct effect on enrollment.

Keywords: Neighbors' Effects, University Access.
JEL classification: I21, I24, R23, R28.

*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. I am also grateful to Karun Adusumilli, Diego Battiston, Dita Eckardt, Felix Koenig, Eui Jung Lee, Claudio Schilter 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 Departamento de Evaluación, Medición y Registros Académicos of the Universidad de Chile and the Instituto Nacional de Estadísticas for giving me access to the administrative data I use in this project.

†London School of Economics and Political Science, Department of Economics, Houghton Street, London WC2A 2AE, U.K. E-Mail: m.barrios-fernandez@lse.ac.uk. Web:andresbarriosf.github.io

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 play an important role in schooling decisions (Hoxby & Avery 2013, Carrell & Sacerdote 2017).[2] These studies also show 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, I am among the first looking at how they influence enrollment in higher education.[4] This question is relevant from a policy perspective because these neighbors' effects would imply that programs that expand access to university generate externalities that we should incorporate into the design and

[1]Figure 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 & 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 more generous funding. This undermatching phenomenon has also been studied by Black et al. (2015), Griffith & Rothstein (2009) and Smith et al. (2013). There is also a vast literature looking at the role of information frictions in schooling investment. Attanasio & 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 & Turner (2015) look at them in the United States, and Oreopoulos & Dunn (2013) in Candada. Carrell & 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 & Hendren (2018*a*) that show that exposure to better neighborhoods increases the probability of college enrollment. Burdick-Will & 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.

evaluation of this type of policies. In addition, by addressing this question I contribute to understanding if neighborhoods effects are driven at least in part by exposure to better peers. This in contrast to being driven by exposure to better institutions (i.e. schools, public infrastructure, security).

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). I exploit the discontinuity generated by this cutoff rule and implement a fuzzy RD using potential applicants' enrollment as outcome and instrumenting their neighbors' enrollment with an indicator of eligibility for student loans.

To conduct this analysis, I create a unique dataset combining 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 allow 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 about 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 increase university enrollment of potential applicants. I study how this indirect effect of 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 in enrollment. Neighbors are not the only peers that may affect potential applicants. I 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 for financial aid. I find no effects on their attendance or on their academic performance during high school.

My main results are consistent with two broad classes of mechanisms. First, neighbors may increase university enrollment of potential applicants by providing them 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, I present some suggestive evidence that information is the mechanism behind the observed 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 & Maurin 2007, Gibbons et al. 2013, 2017), and at the family (Goodman et al. 2015, Dustan 2018, Joensen & Nielsen 2018).[6] Others have studied peer effects in higher education (Sacerdote 2001, Zimmerman 2003, Stinebrickner & Stinebrickner 2006, Foster 2006, Lyle 2007, Carrell et al. 2009, Feld & Zölitz 2017). Few of them find sizeable effects when looking at academic performance, . However, as pointed out by Hoxby & Weingarth (2005) and further discussed by Lavy et al. (2012), Burke & 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 & Katz 1991, Gaviria & Raphael 2001, Sacerdote 2001, Duflo & Saez 2003, Boisjoly et al. 2006, Maurin & Moschion 2009, Mora & 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. Using, for instance, the whole class or looking at neighbors' effects using an extensive 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 from taking full advantage of the education opportunities

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

[6] These papers study the relationship 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 & Nielsen (2018) at the subjects taken during high school. Although related to my work, the focus of these studies is on different margins.

that they have available. The hypotheses most commonly studied for explaining 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 indirect effects of these type of programs (i.e. effects of funding on the peers of the direct beneficiaries).

Finally, it adds to the literature on neighborhood effects. We already know 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 & Hendren 2018*a*,*b*).[8] However, from these results we do not know 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

[7]Examples of papers studying liquidity constraints include Dynarski (2000), Seftor & 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 & Martínez A. (2014), Hastings et al. (2015, 2016), Hoxby & Turner (2015), Oreopoulos & Dunn (2013), Wiswall & Zafar (2013), Booij et al. (2012), Nguyen (2008), Castleman & Page (2015). Carrell & 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 & Dustmann (2014), Fryer & 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 & Ludwig (2010) point out, if neighborhood effects are mainly driven by the quality of local institutions, then educational attainment could be improved investing in these institutions without having to move disadvantaged individuals to different areas.

results of the paper to relevant literature. Finally, the eight section concludes.

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.

2.1 Institutions and Inequality in the System

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

[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 E2 in the appendix illustrates university attendance rates for the whole income distribution. According to the same survey, 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.

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), 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, a group of private universities select their students using a centralized 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 processes.[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 available for university studies require students to score above a cutoff in the

[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] More than 90% 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] In my data, 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.

PSU.

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 for the different types of higher education institutions. The allocation of these benefits is managed by the Ministry of Education. This section, briefly describes the programs that fund university degrees, emphasizing the rules that generate the discontinuities which are later exploited in this paper.

Students that need financial aid have to apply for 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]

Solis (2017) documents that eligibility for student loan creates a discrete jump in the probabilities of university enrollment amongst the financial aid applicants. This is 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 also 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 and institutional requirements.

discontinuity that I exploit in this paper, but this time to study the indirect effects generated by university attendance on the neighbors and siblings of the beneficiaries.

The majority of the scholarship programs are allocated following a similar logic; the 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 change the enrollment probability; 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 offered resources.

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.

3 Data

This section presents the sources of the data used in this project and the sample used to study the effects of neighbors on the potential applicants' probabilities of enrolling in university.

[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.

3.1 Data Sources

This paper combines the 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) for each individual in every grade. In addition, the data 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, 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, 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 and their high school GPA, the data contains information on applications to the universities that participate of the centralized admission system (see Section 2 for more details.). This includes 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 used in this paper are satisfied, and to perform some heterogeneity analyses. These registers also include students' addresses and a unique identifier of parents. This information is used to identify neighbors and siblings.[21]

[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. Some of it can be recovered from the secondary and higher education registers.

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

In order to build the sample used to study neighbors' effects, I combine all these datasets. I create in a similar way the sample used to study siblings' effects.[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.

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 was 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 did 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. Thus, 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*

The baseline specifications do not use controls. Observations with missing values are not used when performing heterogeneity analyses.

[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 A in the appendix.

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 usually correspond to subareas within a municipality.[25] They were defined 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 applied some additional restrictions. I only kept 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 dropped applicant-neighbor pairs in which the applicant completes high school before the neighbor. Finally, I also dropped the pairs in which applicants and neighbors are siblings. These restrictions made me lose around one third of my 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-loan. 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

[24]Even in these regions, it was not possible to identify 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 very particular way, this should work against finding effects.

[25]Standard errors are clustered at this level.

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

affect potential applicants.[27]

The first two columns of table 1 present summary statistics for the sample of potential applicants and their closest neighbors. The third column characterizes all the students registered for taking the PSU between 2007 and 2012 in the country.

Potential applicants and their closest neighbors are very similar. The only relevant differences are in the 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, potential applicants look very similar to the rest of the individuals that appear in the registers of the PSU.

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 by individuals in the PSU. This allows me to implement a fuzzy RD instrumenting neighbors' university enrollment (U_n) with an indicator variable

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

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 their 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 would be 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:

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

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 & Angrist (1994), we need to assume that university enrollment of contemporaneous peers does not affects applicants' own university enrollment (read Section B in the appendix for more details).[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. (2014*b*) 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

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

[29] Considering the timing of the application and enrollment process, individuals have limited scope to respond to university enrollment of their contemporaneous peers.

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. (2014*a*) and Calonico et al. (2017).

Section 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. (2018).[30]

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 D also shows that the results are robust to different bandwidths choices and to the exclusion of observations around the cutoff. It also shows that there are no jumps like the ones observed at the student loans eligibility cutoff in other points where there should not be.

5 Results

This section discusses the main findings of the paper. It uses the definitions introduced in Section 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

[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 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.

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.

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 1, instrumenting neighbors' university enrollment with their eligibility for student loans.

Panel (a) of Figure 2 illustrates the first stage of this exercise. It shows that neighbors' probabilities of going to university increase by around 18 percentage points when 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 student loans 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.

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

[31] Section 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.

to estimate the effect of exposure to a close neighbor going to university on potential applicant's university enrollment. Table 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 not to university.[32] 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,[33], alleviating concerns related to this type of violations to the exclusion restriction.

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 4, but replacing university enrollment of the closest neighbor by university enrollment of the n-th closest neighbor.

In practice, I estimate eight independent specifications to study how each one of the

[32] 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.

[33] 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.

eight closest neighbors affect potential applicants' university enrollment. As discussed in Section 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. 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.[34]

Panel A on Figure 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 seem 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

[34] A more detailed discussion on this is presented in Section 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.

adds an interaction between neighbors' university enrollment and distance.

As illustrated in Panel B of Figure 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 an extensive definition of neighborhoods 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 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.[35] 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 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 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.

5.3 Urban Segregation and Inequality in University Enrollment

As discussed in Section 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 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

[35]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 they belong to the same group. Table E2 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. The effects seem to be 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. Effects are also bigger for females.

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[36] 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. In order to study this, I classified the municipalities in my sample in three groups depending on the university attendance rates observed for potential applicants.[37] Then, I estimated direct and indirect effects independently for each one of this regions using the baseline specification discussed in Section 4 and controlling by a linear polynomial of the running variable.

Figure 5 presents the results of this exercise. The top panel shows the first stage estimates, the panel in the middle the reduced form estimates, and the panel at the bottom the results obtained when combining the previous estimates using 2SLS. These last estimates capture the effects of neighbors' enrollment on potential applicants' enrollment.

The pattern illustrated in this figure shows that 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.[38]

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

[36] According to the results discussed in Section 5.2, in the context of neighbors these spillover effects seem to be very local. Section 6.1 studies indirects effects between siblings.

[37] The map in Figure 4 illustrates the geographic distribution of these three groups for Santiago, Chile's capital city.

[38] 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 38% 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.

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%).[39]

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 this student loans increases 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

[39] 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.

20% of the direct effects.[40] 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 neighbors provide to them. I discuss mechanisms in Section 6.1.

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.

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 3, apart from identifying neighbors, my data allows me to identify siblings. I use this data to study how having an older sibling receiving a student-loan and 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

[40] 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 $IE = 0.03 \cdot \frac{1-\beta^5}{1-\beta}$

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 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. These effects may be bigger in the case of siblings because in this case interactions are presumably more intense than in the case of neighbors.[41]

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 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 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 also affect the household budget constraint. This could explain at least some part of the response observe in the younger siblings.

However, it important to consider that student loans only cover a share of the tuition

[41] The samples used to estimate neighbors and siblings are different. This could also be behind the differences documented between neighbors and siblings.

fees.[42] This means that students and their families still need to pay part of the fees, 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.[43]

There could also be scenarios where the student loans could relax the household budget constraint in a more significant way.[44] Thus, it cannot be ruled out that at least some part of the effects found for siblings is driven by changes in household resources.

