Neighbors' Effects on University Enrollment
Online Appendix

Andrés Barrios Fernández *

March 14, 2021
Latest Version

Abstract

This online appendix is organized in twelve sections. The first provides additional details about the higher education system in Chile. The second section describes the sample used to estimate sibling spillovers. The third section discusses the identification strategy in detail. The fourth section presesents additional results for siblings. The fifth section shows the results of multiple robustness checks. The sixth section studies direct and indirect effects of scholarships on university enrollment. The seventh and eighth sections present additional heterogeneity analyses. The ninth section focuses on other definitions of close neighbors. The tenth section analyses changes in expenditure in higher education for households with one children crossing the student loan eligibility threshold. The eleventh section presents additional evidence of inequality in access to university. Finally, section twelfth illustrates the distribution of the distance between potential applicants and their closest neighbor applying to university one year before them.

Keywords: Neighbors' effects, University access, Spatial spillovers.
JEL classification: I21, I24, R23, R28.

*Centre for Economics Performance (LSE) and VATT Institute for Economic Research. andres.bafer@alumni.lse.ac.uk.

Contents

Appendices **3**

A Higher Education in Chile **3**

A.1 Institutions and Inequality in the System 3

A.2 University Admission System 3

A.3 Financial Aid 4

B Siblings Sample **6**

C Identification Strategy: Further Discussion **8**

D Additional Results Siblings **11**

E Robustness Checks **16**

E.1 Manipulation of the running variable 16

E.2 Discontinuities in potential confounders 17

E.3 Placebo exercises 17

E.4 Different bandwidths 18

E.5 Selection in PSU Registration and Geocoding Success 18

E.6 Statistical Inference Approach 19

F Discontinuities at the Scholarships Eligibility Threshold **36**

G Urban Segregation and Inequality in University Enrollment **41**

H Other Heterogeneity Analyses **47**

I Other Neighbors Definitions **51**

J Older Siblings' Expenditure in Higher Education **54**

K Inequality in Access to Higher Education in Chile **56**

L Distance to closest Neighbor **58**

A Higher Education in Chile

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

A.1 Institutions and Inequality in the System

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

Despite the expansion experienced by the higher education system in recent decades, inequality in access to university remains high. 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.

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

A.2 University Admission System

In Chile, there are public and private universities. All the public universities and 9 of the 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 eight other private universities, select their students through a centralized deferred acceptance admission system that only considers students' performance in high school and in a national level university admission exam (PSU). The PSU assesses students in four areas: language, mathematics, social sciences and natural sciences. To apply to university, students need to take language, mathematics, and at least one of the other sections. Universities are free to set the weights allocated to each sections 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. The raw scores obtained by students in each of these sections are adjusted to obtain a normal distribution of scores, with a mean of 500 and standard deviation of 110. The extremes of the distribution are truncated to obtain a minimum score of 150 and a maximum score of 850 in each section. The PSU is conducted in December, at the end of the Chilean academic year, but students typically need to register before mid-August. Since 2006, all students graduating from public and voucher schools, who roughly represent 93% of high school students in the country, are eligible for a fee waiver that makes the PSU free for them.

Universities that do not participate in the centralized system have their own admission processes. 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 that exist for both students and institutions.[1] For instance, the largest financial aid programs available for university studies require students to score above a cutoff in the PSU.

A.3 Financial Aid

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

Students who need financial aid must apply using an online platform a couple of months before taking the PSU. After verifying the validity of the information provided by the applicants, the Ministry of Education informs them about the benefits they are eligible for. 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 students to consider their funding options before applying and enrolling in higher education.

There are two student loan programs: solidarity fund credit (FSCU) and state guaranteed credit (CAE). The former can be used solely in CRUCH universities, while the latter

[1] Firstly, creating a new test generates 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. Although it was eliminated in 2016, it was in place during the period covered by this study.

can be used in any accredited higher education institution. 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%, for the CAE it varied between 5% and 6%. On top of that, while repayment of the FSCU has always been income contingent, the CAE used to have fixed installments. In order to become eligible for these loans, students need to obtain an average PSU score (language and mathematics) of above 475 and come from households in the bottom 90% of the income distribution.[2]

Solis (2017) documents that eligibility for student loans creates a discrete jump in the probability of enrolling in university. This paper exploits the same discontinuity, but this time to study the effect of having a close neighbor or an older sibling going to university with a student loan.

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 score above 550), and that they are focused on students from more disadvantaged backgrounds. Students eligible for these scholarships are also eligible for student loans. Since scholarships do not need to be repaid, crossing the scholarships' eligibility threshold changes the generosity of the subsidy but not the availability of funding (section F in this document provides additional details and studies direct and indirect effects of scholarships on university enrollment). There are also a few programs that instead of requiring a minimum score in the PSU, allocate funding based on performance in high school. These programs are relatively small, both in terms of beneficiaries and of the support they offer.

Since Chilean universities have complete freedom to decide 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.[3] At the university level, the reference tuition fee roughly covers 80% of the actual fee. This means that students need to fund the additional 20% by using their own resources, by taking private loans or by applying for external support offered by their universities or other private institutions.

[2] The FSCU is available for students from households in the bottom 80% of the income distribution. The CAE, on the other hand, initially focused on students in the bottom 90% of the income distribution; however, since 2014, the loan is available to anyone that satisfies the academic requirements.

[3] 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 actual tuition fee.

B Siblings Sample

Although this paper focuses on neighbors, I also investigate what happens with potential university applicants when an older sibling enrolls in university T years before him/her. The sample used 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 identify 273,806 pairs of siblings. 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. If an older sibling registers more than once, I use the first time he/she takes the PSU. For younger siblings I use the first time they appear in the registers. 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 student-loans 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 sample. Table B.I presents the summary statistics for this sample.

As in the case of the neighbors sample, these students come from relatively low-income households and in the majority of the cases their parents did not attended higher education. Although there are some small differences, potential applicants and their siblings report very similar socioeconomic characteristics. I do not observe important differences in the type of school or educational track chosen by siblings, but older siblings seem to perform better on the PSU. Finally, siblings report some differences in the structure of the household. These differences are consistent with some parents leaving the household.

Table B.I: Summary statistics - Siblings' sample

	Older siblings (1)	Potential applicants (2)
1. Demographic characteristics		
Female	0.55	0.54
Age at PSU registration	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. Household structure		
Household 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 older siblings.

C Identification Strategy: Further Discussion

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

$$U_{at} = \alpha + \sum_{n \in N_a} \beta_{n\tau} U_{n\tau} + \varepsilon_{it} \tag{1}$$

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

As discussed in section 4 of the paper, neighbors decide whether to enroll or not into university before potential university applicants. Thus, their decision should not be affected by what potential university 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).

This paper focuses on the effects of neighbors going to university one year before potential university applicants. To highlight this, equation 1 can be rearranged as follows:

$$U_{at} = \alpha + \beta_{mt-1} U_{mt-1} + \sum_{n \in N_a \setminus U_{mt-T}} \beta_{n\tau} U_{n\tau} + \varepsilon_{it} \tag{2}$$

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 university 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 including all the observed neighbors in the analyses.

Consider the following simplified specification:

$$U_{at} = \alpha + \beta_{mt-1} U_{mt-1} + v_{it} \tag{3}$$

In this case, to consistently estimate β_{mt-1} we need $Cov(U_{mt-1}, v_{it}) = 0$. This means that in addition to the conditions discussed for equation 2, 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.[4] To satisfy the assumption in the third case we would need to assume that neighbors applying two or more years before potential university 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 1, β_n can still be interpreted as a reduced 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 generated on a. This is still a relevant parameter from a policy perspective.

A fuzzy regression discontinuity (RD) design can be thought as a particular case of IV. By abstracting from its local nature, 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 university applicant and his/her neighbor. In my setting, this will only be true around the student loan eligibility threshold 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 of the paper show that this is indeed the case.[5]

A3. Exclusion:
The instrument only affects potential university 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 university applicants.

A4. Monotonicity:
Finally, the monotonicity assumption requires eligibility for student loans to weakly increase neighbors enrollment. In this setting, it is difficult to think of any reasons that would induce individuals to not enroll in university because they are eligible for financial

[4] We are already assuming that younger applicants' decisions are not part of the equation 1.

[5] 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.

aid. Even 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.

According to Imbens and Angrist (1994), under this set of assumptions the IV estimates are consistent and can be interpreted as a local average treatment effect (LATE). In this fuzzy regression discontinuity (RD) design setting, this means that my estimates will have a double local interpretation. First, they are local in the sense that they are valid only for individuals whose neighbors are near the student loan eligibility threshold. Second, they are local in the sense that they are capturing the effect on the population of compliers; this is individuals whose neighbors decide to enroll at university because of their eligibility for funding.

D Additional Results Siblings

Section 5.3 in the paper shows that older siblings eligible for student loans are around 16 pp more likely to enroll in university than those who are not eligible. It also shows that potential applicants with an older sibling crossing the student loan eligibility threshold are more than 2 pp more likely to attend university than those whose older sibling fails to cross it.

Under the identifying assumptions discussed in the main body of the paper, the first stage and reduced form can be combined to estimate the effect of older siblings' loan-induced university enrollment on potential applicants' enrollment. Table D.I summarizes these results. The first two columns present 2SLS estimates, while the third and fourth columns show estimates obtained using the robust approach suggested by Calonico et al. (2014b). According to these figures, having an older sibling going to university with a student loan increases their younger siblings' enrollment by between 12.5 and 16.5 pp.

These 2SLS estimates would represent an upper bound of the effect of older siblings' loan-induced university enrollment on potential applicants if an older sibling's eligibility for student loans directly affected younger siblings' enrollment.

Since siblings usually live together, a potential applicant could learn about the availability of student loans even if his/her older sibling does not enroll in university. However, this is true for potential applicants whose older siblings score marginally above and marginally below the student loan eligibility threshold. In both scenarios, younger siblings are likely to be aware of the existence of funding opportunities and their rules before they need to decide whether to apply or not to university.

While neighbors do not usually share household budgets, siblings do. Thus, an additional concern that arises in this case is that the eligibility of an older sibling for funding could affect the resources available to finance the education of younger siblings. The importance of this threat greatly depends on the generosity of the funding to which older siblings have access. As discussed in section 2 of the paper, student loans only cover a share of the tuition fees. This means that even when older siblings are eligible for a student loan, they and their families have to cover part of the tuition fees as well as commuting, maintenance and study materials costs. Thus, irrespective of the availability of funding, households in which the older sibling enrolls in university are likely to face a tighter budget constraint than those in which the older sibling does not.[6]

[6] In section F of this document, I show that older siblings eligible for a scholarship are not more likely to enroll in university than those eligible for a student loan. Scholarships change the generosity of the subsidy that older siblings receive, but I find no spillovers on younger siblings' enrollment. Apart from working as a placebo test, this result is consistent with the idea that older siblings' funding alone is not enough to change their younger siblings' choices.

Although I cannot completely rule out that the effect is partly driven by changes in household resources, it is unlikely that this is the whole story. There is a significant difference in the share of older siblings going to university at both sides of the cutoff, and as discussed in the previous paragraph, student loans cover only a part of the expenses of sending a child to university.[7] In addition, my results are in line with the findings of Altmejd et al. (2021). In this work, the authors exploit college specific admission cutoffs—instead of student loan eligibility cutoffs—and find even larger sibling spillovers in 4-year college enrollment.

To further investigate how older siblings influence the university choices of potential applicants, in Table D.II, I present the results of additional exercises that look at the type of institutions in which potential applicants enroll. Columns (1) and (2) indicate that around 60% of the difference that I find in university enrollment is driven by potential applicants that otherwise would not have enrolled in any higher education institution; the other 40% corresponds to potential applicants who otherwise would have attended vocational higher education. Column (3) shows that most potential applicants enroll in accredited universities, column (4) shows that roughly half of them choose a university that is part of the CRUCH, and column (5) shows that a similar proportion attends an accredited program. Finally, Columns (6) and (7) indicate that the majority of potential applicants who decide to enroll in university choose the same institution as their older sibling. This last set of results suggests that older siblings not only affect the decision to attend university but also the specific university that their younger siblings attend. This result, however, needs to be interpreted with caution, as part of this increase is a mechanic consequence of the increase in younger siblings' enrollment documented earlier. Nevertheless, the size of the coefficient suggests that older siblings do influence their younger siblings' choice of university. This is consistent with the findings of Altmejd et al. (2021), who address this specific question in more detail.

I conclude this section by showing that the increase in younger siblings' university attendance persists a year later and also leads to an increase in university completion. Columns (1) and (2) of Table D.III look at differences in retention in the university system and in the same institution where they originally enrolled. These estimates are similar in size to the effects documented for first year enrollment, indicating that the majority of younger siblings who decide to go to university following the example of an older sibling, remain enrolled in their second year.[8] In addition, columns (3) and (4) look at the probability of completing higher education or university before 2019. To study this outcome, I restrict

[7] Section J in this document shows that average expenditure in older siblings' higher education fees does not change at the student loan eligibility cutoff. This suggests that on average younger siblings with an older sibling marginally above or below the cutoff come from households that face similar budget constraints.

[8] The outcomes take value 1 for applicants who enroll in t and continue enrolled in $t+1$, and take value 0 for applicants who do not enroll in t or who enroll in t but dropout during the first year.

