Siblings' Spillover Effects on College and Major Choice: Evidence from Chile, Croatia and Sweden*

Adam Altmejd   Andrés Barrios-Fernández   Marin Drlje   Dejan Kovac

Christopher Neilson

First Version : November 1st, 2019†
Latest Version : January 13, 2020

Abstract

While it is widely believed that family and social networks can influence important life decisions, identifying causal effects is notoriously difficult. This paper presents causal evidence from three countries that the educational trajectories of older siblings can significantly influence the college and major choice of younger siblings. We exploit institutional features of centralized college assignment systems in Chile, Croatia, and Sweden to generate quasi-random variation in the educational paths taken by older siblings. Using a regression discontinuity design, we show that younger siblings in each country are significantly more likely to apply and enroll in the same college and major that their older sibling was assigned to. These results persist for siblings far apart in age who are unlikely to attend higher education at the same time. We propose three broad classes of mechanisms that can explain why the trajectory of an older sibling can causally affect the college and major choice of a younger sibling. We find that spillovers are stronger when older siblings enroll and are successful in majors that, on average, have higher scoring peers, lower dropout rates and higher earnings from graduates. The evidence presented shows that the decisions, and even random luck, of your close family members and peer network, can have significant effects on important life decisions such as the choice of specialization in higher education. The results also suggest that college access programs such as affirmative action, may have important spillover effects through family and social networks.

Keywords: Sibling Spillovers, College and Major Choice, Peer Effects.
JEL classification: I21, I24.

*We thank Karun Adusumilli, Josh Angrist, Esteban Aucejo, Christopher Avery, Michal Bauer, Randall K. Filer, Sebastian Gallegos, Alan B. Krueger, Jacob N. Shapiro, Peter Blair, Taryn Dinkelman, Joshua Goodman, Jan Hanousek, Kristiina Huttunen, Xavier Jaravel, Stepan Jurajda, Vasily Korovkin, Camille Landais, Erik Lindqvist, Alexandre Mas, Alan Manning, Sandra McNally, Guy Michaels, Daniel Munich, Andreas Menzel, Christian Ochsner, Tuomas Pekkarinen, Steve Pischke, Mariola Pytliková, Daniel Reck, Steven Rivkin, Matti Sarvimäki, Juanna Schrøter Joensen, Johannes Spinnewijn, Anders Stenberg, Janne Tukiainen, Jan Zápal, Kresimir Zigic and Björn Öckert for their many useful comments. We are also grateful to the participants at the CERGE-EI , Helsinki GSE, LSE, Princeton University and SSE internal seminars, and at the Umag Conference 2017 "Economics in a Changing World" and the EALE 2019. Finally, we thank the Ministries of Education of Chile and Croatia, the DEMRE, ASHE (AZVO), Riksarkivet, UHR and SCB for access to their administrative data.

†This paper was previously circulated as two separate studies that were published in two Ph.D. dissertations (Altmejd, 2018; Barrios-Fernandez, 2019).

1 Introduction

The choice of specialization in higher education is one of the most complex and consequential that an individual can make (Altonji et al., 2012; Oreopoulos and Petronijevic, 2013).[1] Despite its importance for future earnings, employment and life trajectories, we know little about how the preferences and the beliefs that drive this decision are formed and if they can be changed. Recent evidence indicates that family background and social context are important in shaping college and major choices (see for instance Hoxby and Avery, 2013), suggesting that relatives and social networks could significantly influence them. However, it is generally very difficult to establish causally whether a shock to one member of the family group would affect others and whether the observed correlation in behavior across social groups is a product of deeper structural differences.

In this paper, we investigate how college applications and enrollment decisions are influenced by the higher education trajectories of one of the most important social peers a person has when growing up: older siblings. Using a regression discontinuity design, we show that younger siblings are significantly more likely to apply and enroll in the same major and college that their older sibling was assigned to. We document this significant within-family spillover effect in three countries with different education systems, culture and levels of economic development: Chile, Croatia, and Sweden.

Establishing the existence of these family spillovers has important policy implications. First, they could help to explain inequality in education uptake and trajectories across families and socio-economic groups. Second, policies that change the pool of students admitted to specific programs and institutions, such as affirmative action, would have an indirect multiplier effect on members of the social network of their beneficiaries. Finally, if the reason why individuals respond to their older siblings' choices is incomplete information, there is scope to improve the match of students and educational programs through information provision.

To causally identify spillover effects, we exploit the fact that all three countries have centralized admission systems that employ a deferred acceptance (DA) mechanisms to allocate applicants to majors depending on their stated preferences and previous academic performance. These selection systems give rise to admission cutoffs in all oversubscribed majors. Taking advantage of the quasi-random variation generated by these cutoffs, we implement a fuzzy Regression Discontinuity Design to investigate how having an older sibling enrolling in a specific major, college or field of study affects individuals' probabilities of applying and enrolling in them.

[1] Average returns to higher education can be substantial, but there is considerable heterogeneity in earnings by both institution and field of study. Growing empirical evidence shows that these differential returns have an important causal component (see for example Hastings et al. (2013); Kirkebøen et al. (2016)), highlighting the relevance of the college and major choice. However, as pointed out by Oreopoulos and Petronijevic (2013), choosing the right institution and field of study can be extremely complex. Optimal decisions are different for each applicant, who in order to make the best decision should be able to anticipate future labor market earnings, the likelihood of completion, and the costs and funding opportunities available.

A key challenge for the identification of peer effects is to distinguish between social interactions and correlated effects. In our setting, correlated effects arise because siblings share genetic characteristics and a social environment. Thus, it is not surprising that their outcomes are correlated. Our empirical strategy compares individuals whose older siblings are marginally admitted or rejected from specific majors. Since these individuals are very similar both in their observable and unobservable characteristics, we can isolate the social interaction effect. In addition, if siblings simultaneously affect each other's decision, the so called reflection problem (Manski, 1993) arises. But since siblings apply and enroll in college sequentially, the lagged structure of their decisions and the fact that the variation that we exploit in older siblings' enrollment comes only from admission cutoffs allow us to abstract from this issue.

Despite the differences that exist between Chile, Croatia and Sweden, we find similar spillover magnitudes in all three countries. Having an older sibling marginally enrolling[2] in their preferred alternative (major-college combination) increases the likelihood of applying there between 1 and 4 percentage points. We also show that Individuals are between 10 and 16 percentage points more likely to apply to the college where their sibling is enrolled, and between 4 and 9 percentage points more likely to enroll there.

The effects that we document are stronger when individuals resemble their older siblings in terms of gender and academic potential. They seem to be driven by individuals whose older siblings "marginally enroll" in relatively selective institutions and persist even when the age difference between siblings makes it unlikely that they will be attending university at the same time.

Our main results are consistent with three broad classes of mechanisms. First, the effects could be driven by a change in the cost of attending specific majors and colleges. Alternatively, they could be driven by changes in individuals' preferences. Finally, the effects could be driven by changes in the choice set of individuals, something that could be triggered by salience or by information transmission. We discuss all of these alternatives, and present suggestive evidence that information is an important driver of our results.

Despite all the research on family and peers effects in education, little is known about how siblings affect human capital investment decisions.[3] Recent evidence shows that older siblings can affect high school related choices. Dustan (2018) uses an approach similar to ours and finds that older siblings' influence the choice of high school in Mexico. Joensen and Nielsen (2018), on the other hand, exploit quasi-random variation induced by a policy change in Denmark and find that siblings affect participation in advanced mathematics and science courses.