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 again employ the fuzzy RD used in previous sections, but this time to investigate how these other outcomes change when a close neighbor goes to university.

According to the results presented in table 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; [45] 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, [46] and an important part of the documented improvement on PSU scores is driven by the extensive margin response mentioned earlier.[47]

[42] student loans cover up to the reference tuition fee set by the government. See Section 2 for more details.

[43] 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.

[44] 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.

[45] I only observe applications to universities that use the centralized admission system described in section 2. These are the applications I use as outcome.

[46] 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.

[47] 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

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 university enrollment documented in previous sections is driven by a change in the decision to apply. This is consistent with the results on undermatching discussed by Hoxby & Avery (2013) and Black et al. (2015), suggesting that there are students who despite having the potential to be admitted into university and receive funding are not even trying to go.

7 Discussion

The results presented in this paper show that exposure to close peers going to university increases potential applicants' university 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 and benefits 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 providing 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 & 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

(i.e. they are less likely to have -475 points in the PSU).

information constraints.[48]

An alternative way in which neighbors and siblings could affect potential applicants' enrollment is by affecting the costs and benefits of going to university. This would be the case for instance if they face a social sanction for going to university. Austen-Smith & 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 & 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. These costs and benefits 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.[49]

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 go to college, something that does not seem to be occurring. The results seem more consistent with the transmission of information.

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 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

[48]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.

[49]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.

increases by a similar magnitude to the one documented in this paper. Also in the United States, Carrell & 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,[50] but it might also be more relevant because it comes from someone closer.[51]

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 mechanisms behind my results.

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 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 play 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

[50] 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 & 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.

[51] 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.

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 represent 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 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 gain a better understanding of the mechanisms behind peers effects in this and other settings.

References

Ammermueller, A. & Pischke, J. (2009), 'Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study', *Journal of Labor Economics* **27**(3), 315–348.

Angrist, J. D. & Lang, K. (2004), 'Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program', *American Economic Review* **94**(5), 1613–1634.

Attanasio, O. & Kaufmann, K. (2014), 'Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender', *Journal of Development Economics* **109**(C), 203–216.

Austen-Smith, D. & Fryer, R. G. (2005), 'An economic analysis of "acting white"', *The Quarterly Journal of Economics* **120**(2), 551–583.

Bettinger, E. P., Long, B. T., Oreopoulos, P. & Sanbonmatsu, L. (2012), '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.

Bifulco, R., Fletcher, J. M., Oh, S. J. & Ross, S. L. (2014), 'Do high school peers have persistent effects on college attainment and other life outcomes?', *Labour Economics* **29**, 83–90.

Black, S. E., Cortes, K. E. & Lincove, J. A. (2015), 'Academic undermatching of high-achieving minority students: Evidence from race-neutral and holistic admissions policies', *American Economic Review* **105**(5), 604–10.

Boisjoly, J., Duncan, G. J., Kremer, M., Levy, D. M. & Eccles, J. (2006), 'Empathy or antipathy? the impact of diversity', *American Economic Review* **96**(5), 1890–1905.

Booij, A. S., Leuven, E. & Oosterbeek, H. (2012), 'The role of information in the take-up of student loans', *Economics of Education Review* **31**(1), 33–44.

Boozer, M. A. & Cacciola, S. E. (2001), Inside the 'Black Box' of Project STAR: Estimation of Peer Effects Using Experimental Data, Working Papers 832, Economic Growth Center, Yale University.

Burdick-Will, J. & Ludwig, J. (2010), Neighborhood and Community Initiatives, *in* 'Targeting Investments in Children: Fighting Poverty When Resources are Limited', NBER Chapters, National Bureau of Economic Research, Inc, pp. 303–321.

Burke, M. A. & Sass, T. R. (2013), 'Classroom Peer Effects and Student Achievement', *Journal of Labor Economics* **31**(1), 51–82.

Bursztyn, L. & Jensen, R. (2015), 'How Does Peer Pressure Affect Educational Investments?', *The Quarterly Journal of Economics* **130**(3), 1329–1367.

Busso, M., Dinkelman, T., Claudia Martínez, A. & Romero, D. (2017), 'The effects of financial aid and returns information in selective and less selective schools: Experimental evidence from Chile', *Labour Economics* .

Calonico, S., Cattaneo, M. D., Farrell, M. H., Titiunik, R., Bonhomme, S., Canay, I., Drukker, D., Imai, K., Jansson, M., Kilian, L., Kline, P., Ma, X., Santos, A. & Vazquez, G. (2017), 'Regression Discontinuity Designs Using Covariates'.

Calonico, S., Cattaneo, M. D. & Titiunik, R. (2014*a*), 'Robust data-driven inference in the regression-discontinuity design', *Stata Journal* .

Calonico, S., Cattaneo, M. D. & Titiunik, R. (2014*b*), 'Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs', *Econometrica* .

Carrell, S. E., Hoekstra, M. & Kuka, E. (2018), 'The long-run effects of disruptive peers', *American Economic Review* **108**(11), 3377–3415.

Carrell, S., Fullerton, R. & West, J. (2009), 'Does Your Cohort Matter? Measuring Peer Effects in College Achievement', *Journal of Labor Economics* **27**(3), 439–464.

Carrell, S. & Sacerdote, B. (2017), 'Why do college-going interventions work?', *American Economic Journal: Applied Economics* **9**(3), 124–51.
URL: *http://www.aeaweb.org/articles?id=10.1257/app.20150530*

Case, A. & Katz, L. (1991), The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths, Technical report, National Bureau of Economic Research, Cambridge, MA.

Castleman, B. L. & Page, L. C. (2015), '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.

Cattaneo, M. D., Jansson, M. & Ma, X. (2018), 'Manipulation Testing based on Density Discontinuity', *Stata Journal* **18(1)**, 234–261.

Chetty, R. & Hendren, N. (2018*a*), 'The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects*', *The Quarterly Journal of Economics* **133**(3), 1107–1162.

Chetty, R. & Hendren, N. (2018*b*), 'The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates*', *The Quarterly Journal of Economics* **133**(3), 1163–1228.

Chetty, R., Hendren, N. & Katz, L. F. (2016), 'The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment', *American Economic Review* **106**(4), 855–902.

Chetty, R., Hendren, N., Kline, P. & Saez, E. (2014), 'Where is the Land of Opportunity? The Geopgraphy of Intergenerational Mobility in the United States', *Quarterly Journal of Economics* **129**(4), 1553–1623.

Coleman, J. S. (1966), 'Equality of Educational Opportunity Study'.

Dahl, G. B., Løken, K. V. & Mogstad, M. (2014), 'Peer effects in program participation', *American Economic Review* .

Damm, A. P. & Dustmann, C. (2014), 'Does growing up in a high crime neighborhood affect youth criminal behavior? †', *American Economic Review* **104**(6), 1806–1832.

Dinkelman, T. & Martínez A., C. (2014), 'Investing in Schooling In Chile: The Role of Information about Financial Aid for Higher Education', *Review of Economics and Statistics* **96**(2), 244–257.

Duflo, E. & Saez, E. (2003), '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.

Dustan, A. (2018), 'Family networks and school choice', *Journal of Development Economics* **134**(June 2017), 372–391.

Dynarski, S. (2000), 'Hope for Whom? Financial Aid for the Middle Class and its Impact on College Attendance', *NBER(National Bureau of Economic Research)* (7756).

Dynarski, S. M. (2003), 'Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion', *American Economic Review* **93**(1), 279–288.

Feld, J. & Zölitz, U. (2017), 'Understanding peer effects: On the nature, estimation, and channels of peer effects', *Journal of Labor Economics* **35**(2), 387–428.

Foster, G. (2006), '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.

Fryer, Roland G., J. & Katz, L. F. (2013), 'Achieving escape velocity: Neighborhood and school interventions to reduce persistent inequality', *American Economic Review* **103**(3), 232–37.

Gaviria, A. & Raphael, S. (2001), 'School-based peer effects and juvenile behavior', *Review of Economics and Statistics* **83**(2), 257–268.

Gibbons, S., Silva, O. & Weinhardt, F. (2013), 'Everybody Needs Good Neighbours? Evidence from Students' Outcomes in England', *The Economic Journal* **123**(571), 831–874.

Gibbons, S., Silva, O. & Weinhardt, F. (2017), 'Neighbourhood Turnover and Teenage Attainment', *Journal of the European Economic Association* **15**(4), 746–783.

Goodman, J., Hurwitz, M., Smith, J. & Fox, J. (2015), 'The relationship between siblings' college choices: Evidence from one million sat-taking families', *Economics of Education Review* **48**, 75 – 85.

Goux, D. & Maurin, E. (2007), 'Close Neighbours Matter: Neighbourhood Effects on Early Performance at School', *The Economic Journal* **117**(523), 1193–1215.

Griffith, A. L. & Rothstein, D. S. (2009), 'Can't get there from here: The decision to apply to a selective college', *Economics of Education Review* **28**(5), 620–628.

Hanushek, E. A., Kain, J. F., Markman, J. M. & Rivkin, S. G. (2003), 'Does peer ability affect student achievement?', *Journal of Applied Econometrics* **18**(5), 527–544.

Hastings, J., Neilson, C. & Zimmerman, S. (2015), The Effects of Earnings Disclosure on College Enrollment Decisions, Technical report, National Bureau of Economic Research, Cambridge, MA.

Hastings, J. S., Neilson, C. A., Ramirez, A. & Zimmerman, S. D. (2016), '(Un)informed college and major choice: Evidence from linked survey and administrative data', *Economics of Education Review* **51**, 136–151.

Hoxby, C. (2000), Peer Effects in the Classroom: Learning from Gender and Race Variation, NBER Working Papers 7867, National Bureau of Economic Research, Inc.

Hoxby, C. (2002), 'The Power of Peers: How does the makeup of a classroom influence achievement?', *Education next* **2**(2), 57–63.

Hoxby, C. & Avery, C. (2013), ‘The Missing ”One-Offs”: The Hidden Supply of High-Achieving, Low-Income Students’, *Brookings Papers on Economic Activity* **2013**(1), 1–65.

Hoxby, C. M. & Turner, S. (2013), ‘Informing Students about Their College Options : A Proposal for Broadening the Expanding College Opportunities Project’, *The Hamilton Project* (June).

Hoxby, C. M. & Turner, S. (2015), ‘What High-Achieving Low-Income Students Know About College’, *American Economic Review* **105**(5), 514–517.

Hoxby, C. M. & Weingarth, G. (2005), ‘Taking race out of the equation: School reassignment and the structure of peer effects’, *Mimeograph* **18**(2005), 2007.