Table D.I: Effect of older siblings on potential applicants' university enrollment

	2SLS-1 (1)	2SLS-2 (2)	CCT-1 (3)	CCT-2 (4)
Sibling goes to university (t-T)	0.126	0.165	0.140	0.165
	(0.053)	(0.068)	(0.064)	(0.079)
First stage	0.170	0.155	0.158	0.161
	(0.009)	(0.010)	(0.011)	(0.013)
Reduced form	0.021	0.026		
	(0.009)	(0.011)		
Year fixed effects	Yes	Yes	Yes	Yes
N. of students	57,713	95,969	57,713	95,969
PSU Polynomial	1	2	1	2
Bandwidth	(37.0-74.5)	(60.0 - 132.0)	(37.0-74.5)	(60.0 - 132.0)
Kleibergen-Paap F statistic	362.60	223.08		
Outcome mean	0.37	0.40	0.37	0.40

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 the robust approach suggested by Calonico et al. (2014b). Optimal bandwidths are used in all the specifications. In parenthesis, standard errors clustered at family level.

the sample to observations in which the younger sibling registers for the PSU no later than 2013. These results show that potential applicants with an older sibling going to university before them, are 12.3 pp more likely to complete a university degree before 2019. I find no difference in the probability of completing higher education, which suggests that potential applicants whose older siblings do not enroll in university are more likely to attain a vocational higher education degree.

Table D.II: Effect of older siblings on potential applicants' enrollment by type of institution

	Pr. of Enrolling in:						
	Any HEI (1)	Vocational HEI (2)	Accredited university (3)	CRUCH university (4)	Accredited program (5)	Sibling's university (6)	Other university (7)
Older sibling goes to university (t-T)	0.073	-0.053	0.118	0.069	0.058	0.098	0.028
	(0.057)	(0.046)	(0.052)	(0.044)	(0.039)	(0.020)	(0.051)
Reduced form	0.012	-0.009	0.020	0.012	0.010	0.017	0.005
	(0.010)	(0.008)	(0.009)	(0.008)	(0.007)	(0.004)	(0.009)
First stage	0.170	0.170	0.170	0.170	0.170	0.170	0.170
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N. of students	57,713	57,713	57,713	57,713	57,713	57,713	57,713
PSU Polynomial	1	1	1	1	1	1	1
Bandwidth	(37.0-74.5)	(37.0-74.5)	(37.0-74.5)	(37.0-74.5)	(37.0-74.5)	(37.0-74.5)	(37.0-74.5)
Kleibergen-Paap F statistic	362.60	362.60	362.60	362.60	362.60	362.60	362.60
Counterfactual mean	0.56	0.19	0.36	0.17	0.22	0.06	0.31

Notes: The table presents the estimated effects of neighbors on potential applicants' enrollment in any higher education institution (column 1), in vocational higher education institutions (column 2), in accredited universities (column 3), in CRUCH universities (column 4), in accredited programs (column 5), in the neighbors' university (column 6), and in any other university (column 7). All specifications include a linear polynomial of the PSU which slope is allowed to change at the cutoff. Bandwidths are the same used in Table VII. In parenthesis, standard errors clustered at family level.

Table D.III: Effect of siblings on potential applicants' second year enrollment and university completion

	Pr. of remaining in the:		Pr. of completing:	
	University system	Same institution	Higher education	University
	(1)	(2)	(3)	(4)
Sibling goes to university (t-1)	0.097	0.083	0.025	0.123
	(0.046)	(0.045)	(0.068)	(0.062)
Reduced form	0.016	0.014	0.004	0.021
	(0.008)	(0.008)	(0.012)	(0.011)
First stage	0.170	0.170	0.172	0.172
	(0.009)	(0.009)	(0.011)	(0.011)
Years fixed effects	Yes	Yes	Yes	Yes
N. of students	57,713	57,713	36,923	36,923
Kleibergen-Paap F statistic	362.60	362.60	243.30	243.30
Outcome mean	0.25	0.23	0.59	0.32

Notes: The table presents estimated effects of siblings on potential applicants' permanence in the system and in the university where they start one year after enrollment. It also present estimated effects on their probability of completing a higher education and a university degree. Column 1 looks at permanence in any university, column 2 at permanence in the same university in which potential applicants enrolled in their first year, column 3 at the probability of completing any higher education degree, and column 4 at the probability of completing a university degree. When looking at potential applicants' permanence, the outcome is 1 for applicants who enroll and remain enrolled one year later; it is 0 for applicants who do not enroll at all or who enroll but dropout after their first year. When looking at degree completion I focus on potential applicants who register for the PSU no later than 2013. 2SLS estimates come from specifications that control for a linear polynomial of PSU which slopes are allowed to change at the cutoff. Bandwidths are the same used in the specifications presented in Table VII. In parenthesis, standard errors clustered at family level.

E Robustness Checks

In this section, I study whether the identification assumptions of the empirical strategy used in the paper are satisfied. I start by investigating if there is evidence of manipulation of the running variable. Then, I check whether other variables that could be related to the decision of enrolling in university present jumps around the student loan eligibility threshold. I continue by showing the results of different placebo exercises and the robustness of my estimates to different bandwidths choices. Next, I discuss concerns related to endogeneity in PSU registration and in geocoding success. I finish this Section presenting figures that illustrate reduced form results using a second degree polynomial of the running variable.

E.1 Manipulation of the running variable

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

In this case, it would mean that potential university applicants have the ability to affect the average PSU score of their older neighbors and siblings. As discussed in section 2 of the paper, the PSU is a national level exam whose application and marking processes are completely centralized. This means that the teachers or the high school of a potential university applicant do not play any role in the process. 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 would need to score above the eligibility cutoff.

All this makes manipulating scores around the threshold very difficult, even for individuals taking the exam. Considering this, it seems very unlikely that potential university applicants could strategically affect it.

In the context of this paper, a way in which potential university applicants could manipulate the score obtained by their neighbors would be to move to a different neighborhood. However, the results on movers and no-movers presented in section 5 of the paper 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 university 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 choose their new neighborhood.

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 E.I and E.II 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 0.7759, whereas in the case of siblings it is 0.5968. Therefore, the results that I find do not seem to be driven by manipulation of the running variable.

E.2 Discontinuities in potential confounders

A second concern in the context of a fuzzy regression discontinuity design (RD), is the existence of discontinuities in potential confounders around the cutoff that could explain the differences that we observe in the outcome of interest.

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

Figure E.III summarizes these results for neighbors, and figure E.IV for siblings. They illustrate the estimated discontinuities at the cutoff and their 95% confidence intervals. To estimate these discontinuities, I use the optimal bandwidths estimated for the main specification following Calonico et al. (2014a). In both figures, the left panel looks at characteristics of potential university applicants, and the right panel at characteristics of their older peers (i.e., neighbors or siblings).

I do not find any significant difference in potential university applicants', neighbors' or older siblings' characteristics around the threshold. In addition, the magnitudes of the coefficients are small in all cases.

E.3 Placebo exercises

This section presents the results of a set of placebo exercises designed to investigate if responses like the ones documented in the main body of the paper arise in cases in which they should not.

I start by investigating if university enrollment of a younger applicant has any effect on older neighbors or siblings. Since older peers apply and decide to enroll in university before potential university applicants, their decision to enroll in university should not be affected by what potential university applicants do.

Figures E.V and E.VI illustrate the results of an exercise in which I study whether potential university applicants' eligibility for student loans changes the probability of going to university of their older neighbors and siblings. As expected, I find no discontinuity in older peers' university enrollment at the eligibility threshold; both the levels and slopes seem to be continuous around it.

The second placebo exercise that I implement consists in studying whether significant discontinuities can be found in points different to the student loan eligibility threshold. Since in these points there is no first stage (i.e., older peer's probability of going to university does not change), we should not find jumps around these placebo cutoffs. Figure E.VII presents these results for neighbors and siblings. None of the jumps at placebo cutoffs is statistically different from 0.

Finally, I investigate whether there are discontinuities around the student loan eligibility threshold for potential university applicants whose closest neighbor does not apply for funding. Since the neighbor does not apply for funding, being above or below the eligibility threshold does not change his/her likelihood of going to university. I show that this is indeed the case in Table E.I. As can be appreciated, there is no first stage. As expected, in the absence of a first stage I find no effect on potential university applicant's applications, enrollment or academic performance.

E.4 Different bandwidths

In this section, I study how sensitive my results are to the choice of bandwidth. 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 far from the relevant cutoff.

Figures E.VIII and E.IX present the estimated coefficients using bandwidths that go from 0.5 to 1.5 times the optimal bandwidths computed according to Calonico et al. (2014b). These results correspond to specifications that control for a first degree polynomial of the running variable whose slope is allowed to change at the cutoff. As shown in the figures, the estimated effects do not experience important changes when varying bandwidths.

E.5 Selection in PSU Registration and Geocoding Success

This section discusses threats related to endogenous registration in the PSU and endogenous geocoding success. As explained in section 3 of the paper, I identify potential university entrants and their close neighbors using the information that individuals provide when registering for the PSU. Thus, if the university enrollment of a close neighbor affects the PSU registration of potential university applicants, the estimated effects could be biased. Something similar could happen if the university enrollment of a close neighbor affects the probability of successfully geocoding an address.

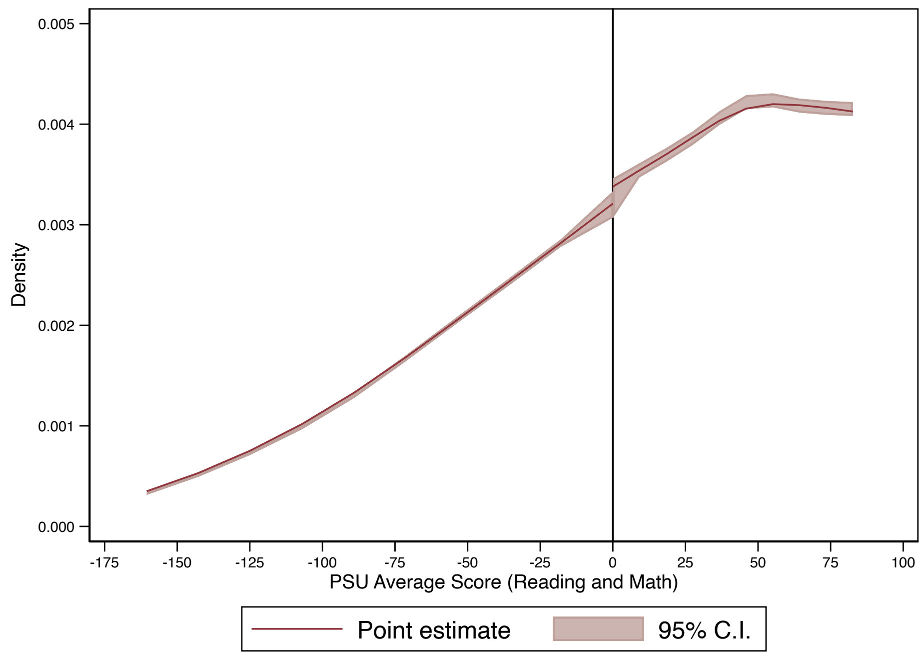
A first element that attenuates concerns respect endogenous PSU registration is that registering for the PSU is free for students completing secondary education in subsidized schools (93% of high school graduates). This results in more than 85% of high school graduates registering for the PSU even if they end not taking it. I formally investigate whether university enrollment driven by funding eligibility affects PSU registration or geocoding success in Table E.II and Figure E.XII. These exercises study whether a neighbor marginal eligibility for student loans changes the distance to the closest individual registered for the PSU the following year, and the number of individuals registered for the PSU the following year at 50m, 100m, 150m and 200m from the neighbors' address. To implement this exercise, I created a new sample using older neighbors as reference and identifying all the potential university entrants living at 200m or less from them and who appear in the PSU registers one year after the older neighbors. As shown in Table E.II, there is no significant difference in the distance between older neighbors and their closest potential university applicant at the cutoff. Similarly, Figure E.XII shows that the number of potential university applicants registered for the PSU living at 50m, 100m, 150m and 200m from the neighbor does not changes at the cutoff. These results suggest that older neighbors' eligibility for funding does not change PSU registration nor affects the probability of successfully geocoding addresses.

To further study how differences in geocoding success rates could affect my results, I present an additional exercise that replicates the main analysis just focusing on the Metropolitan Region of Santiago, as in this area the geocoding rate of success was higher than in the other two studied regions. Table E.III presents the results of this exercise. The obtained estimates are slightly larger than the ones I present in the main body of the paper.

E.6 Statistical Inference Approach

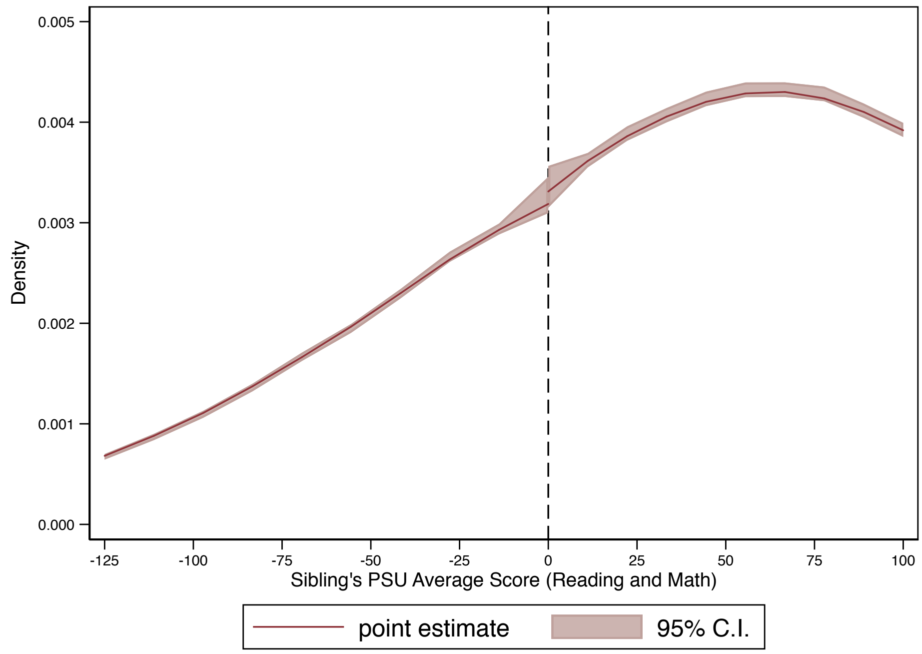
The results presented in the main body of the paper cluster standard errors at the neighborhood unit level. As explained in section 3 of the paper, neighborhood units correspond to subareas within a municipality and were defined by the Ministry of Social Development to decentralize certain local matters and to foster citizen participation and community-based management. In Table E.IV I show that the precision of the estimates does not suffer major changes when modifying the clustering level. Column (1) replicates the results presented in the paper. In the rest of the columns standard errors are computed clustering at the closest neighbor (column 2), potential applicants' high school (column 3), and potential applicants' municipality level (column 4). In all cases the estimated effects are statistically different from zero.