Much less is known about the role of siblings in higher education specialization choices. Goodman et al. (2015) investigate the relationship between siblings' college choices in the United States and find

[2]We use the term *marginal enrollment* to highlight the fact that these results come from a fuzzy RD that compares individuals whose older siblings were marginally admitted or rejected from specific majors.

[3]Björklund and Salvanes (2011) and Black and Devereux (2011) review the literature studying the role of family, while Sacerdote (2011) and Sacerdote (2014) review the literature on peers effects in education.

that the correlation between siblings' applications is much stronger than among similar classmates.[4] Barrios-Fernandez (2018) studies spillovers from both neighbors and siblings in the access to university in Chile, and finds that having a close neighbor or sibling going to university increases the probability of reaching this level of education, especially in areas where university attendance is traditionally low. Our paper complements this work by exploiting a different source of variation and by focusing on the choice of college and major, rather than in the decision to attend college. Aguirre and Matta (2019) and Goodman et al. (2019), two contemporaneous working papers, also investigate siblings' spillovers in college choices in Chile and the US and provide similar results.[5]

More generally, this paper also contributes to the literature that studies how individuals choose colleges and majors. This has been an active area of research in recent decades that has investigated the role of costs, information, and, more recently, of some behavioral barriers.[6] This paper adds a new element by analyzing the role of family networks on these choices.

The rest of the paper is organized in seven sections. Section 2 describes the higher education systems of Chile, Croatia and Sweden, Section 3 the data, and Section 4 the empirical strategy and the samples that we use. Section 5 presents the main results and Section 6 places them in the context of previous findings and discusses potential mechanisms. Finally, Section 7 concludes.

[4]In Sociology, Kaczynski (2011) presents a qualitative analysis in line with our findings. She argues that educational experience can decrease the choice set due to fear of competition, but also increase it through transmission of institution-specific knowledge and general encouragement. Shahbazian (2018) studies the correlation of siblings' education choices in Sweden, focusing on gender differences in STEM subjects. He reports a positive association in STEM education, especially for girls.

[5]Our paper was previously circulated as two separate studies that were published in two Ph.D. dissertations (Altmejd, 2018; Barrios-Fernandez, 2019).

[6]The role of funding and liquidity constraints has been investigated by Dynarski (2000), Seftor and Turner (2002), Dynarski (2003), Long (2004), van der Klaauw (2002), and Solis (2017). Misinformation and biased beliefs can also be important determinants of college and major choices Wiswall and Zafar (2015). Hoxby and Avery (2013) show that low-income, high-achieving students do not apply to selective colleges in the US, even if they are likely to be admitted and would receive more generous funding than they receive from the non-selective colleges to which they currently apply. Mismatches in higher education have also been studied by Griffith and Rothstein (2009), Smith et al. (2013), Black et al. (2015) and Dillon and Smith (2017). Hoxby and Turner (2013) find that providing low-income students with targeted information on their college options, the application process and funding opportunities significantly increased their applications and actual enrollment in selective institutions. In the context of Chile, Hastings et al. (2016) and Hastings et al. (2015) respectively show that students are uninformed about the costs and benefits of majors and colleges, and that individuals from lower socioeconomic backgrounds are more likely to choose majors with lower earnings. The latter also shows that providing disadvantaged applicants with information about the labor market outcomes of graduates in different programs changed their applications towards majors with higher net of costs earnings. Similarly, Busso et al. (2017) finds that information on funding and labor market opportunities improves the quality of the majors to which Chilean students apply in comparison to their baseline preferences. However, there is also research indicating that only providing information is not enough to change applicants decisions. Bettinger et al. (2012) finds that a pure information intervention in the US does not increase college applications or enrollment, and Pekkala Kerr et al. (2015) finds that information on labor market prospects of postsecondary education programs does not significantly affect Finnish students' applications or enrollment decisions. Lavecchia et al. (2016); French and Oreopoulos (2017) discuss a host of frictions and behavioral barriers that could explain why some individuals do not take full advantage of educational opportunities. Along this line, Carrell and Sacerdote (2017) argue that college-going interventions work not because of their information component, but because they compensate for the lack of support that disadvantaged students receive from their families and schools.

2 Institutions

This section describes the college admission systems of Chile, Croatia and Sweden, emphasizing the rules that generate the discontinuities that we later exploit to identify spillovers among siblings. Despite the differences that exist among these three countries in terms of size, economic development and inequality (Table 1), a common feature is that a significant share of each countries' universities select students using centralized admission systems that allocate applicants to majors only considering their preferences and previous academic performance. These systems generate sharp admission cutoffs in all oversubscribed programs that we later exploit to identify siblings' spillovers.

Table 1: Differences across Countries

	Chile (1)	Croatia (2)	Sweden (3)
		A. Countries Characteristics	
Population	17,969,353	4,203,604	9,799,186
Area (km^2)	756,700	56,590	447,430
GDP per Capita	$22,688,01	$23,008.21	$48,436.98
GDP Growth (2000-2015)	285.60%	227.47%	185.25%
GINI Index	47.7	31.1	29.2
Human Development Index	0.84	0.827	0.929
Adults w/ Postsecondary Ed.	15.2%	18.3%	34.6%
Main Religious Affiliation	Christian (78%)	Christian (91%)	Christian (69%)
Official Language	Spanish	Croatian	Swedish
		B. University System Characteristics	
Colleges	33/60	49/49	35/36
Majors	1,423	564	2,421
Tuition Fees	Yes	Yes	No
Funding	Student loans and scholarships	Fee waiver when accepting offer*.	NA

Notes: The statistics presented in Panel A come from the World Bank (`https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD`) and from the United Nations (`http://hdr.undp.org/en/data`) websites. All the statistics reported in the table correspond to the values observed in 2015, the last year for which we observe applications in Chile (in Croatia we observe them until 2018 and in Sweden until 2016). The only exceptions are the share of adults with complete postsecondary education and religious affiliation. We only observe these statistic in 2011 for the three countries. The share of adults with complete postsecondary education is computed by looking at the level of education completed by individuals who were at least 25 years old in 2011. In the row "Colleges"the first number refers to colleges selecting students through the centralized admission system, while the second to the total number of colleges in the system. The row "Majors"on the other hand, reports the total number of major-college combinations available for students through the centralized admission system in 2015. (*) Although in Croatia there are tuition fees, all students accepting the offer they receive the first time that they apply to university receive a fee waiver. They only loss the fee waiver if they reject the offer.

2.1 College Admission System in Chile

In Chile, all of the public universities and 9 of the 43 private universities are part of the Council of Chilean Universities (CRUCH).[7] All CRUCH institutions, and since 2012 an additional eight private colleges, select their students using a centralized deferred acceptance admission system that only takes into account students' academic performance in high school and in a college admission exam similar to the SAT (Prueba de Selección Universitaria, PSU).[8] Students take the PSU in December, at the end of the Chilean academic year, but they typically need to register before mid-August.[9] As of 2006, all public and voucher school graduates are eligible for a fee waiver that makes the PSU free for them.[10]

Colleges publish the list of majors and vacancies offered for the next academic year well in advance of the PSU examination date. Concurrently, they inform the weights allocated to high school performance and to each section of the PSU to compute the application score for each major.