Imbens, G. W. & Angrist, J. D. (1994), ‘Identification and estimation of local average treatment effects’, *Econometrica* **62**(2), 467–475.

Imberman, S. A., Kugler, A. D. & Sacerdote, B. I. (2012), ‘Katrina’s Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees’, *American Economic Review* **102**(5), 2048–2082.

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

Joensen, J. S. & Nielsen, H. S. (2018), ‘Spillovers in education choice’, *Journal of Public Economics* **157**(November 2015), 158–183.

Kling, J. R., Liebman, J. B. & Katz, L. F. (2007), ‘Experimental analysis of neighborhood effects’, *Econometrica* **75**(1), 83–119.

Kling, J. R., Ludwig, J. & Katz, L. F. (2005), ‘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.

Lavecchia, A. M., Liu, H. & Oreopoulos, P. (2016), Behavioral economics of education: Progress and possibilities, *in* 'Handbook of the Economics of Education', Vol. 5, Elsevier, pp. 1–74.

Lavy, V., Paserman, M. D. & Schlosser, A. (2012), '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.

Lavy, V. & Schlosser, A. (2011), 'Mechanisms and Impacts of Gender Peer Effects at School', *American Economic Journal: Applied Economics* **3**(2), 1–33.

Long, B. T. (2004), '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.

Ludwig, J., Duncan, G. J., Gennetian, L. A., Katz, L. F., Kessler, R. C., Kling, J. R. & Sanbonmatsu, L. (2012), 'Neighborhood effects on the long-term well-being of low-income adults', *Science* **337**(6101), 1505–1510.

Lyle, D. S. (2007), '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.

Manski, C. F. (1993), 'Identification of Endogenous Social Effects: The Reflection Problem', *The Review of Economic Studies* .

Maurin, E. & Moschion, J. (2009), 'The Social Multiplier and Labor Market Participation of Mothers', *American Economic Journal: Applied Economics* **1**(1), 251–272.

Mendolia, S., Paloyo, A. R. & Walker, I. (2018), 'Heterogeneous effects of high school peers on educational outcomes', *Oxford Economic Papers* **70**(3), 613–634.

Mora, T. & Oreopoulos, P. (2011), '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.

Nguyen, T. (2008), 'Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar *'.

Oreopoulos, P. & Dunn, R. (2013), 'Information and College Access: Evidence from a Randomized Field Experiment', *Scandinavian Journal of Economics* **115**(1), 3–26.

Sacerdote, B. (2001), 'Peer Effects with Random Assignment : Results for Dartmouth Roommates', *The Quarterly Journal of Economics* **116**(2), 681–704.

Seftor, N. S. & Turner, S. E. (2002), 'Back to School: Federal Student Aid Policy and Adult College Enrollment', *The Journal of Human Resources* **37**(2), 336.

Smith, J., Pender, M. & Howell, J. (2013), 'The full extent of student-college academic undermatch', *Economics of Education Review* **32**, 247–261.

Sojourner, A. (2013), 'Identification of Peer Effects with Missing Peer Data: Evidence from Project STAR*', *The Economic Journal* **123**(569), 574–605.

Solis, A. (2017), 'Credit Access and College Enrollment', *Journal of Political Economy* **125**(2), 562–622.

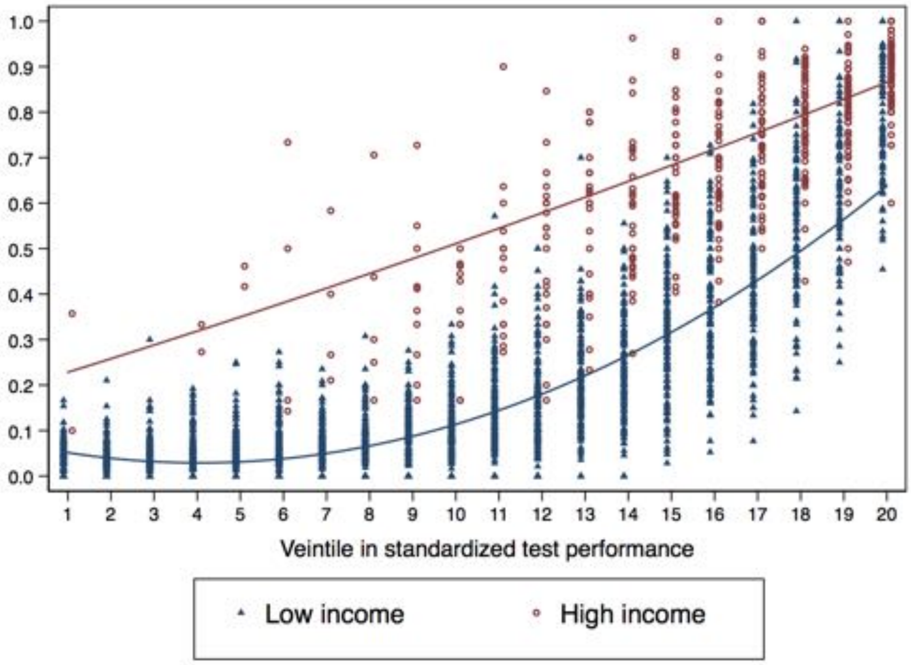
Stinebrickner, R. & Stinebrickner, T. R. (2006), 'What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds', *Journal of Public Economics* **90**(8-9), 1435–1454.

van der Klaauw, W. (2002), 'Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach*', *International Economic Review* **43**(4), 1249–1287.

Wiswall, M. & Zafar, B. (2013), 'Determinants of college major choice: Identification using an information experiment', *Review of Economic Studies* **82**(2), 791–824.

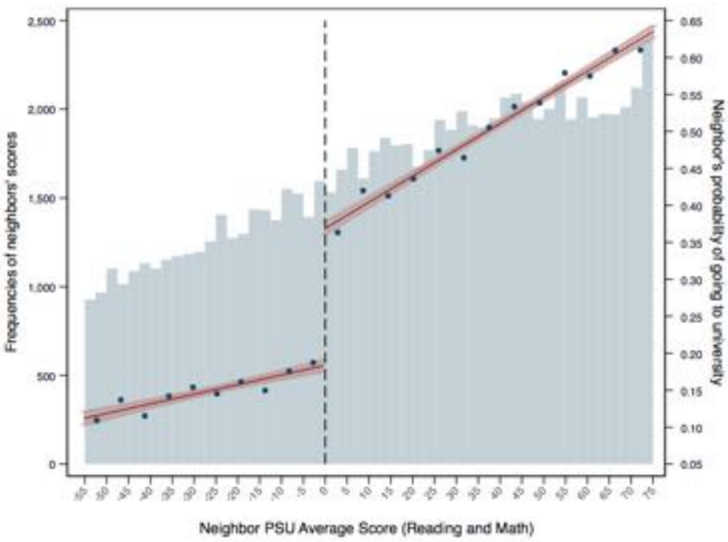
Zimmerman, D. J. (2003), 'Peer Effects in Academic Outcomes: Evidence from a Natural Experiment', *Review of Economics and Statistics* **85**(1), 9–23.

Figure 1: University Enrollment by Household Income, Ability Level and Municipality

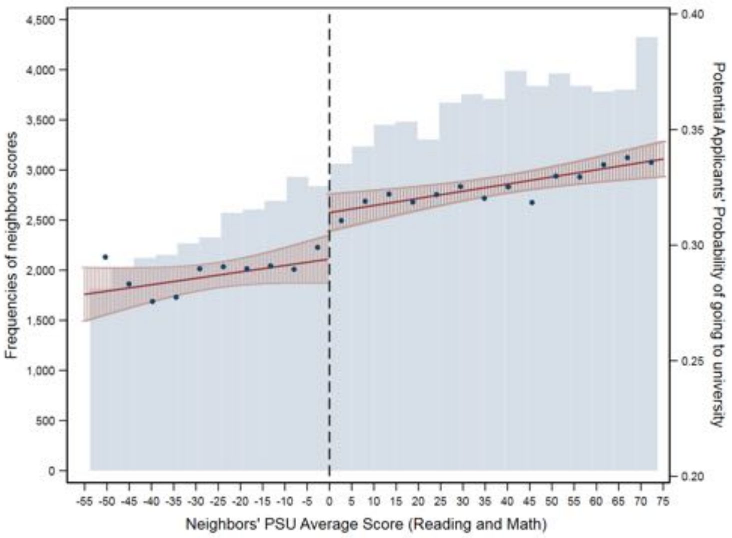


Notes: This figure illustrates the share of low and high income students enrolling in the university by ability level and municipality. Blue triangles represent the shares of low-income students, while red circles represent the shares of high-income students. The figure also presents quadratic fits of 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: First Stage and Reduced Form of Neighbors' RD



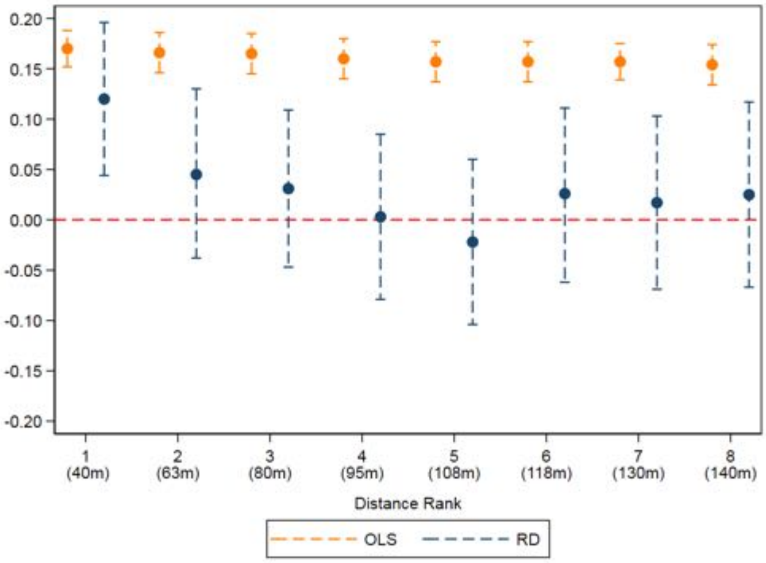
(a) First Stage: Neighbors' Probability of going to University



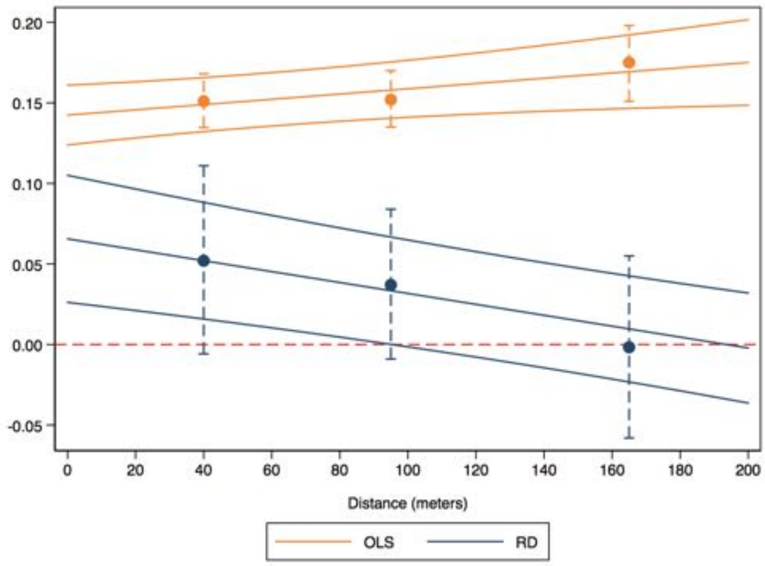
(b) Reduced Form: Potential Applicants' Probability of going to University

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 that were computed following Calonico et al. (2014*b*).