Figure E.I: Density of neighbors' PSU scores around the student loans eligibility threshold)



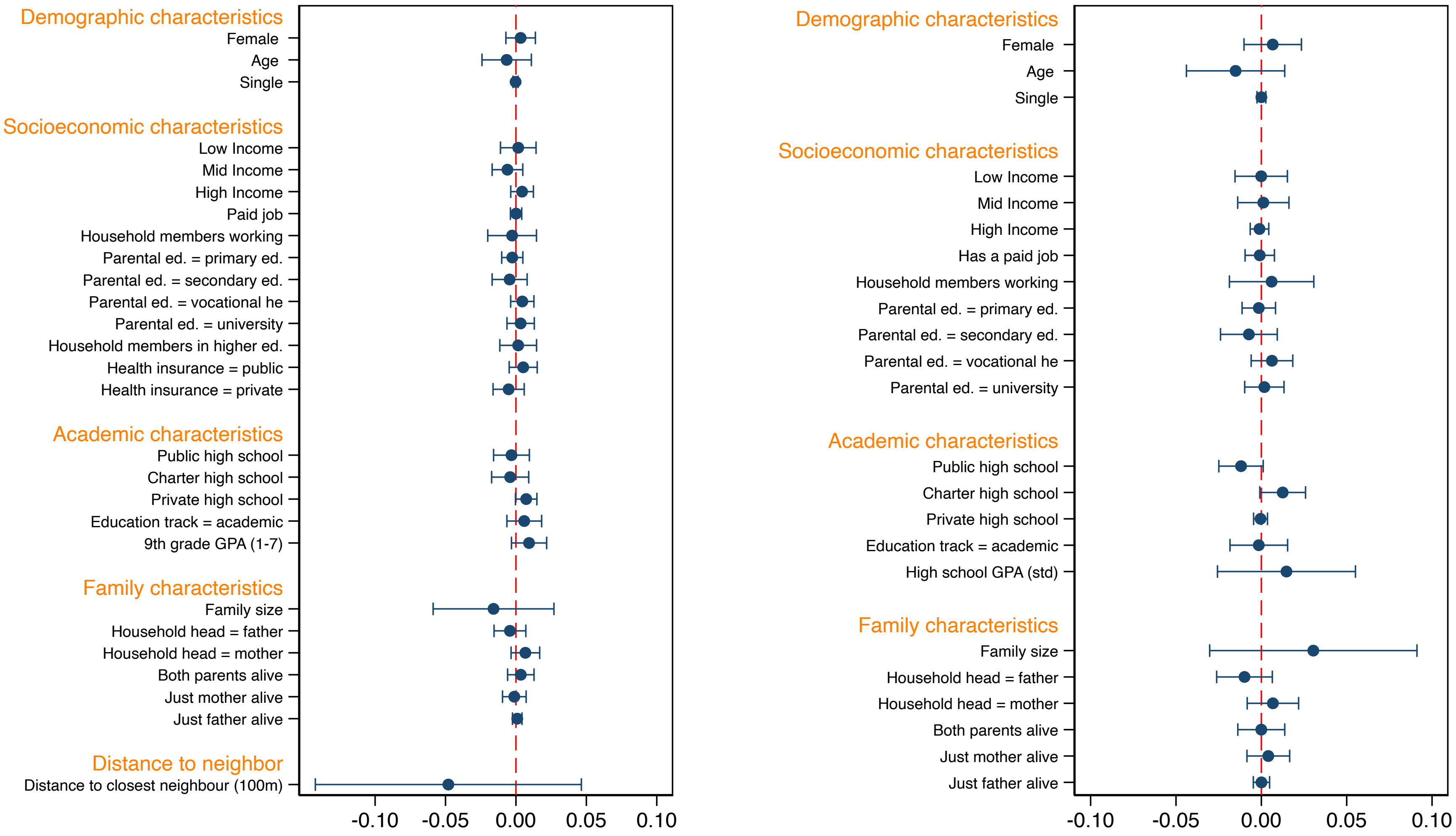
Notes: This figure illustrates the density of neighbors PSU scores around the student loans eligibility thresholds. The density and its confidence intervals on each side of the cutoff were estimated following Cattaneo et al. (2018). This chart complements the formal test they suggest to study discontinuities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

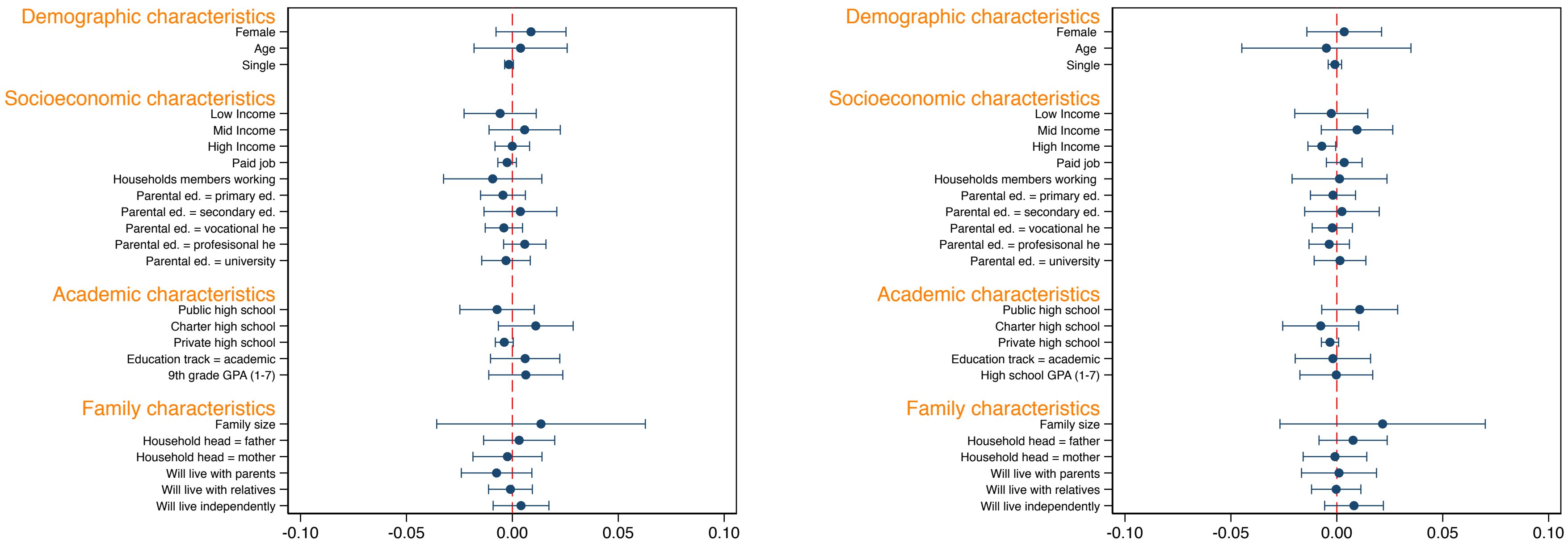
Figure E.III: Discontinuities in potential confounders at the cutoff (neighbors)



(a) Potential applicants (b) Neighbors

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

Figure E.IV: Discontinuities in potential confounders at the cutoff (siblings)

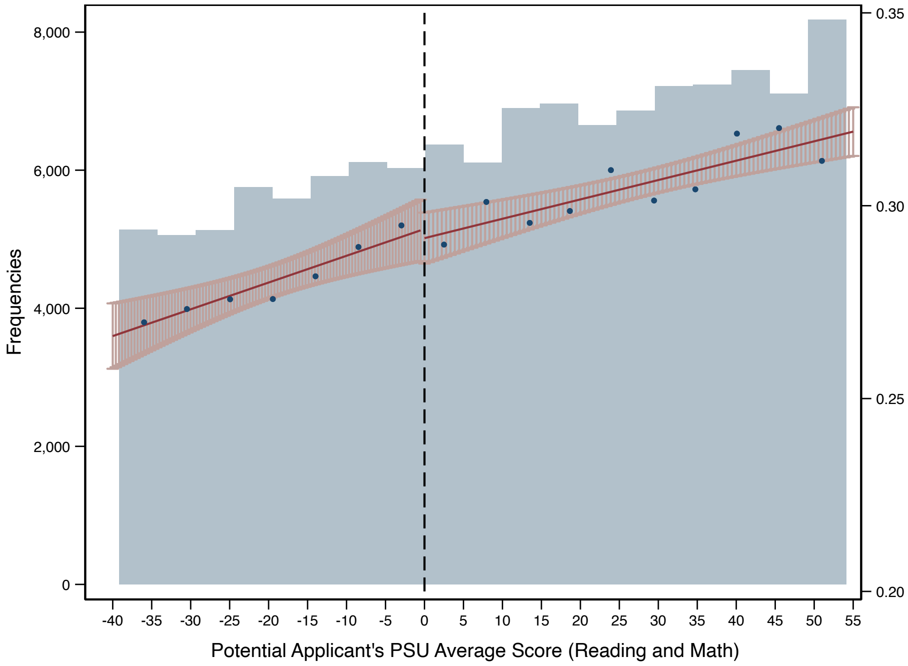


(a) Potential applicants

(b) Siblings

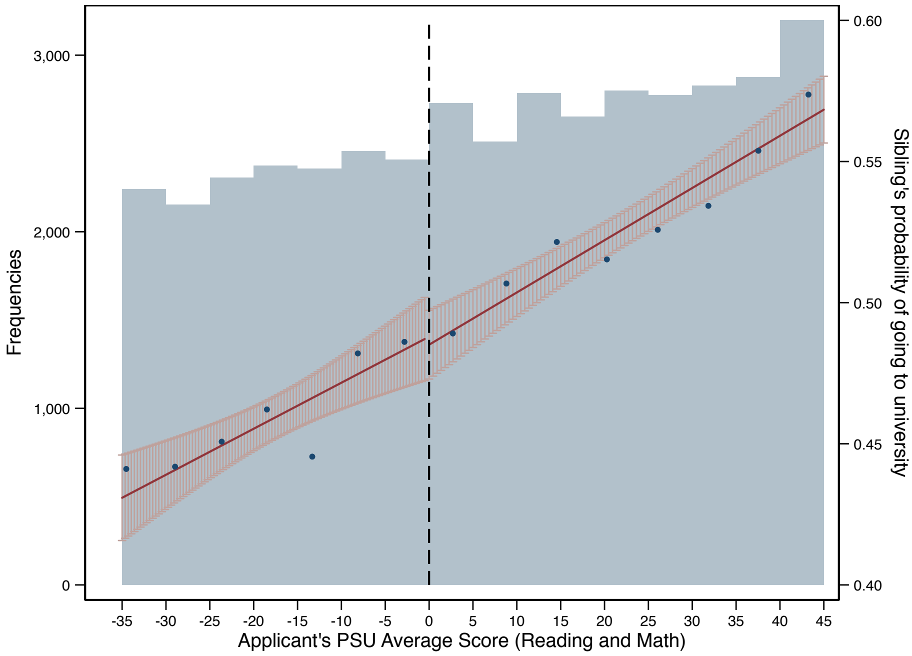
Notes: This figure illustrates the coefficients obtained when studying discontinuities in other variables that could potentially affect the outcome of interest. The left panel presents the results for potential applicants, while the right panel for siblings. Apart from the coefficients, the figures illustrate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

Figure E.V: Placebo exercise: Effect of potential applicants (t) on neighbors (t-1)



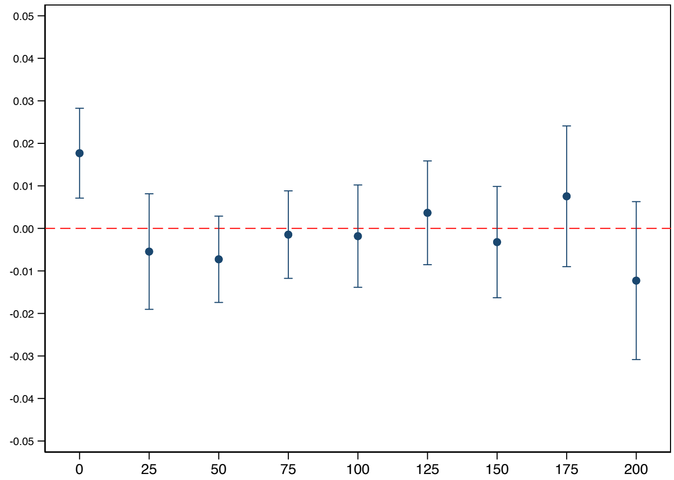
Notes: This figure illustrates the reduced form of a placebo exercise. It shows how neighbors' probability of going to university evolves with the PSU score of potential applicants. The PSU score is centered around the student-loans eligibility threshold. Each dot represents the share of neighbors going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

Figure E.VI: Placebo exercise: Effect of potential applicants (t) on older siblings (t-T)