With this information available and after receiving their PSU scores, students apply to their majors of interest using an online platform. They are asked to rank up to 10 majors according to their preferences. Places are then allocated using an algorithm of the Gale-Shapley family that matches students to majors using their preferences and scores as inputs. Once a student is admitted to one of her preferences, the rest of her applications are dropped. As shown in panel (a) of Figure 1, this system generates a sharp discontinuity in admission probabilities in each major with more applicants than vacancies.

Colleges that do not use the centralized system have their own admission processes.[11] Although they could use their own entrance exams, the PSU still plays an important role in the selection of their students, mostly due to the existence of strong financial incentives for both students and institutions.[12] For instance, the largest financial aid programs available for university studies require students to score above a certain threshold in the PSU.

The coexistence of these two selection systems means that being admitted to a college that uses the centralized platform does not necessarily translate into enrollment. Once students receive an offer

[7] The CRUCH is an organization that was created to improve coordination and to provide advice to the Ministry of Education in matters related to higher education.

[8] The PSU has four sections: language, mathematics, social sciences and natural sciences. The scores in each section are adjusted to obtain a normal distribution of scores with a mean of 500 and a 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 order to apply to university, individuals need to take the language, and the mathematics sections and at least one of the other sections. Universities set the weights allocated to these instruments for selecting students in each program.

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

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

[11] From 2007, we observe enrollment at all colleges in Chile independent of the admission system they use.

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

from a college they are free to accept or reject it without any major consequence. This also makes it possible for some students originally rejected from a program to receive a later offer. Panel (d) of Figure 1 illustrates how the admission to a major translates into enrollment.

2.2 College Admission System in Croatia

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

As in Chile, NISpVU uses a deferred acceptance admission system that focuses primarily on students' high-school performance and in a national level university exam.[13] The national exam is taken in late June, approximately one month after the end of the Croatian academic year. However, students are required to submit a free-of-charge online registration form by mid-February.

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

Once registered, students are able to submit a preference ranking of up to 10 majors. The system allows them to update these preferences until mid-July. At this point students are allocated to programs based on their current ranking. As in Chile, vacancies are allocated using a Gale-Shapley algorithm, giving rise to similar discontinuities in admission probabilities (Figure 1).

Before the final deadline, the system allows students to learn their position in the queue for each of the majors to which they applied. This information is regularly updated to take into account the changes that applicants make in their list of preferences. In this paper, we focus on the first applications submitted by students after receiving their scores on the national admission test. Since some of them change their applications before the deadline, admission based on these applications does not translate one-to-one into enrollment (Figure 1).[14]

There are two important differences between the Chilean and Croatian systems. First, all Croatian colleges use the centralized admission system and second, rejecting an offer is costly since it invalidates eligibility for the enrollment fee waiver.

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

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

2.3 Higher Education Admission System in Sweden

Almost all higher academic institutions in Sweden are public. Neither public nor private institutions are allowed to charge tuition or application fees. Our data include 40 academic institutions, ranging from large universities to small specialized schools.[15]

Each institution is free to decide which majors and courses to offer, and the number of students to admit in each alternative. As in Chile and Croatia, the admission system is centrally managed and students are allocated to programs using a deferred acceptance admission system.

The Swedish admission system has a few important differences compared to the Chilean and Croatian systems. For one thing, the same system is open to applications to full majors and shorter courses alike. To simplify, we will henceforth refer to all these alternatives as *majors*. Moreover, applicants are ranked by different scores separately in a number of *admission groups*. Their best ranking is then used to determine their admission status.[16] Finally, the Swedish admission system has two rounds. After the first round, applicants learn their admission status and they place in the waiting list for all their applications. At this point, they can decide wether to accept the best offer they have or to wait and participate in a second application round. Their scores and lists of preferences do not change between the two rounds, but the cutoffs might. In this project we focus on the variation generated by the cutoff of the second round. Since some applicants decide to accept the offers they received after the first round instead of waiting for the second round, not all applicants above the second round admission cutoff end receiving an offer. Those who dropout from the waiting list after the first round cannot receive a second round offer, even if their score was above the final admission cutoff. This explains why in the case of Sweden the jump in older siblings' admission and enrollment probabilities is smaller than in the other two countries (see Figure 1).

For each program, at least a third of the vacancies are reserved for the high school GPA admission group. No less than another third is allocated based on results from the Högskoleprovet exam. The remaining third of vacancies are mostly also assigned by high school GPA, but can sometimes be used for custom admission.[17]

Högskoleprovet is a standardized test, somewhat similar to the SAT. Unlike the college admission exams of the other countries, Högskoleprovet is voluntary. Taking the test does not affect admission probabilities in the other admission groups, and therefore never decreases the likelihood of acceptance.

Students can apply to majors starting in the fall or spring semester, and the application occurs

[15] We exclude from our sample small art schools and other specialized institutions with non-standard admission systems.

[16] Admission is essentially determined by a max function of high school GPA and Högskoleprovet score, as compared to a weighted average in Chile and Croatia. In the analysis, we collapse these admission groups and use as our running variable the group-standardized score from the admission group where the applicant performed the best.

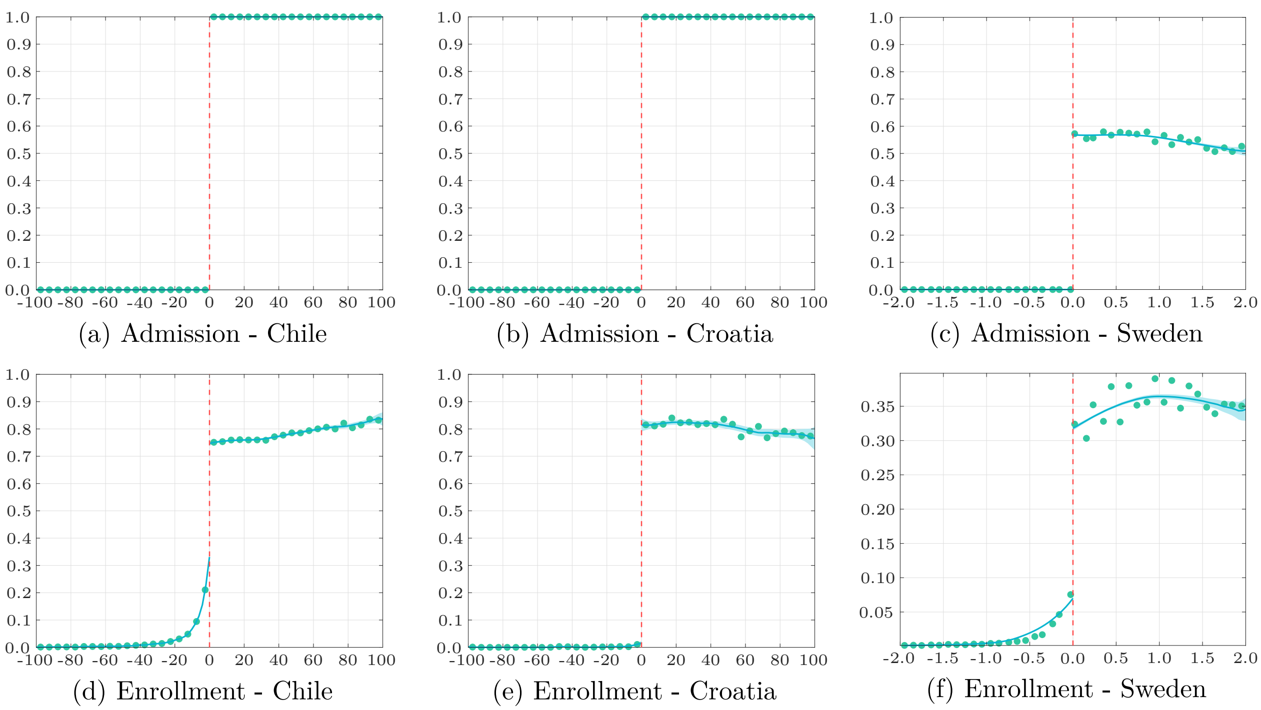
[17] This is the case in some highly selective majors, where an additional test or an interview is sometimes used to allocate this last third of vacancies. We do not include admissions through such groups in our analysis.