Figure 3: Effect of Neighbors on Potential Applicants by Distance



(a) Evolution of Effects by Distance Rank



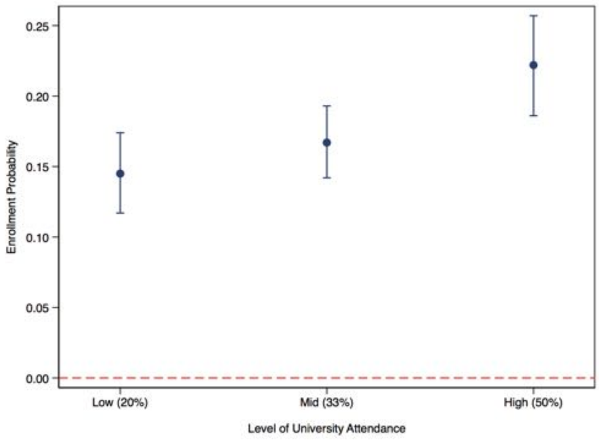
(b) Evolution of Effects by Physical Distance

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. (2014*b*) for the main specification.

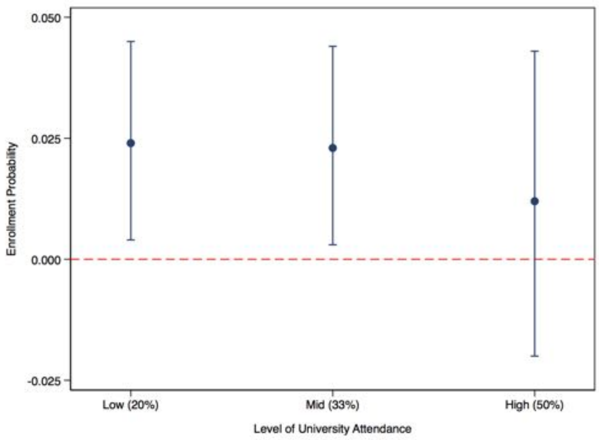
Figure 4: University Attendance across Municipalities in Santiago

Notes: The figure illustrates the share of potential applicants going to university in different municipalities of Santiago between 2007 and 2013. In red areas the average attendance is 20%, in yellow areas 33% and in green areas 50%.

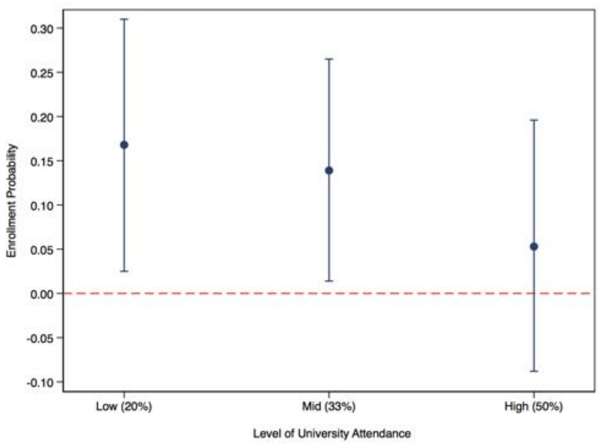
Figure 5: Neighbors' Effects by Municipality Level of Attendance



(a) First Stage



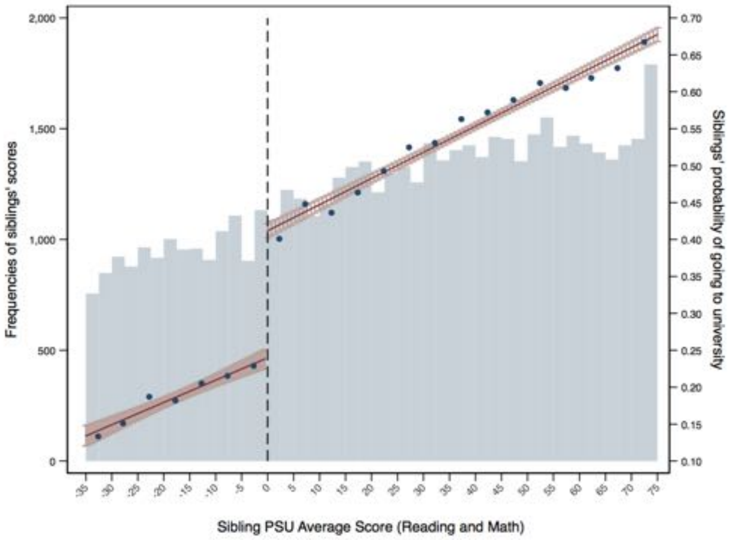
(b) Reduced Form



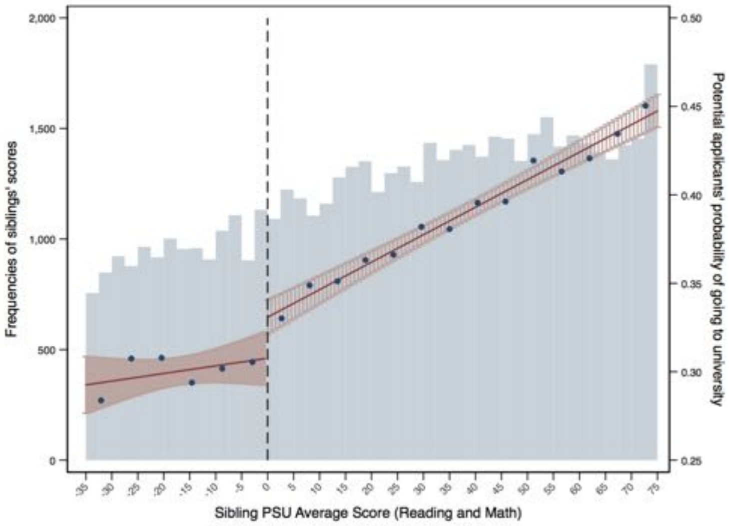
(c) 2SLS Estimates

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 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. (2014*b*).

Table 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: 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***	0.134***	0.118**	0.104*
	(0.037)	(0.047)	(0.053)	(0.062)
Constant	0.240***	0.223***		
	(0.019)	(0.023)		
First stage coefficient	0.178***	0.168***	0.171***	0.172***
	(0.009)	(0.010)	(0.011)	(0.013)
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. (2014*b*). 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 3: Effect of Neighbors on Potential Applicants by Social Distance

	Socioeconomic Status		Gender		Age	
	Same	Different	Same	Different	$\leq$ 2 years	$>$ 2 years
	(1)	(2)	(3)	(4)	(5)	(6)
Neighbor goes to university (t-1)	0.167***	0.044	0.142***	0.095*	0.110**	0.089
	(0.047)	(0.065)	(0.053)	(0.052)	(0.048)	(0.063)
Constant	0.190***	0.315***	0.215***	0.256***	0.303***	0.245***
	(0.024)	(0.034)	(0.027)	(0.027)	(0.051)	(0.034)
First stage coefficient	0.179***	0.169***	0.180***	0.171***	0.167***	0.227***
	(0.011)	(0.011)	(0.011)	(0.010)	(0.009)	(0.017)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N. of potential applicants	49,046	34,261	40,729	46,748	63,696	21,126
Kleibergen-Paap Wald F Statistic	282.52	269.88	279.54	284.83	315.64	175.77

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. (2014*b*). In parenthesis, standard errors clustered at neighborhood unit level. *p-value<0.1 **p-value<0.05 ***p-value<0.01

Table 4: Effect of Neighbors on Potential Applicants by Time at the Neighborhood

	Time at the neighborhood		Neighbors remain-leave		Mother works outside the hh.	
	< 4 years	≥ 4 years	Remain	Leave	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)
Neighbor goes to university (t-1)	0.049	0.141***	0.186***	0.098*	0.169***	0.049
	(0.097)	(0.050)	(0.061)	(0.056)	(0.049)	(0.082)
Constant	0.249***	0.225***	0.203***	0.255***	0.220***	0.265***
	(0.046)	(0.027)	(0.030)	(0.028)	(0.025)	(0.039)
First stage coefficient	0.211***	0.189***	0.179***	0.169***	0.180***	0.171***
	(0.023)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N. of potential applicants	9,056	93,818	30625	36476	48519	18582
Kleibergen-Paap Wald F Statistic	83.68	273.97	203.79	247.61	261.27	102.91

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 study 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 will remain or leave the neighborhood if 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. (2014*b*). In parenthesis, standard errors clustered at neighborhood unit level. *p-value<0.1 **p-value<0.05 ***p-value<0.01

Table 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)	0.138**	0.179**	0.169**	0.197**
	(0.055)	(0.070)	(0.067)	(0.082)
Constant	0.279***	0.199***		
	(0.036)	(0.038)		
First stage coefficient	0.169***	0.154***	0.157***	0.160***
	(0.009)	(0.011)	(0.011)	(0.013)
Year fixed effects	Yes	Yes	Yes	Yes
N. of students	54,142	93,492	54,142	93,492
PSU Polynomial	1	2	1	2
Bandwidth	(35.5-75.5)	(60.5 - 138.5)	(35.5-75.5)	(60.5 - 138.5)
Kleibergen-Paap F statistic	322.17	207.11		

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. (2014*b*). 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 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	0.033	0.096
	(0.061)	(0.063)
High school attendance	0.009	0.010
	(0.001)	(0.014)
PSU Performance	34.970**	33.880*
	(15.405)	(17.362)
	Panel B - Application	
Take PSU	0.065**	0.047*
	(0.024)	(0.028)
Apply to financial aid	0.067	0.139***
	(0.046)	(0.052)
Eligible for financial aid	0.082*	0.092*
	(0.045)	(0.052)
Take up financial aid	0.087**	0.117*
	(0.036)	(0.049)
Apply to CRUCH universities	0.074*	0.185***
	(0.043)	(0.056)
Active application to CRUCH universities	0.074*	0.106*
	(0.042)	(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 and column 2 for siblings. 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. (2014*b*). In parenthesis, standard errors clustered at neighborhood unit level. *p-value<0.1 **p-value<0.05 ***p-value<0.01

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 E1 presents the summary statistics for this sample.