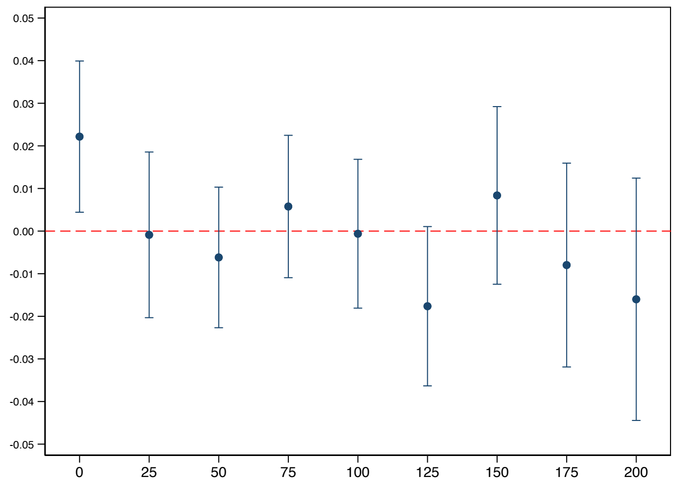


Notes: This figure illustrates the reduced form of a placebo exercise. It shows how siblings' probability of going to university evolves with the PSU score of potential applicants. The PSU score is centered around the student-loans eligibility threshold. Each dot represents the share of siblings going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

Figure E.VII: Neighbors' and siblings' effects at placebo cutoffs



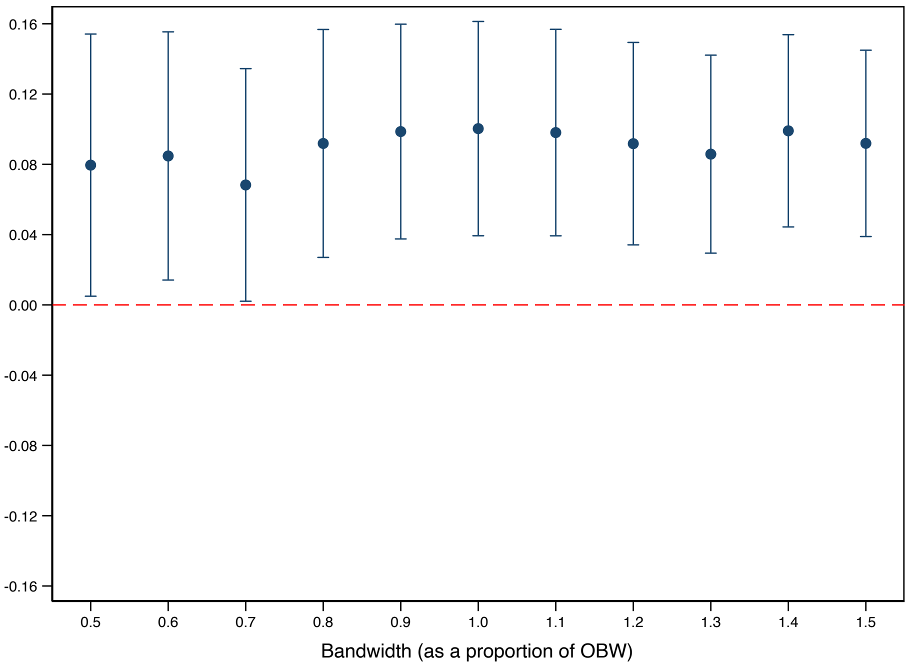
(a) Neighbors



(b) Siblings

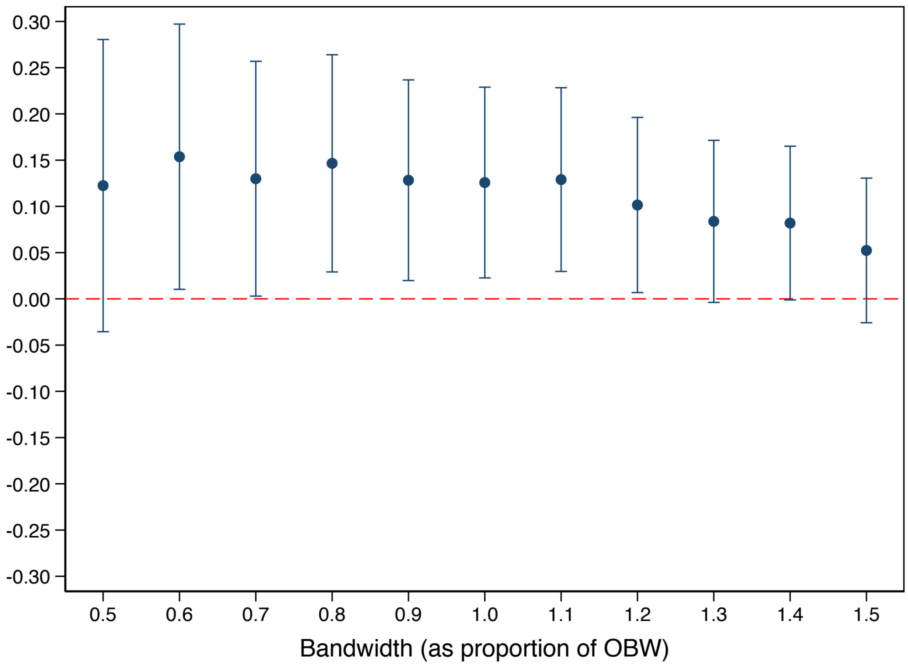
Notes: This figure illustrates the reduced form coefficients for the different cutoffs. The top panel illustrates the results for neighbors, and the panel at the bottom for siblings. Apart from the coefficients, the figures illustrate 95% confidence intervals. Standard errors are clustered at the neighborhood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

Figure E.VIII: Neighbors' effects on potential applicants' university enrollment using different bandwidths



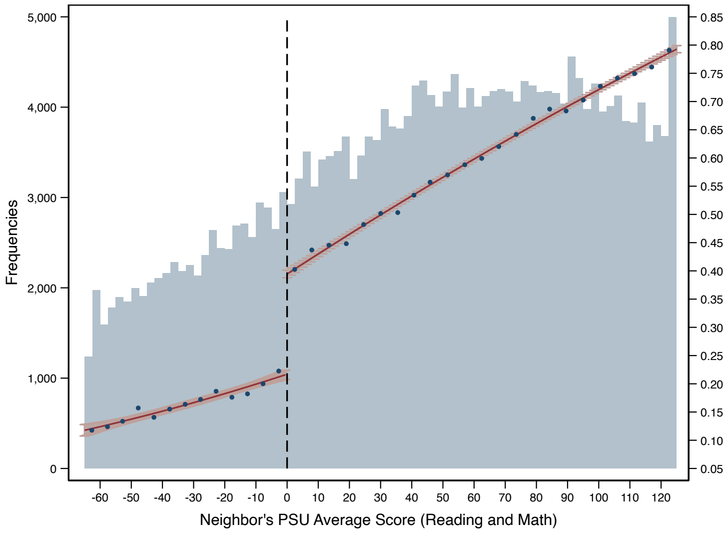
Notes: This figure illustrates the coefficients obtained when studying neighbors' effects using different bandwidths. The dots represent the coefficients, and the lines illustrate 95% confidence intervals.

Figure E.IX: Older siblings' effects on potential applicants' university enrollment using different bandwidths

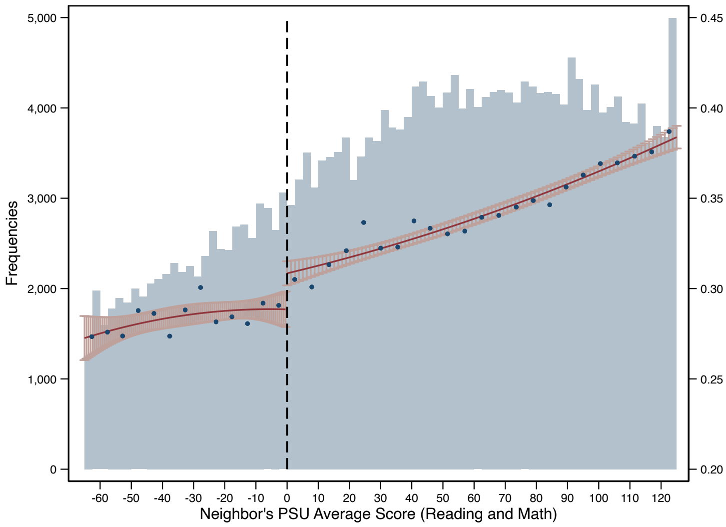


Notes: This figure illustrates the coefficients obtained when studying siblings' effects using different bandwidths. The dots represent the coefficients, and the lines illustrate 95% confidence intervals.

Figure E.X: Effect of neighbors' eligiblity for student loans on their own and on potential applicants' enrollment (second degree polynomial)



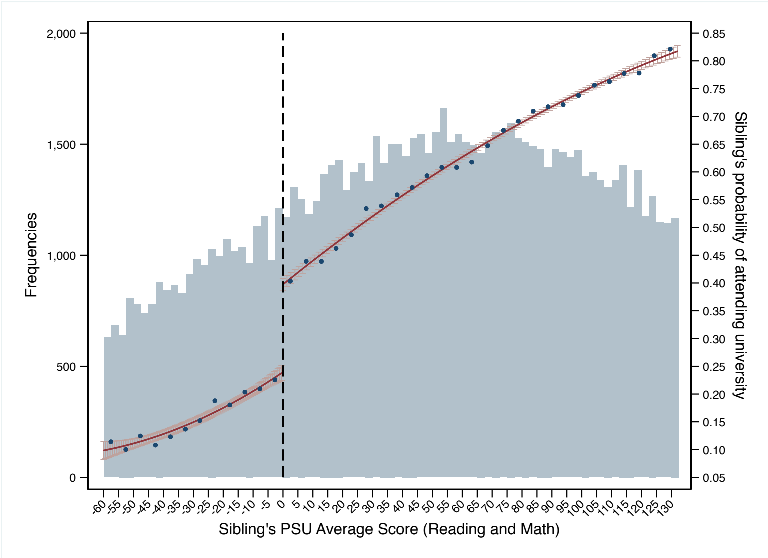
(a) First stage: Neighbors' own probability of going to univeristy



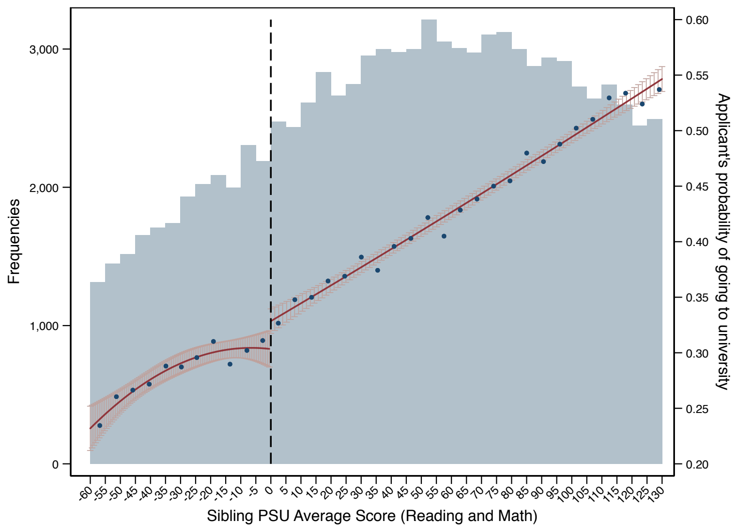
(b) Reduced form: Potential applicants' probability of going to university

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

Figure E.XI: Effect of older siblings' eligiblity for student loans on their own and on potential applicants' enrollment (second degree polynomial)



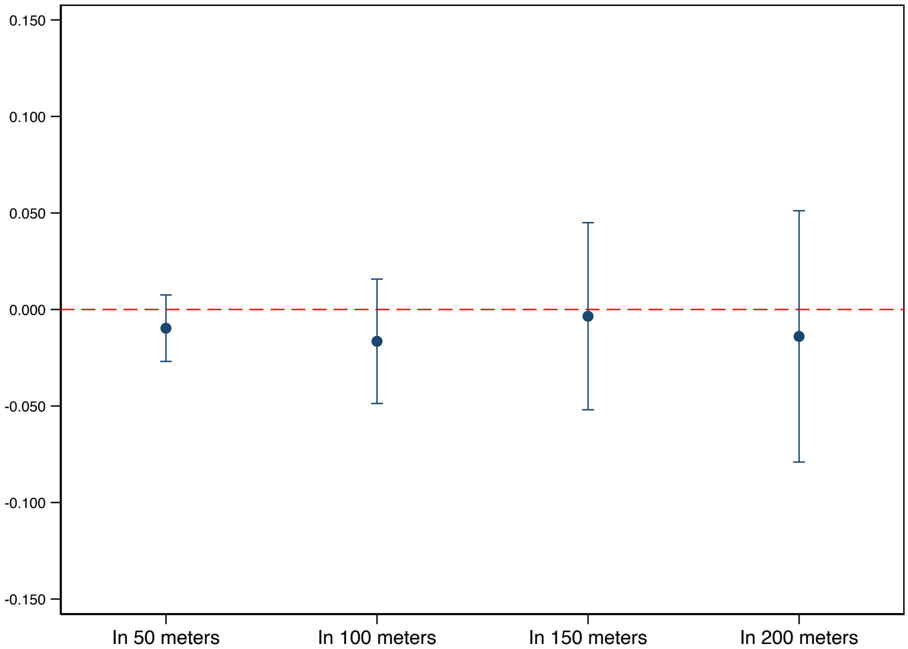
(a) First Stage: Siblings' Probability of going to Univeristy



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

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

Figure E.XII: Effect of neighbor's eligibility for student loans on the number of potential applicants registered for the PSU in $t + 1$



Notes: This figure illustrates the effect of the closest neighbor eligibility for funding on the number of potential applicants registered for taking the PSU within 50m, 100m, 150m and 200m. The dots illustrate the coefficients and the bars 95% confidence intervals. Standard errors are clustered at the neighborhood unit level. Each coefficient was independently estimated and optimal bandwidths were computed following Calonico et al. (2014b).