in the previous semester. In each application they rank up to 20 alternatives (students were able to rank 12 alternatives until 2005). Full-time studies correspond to 30 credits per semester, but students who apply to both full-time majors and courses in the same application receive offers for the highest-ranked 45 credits in which they are above the threshold.

After receiving an offer, applicants can either accept or decide to stay on the waiting list for choices for which they have not yet been admitted to. Should they decide to wait, admissions after the second round will again only include the highest-ranked 45 ECTS, and all lower-ranked alternatives will be discarded, even those that they were previously admitted to.[18]

Finally, the running variables used in the Swedish admission are far coarser than those in Chile and Croatia. This generates a lot of ties in student rankings. In some cases, ties exactly at the cutoff are broken by lottery.

Figure 1: Older Siblings' Admission and Enrollment Probabilities in Target Major-College at the Admission Cutoff (First Stage)



(a) Admission - Chile (b) Admission - Croatia (c) Admission - Sweden

(d) Enrollment - Chile (e) Enrollment - Croatia (f) Enrollment - Sweden

This figure illustrates older siblings' admission and enrollment probabilities around the admission cutoffs of their target majors in Chile, Croatia and Sweden. Figures (a) and (d) illustrate these probabilities for the case of Chile, figures (b) and (e) for Croatia and figures (c) and (f) for Sweden. Blue lines and the shadows in the back of them represent local linear polynomials and 95% confidence intervals. Green dots represent sample means of the dependent variable at different values of older siblings' own application score.

[18] As in Croatia, we focus on first-round submissions. As many applicants stay on the waiting list for the second round and are admitted to higher ranked alternatives, Sweden has a substantially lower first stage compared to the other two countries.

3 Data

In this paper we exploit administrative data provided by various public agencies in Chile, Croatia and Sweden. In these three countries, the main data sources are the agencies in charge of the centralized college admission system: DEMRE in Chile, NISpVU and ASHE (AZVO) in Croatia, and UHR in Sweden.

From DEMRE we get individual-level data on all the students registered to take the PSU between 2004 and 2015. These datasets contain information on students' performance in high school and in the different sections of the college admission exam. It also contains student-level demographic and socioeconomic characteristics, information on their application, college acceptances through the centralized application system, and college enrollment. To identify siblings, we exploit the fact that when registering for the exam, students provide the national id number of their parents. Using this unique identifier we can match all siblings that correctly reported this number for at least one of their parents.[19]

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

NISpVU and ASHE provided us with similar data for Croatia. These individual registers contain information on students' performance in high school and in the various sections of the college admission exam, and on applications and enrollment at all Croatian colleges between 2012 and 2018. These registers include the home address of students and their surnames, information that we exploit to identify siblings. We define as siblings two individuals if they have the same surname and if they live at exactly the same address at the moment of registration for the college admission exam.

The data for Sweden comes from the Swedish National Archives, the Swedish Council for Higher Education (UHR) and Statistics Sweden (SCB).

The Swedish application data consists of two parts. We get data on applications from the modern system, for the years 2008 to 2016, directly from the Swedish Council for Higher Education (UHR). Applications for the years 1992–2005 are from an older system and are obtained from the Swedish National Archives (Riksarkivet). While the modern system contains the universe of applications to

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

higher education in Sweden, institutions were not required to participate in centralized admissions before 2006.[20] Family connections and all demographic and socioeconomic variables that we use are provided by Statistics Sweden.

Using these data, we identify around $83,000$, $17,000$, and $301,967$ pairs of siblings in Chile, Croatia, and Sweden respectively where the older sibling had at least one active application to an oversubscribed major with an application score within the minimum bandwidth used in each country. Table 2 presents summary statistics for these subsets of siblings and also for the full set of potential applicants.[21]

In the three countries, the sample of siblings is very similar to the rest of the applicants in terms of gender. Individuals with older siblings who already applied to higher education seem slightly younger at application than the rest of the applicants and, not surprisingly, they come from bigger households. Greater differences arise when looking at socioeconomic and academic variables. In Chile and Sweden, where we observe socioeconomic characteristics, the individuals in our sample come from wealthier and more educated households than the rest of the potential applicants. This difference is clearer in Chile, where the "Whole Sample"column consists of all students who registered for the admission exam, irrespective of whether they end up applying to college or not. In Chile and Croatia, we observe that individuals with older siblings applying to university are more likely to have followed the academic track in high school. Finally, in all three countries these individuals perform better in high school and in the college admission test than the rest of the applicants.

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

[20] Institutions with local admission are not included in our data. Most of these programs had special admission groups and would have been excluded from our analysis in any case. The only larger exception is Stockholm University, where admissions to some of the larger programs were managed locally for almost the whole period. It is unlikely that this fact has any strong bearing on our results. The results do not change much qualitatively when the sample is restricted to only include the later period.

[21] In the case of Chile "All potential applicants" includes all students registered for the university admission exam (they do not necessarily take it). In Croatia and Sweden the column includes all students applying to college or higher education respectively.

Table 2: Summary Statistics

	Chile		Croatia		Sweden	
	Siblings Sample	Whole Sample	Siblings Sample	Whole Sample	Siblings Sample	Whole Sample
	(1)	(2)	(3)	(4)	(5)	(6)
			A. Demographic characteristics			
Female	0.521	0.520	0.572	0.567	0.579	0.595
	(0.500)	(0.499)	(0.494)	(0.495)	(0.493)	(0.490)
Age when applying	18.783	19.829	18.878	19.158	20.589	20.872
	(0.604)	(2.484)	(0.621)	(0.963)	(2.374)	(2.562)
Household size[1]	4.782	4.625	2.784	1.925	3.086	2.946
	(1.498)	(1.607)	(1.287)	(1.198)	(1.142)	(1.186)
			B. Socioeconomic characteristics			
High income[2]	0.287	0.128			0.349	0.339
	(0.452)	(0.334)			(0.477)	(0.473)
Mid income[2]	0.398	0.325			0.262	0.290
	(0.490)	(0.469)			(0.440)	(0.454)
Low income[2]	0.315	0.546			0.389	0.371
	(0.464)	(0.498)			(0.488)	(0.483)
Parental ed: < high school	0.094	0.254			0.038	0.056
	(0.292)	(0.435)			(0.191)	(0.229)
Parental ed: high school	0.331	0.386			0.339	0.481
	(0.471)	(0.487)			(0.471)	(0.481)
Parental ed: vocational HE	0.146	0.115			0.067	0.063
	(0.354)	(0.319)			(0.250)	(0.244)
Parental ed: university	0.419	0.234			0.562	0.517
	(0.493)	(0.423)			(0.496)	(0.500)
			C. Academic characteristics			
High school track: academic[3]	0.846	0.673	0.439	0.416		
	(0.361)	(0.469)	(0.496)	(0.496)		
High school: vocational[3]	0.154	0.327	0.561	0.584		
	(0.361)	(0.469)	(0.496)	(0.496)		
Takes admission test	0.956	0.868	0.865	0.835	0.679	0.628
	(0.205)	(0.338)	(0.342)	(0.372)	(0.467)	(0.483)
High school GPA score	-0.080	-0.465	268.373	265.298	0.673	0.432
	(1.231)	(1.357)	(65.766)	(66.600)	(0.766)	(0.773)
Admission test avg. score	0.261	-0.512	312.800	286.247	0.281	-0.061
	(1.283)	(1.708)	(102.568)	(112.787)	(0.991)	(1.000)
Applicants	83,379	2,823,897	16,721	199,475	301,967	3,822,188