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:

$$U_{at} = \alpha + \sum_{n \in N_a} \beta_{n\tau} U_{n\tau} + \varepsilon_{it} \tag{2}$$

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 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).[52]

This paper focuses on the effects of neighbors going to university one year before potential applicants. To highlight this, equation 2 can be rearranged as follows:

$$U_{at} = \alpha + \beta_{mt-1}U_{mt-1} + \sum_{n \in N_a \setminus U_{mt-T}} \beta_{n\tau}U_{n\tau} + \varepsilon_{it} \tag{3}$$

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:

$$U_{at} = \alpha + \beta_{mt-1}U_{mt-1} + v_{it} \tag{4}$$

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 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.[53]

[52] If younger applicants' decision enter equation 2, instrumenting enrollment of the older neighbor with student-loan eligibility would be enough to solve the reflection problem.

[53] We are already assuming that younger applicants' decision are not part of equation 2.

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, 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-stage regressions in section 5 show that this is indeed the case.[54]

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

[54] 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.

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.[55]

According to Imbens & 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.

C Other Neighbors Definitions

The results discussed on section 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 **??** 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

[55] 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.

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.

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.

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, 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 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. (2018). Figures E3 and E10 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.

Finally, I investigate how sensible are the results to the omission of potential applicants whose older neighbors and siblings obtain a PSU score very close to the student loans eligibility threshold. In the presence of manipulation, we would expect these individuals who are very close to the cutoff to be the ones creating more problems. However, as shown in figure E15 leaving them out of the sample does not change the conclusions of the paper.

Therefore, the results I find do not seem to be driven by manipulation of the running variable.

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 E4 summarizes these results for neighbors, and figure E11 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. (2014*a*). 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 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. [56]

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 E5 and E12 illustrate the results when performing this exercise in the same sample I use in when estimating the main results. Table E3 presents the estimated coefficients of this exercise; table E4 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 5, tables E3 and E4 present the estimated coefficient using two stages least squares and local polynomials. The coefficients are small and never significant.

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

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 E6 presents these results for neighbors and siblings. As can be appreciated, none of these jumps is significant.

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 E7 and E13 presents the estimated c oefficients 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.

D.5 Missing students

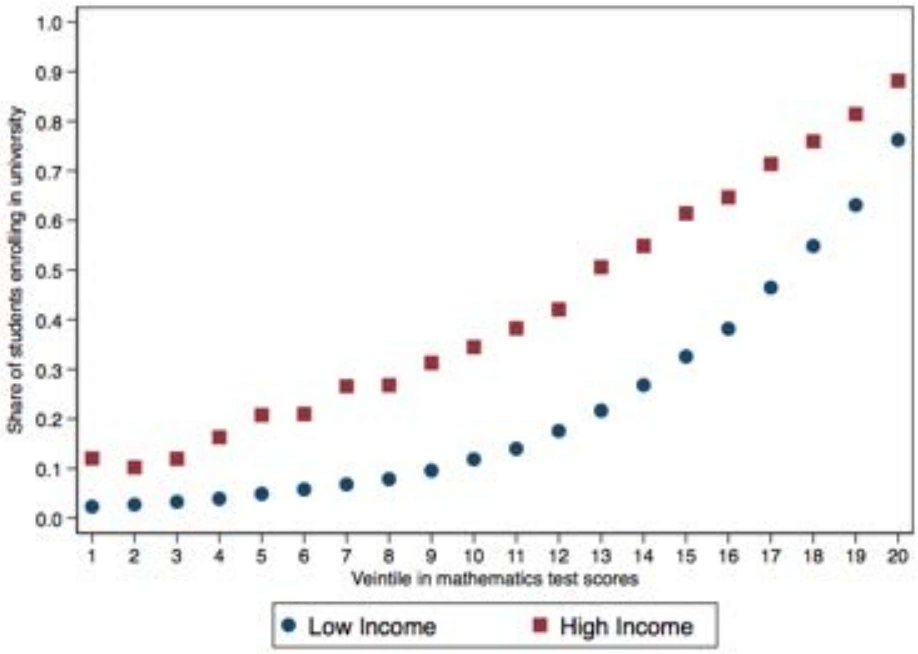
In this section, I discuss how missing information about applicants and their older peers could affect my results. As mentioned in 3, to identify neighbors I rely on the geocoding process of addresses; since the addresses I use do not include postcode, finding them was not always possible. This is especially 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 E5 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.

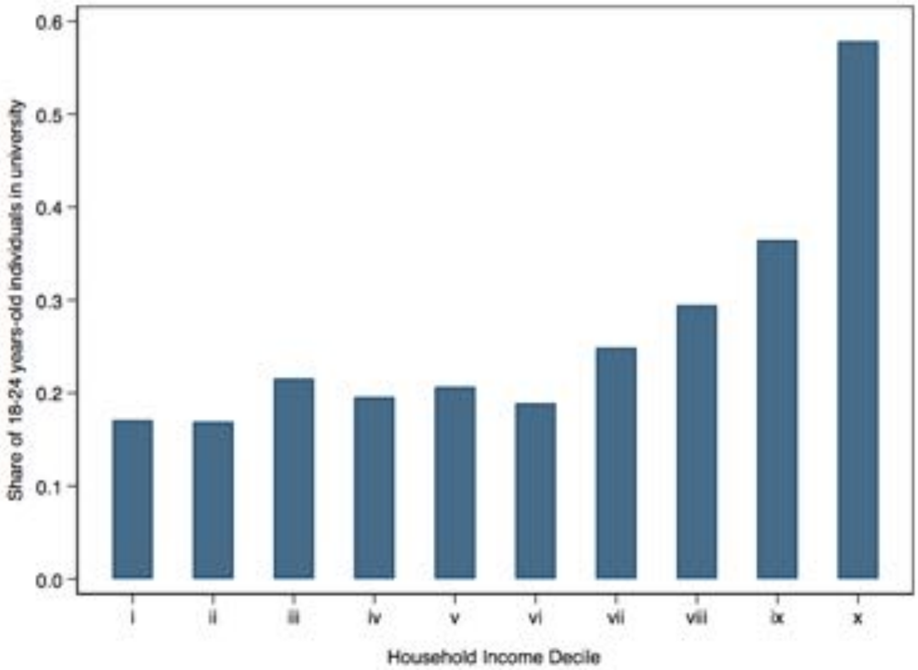
E Additional Figures and Tables

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



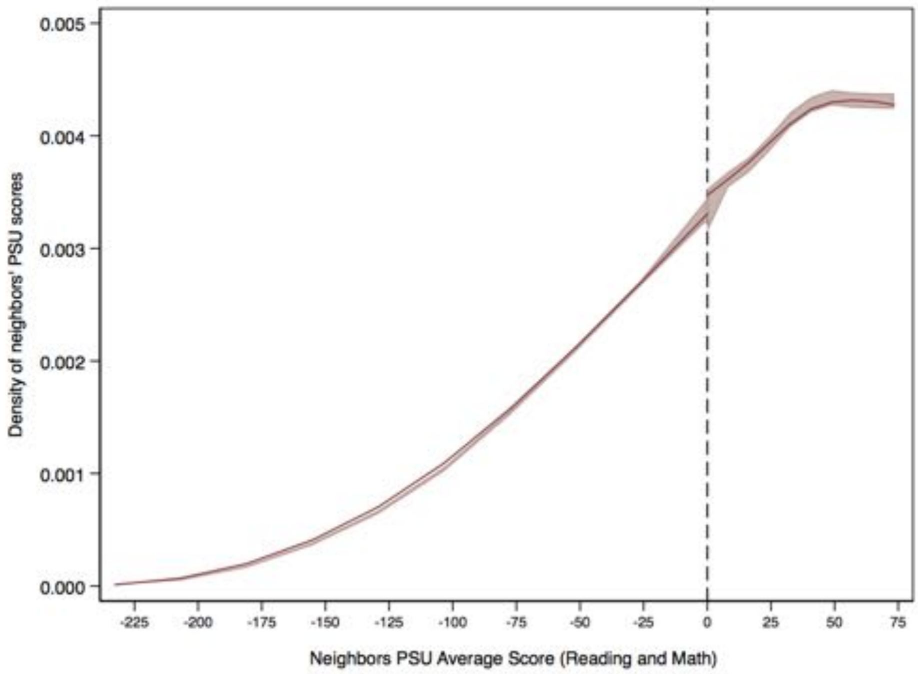
Notes: This figure illustrates how the gap in university enrollment observed across income groups evolves with ability. Ability is measured by students performance in grade 10 mathematics standardized test. University enrollment is measured 3 years later; if students do not repeat or dropout, this is 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 E2: Share of Students going to University vs Household Income (2015)



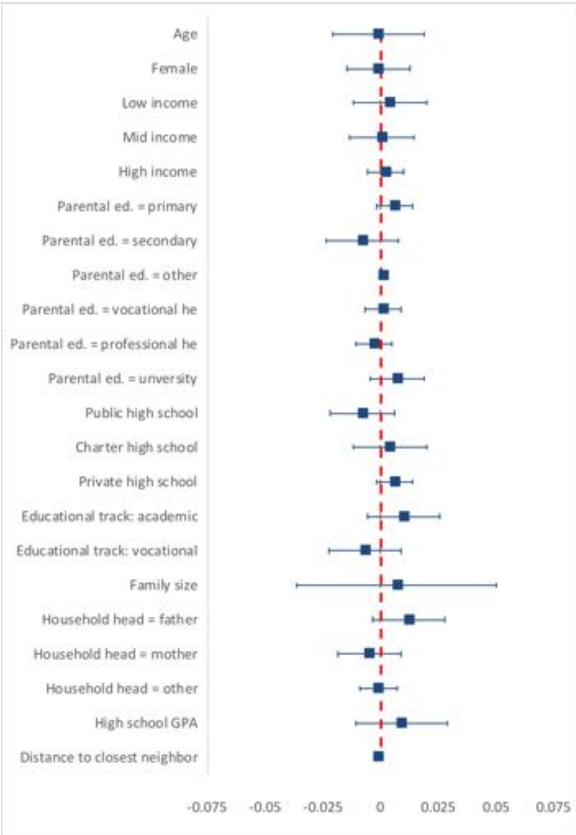
Notes: This figure illustrates the relationship between the share of 18 to 24 years old individuals going to university in 2015 and their household income. It was build using data from the Chilean national household survey, CASEN (http://observatorio.ministeriodesarrollosocial.gob.cl/casen-multidimensional/casen/basedatos.php).