Table E.I: Placebo effect of neighbors on potential applicants' outcomes (Neighbors not applying for student loans)

	Attends university (1)	Takes the PSU (2)	Applies for financial aid (3)	PSU score | Taking the PSU (4)	High school GPA (5)
Neighbor scores above student loans cutoff (t-1)	-0.003	0.004	0.003	0.039	0.002
	(0.006)	(0.004)	(0.006)	(1.316)	(0.008)
First Stage (neighbor enrolls in university)	0.007	0.007	0.007	0.007	0.007
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Year fixed effects	Yes	Yes	Yes	Yes	Yes
N. of students	142311	142311	142311	129027	135354
PSU Polynomial	1	1	1	1	1
Bandwidth	(67.34-56.65)	(67.34-56.65)	(67.34-56.65)	(67.34-56.65)	(67.34-56.65)
Counterfactual mean	0.33	0.89	0.59	21.73	5.54

Notes: The table presents the estimated effects of neighbors scoring above the financial aid threshold on potential applicants' enrollment in university (column 1), probability of taking the PSU (column 2), probability of applying for financial aid (column 3), performance in the PSU (column 4), and performance in high school (column 5). The sample only includes older neighbors not applying for financial aid; thus, scoring above the student loans eligibility threshold does not change their enrollment status. All specifications include a linear polynomial of the PSU which slope is allowed to change at the cutoff. Optimal bandwidths are used in all the specifications. In parenthesis, standard errors clustered at neighborhood unit level.

Table E.II: Effects of close neighbors' eligibility for student loans on distance to the closest potential applicant registered for the PSU in $t+1$

	Distance to closest potential applicant
Neighbor eligible for student loans	0.0015
	(0.002)
Running variable polynomial	Yes
Bandwidth	(46.75-76.39)
N. of students	73208
Outcome mean	0.099

Notes: The table presents results for a specification that studies how the distance to the closest potential applicant registered for the PSU changes at the cutoff. It controls for a linear polynomial of PSU which slope is allowed to change at the cutoff. Optimal bandwidths computed according to Calonico et al. (2014b) are used. In parenthesis, standard errors clustered at neighborhood unit level.

Table E.III: Effect of neighbors on potential applicants' university enrollment (Metropolitan Region of Santiago)

	2SLS-1 (1)	2SLS-2 (2)	CCT-1 (3)	CCT-2 (4)
Neighbor goes to university (t-1)	0.114	0.135	0.119	0.143
	(0.041)	(0.052)	(0.060)	(0.069)
First stage coefficient	0.160	0.154	0.161	0.161
	(0.010)	(0.013)	(0.013)	(0.017)
Reduced form coefficient	0.018	0.021		
	(0.006)	(0.008)		
Year fixed effects	Yes	Yes	Yes	Yes
N. of students	97,104	174,469	97,104	174,469
PSU Polynomial	1	2	1	2
Bandwidth	(54.48-63.47)	(70.30-124.78)	(54.48-63.47)	(70.30-128.78)
Kleibergen-Paap F statistic	236.33	142.19		
Outcome mean	0.29	0.30	0.29	0.30

Notes: The table presents the results of analysis similar to those presented in table II of the paper but only focusing in the Metropolitan Region of Santiago. 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 the robust approach suggested by Calonico et al. (2014b). Optimal bandwidths are used in all the specifications. In parenthesis, standard errors clustered at neighborhood unit level.

Table E.IV: Effect of neighbors on potential applicants' university enrollment (Different clustering levels)

	Neighborhood Unit	Closest Neighbor	Potential Applicant's High School	Potential Applicant's Municipality
	(1)	(2)	(3)	(4)
Neighbor goes to university (t-1)	0.104	0.104	0.104	0.104
	(0.031)	(0.032)	(0.028)	(0.028)
Reduced form coefficient	0.019	0.019	0.019	0.019
	(0.005)	(0.006)	(0.005)	(0.005)
First stage coefficient	0.178	0.178	0.178	0.178
	(0.009)	(0.009)	(0.009)	(0.009)
Year fixed effects	Yes	Yes	Yes	Yes
N. of students	97,104	174,469	97,104	174,469
PSU Polynomial	1	1	1	1
Bandwidth	(49.14-64.35)	(49.14-64.35)	(49.14-64.35)	(49.14-64.35)
Kleibergen-Paap F statistic	449.63	396.47	1077.31	426.04
Outcome mean	0.31	0.31	0.31	0.31

Notes: The table presents the results of analysis similar to those presented in Table II but using different statistic inference approaches. Columns 1 replicates the main results and clusters at the neighborhood unit level, column 2 clusters at the closest neighbor level, column 3 at the potential applicant's high school level, and column 4 at the potential applicant's municipality level.

F Discontinuities at the Scholarships Eligibility Threshold

The main results of the paper exploit variation generated by eligibility for student loans. As explained in section 2 of the paper, to be eligible for a student loan individuals need an average score of 475 or more in the PSU (average between reading and math). Apart from student loans, the government offers a variety of scholarships. Eligibility for most of them depends on an eligibility rule similar to the one used for student loans. The main difference is that the cutoff that determines eligibility for scholarships is higher (i.e., 550).[9] This means that individuals marginally missing the scholarships cutoff are still eligible for student loans. Thus, crossing the scholarships cutoff changes the generosity of the subsidy for which individuals are eligible, but not their overall eligibility for government funding.

Since Chilean universities have complete freedom to define their tuition fees, the government sets a reference tuition fee for each program and institution that defines the maximum amount of funding that a student can receive from the government.[10] At the university level, the reference tuition fee covers around 80% of the actual fee. This means that students need to cover the additional 20% using their own resources, by taking a private loan, or by applying to scholarships offered at their higher education institutions if available.

In this section, I first study how crossing the scholarship eligibility threshold affects older neighbors' and older siblings' own outcomes. Then, as in the main body of the paper, I study whether it affects potential university applicants as well.

Figure F.I illustrates reduced form results for neighbors. Panel (a) indicates that neighbors eligible for a scholarship rely significantly less in student loans to fund their university studies. Some of them still use student loans, but since part of their funding is a scholarship, they are likely to accumulate a smaller debt. However, as shown in Panel (b) this change in the generosity and structure of the funding does not affect neighbors' own enrollment in university. This result is not surprising. If the expected returns to university studies accounting for the costs of student loans are positive, then crossing the scholarships eligibility threshold should not affect enrollment. Panels (c) and (d) focus instead on potential applicants' outcomes. They show that having an older neighbor marginally eligible for a scholarship does not affect potential applicants' probability of

[9] 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 of the support they offer.

[10] The only exception to this rule is given by the CAE. In this case, students can still receive at most an amount equal to the reference tuition fee through the CAE loan, but they can use it to complement scholarships or the FSCU loan, up to the actual tuition fee.

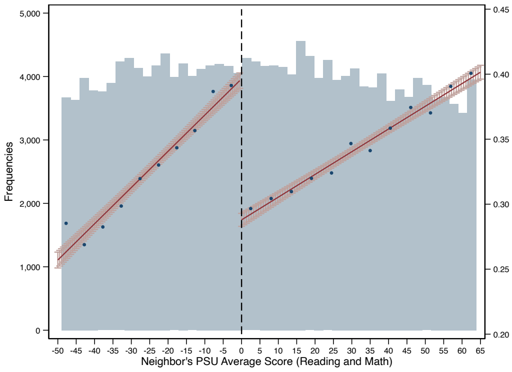
applying for financial aid or of attending university. This finding is consistent with the idea that learning about funding opportunities alone does not change enrollment.

Figure F.II replicates these results, but this time focusing on siblings. Panel (a) of Figure F.II shows that older siblings eligible for a scholarship are less likely to use a student loan to pay for their studies. Despite the change they experience in the generosity and structure of the funding, crossing the scholarships threshold does not make them more likely to enroll in university (Panel (b)). When focusing on the outcomes of potential university applicants (i.e., younger siblings), I find that having an older sibling eligible for a scholarship does not affect potential university applicants' applications for funding or enrollment at university (See Panels (c) and (d)). This seems to suggest that learning about funding opportunities alone does not change enrollment decisions of younger siblings.

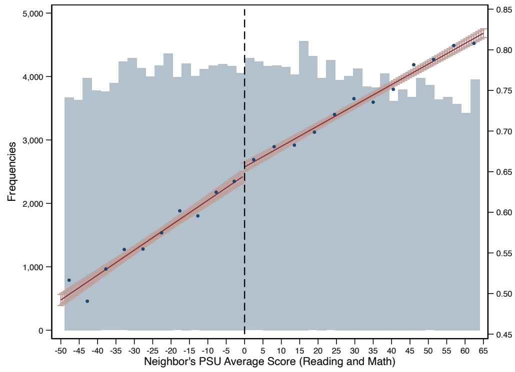
To further explore the role of funding on the effects documented in the main body of the paper, I present next an analysis studying how the responses vary depending on potential applicants' eligibility for student loans. Considering that that student loan eligibility is a potential outcome of the treatment this exercise has some problems, but it is still useful to shed some light about the drivers of the effect. An additional consideration worth having in mind is that the loan eligibility cutoff is quite low—i.e., percentile 40 of the PSU distribution—and therefore the number of students with scores below this cutoff that are admitted into university quickly decreases. With these caveats in mind, in Table F.I I present the results of an exercise that focus on potential applicants scoring below 0, -10, -20, -30, -40 and -50 (i.e., non-eligible for university funding). The last figure corresponds to the percentile 24 of the PSU distribution.

According to these results even potential applicants who are not eligible for funding are affected by having a neighbor going to university with funding. The effect decreases along columns, but this is not surprising. When we move from columns (1) to (6) we leave out of the sample students close to the percentile 40 of the PSU distribution and we give more importance to students in lower percentiles. Since PSU scores affects university admissions, it is natural to observe that the effects decrease. The fact that even the coefficients in the first columns are smaller than the ones obtained using the full sample is not surprising. First, the restrictions I applied leave out of the sample candidates that are more attractive for universities. Second, universities in Chile are relatively expensive. This means that many individuals need support from government in order to enroll. By focusing on individuals who are not eligible for government funding, we are focusing on a group of students for whom it is more difficult to enroll even if they want to. Even for individuals in this group I find large effects, especially if we consider that the baseline probability of attending university for them is much lower than in the whole sample.

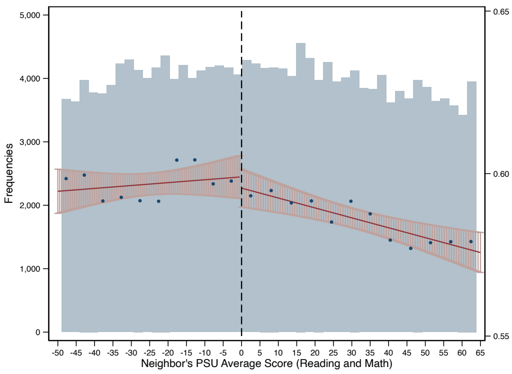
Figure F.I: Changes in neighbor's and applicant's outcomes at the scholarships eligibility cutoff



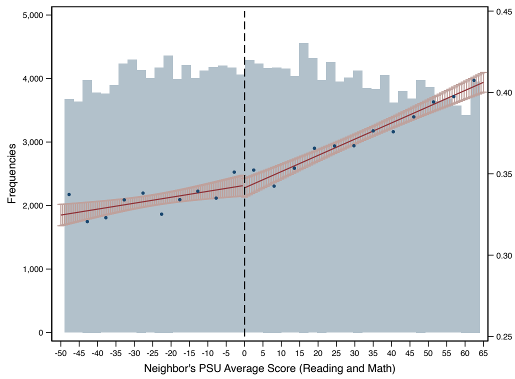
(a) Pr. of using a student loan (n)



(b) Pr. of going to university (n)



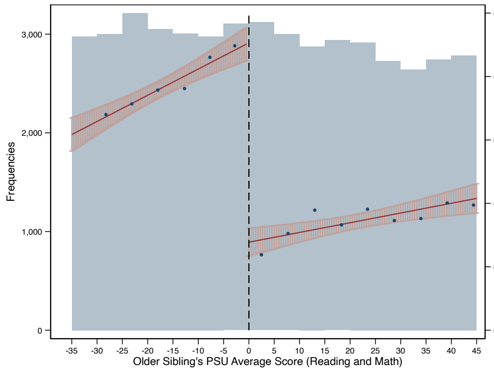
(c) Pr. of applying for financial aid (a)



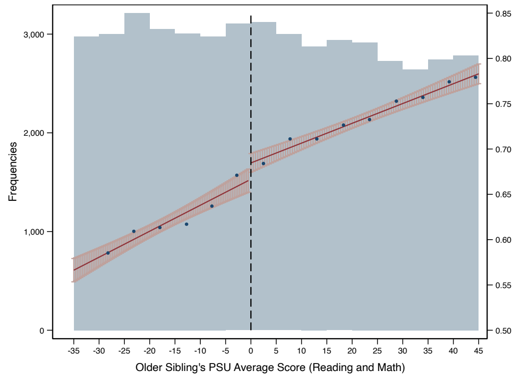
(d) Pr. of going to university (a)

This figure illustrates how neighbor's (n) and potential applicant's (a) outcomes change around the cutoff that defines eligibility for the largest scholarship programs in Chile. This cutoff is higher than the one defining eligibility for student loans, what means that individuals below the scholarship cutoff still qualify for other sources of funding. Panel (a) illustrates the drop in the share of neighbors funding their university studies with student loans at the scholarship threshold, while Panel (b) shows that neighbor's enrollment remains unchanged. Panel (c) illustrates how potential applicants' probability of applying for funding changes when a close neighbor qualify for a scholarship and Panel (d) does something similar but focusing on potential applicants' enrollment probability. Red lines and the shadows in the back of them represent linear polynomials and 95% confidence intervals. Blue dots represent sample means of the dependent variable at different values of neighbors' average score in the PSU.