Notes: The table present summary statistics for Chile, Croatia and Sweden. Columns (1), (3) and (4) describe individuals in the siblings samples used in this paper, while columns (2), (4) and (6) describe all potential applicants. While in Chile "potential applicants"include all students who register for the admission exam, even if they end not taking it, in Croatia and Sweden the term refers to all students applying to higher education.

[1] In Croatia, *Household Size* only refers to the number of siblings within a household.

[2] In Chile, we only observe income brackets. The High Income category includes households with monthly incomes greater or equal than CLP 850K (USD 2,171 of 2015 PPP); the Mid Income category includes households with monthly incomes between CLP 270K - 850K; and the Low Income category includes households with monthly incomes below CLP 270K (USD 689.90 of 2015 PPP). In Sweden, the High Income category includes households in the top quintile of the income distribution; the Mid Income category includes households in quintiles 3 and 4; and the Low Income category households in quintiles 1 and 2. The average disposable income in the Swedish sibling sample is USD 5,664 (2015 PPP), while in the whole set of applicants USD 5,265 (2015 PPP).

[3] In Croatia, high school academic performance is only available from 2011 to 2015. This sample has 155,587 observations (the corresponding siblings sample has 8,398 observations).

4 Empirical Strategy

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

To overcome these identification challenges, we exploit thousands of cutoffs generated by the deferred acceptance admission (DA) systems that Chilean, Croatian and Swedish universities use to select their students. Taking advantage of the discontinuities created by these cutoffs on admission, we use a Regression Discontinuity (RD) design to investigate how older siblings' admission to their target major affects the probability that their younger siblings will apply and enroll in the same major, college or field of study.[22]

Since individuals whose older siblings are marginally admitted or rejected from a specific major are very similar, the RD allows us to rule out the estimated effects being driven by differences in individual or family characteristics, eliminating concerns about correlated effects. Moreover, considering that the variation that we exploit in the major-college in which older siblings enroll comes only from their admission status and cannot be affected by the choices that their younger siblings will make in the future, we can abstract from the reflection problem.[23]

As discussed in Section 2, rejecting an offer does not have any major consequence for Chilean students. As a result, there is a non-negligible share of applicants who, despite being admitted to a particular college or major, decide not to enroll. Thus, when studying how older siblings' actual enrollment affects their younger siblings, we use a fuzzy RD in which older siblings' enrollment in a specific major is instrumented with an indicator of admission.

We follow a similar approach for Croatia. Although in this setting rejecting an offer is costly, we use a fuzzy and not a sharp RD because, as explained in Section 2, we focus our attention on the first application students submit after receiving their results in the college admission exam. Since

[22]We define a major as a specific combination of major and college. For brevity we refer to this combination simply as major. On the other hand, we define a field of study as the three digit-level ISCED category to which a major belongs. If we consider economics for instance, its ISCED code is 0311. Thus, an individual whose older sibling enrolls in economics at the University of Chile is said to choose the same field of study as her older sibling if she applies in economics (0311) in any college. She is said to choose the same major as her older sibling only if she applies to economics at the University of Chile.

[23]We show that this is indeed the case in a series of placebo exercises that we present in Appendix B.

some individuals modify their applications in the weeks following the exam results, admission to the first set of preferences does not translate one-to-one into enrollment.[24]

For Sweden, we focus our attention on the applications submitted during the first round of the admission process. Since students can reject these offers there is no perfect compliance either.[25] Thus, as in the previous two cases, we also use a fuzzy-RD to identify the siblings' spillovers.

This paper investigates how individuals' probabilities of applying and enrolling in specific majors, colleges and fields of study change when their older siblings are marginally admitted and enroll in them. The basic idea behind our empirical design consists in defining for each major, college and field of study the sample of older siblings marginally admitted and marginally rejected from them, and then compare how this affects their younger siblings' choices. Therefore, each observation in our estimation sample corresponds to a pair of siblings in which the older one is close enough to the admission cutoff of a specific major. Given that in the three countries individuals are allowed to apply to multiple programs, this means that the same pair of siblings could eventually appear several times in the sample.

We define major as a specific combination of major and college, and field of study as the three digit-level ISCED code of these majors.[26] This means that in each country we consider around 80 different fields of study.

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

4.1 Major Sample

This section describes the restrictions applied to the data in order to build the sample used to study how older siblings' marginal admission and enrollment in their target majors affects their younger siblings' probabilities of applying and enrolling in the same major.

As discussed earlier, the assignment mechanism used in Chile, Croatia and Sweden results in cutoff scores for each major with more applicants than available places; these cutoffs correspond to the lowest score among the admitted students. Let c_{jfut} be the cutoff for major j belonging to field of study f in college u in year t. If the major j of field f offered in college u is ranked before the major j' of field f' offered by college u' in student i's preference list, we write $(j, f, u) \succ (j', f', u')$.[27]

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

[25] In addition, in the Swedish setting ties at the cutoff are decided through lotteries. When implementing the RD we modify the score of students at the cutoff by $score - \varepsilon$ for individuals who lose the lottery. We set ε to the minimum computer detectable number.

[26] In the case of Sweden, the definition of major is slightly different. We pool together all the programs in the same field and define a major as the combination of field-institution.

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

Denoting the application score of individual i as a_{ijfut}, we can define marginal students in the major sample as those whose older siblings:

1. listed major j of field f offered in college u as a choice, such that all majors preferred to j had a higher cutoff score than j (otherwise assignment to j is impossible):
   $c_{jfut} < c_{j'f'u't} \; \forall \; (j', f', u') \succ (j, f, u)$.
2. had a score sufficiently close to j's cutoff score to be within a given bandwidth bw around the cutoff:
   $|a_{ijfut} - c_{jfut}| \leq bw$.

This means that in the major sample, the field and college attended by older siblings does not necessarily change by being above or below the admission cutoff. As far as the exact major-college combination in which they are admitted changes, they will be in the sample.

Note that this sample includes individuals whose older siblings were rejected from (j, u) ($a_{ijfut} < c_{jfut}$) and those whose older siblings scored above the admission cutoff ($a_{ijfut} \geq c_{jfut}$). Since the application list in general contains more than one preference, this means that the same individual may belong to more than one major-college marginal group. Figure 1 illustrates the probability of admission and enrollment in a given major around the admission cutoff in Chile, Croatia and Sweden.