Figure E3: Density of Neighbors' PSU Scores around the Student Loans Eligibility Threshold)

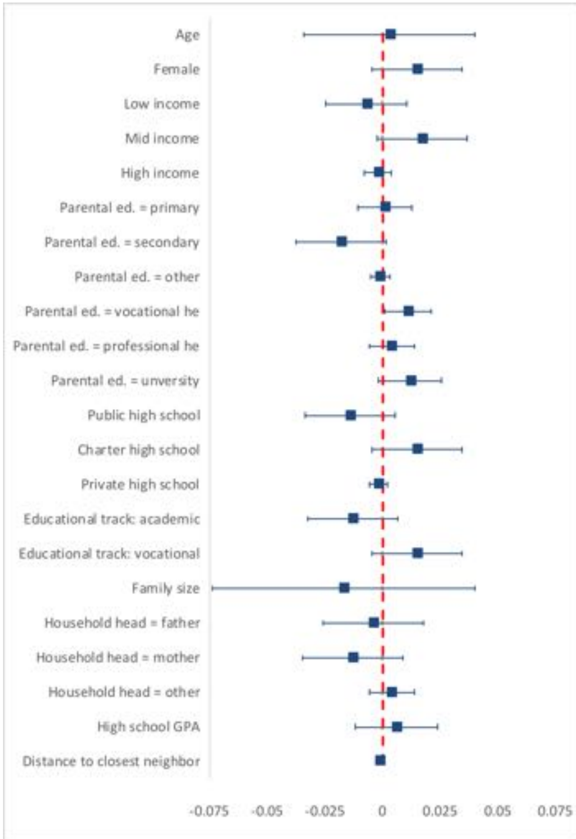


Notes: This figure illustrates the density of neighbors PSU scores around the student loans eligibility thresholds. The density and its confidence intervals on each side of the cutoff were estimated following Cattaneo et al. (2018). This chart complements the formal test they suggest to study discontinuities 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 E4: Discontinuities in other Variables at the Cutoff



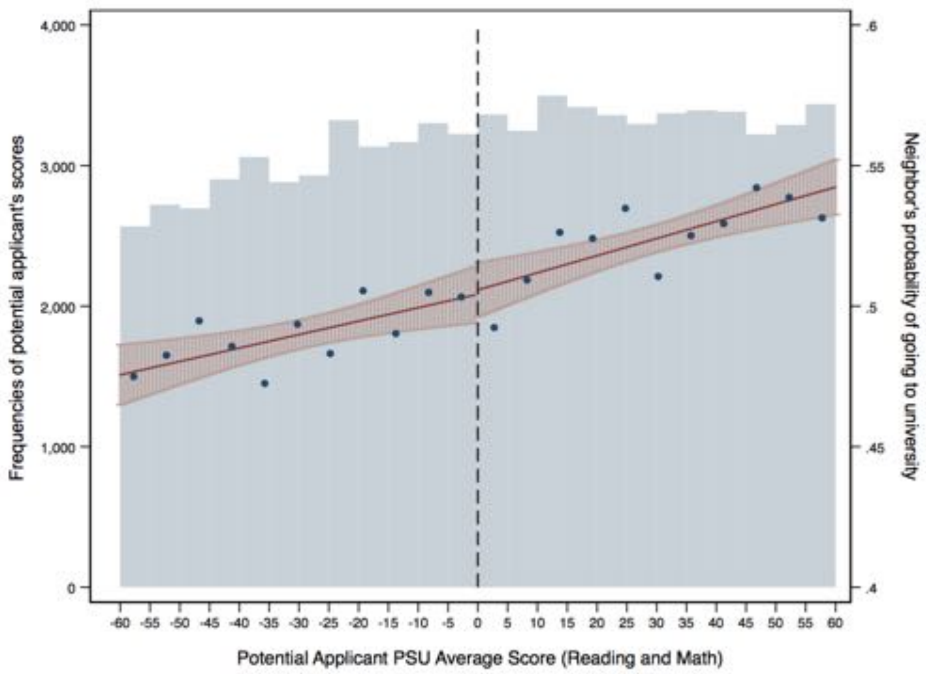
(a) Potential Applicants



(b) Neighbors

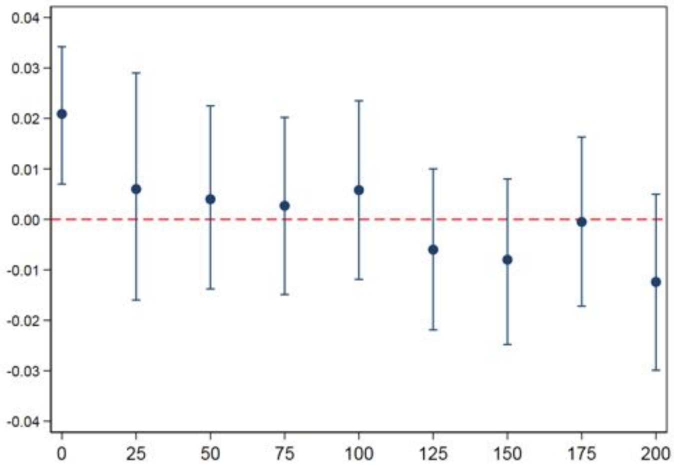
Notes: This figure illustrates the coefficients obtained when studying discontinuities in other variables that could potentially affect the outcome of interest. The left panel presents the results for potential applicants, while the right panel for neighbors. Apart from the coefficients, the figures 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. (2014*b*).

Figure E5: Placebo Exercise: Effect of Potential Applicants (t) on Neighbors (t-1)

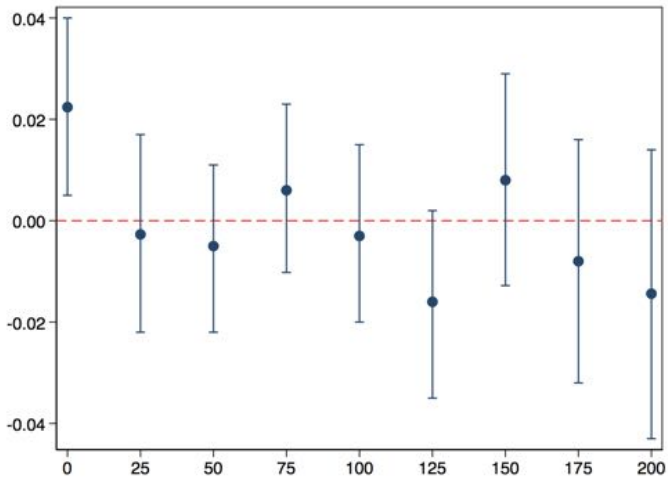


Notes: This figure illustrates the reduced form of a placebo exercise. It shows how neighbors' probability of going to university evolves with the PSU score of potential applicants. The PSU score is centered around the student-loans eligibility threshold. Each dot represents the share of neighbors going 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. (2014*b*).

Figure E6: Neighbors and siblings placebo cutoffs



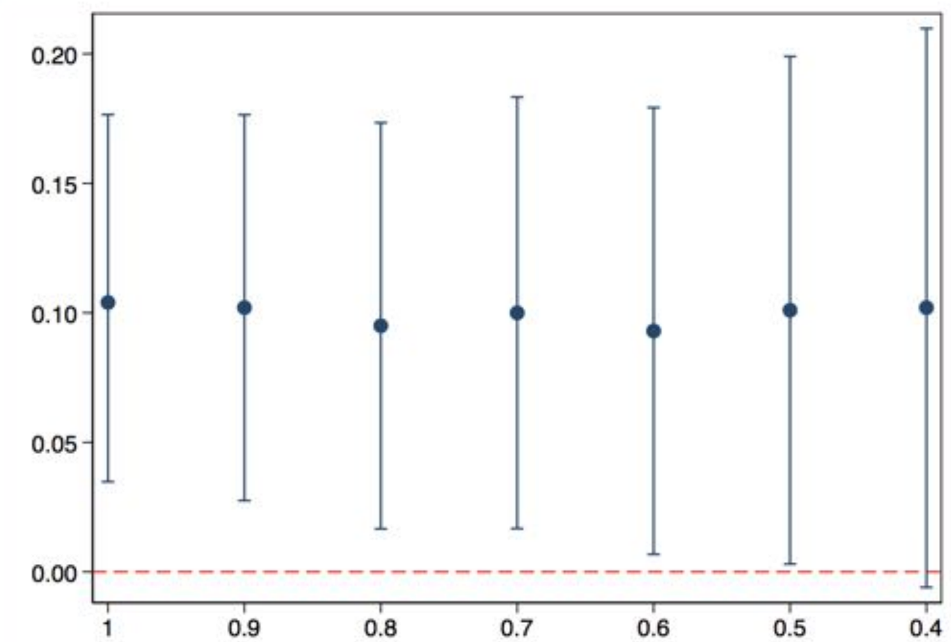
(a) Neighbors



(b) Siblings

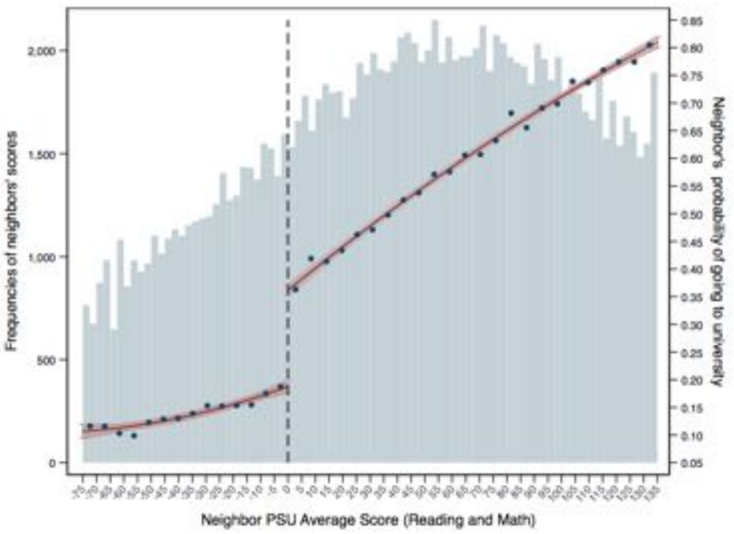
Notes: This figure illustrates the reduced form coefficients for the different cutoffs. The top panel illustrates the results for neighbors, and the panel at the bottom for siblings. Apart from the coefficients, the figures illustrate 95% confidence intervals. Standard errors are clustered at the neighborhood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

Figure E7: Estimated Neighbors' Effects with Different Bandwidths

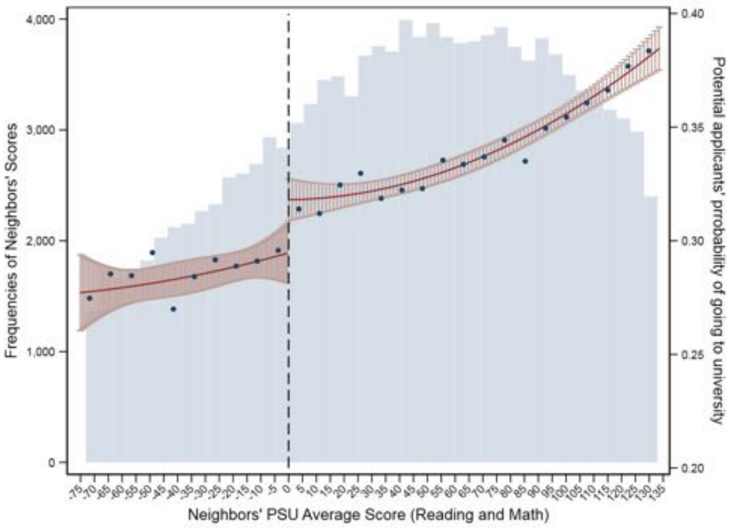


Notes: This figure illustrates the coefficients obtained when studying neighbors' effects using different bandwidths. The dots represent the coefficients, and the lines illustrate 95% confidence intervals.