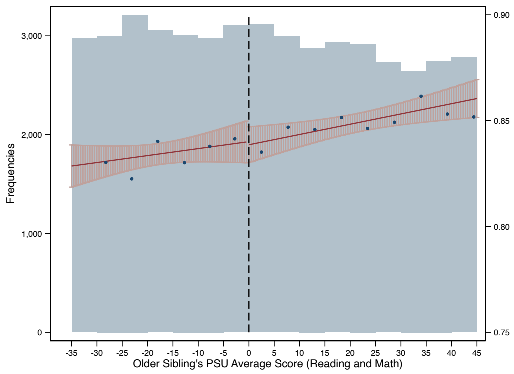
Figure F.II: Changes in older sibling's and applicant's outcomes at the scholarships eligibility cutoff



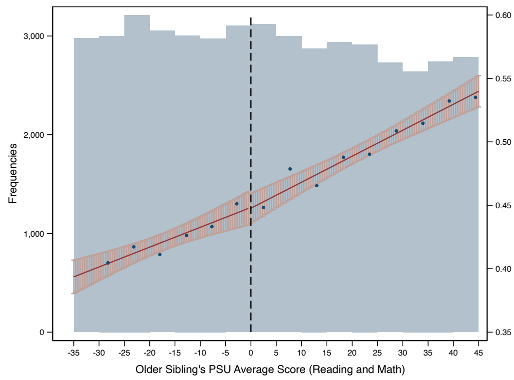
(a) Pr. of using a student loan (o)



(b) Pr. of going to university (o)



(c) Pr. of applying for financial aid (y)



(d) Pr. of going to university (y)

This figure illustrates how older (o) and younger (y) siblings' outcomes change around the cutoff that defines eligibility for the largest scholarship programs in Chile. This cutoff is higher than the one defining eligibility for student loans, what means that individuals below the scholarship cutoff still qualify for other sources of funding. Panel (a) illustrates the drop in the share of older siblings funding their university studies with student loans at the scholarship threshold, while Panel (b) shows that older siblings' enrollment remains unchanged. Panel (c) illustrates how younger siblings' probability of applying for funding changes when their older sibling qualifies for a scholarship, and Panel (d) does something similar but focusing on younger siblings' enrollment probability. Red lines and the shadows in the back of them represent linear polynomials and 95% confidence intervals. Blue dots represent sample means of the dependent variable at different values of older siblings' average score in the PSU.

Table F.I: Neighbor Effects on University Enrollment for Potential Applicants Non-Eligible for Funding

	Potential Applicant's Score					
	< 0	< -10	< -20	< -30	< -40	< -50
	(1)	(2)	(3)	(4)	(5)	(6)
Closest neighbor enrolls in university (t-1)	0.046	0.046	0.041	0.050	0.032	0.015
	(0.022)	(0.022)	(0.023)	(0.024)	(0.023)	(0.022)
Closest neighbor is eligible for funding (t-1)	0.008	0.008	0.007	0.008	0.005	0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
First stage	0.181	0.180	0.176	0.167	0.166	0.171
	(0.010)	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)
Observations	65,444	60,376	50,512	51,609	48,829	46,040
Kleibergen-Paap Wald F-Statistic	316.75	274.35	223.01	204.08	213.80	227.63
Outcome mean	0.07	0.06	0.06	0.05	0.05	0.04

Notes: The table presents the estimated effects of neighbors on potential applicants' university enrollment. All specifications control for a linear polynomial of the running variable which slope is allowed to change at the cutoff. Optimal bandwidths computed according to Calonico et al. (2014b) are used in all the specifications. In parenthesis, standard errors clustered at neighborhood unit level.

G Urban Segregation and Inequality in University Enrollment

As discussed in section 3 of the paper, access to university is very unequal in Chile. Given the high levels of urban segregation in the country, this also translates into spatial inequality. The map in Figure G.I illustrates this for Santiago, Chile's capital city. Figures G.II and G.III present similar maps for Valparaíso and Concepción, the two major cities of the other regions studied in the paper.

Since I do not have a formal definition of neighborhood, in order to create these areas I use a k-cluster algorithm to classify individuals according to their geographic coordinates in 1150 clusters (i.e., an average of 10 neighborhoods per each municipality). Then, using university attendance rates of individuals that could have gone to university before the first cohort of potential university applicants in my sample, I classify these areas in three groups.[11] The red areas in the maps correspond to neighborhoods where on average 33.0% of potential applicants go to university, yellow areas to neighborhoods where on average 52.2% of individuals go to university, and green areas to neighborhoods where more than 72% of potential applicants go to university.

The results discussed in the main body of the paper indicate that 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 more than 10% of their direct effect. In order to estimate the full extent of these indirect effects, we would need to investigate whether they also emerge among other peers[12] 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. Although, the results presented in section I of this document suggest that at least in the case of neighbors, these effects quickly decay with time.

So far, the analyses have assumed that direct and indirect effects are constant across different areas. However, they may change depending on the level of exposure to individuals going to university. To investigate this in greater detail, I estimate the direct and indirect effect of student loans independently for low, mid and high exposure neighborhoods.

Figure G.IV presents the results of this exercise. The top panel shows the first stage

[11] The cohorts used to build the measures of attendance are not included in the main analyses of the paper because these old cohorts did not have the main loan program available. Thus, I do not have exogenous variation on their university enrollment.

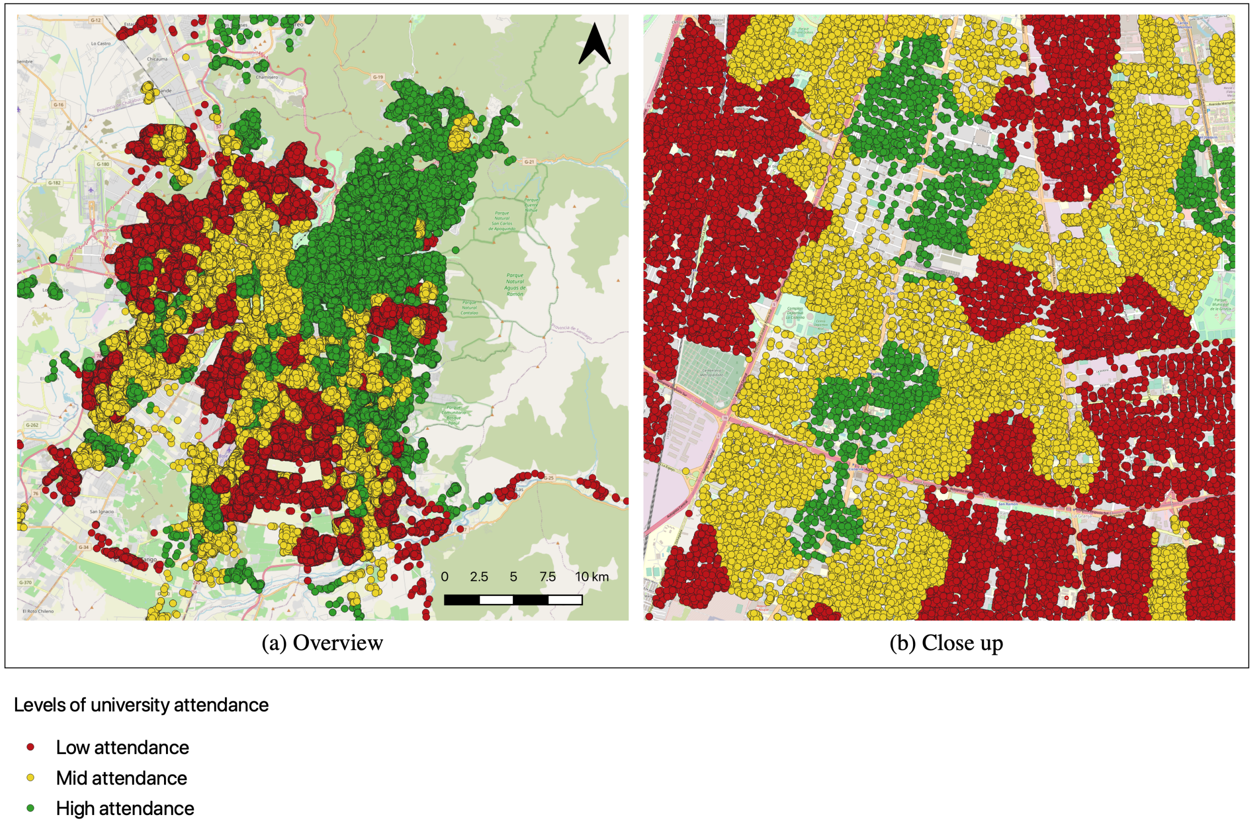
[12] According to the results discussed in section 5.2 of the paper, in the context of neighbors these spillover seem to be very local. Section 5.3 in the paper shows that similar indirect effects arise between siblings.

estimates, the middle panel the reduced form estimates, and the bottom panel the results obtained when combining them to obtain 2SLS estimates. Under the assumptions discussed in section 4 of the paper, these last estimates capture the effects of neighbors' enrollment on potential applicants' enrollment.

The pattern illustrated in this figure shows that the direct effect (i.e., the share of individuals who take up student loans and go to university) does not change much across the three types of neighborhoods. However, the reduced form results and the 2SLS estimates seem stronger in low and mid attendance areas. Indeed, in high attendance areas these coefficients are small and not statistically different from 0.

Although the standard errors of these estimates do not allow me to conclude that they are statistically different, these results show that indirect effects are relevant in low and mid attendance areas. This suggests that in areas where university attendance is relatively low, policies expanding university access would not only affect their direct beneficiaries, but also other individuals living close to them.

Figure G.I: University attendance across neighborhoods in Santiago

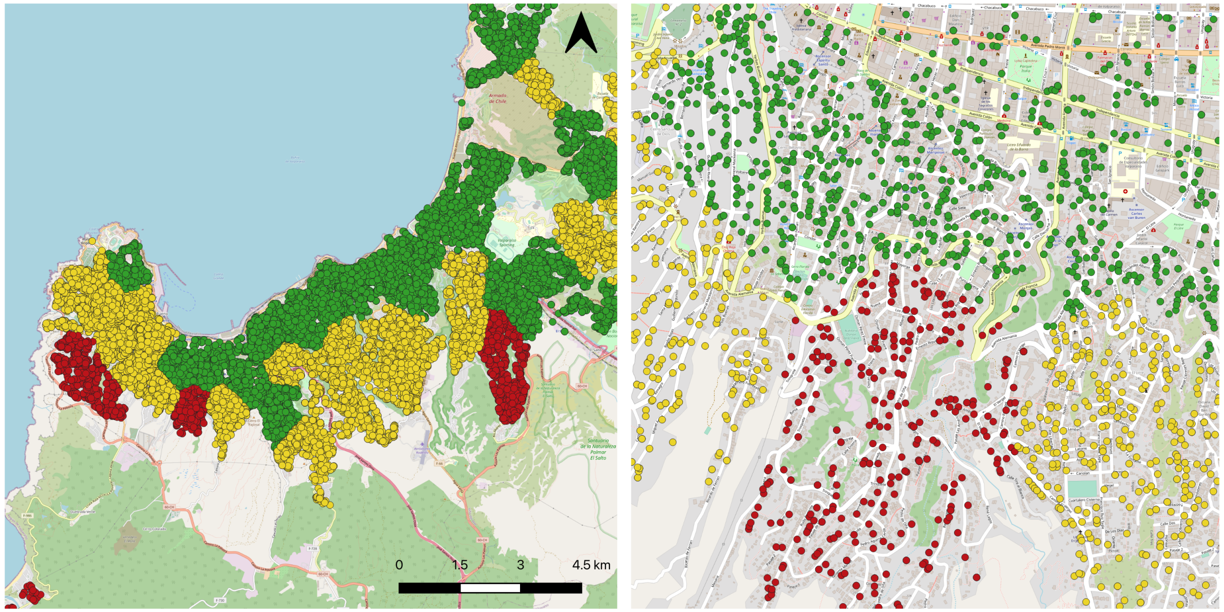


(a) Overview

(b) Close up

Notes: The figure illustrates potential applicants from Santiago and classifies them in three groups according to the share of individuals from their neighborhood going to university before the first cohort of potential applicants that I observe in my sample. The figure in the left presents an overview of the whole city, while the figure in the right zooms in around a specific area. In red neighborhoods average university attendance is 33.0%, in yellow neighborhoods 52.2% and in green neighborhoods 72.2.1%. Neighborhoods were defined using a k-cluster alorithm that grouped individuals according to the geographic coordinates of their household addresses in 1,150 clusters (i.e., on average 10 per municipality in the sample)

Figure G.II: University attendance across neighborhoods in Valparaíso and Viña del Mar