4.2 College Sample

In addition to studying the effect older siblings on the choice of major, we study how individuals' probability of applying and enrolling in a specific college changes when an older sibling is marginally admitted and enrolls in that college. The sample used in this case is similar to the one described in the previous section, but in this case we need to add an additional restriction. Thus, we define marginal students in the college sample as those whose older siblings apart from restrictions 1 and 2, also:

3.A. listed major j in college u as a choice, such that majors not preferred to j are dictated by an institution different from u (otherwise being above or below the cutoff would not generate variation in the college attended).

4.3 Field of Study Sample

Finally, we also study how the field of study to which the older siblings' major belongs affects the field of study chosen by younger siblings.

To generate the sample used to study this margin, we follow the same logic behind the creation of the college sample, but we slightly modify the third restriction to the one below:

3.B. listed major j in field f as a choice, such that majors not preferred to j belong to a field different from f (otherwise being above or below the cutoff would not generate variation in the field of study attended).

This means that the field sample only contains individuals whose older siblings marginal admission or rejection from their target major changes the field of study to which they are allocated.

4.4 Identifying Assumptions

As in any other RD setting, the validity of our estimates relies on two key assumptions. First, individuals should not be able to manipulate their application scores around the admission cutoff. The structures of the admission systems in Chile, Croatia and Sweden make the violation of this assumption unlikely. However, to confirm this, we show that the distribution of the running variable (i.e. older sibling's application score) is continuous at the cutoff (see Appendix B for more details).

Second, in order to interpret changes in individuals' outcomes as a result of the admission status of their older siblings, there cannot be discontinuities in other potential confounders at the cutoff (i.e. the only relevant difference at the cutoff must be older siblings' admission). Appendix B shows that this is indeed the case for a rich set of socioeconomic and demographic characteristics.

As previously mentioned, we use a fuzzy RD to study the effect of older siblings' enrollment (instead of admission) on younger siblings' outcomes. This approach can be thought of as an IV strategy, meaning that in order to interpret our estimates as a local average treatment effect (LATE) we need to satisfy the assumptions discussed by Imbens and Angrist (1994).[28] In this setting, in addition to the usual IV assumptions, we also need to assume that receiving an offer for a specific major does not make the probability of enrolling in a different major bigger than in the absence of the offer. [29] Given the structure of the admission systems that we study, this additional assumption does not seem very demanding.[30]

An additional issue related to the interpretation of our estimates is that as noted by Cattaneo et al. (2016), by pooling together different cutoffs, our estimates correspond to a weighted average of LATEs across programs. This weighted average gives more importance to programs with more

[28]Independence, relevance, exclusion and monotonicity. In this setting, independence is satisfied around the cutoff. The existence of a first stage is shown in Figure 1. The exclusion restriction implies that the only way through which older siblings' admission to a major affects younger siblings' outcomes is by the increase it generates in older siblings' enrollment in that major. Finally, the monotonicity assumption means that admission to a major weakly increases the probability of enrollment in that major (i.e. being admitted into a major does not reduce the enrollment probability in that major).

[29]Appendix A presents a detailed discussion of the the identification assumptions.

[30]In Chile, where not all colleges use the centralized admission system and rejecting an offer is not costly for students, this assumption could be violated if, for instance, colleges that do not use the centralized admission system were able to offer scholarships or other types of incentives to attract students marginally admitted to colleges that do use it. Although it does not seem very likely that colleges outside the centralized system would define students' incentives based on marginal offers to other institutions, we cannot completely rule out this possibility. In the case of Croatia —where students lose their funding in case of rejecting an offer— and Sweden —where there are no tuition fees— violations of this assumption seem unlikely.

applicants in the vicinity of the admission cutoff. Since there could be heterogeneity in the characteristics of individuals around each admission cutoff, and also on the effect of admission and enrollment at each admission cutoff, we need to be careful with the interpretation of this weighted averages. [31]

A final consideration for the interpretation of our results relates to the findings of Barrios-Fernandez (2018). According to these, the probability of attending university increases with close peers' enrollment. If marginal admission to the programs that we study translates into an increase in total university enrollment, then our estimated results could simply reflect that individuals whose older siblings attend college are more likely to enroll. We address this concern in Appendix B where we show that older siblings' marginal admission to their target majors does not generate a difference in younger siblings' total enrollment. [32]

Appendix B presents multiple additional robustness checks. We show that, as expected, changes in the admission status of younger siblings do not have an effect on older siblings; that our estimates are robust to different bandwidth choices and that placebo cutoffs do not significantly effect any of the outcomes that we study.

5 Results

This section begins by providing additional details about the empirical approach used to estimate the effects of interest. It then discusses how the probabilities of applying and enrolling in a specific major-college combination change when an older sibling is marginally admitted and enrolls in it. The section continues by investigating how college and field of study choices are affected. Next it discusses how these responses vary depending on siblings and majors characteristics, and concludes by looking at effect on individuals' academic performance.

5.1 Method

In all of the specifications used in this paper, we pool together observations from all over-subscribed majors and center older siblings' application scores around the relevant admission cutoff. The

[31] In order to understand what is driving our results we perform a detailed heterogeneity analysis along multiple dimensions including both individual and program characteristics. In Appendix B we study how our results vary when we re-weight observations around each cutoff by the inverse of the total number of applicants around it. Although the estimates are slightly smaller, the main conclusions still hold.

[32] In Chile, we find a small increase in the total enrollment of older siblings. This result is not surprising. As discussed in Section 2, the colleges that use the centralized admission system in Chile are, on average, more selective than the rest. This means that individuals rejected from these institutions still have many other alternatives available. In Croatia, we find that marginal admission translates into a more significant increase in older siblings total enrollment. However, we do not find an extensive margin response among younger siblings. Finally, in Sweden we once again find a small increase in older siblings' total enrollment, but as in the previous cases it does not translate into any significant difference in the total enrollment of their younger siblings.

following expression describes our baseline specification:

$$y_{ijut\tau} = \beta admitted_{iju\tau} + f(a_{iju\tau}; \gamma) + \mu_t + \mu_{ju\tau} + \varepsilon_{ijut\tau} \tag{1}$$

where,

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

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

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

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

We estimate two versions of this specification. In both cases, $f(a_{ijut\tau}; \gamma)$ corresponds to a linear or a quadratic polynomial of $a_{iju\tau}$ which slope is allowed to change at the admission cutoff. However, while in one specification we use a uniform kernel, in the second one we use instead a triangular kernel to give more weight to observations close to the cutoff.[33] Our analysis of younger siblings responses to older siblings' marginal enrollment focuses on three levels: first preference in the application list, all the preferences in the application list, and enrollment. Depending on the margin of interest (i.e. major, college or field) we use one of the samples described in Section 4. We compute optimal bandwidths according to Calonico et al. (2014) for each sample and level being investigated , but then we use a single bandwidth per sample: the smallest one among the three computed.[34]

Since all the specifications that we use focus on individuals whose older siblings are near an admission cutoff, our estimates represents the average effect of older siblings' marginal admission compared to the counterfactual of marginal rejection from a target major.[35]

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

[33] In Appendix Tables B5 , B6, and B7 we also present a specification in which we allow the slope of the running variable to be different for each admission cutoff. The estimation of these specifications is costly in computing time. In addition to the fixed effects included in the baseline specification, we need to include interactions between the running variable $a_{iju\tau}$ and $\mu_{ju\tau}$, and also between $a_{iju\tau}$, $\mu_{ju\tau}$ and $admitted_{ijut\tau}$. The estimates obtained with this specification are very similar to the ones discussed in this section.