Figure E8: First Stage and Reduced Form of Neighbors RD



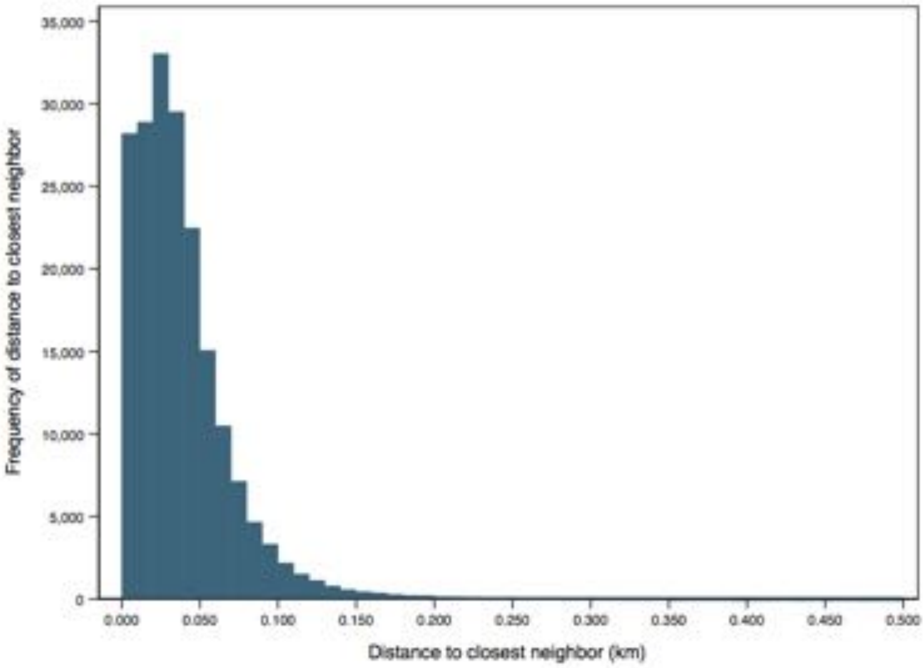
(a) First Stage: Neighbors' Probability of going to Univeristy



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

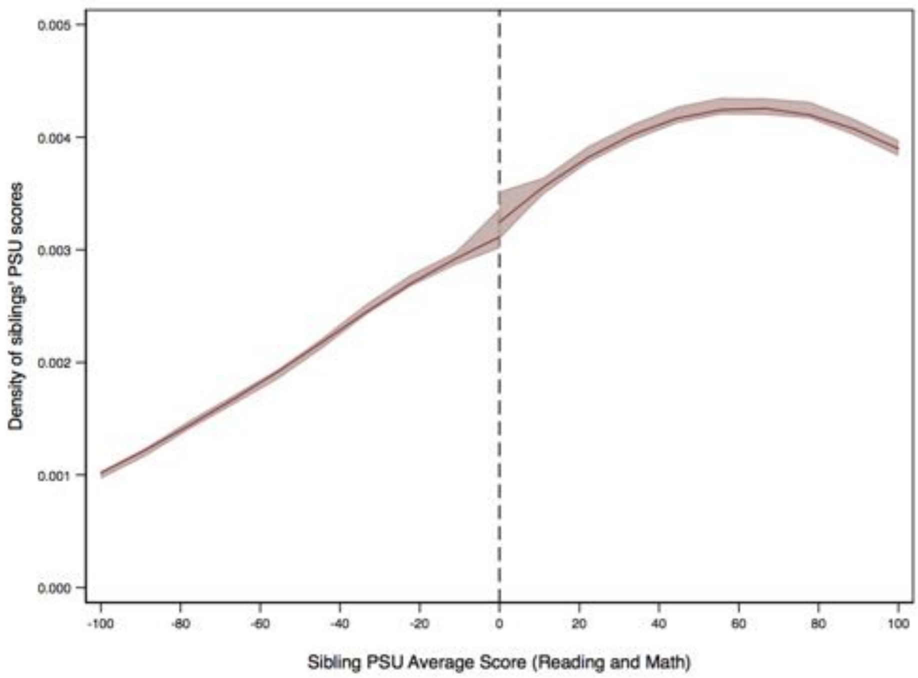
Notes: This figure illustrates the first stage and reduced form of the neighbors rd. The first panel shows how neighbors' probability of going to university evolves with the score they obtain in the PSU. The second panel shows how potential applicants' probability of going to university evolves with 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 **?**.

Figure E9: Distribution of Distance between Potential Applicant and Closest Neighbor



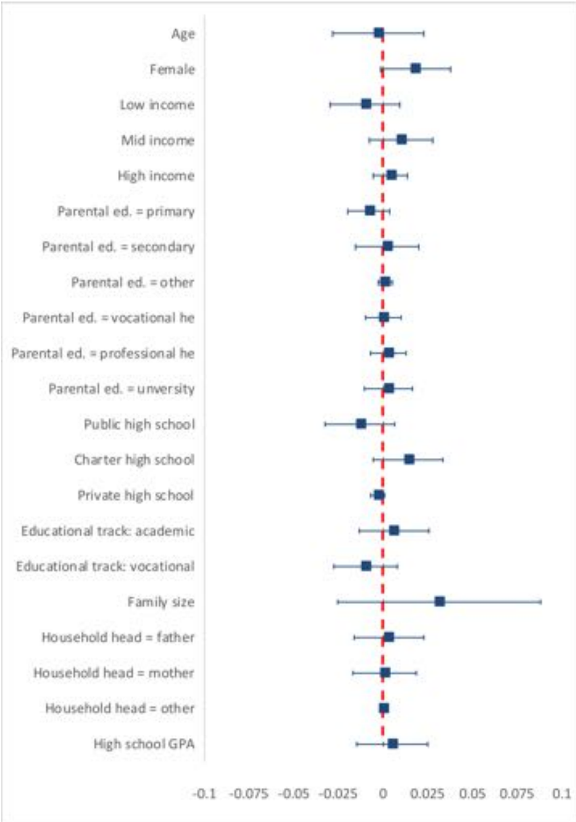
Notes: This figure illustrates the distribution of distance between potential applicants' household and their closest neighbor. Potential applicants are individuals that appear in the PSU registers between 2007 and 2012. Their neighbors are individuals that appear in the PSU registers one year before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

Figure E10: Density of Siblings' PSU Scores around the Student Loans Eligibility Threshold)

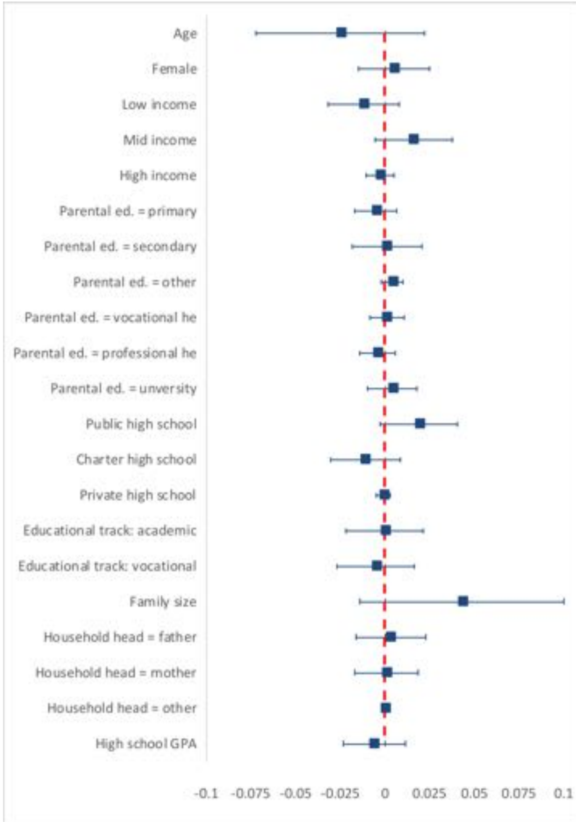


Notes: This figure illustrates the density of siblings PSU scores around the student loans eligibility thresholds. The density and its confidence intervals on each side of the cutoff were estimated following (Cattaneo et al. 2018). This chart complements the formal test they suggest to study discontinuities 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 E11: Discontinuities in other Variables at the Cutoff - Siblings