(a) Overview

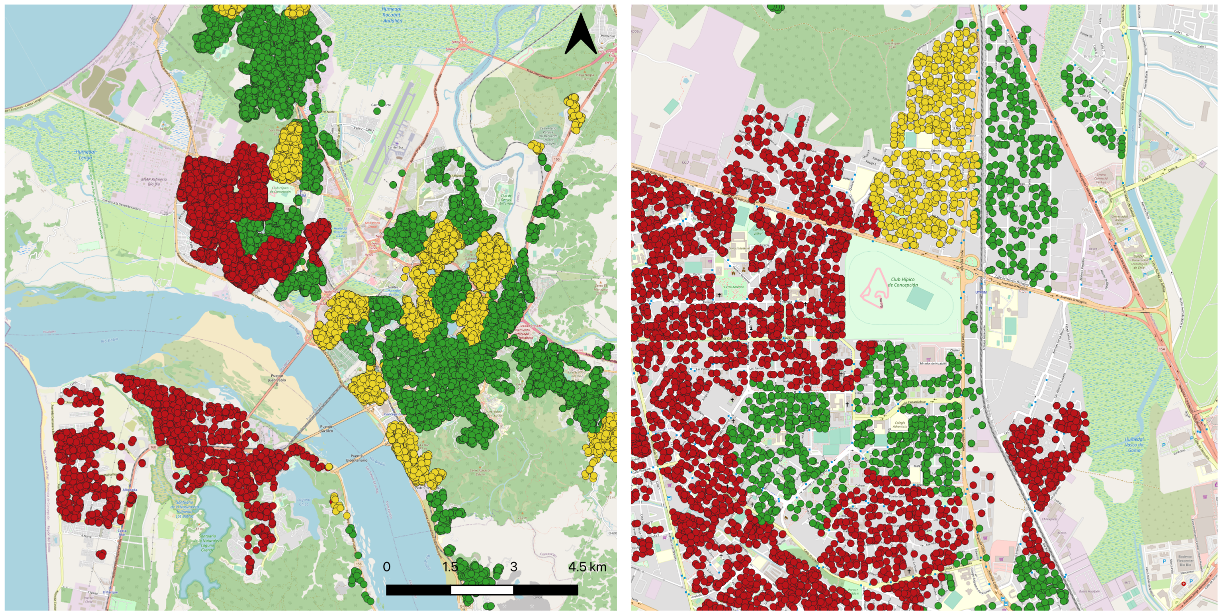
(b) Close up

Levels of university attendance

- Low attendance
- Mid attendance
- High attendance

Notes: The figure illustrates potential applicants from Valparaíso and Viña del Mar and classifies them in three groups according to the share of individuals from their neighborhood going to university before the first cohort of potential applicants that I observe in my sample. The figure in the left presents an overview of the whole city, while the figure in the right zooms in around a specific area. In red neighborhoods average university attendance is 33.0%, in yellow neighborhoods 52.2% and in green neighborhoods 72.2.1%. Neighborhoods were defined using a k-cluster alorithm that grouped individuals according to the geographic coordinates of their household addresses in 1,150 clusters (i.e., on average 10 per municipality in the sample)

Figure G.III: University attendance across neighborhoods in Concepción and Talcahuano



(a) Overview

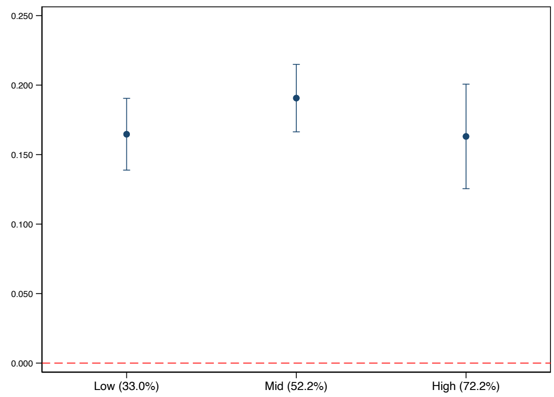
(b) Close up

Levels of university attendance

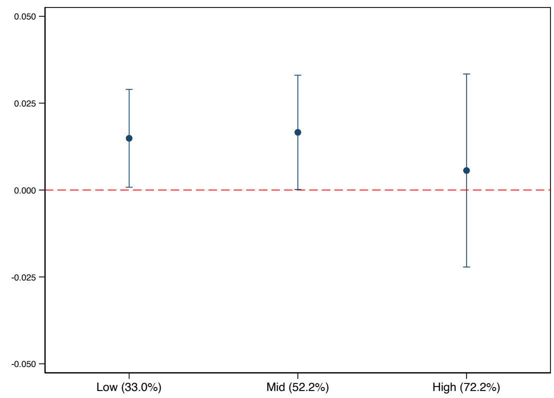
- Low attendance
- Mid attendance
- High attendance

Notes: The figure illustrates potential applicants from Concepción and Talcahuano and classifies them in three groups according to the share of individuals from their neighborhood going to university before the first cohort of potential applicants that I observe in my sample. The figure in the left presents an overview of the whole city, while the figure in the right zooms in around a specific area. In red neighborhoods average university attendance is 33.0%, in yellow neighborhoods 52.2% and in green neighborhoods 72.2.1%. Neighborhoods were defined using a k-cluster alorithm that grouped individuals according to the geographic coordinates of their household addresses in 1,150 clusters (i.e., on average 10 per municipality in the sample)

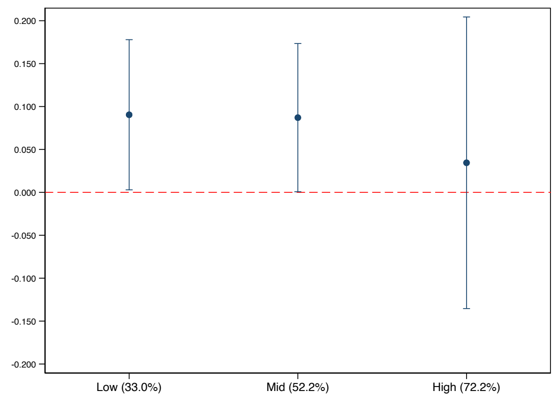
Figure G.IV: Neighbors' effects on potential applicants university enrollment by attendance level in the neighborhood



(a) First Stage



(b) Reduced Form



(c) 2SLS Estimates

Notes: The figure illustrates how neighbors' effects evolve depending on the level of university attendance of the neighborhood of potential applicants before they decide whether or not to apply. The dots represent coefficients from three different samples: low, mid and high attendance neighborhood. The lines represent 95% confidence intervals. The specification used for this exercise controls for a linear polynomial of the running variable which slope is allowed to change at the cutoff. The bandwidth correspond to optimal bandwidths computed according to Calonico et al. (2014b) for the whole sample. Standard errors are clustered at the neighborhood unit level.

H Other Heterogeneity Analyses

This section extends the heterogeneity analyses presented in the paper.

First, I study if the effects differ by potential applicants' household income, high school track, and gender. The table also looks at heterogeneous effects depending on the difference in academic potential between potential applicants and their closest neighbors. The difference in academic potential is computed using GPA in grade 9.[13] According to the results in Table H.I, potential university applicants from households with very low monthly incomes are less responsive than those coming from middle income households. Indeed, potential university applicants from households with monthly incomes between CLP 270,000 and CLP 834,000 seem to be the ones driving the effects.[14]. There are not many potential applicants from the top income category in my estimation sample, which results in very imprecise estimates for this category. Potential applicants in the vocational track of high school seem less responsive than those in the academic track. This suggests that potential applicants who are better prepared for the PSU and for university in general are more likely to successfully respond. There are no major differences by gender, and when looking at academic potential the effects seem slightly larger when potential applicants perform better than the neighbors in high school.

Second, I expand the analyses of heterogeneity by potential applicants' and neighbors' gender. According to the results in Table H.II, independently of their gender, potential university applicants seem to be more responsive to male than to female neighbors. This difference is more clear for male potential university applicants, who are 10 pp more likely to follow a male than a female neighbor. The difference for female potential applicants is smaller (i.e., 3 pp) and not statistically significant.

I conclude this section by studying whether the influence of older siblings on potential applicants depends on the age difference between siblings. To study this I split the sample in to groups of similar size. The first one includes siblings who were born no more than four years apart, while the second includes siblings who were born between four and twelve years apart. Table H.III indicates that the effects are very similar for both groups of siblings. If anything, the effect seems larger for siblings with larger age differences. This is the group of siblings less likely to attend university at the same time.

[13] I do not use the GPA in grade 12 because it could be affected by learning that a close neighbor enrolls in university. Students' grades in high school depend on their teachers and on grade policies within establishments. Considering that only 6% of potential applicants attend their closest neighbor's high school, their GPA are not directly comparable. Unfortunately, I do not observe any standardized measure of ability that could be used in this exercise.

[14] This income range is equivalent to around USD 280 to USD 1170 in 2015

Table H.I: Heterogeneity in the effects of closest neighbor on potential applicants' university enrollment

	Household income			High school track		Gender		Difference in academic ability	
	$\leq$ CLP 270K	CLP 270K - CLP 834K	$>$ CLP 834K	Academic	Vocational	Male	Female	≥ 0	< 0
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Neighbor goes to university (t-1)	0.051	0.234	-0.050	0.108	0.067	0.100	0.109	0.117	0.083
	(0.030)	(0.061)	(0.148)	(0.044)	(0.030)	(0.040)	(0.043)	(0.047)	(0.034)
Reduced form	0.010	0.037	-0.007	0.019	0.012	0.018	0.020	0.020	0.016
	(0.006)	(0.009)	(0.021)	(0.008)	(0.005)	(0.007)	(0.008)	(0.008)	(0.006)
First stage	0.192	0.160	0.142	0.179	0.176	0.175	0.181	0.169	0.188
	(0.009)	(0.011)	(0.030)	(0.010)	(0.011)	(0.010)	(0.010)	(0.011)	(0.011)
N. of potential applicants	84689	48512	11507	86445	58279	77385	67339	74347	70377
Kleibergen-Paap F statistic	422.13	203.06	21.73	325.32	260.81	331.00	343.61	251.97	312.32
Outcome mean	0.22	0.39	0.67	0.45	0.12	0.31	0.32	0.43	0.19

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 household income of potential applicants. Columns 4 and 5 do the same, but distinguishing by the high school track followed by potential applicants. Columns 6 and 7 look at heterogeneous effects by gender. Finally, columns 8 and 9 look at heterogenous effects depending on the difference in grade 9 gpa between potential applicants and their closest neighbor. All specifications include years fixed effects and a linear polynomial of the closest neighbor or sibling PSU score which slope is allowed to change at the cutoff. Bandwidths are the same used in Table II. In parenthesis, standard errors clustered at neighborhood unit level.

Table H.II: Effects of close neighbors on potential applicants' university enrollment by gender

	Older Neighbor: Female		Older Neighbor: Male	
	Potential Applicant:		Potential Applicant	
	Female	Male	Female	Male
	(1)	(2)	(3)	(4)
Neighbor goes to university (t-1)	0.091	0.077	0.121	0.171
	(0.045)	(0.049)	(0.088)	(0.082)
Reduced form	0.018	0.015	0.016	0.027
	(0.009)	(0.010)	(0.011)	(0.013)
First Stage	0.201	0.198	0.131	0.155
	(0.013)	(0.013)	(0.015)	(0.016)
Year fixed effects	Yes	Yes	Yes	Yes
N. of students	45942	39741	31443	27598
Bandwidth	(49.09-64.35)	(49.09-64.35)	(49.09-64.35)	(49.09-64.35)
Kleibergen-Paap F statistic	244.16	244.99	71.75	100.01
Outcome mean	0.31	0.32	0.31	0.32

Notes: The table presents results for specifications that study the effect of close neighbors on potential applicants' university enrollment depending on gender. All specifications include a linear polynomial of PSU which slope is allowed to differ at both sides of the cutoff. Bandwidths are the same used in Table II. In parenthesis, standard errors clustered at neighborhood unit level.

Table H.III: Effects of older siblings on potential applicants' university enrollment by age difference

	Age Difference < 5 (1)	Age Difference ≥ 5 (2)
Older sibling goes to university (t-1)	0.114	0.136
	(0.070)	(0.067)
Reduced form	0.018	0.025
	(0.012)	(0.013)
First Stage	0.157	0.181
	(0.012)	(0.012)
Year fixed effects	Yes	Yes
N. of students	28615	29098
Bandwidth	(37.0-74.5)	(37.0-74.5)
Kleibergen-Paap F statistic	175.47	220.43
Outcome mean	0.36	0.38

Notes: The table presents results for specifications that study the effect of older siblings on potential applicants' university enrollment depending on age difference. All specifications include a linear polynomial of the older sibling's PSU which slope is allowed to change at the cutoff. Bandwidths are the same used in Table VII. In parenthesis, standard errors clustered at family level.

I Other Neighbors Definitions

The results discussed in section 5 of the paper focus on the closest neighbor applying to university one year before the potential university applicant. However, there could be other neighbors also affecting potential university applicants' decisions. Here, I expand the results discussed in the paper by looking at the effects of close neighbors applying to university two or more years before, the year before, the same year, one year after, and two or more years after the potential applicant. Figure I.I summarizes these results. As expected, college applications in the future do not affect choices today ($T+1, \geq T+2$). Given the nature of the exploited variation, not finding contemporaneous effects is not surprising either ($T+0$). The shock on the neighbor's education trajectory takes place at a point in the academic year in which the potential applicants have limited ability to respond. When the shock affecting the neighbor takes place one year before the potential applicant could apply, the effects are large and significant ($T+1$). However, they decline and become non-significant when looking at neighbors applying two or more years before. This suggests that age plays a particularly important role in social interactions among young neighbors, but it could also indicate that individuals only pay attention to this type of shocks when they are very close to deciding whether or not to enroll in college.[15]

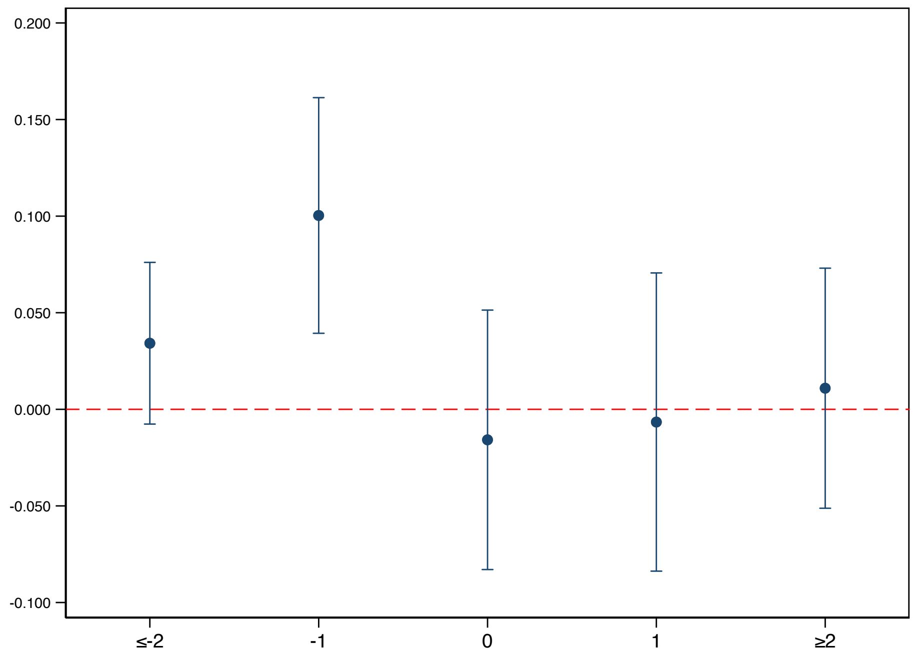
I further investigate how the effects evolve depending on the different definitions of close neighbors. The main specification in the paper focuses on the closest one. Here I look at the effect of the best neighbor applying to university in $T-1$ within different radius (i.e., best within 100m, 125m, 150m, 175m, 200m). The best neighbor is defined as the one for whom the running variable (i.e. average PSU score) takes the highest value.