[34] In principle, optimal bandwidths should be estimated for each admission cutoff independently. However, given the number of cutoffs in our sample, doing this would be impractical. Therefore, we compute optimal bandwidths pooling together all the cutoffs. Appendix Figures B4, B5 and B6 illustrates how sensitive our estimates are to the choice of bandwidth.

[35] Strictly speaking, our estimates represent a weighted average of multiple LATEs. See Section 4.4 for additional details. In addition, Appendix Tables B8, B9 and B10 present the results of an additional specification that controls by target major × counterfactual major fixed effect. The effects are very similar to the ones presented in the main section of the paper.

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

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

5.2 Effects of Older Siblings on Major Choice

This section discusses how older siblings' admission and enrollment in a specific major-college combination affect their younger siblings' probabilities of applying to and enrolling in it. To investigate changes in this margin, we use the Major Sample defined in Section 4.2.

The RD estimates illustrated in Figure 2 provide consistent causal evidence that students are more likely to apply to and enroll in a major if an older sibling was admitted to it before.[36]

As discussed in Section 4, receiving an offer for a specific major does not translates one-to-one into enrollment in any of the settings that we study. Thus, in order to estimate the effect of older siblings' enrollment on individuals applications and enrollment decisions, we combine the reduced form results discussed in the previous paragraph with the respective first stages illustrated in Figure 1, and obtain the fuzzy-RD estimates presented in Table 3. Under the identification assumptions discussed in Section 4, these fuzzy-RD provide consistent estimates for the effects of interest.

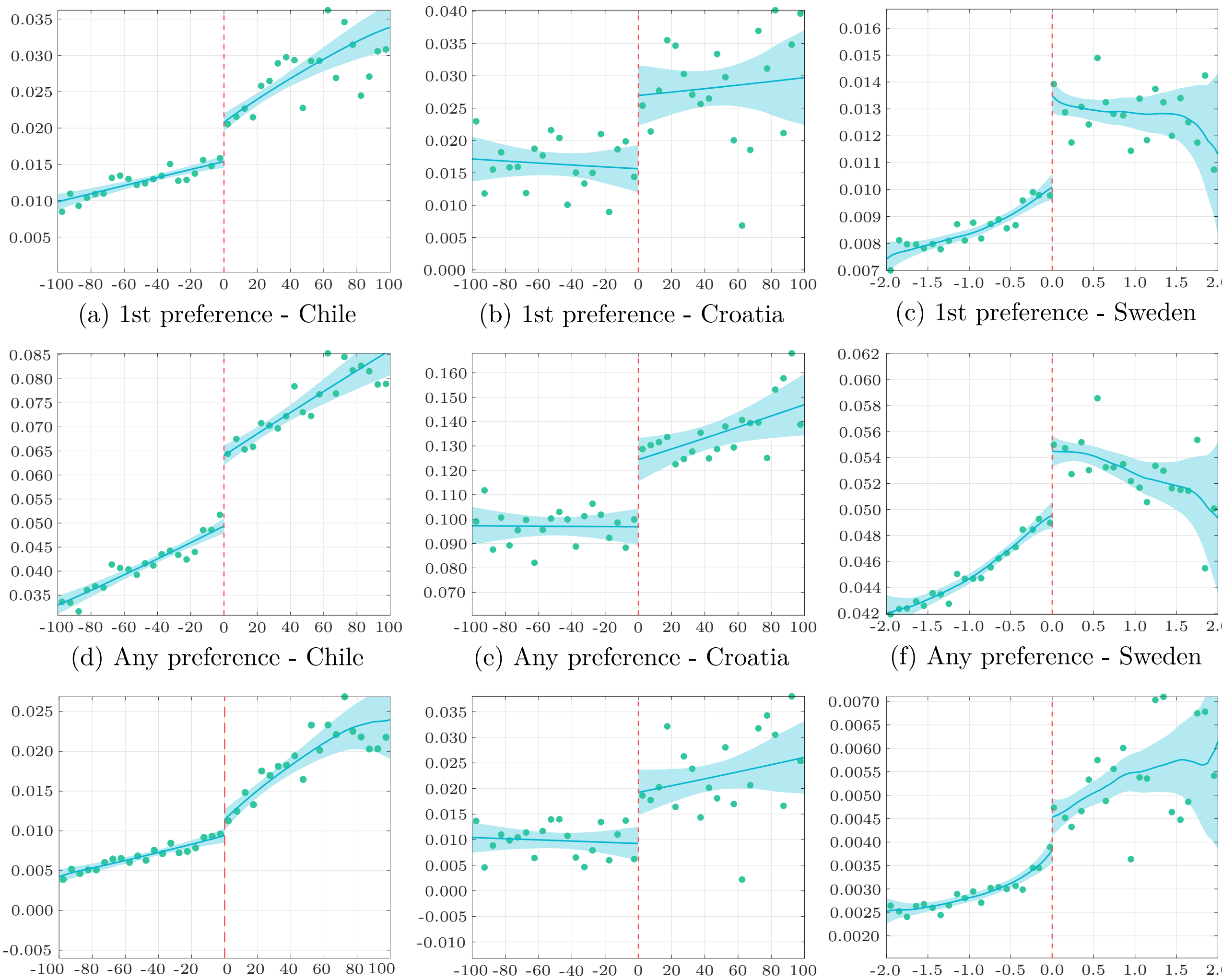
We find that in Chile, having an older sibling "marginally enrolling"[37] in a specific major increases the likelihood of applying to that major in the first preference by 0.8 percentage points (40%) and in any preference by around 2.8 pp (55%). These changes in applications also translate into an increase of around 0.3 pp (30%) in enrollment (although this last figure is not statistically significant). The results for Croatia are very similar. Individuals are 1.4 pp (45%) more likely to apply to their older siblings' target major in the first preference, 3.4 pp (33%) more likely to apply to it in any preference and 1.4 pp (58%) more likely to enroll in it. Finally, in Sweden, the likelihood of ranking older siblings' target major in the first place increases by around 2 pp (180%), while the likelihood of ranking it in any position increases by around 3 pp (63.8%). We also show that enrollment in older siblings' major increases by roughly 0.4 pp (100%).

Since in the three settings that we investigate, applicants know their scores before submitting their applications, their responses may depend on how likely they believe it is to be admitted in their

[36] In the case of Sweden, ties at the cutoff are broken through lotteries. For estimation and illustration purposes, we subtracted ε from the running variable of lotteries' losers. We set ε at the smallest machine detectable number.

[37] "marginally enrolling"means that the individual was marginally admitted to the major in which she enrolled. We emphasize this to remind the reader that the estimates come from comparing individuals whose older siblings were marginally admitted and marginally rejected from specific majors.