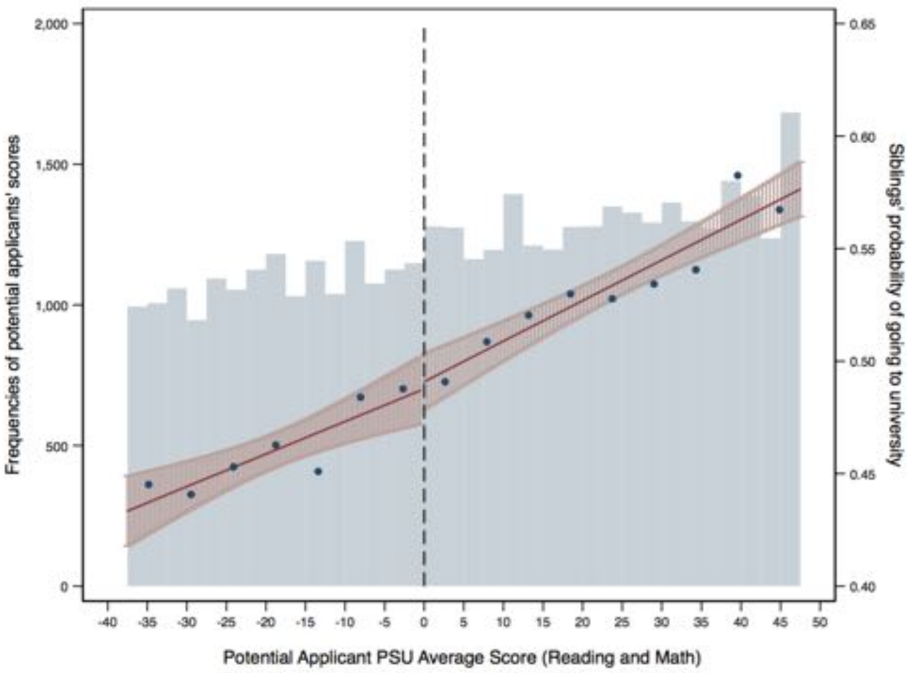
(a) Potential Applicants



(b) Siblings

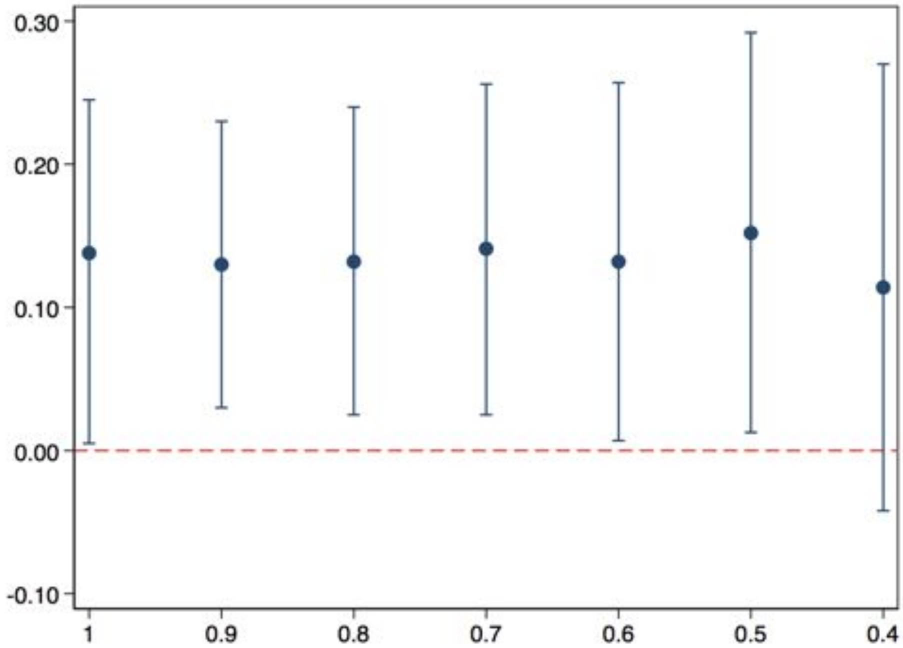
Notes: This figure illustrates the coefficients obtained when studying discontinuities in other variables that could potentially affect the outcome of interest. The left panel presents the results for potential applicants, while the right panel for siblings. Apart from the coefficients, the figures illustrate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

Figure E12: Placebo Exercise: Effect of Potential Applicants (t) on Siblings (t-T)



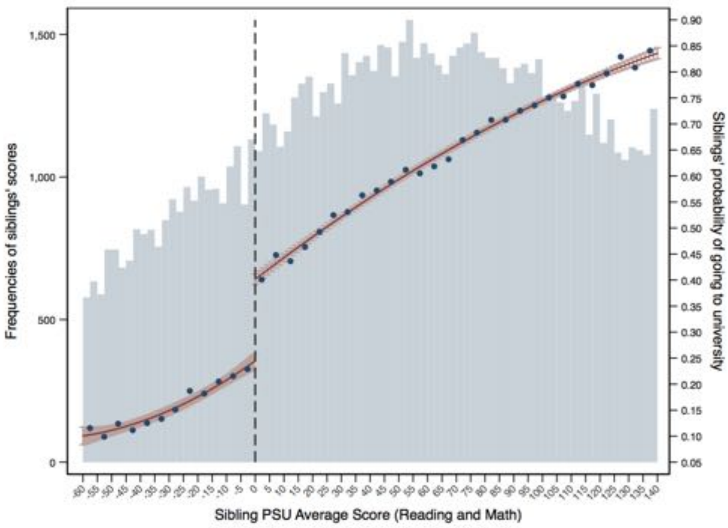
Notes: This figure illustrates the reduced form of a placebo exercise. It shows how siblings' probability of going to university evolves with the PSU score of potential applicants. The PSU score is centered around the student-loans eligibility threshold. Each dot represents the share of siblings going 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. (2014*b*).

Figure E13: Estimated Siblings' Effects with Different Bandwidths

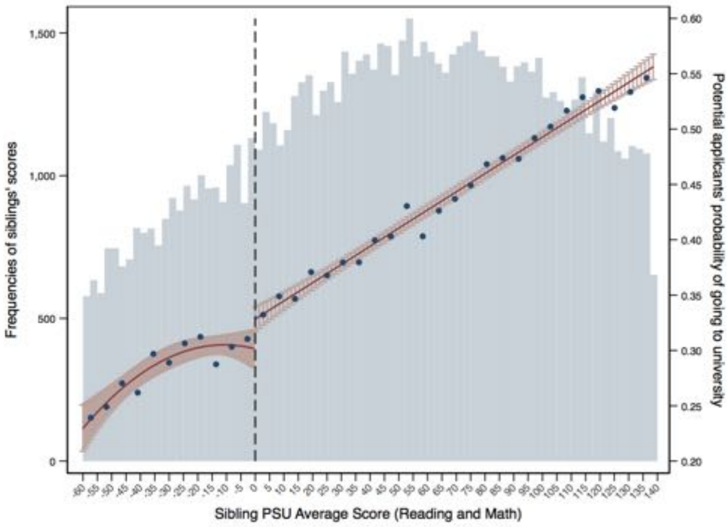


Notes: This figure illustrates the coefficients obtained when studying siblings' effects using different bandwidths. The dots represent the coefficients, and the lines illustrate 95% confidence intervals.

Figure E14: First Stage and Reduced Form of Siblings RD



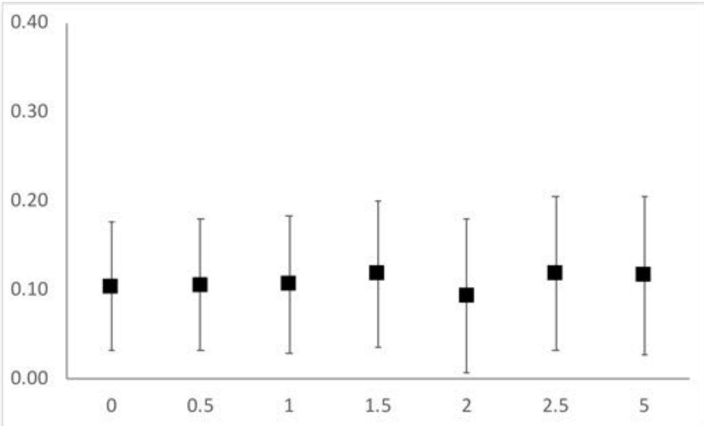
(a) First Stage: Siblings' Probability of going to Univeristy



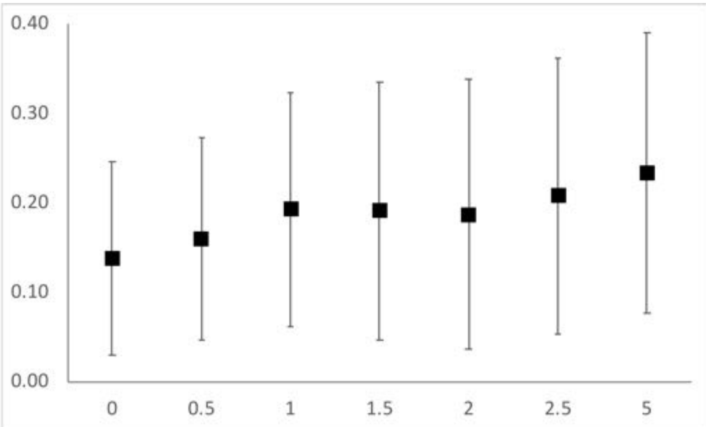
(b) Reduced Form: Potential Applicants' Probability of going to University

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 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. (2014*b*).

Figure E15: Effects of neighbors and siblings on potential applicants enrollment excluding observations around the cutoff



(a) Neighbors



(b) Siblings

Notes: This figure illustrates the estimated effects for neighbors and siblings when ommiting observations around the eligibility threshold. The coefficient at the left extreme corresponds to the one obtained using the whole sample. The rest were obtained ommiting potential applicants whose older peers obtained scores within 0.5, 1, 1.5 2, 2.5 and 5 the student-loans eligibility threshold. 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 clustered at the neighborhood unit or household level. The estimates were obtained using optimal bandwidths computed following Calonico et al. (2014*b*).

Table E1: 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.

Table E2: Heterogeneity in Effects of Closest Neighbor on Potential Applicants

	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)	0.037	0.122**	0.105	0.112**	0.048	0.073	0.160***
	(0.041)	(0.054)	(0.098)	(0.055)	(0.038)	(0.049)	(0.056)
Constant	0.092***	0.299***	0.350***	0.353***	0.092***	0.263***	0.212***
	(0.020)	(0.028)	(0.052)	(0.030)	(0.018)	(0.026)	(0.028)
First stage coefficient	0.180***	0.180***	0.155***	0.175***	0.178***	0.189***	0.164***
	(0.013)	(0.011)	(0.015)	(0.011)	(0.012)	(0.011)	(0.010)
N. of potential applicants	23,917	43,352	24,273	49,447	28,206	41,777	43,113
Kleibergen-Paap Wald F Statistic	182.64	279.27	109.22	255.74	205.15	314.29	260.73

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. (2014*b*). In parenthesis, standard errors clustered at neighborhood unit or household level. *p-value<0.1 **p-value<0.05 ***p-value<0.01

Table E3: Placebo - 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. (2014*b*). 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 E4: 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. (2014*b*). 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 E5: Effect of Neighbors on Potential Applicants University Enrollment (RM)

	2SLS-1 (1)	2SLS-2 (2)	CCT-1 (3)	CCT-2 (4)
Neighbor goes to university (t-1)	0.128**	0.168**	0.151**	0.128
	(0.060)	(0.075)	(0.075)	(0.092)
Constant	0.193***	0.173***		
	(0.028)	(0.034)		
First stage coefficient	0.128***	0.130***	0.137***	0.140***
	(0.060)	(0.014)	(0.009)	(0.012)
Year fixed effects	Yes	Yes	Yes	Yes
N. of students	57,110	75,120	57,110	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	152.08	88.51		

Notes: The table presents the results of analysis similar to those presented in table 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. (2014*b*). 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