When implementing these exercises, the sample size decreases with the size of the group being analyzed. The student loans cutoff is relatively low (percentile 40 in the PSU distribution); making it more difficult to find individuals that are at the same time the best of a group and 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 within 100m is close enough to the cutoff could be different from those where the best neighbor within 200m is close to the cutoff. By expanding the radius, the average distance to the neighbor also changes. However, since the composition of the sample is also changing, these results do not tell us much about how neighbors effects evolve with distance. Table I.I presents the results of these analysis. I find effects similar—if anything slightly larger—than the ones documented in the main body of the paper.

[15] Each coefficient comes from an independent sample focusing on potential university applicants and their closest neighbors applying to college in $T \leq -2, T-1, T, T+1, T \geq T+2$. Since for neighbors I only observe applications and enrollment in university between 2006 and 2012, I use a different group of cohorts in each specification.

Figure I.I: Neighbors' effects on potential applicants university enrollment by differences in the application year



Notes: This figure illustrates the effect of the closest neighbor applying to university between two years before and two years after the potential applicant. The dots represent 2SLS coefficients and the bars 95% confidence intervals. As in the rest of the paper standard errors are clustered at the neighborhood unit level. Each coefficient was independently estimated and optimal bandwidths were computed following Calonico et al. (2014b).

Table I.I: Effects of other close neighbors on potential applicants' university enrollment

	Best neighbor within:				
	100m	125m	150m	175m	200m
	(1)	(2)	(3)	(4)	(5)
Neighbor goes to university (t-1)	0.136	0.172	0.125	0.135	0.106
	(0.053)	(0.054)	(0.058)	(0.066)	(0.073)
Reduced form	0.023	0.033	0.026	0.028	0.023
	(0.009)	(0.010)	(0.012)	(0.014)	(0.016)
First Stage	0.173	0.189	0.205	0.210	0.218
	(0.012)	(0.014)	(0.016)	(0.018)	(0.021)
Year fixed effects	Yes	Yes	Yes	Yes	Yes
N. of students	64504	56905	47367	38695	31543
Bandwidth	(49.09-64.35)	(49.09-64.35)	(49.09-64.35)	(49.09-64.35)	(49.09-64.35)
Kleibergen-Paap F statistic	204.19	190.46	169.83	128.98	108.88
Outcome mean	0.30	0.30	0.31	0.32	0.32

Notes: The table presents results for specifications that study the effect of other close neighbors on potential applicants' university enrollment. Column 1 looks at the effect of the best neighbor within 100m, column 2 at the best within 125m, column 3 at the best within 150m, column 4 at the best within 175m and column 5 at the best within 200m. All specifications include a linear polynomial of PSU which slope is allowed to differ at both sides of the cutoff. Bandwidths are the same used in Table II. In parenthesis, standard errors clustered at neighborhood unit level.

J Older Siblings' Expenditure in Higher Education

This section investigates how the household budget constraint is affected when an older sibling becomes eligible for a student loan. As discussed in section 2 of the paper, in Chile universities set their own tuition fees. To control public expenditure the Ministry of Education sets a reference tuition fee that limits the maximum amount of funding that an individual can receive from government. This reference tuition fee is specific to each college and program, and at university level represents roughly an 80% of the actual tuition fees. Thus, even if an individual is eligible for financial aid, families typically have to finance a share of the tuition fees, in addition to study materials, and commuting and living expenses.

Unfortunately, I do not have information on all these costs. I do observe, however, reference and actual tuition fees from 2008 onward. I also observe an additional fee that some institutions charge to their students when they enroll in first year. By combining this information with the registers on funding recipients and higher education enrollment I can study how expenditure in tuition fees changes at the student loan eligibility cutoff. For this analysis, I focus on older siblings who appear in the main estimation sample and apply to higher education after 2007. If they do not enroll in higher education, I assume their expenditure in tuition fees is 0.

Table J.I summarizes the results of this exercise. First, it shows that being eligible for a student loan significantly increases attendance to higher education. It also shows that having access to a student loan for university moves some individuals from vocational higher education to universities. This explains why the effect of student loans on university enrollment is twice their effect on higher education enrollment.

Eligibility for student loans and scholarships to fund vocational higher education does not depend on PSU scores. In this level, most benefits are allocated based on high school performance. This explains why crossing the student loans university threshold results in a small decrease in take up of scholarships. This result reflects that some of the individuals who choose to take up a loan and enroll in university were eligible for scholarships in vocational higher education institutions.

The changes in enrollment decisions discussed in the previous paragraphs result on no significant differences in tuition fees expenditure at the cutoff. If anything, the households of individuals who are eligible for a student loan spend more in tuition fees than the households of individuals who are non-eligible. This difference reflects that individuals to the right of the eligibility threshold are more likely to enroll in higher education, and to attend more expensive institutions (i.e., universities). Although not statistically significant, this difference is likely to represent a lower bound. It ignores all costs apart from tuition fees and to compute it I focused only on the first year of studies. University degrees, however, are longer than vocational higher education degrees which implies that the difference in total expenditure will be larger.

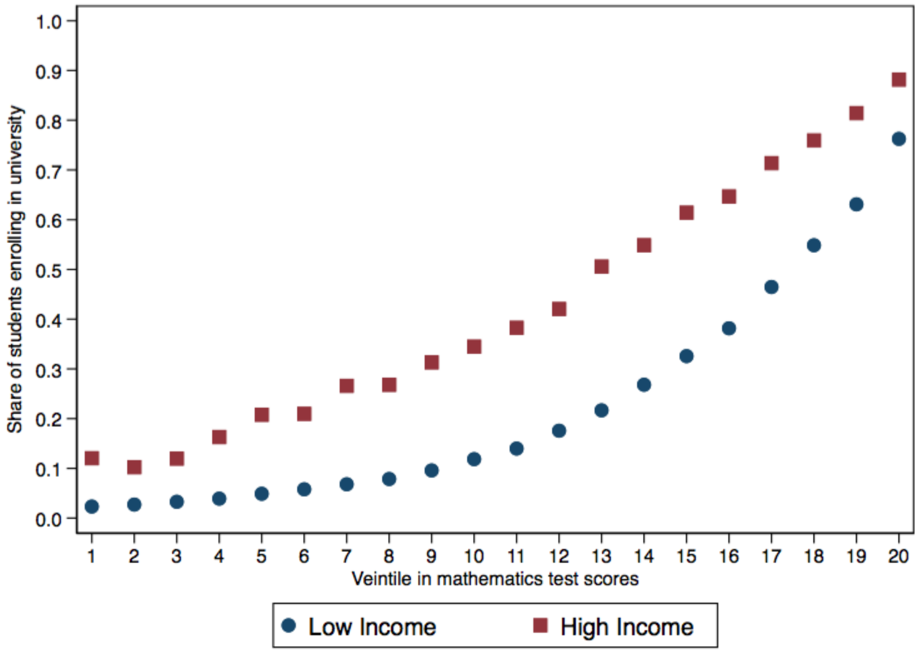
Table J.I: Effect of older siblings' eligibility for funding on older siblings' own enrollment and education expenditure

	Enrolls in higher ed.	Enrolls in vocational higher ed.	Enrolls in university	Takes up a scholarship	Annual expenditure in tuition fees (000 CLP)	Annual expenditure in tuition and enrollment fees (000 CLP)
	(1)	(2)	(3)	(4)	(5)	(6)
Older sibling is eligible for a loan	0.077	-0.066	0.143	-0.031	15.569	25.014
	(0.010)	(0.010)	(0.010)	(0.008)	(20.533)	(21.898)
Observations	37504	37504	37504	37504	37504	37504
Outcome mean	0.69	0.27	0.42	0.17	714.835	815.897

Notes: The table presents estimates of the effect of older siblings' eligibility for university student loans on their own enrollment and on the implied expenditure in tuition and enrollment fees. All specifications control for a linear polynomial of the running variable which slope is allowed to change at the cutoff. Bandwidths are the same used in Table VII. In parenthesis, standard errors clustered at family level.

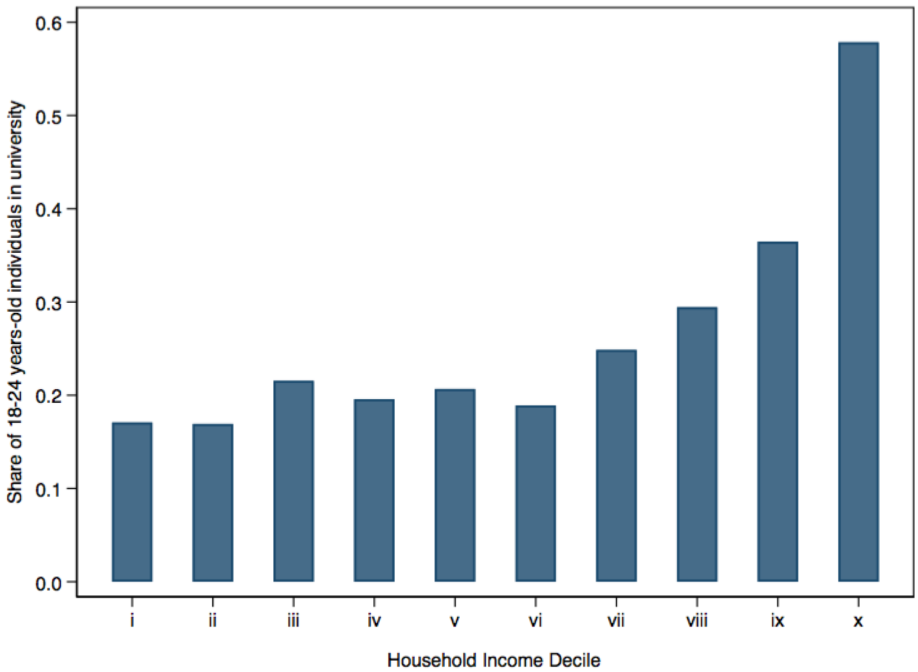
K Inequality in Access to Higher Education in Chile

Figure K.I: 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.

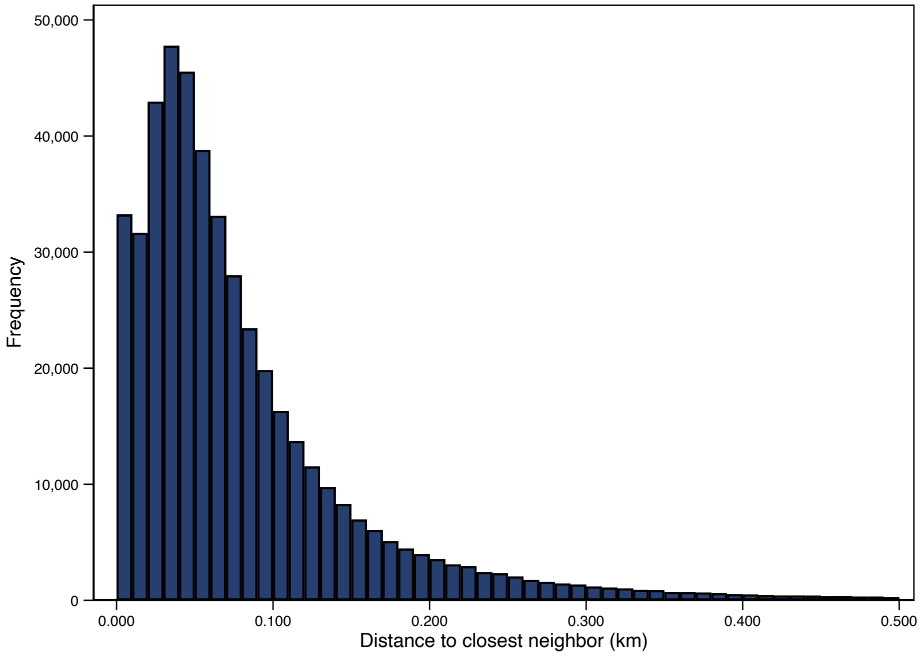
Figure K.II: Share of students going to university by household income (2015)



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).

L Distance to closest Neighbor

Figure L.I: Distribution of distance between potential applicants and their closest neighbor



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

References

Altmejd, A., A. Barrios-Fernandez, M. Drlje, J. Goodman, M. Hurwitz, D. Kovac, C. Mulhern, C. Neilson, and J. Smith (2021). O Brother, Where Start Thou? Sibling Spillovers on College and Major Choice in Four Countries. *The Quarterly Journal of Economics*, forthcoming.

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

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

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

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

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