Figure 2: Probabilities of Applying and Enrolling in Older Sibling's Target Major-College



(a) 1st preference - Chile (b) 1st preference - Croatia (c) 1st preference - Sweden

(d) Any preference - Chile (e) Any preference - Croatia (f) Any preference - Sweden

(g) Enrolls - Chile (h) Enrolls - Croatia (i) Enrolls - Sweden

This figure illustrates the probabilities that younger siblings apply to and enroll in the target major of their older siblings in Chile, Croatia and Sweden. Figures (a), (d) and (e) illustrate the case of Chile, figures (b), (e) and (h) the case of Croatia, while figures (c), (f) and (i) the case of Sweden. Blue lines and the shadows in the back of them correspond to local polynomials of degree 1 and 95% confidence intervals. Green dots represent sample means of the dependent variable at different values of older sibling's admission score.

older siblings' target major once they learn their application score. In Table 4 we present additional results that come from specifications that expand the baseline specification by adding an interaction between older siblings' marginal enrollment and a proxy of younger siblings' eligibility for their older sibling's target major.[38]. According to the results presented in columns (1) to (3) of Table 4, younger siblings are more likely to apply and enroll in their older siblings' target major if they are eligible for it.[39]

In order to gain a deeper understanding about what is behind this "major following"behavior, in columns (4) to (6) of Table 4 we estimate the same specifications just discussed, but this time focusing on the sub-sample of older siblings whose target and counterfactual majors were offered by the same college. For these older siblings, being rejected from their target major does not change the college in which they end being admitted. Finding that even in this restricted sample younger siblings are more likely to apply to and enroll in their older siblings target major, suggests that the effects discussed in this section are not only driven by an increase in applications and enrollment in the older sibling's target college.

Despite the differences that exist among the three countries that we study, the results of this section are pretty consistent. They indicate that especially when younger siblings are eligible for their older siblings' specific major-college combination, they are more likely to apply and enroll in it.

5.3 Effects of Older Siblings on College and Field of Study Choices

While the focus of the previous section was on the specific major-college choice, this section independently investigates how younger siblings' choices of college and field of study are affected by older siblings. To study these margins we slightly modify the baseline specification of the previous section by replacing the outcome for a dummy variable that indicates if the younger sibling applies or enrolls in the target college or in the target field of study of the older sibling.[40] Depending on the margin being investigated, we focus our attention on the College Sample or on the Field Sample defined in Section 4.2.[41]

Table 5 summarizes the results of siblings' spillovers on the choice of college. In Chile, individuals are 7.2 pp (45%) more likely to rank their older siblings' target college first and 10.1 pp (30%)

[38]These specifications also control by the main effect of the eligibility proxy. In Chile and Croatia the eligibility proxy is an indicator that takes value 1 if the younger sibling average score in the admission exam is equal or greater than the average score obtained by the older sibling. In Sweden, given that the scale of the GPA and of the admission exam change during the period that we study, we use instead a variable that indicates if given their high school GPA, younger siblings are likely to be admitted in the target program of their older siblings.

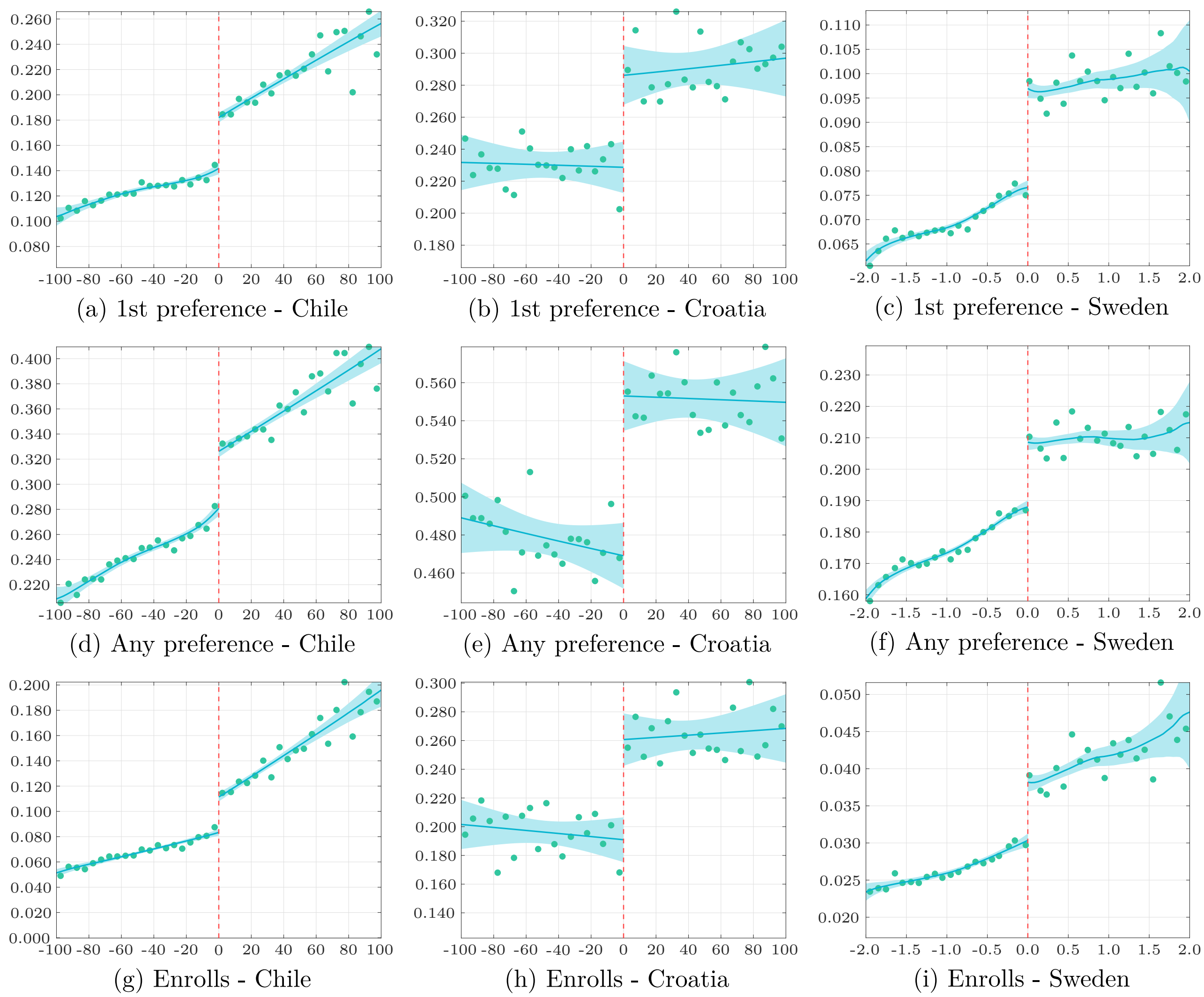
[39]In section 5.8, we show that older siblings' enrollment on their target major does not increase younger siblings' academic performance in high school or in the university admission exam. These results attenuate selection concerns that could have arisen by adding eligibility into the analisys.

[40]We define target college as the college offering the target major of the older sibling. Similarly, we define target field as the 3-digits ISCED code category to which the older sibling's target major belongs.

[41]Note that by changing the sample, we change the type of individuals that enter the estimations, something that could potentially affect the comparability of our results across samples.

more likely to apply to it in any preference. They are also 4.4 pp (44%) more likely to enroll in that college. For Croatia, the same figures are 7.5 pp (23%), 10.9 pp (19%) and 8.4 pp (29%) respectively, and for Sweden they are 15 pp (170%), 15.3 pp (79%) and 6.4 pp (188%).

Figure 3: Probabilities of Applying and Enrolling in Older Sibling's Target College

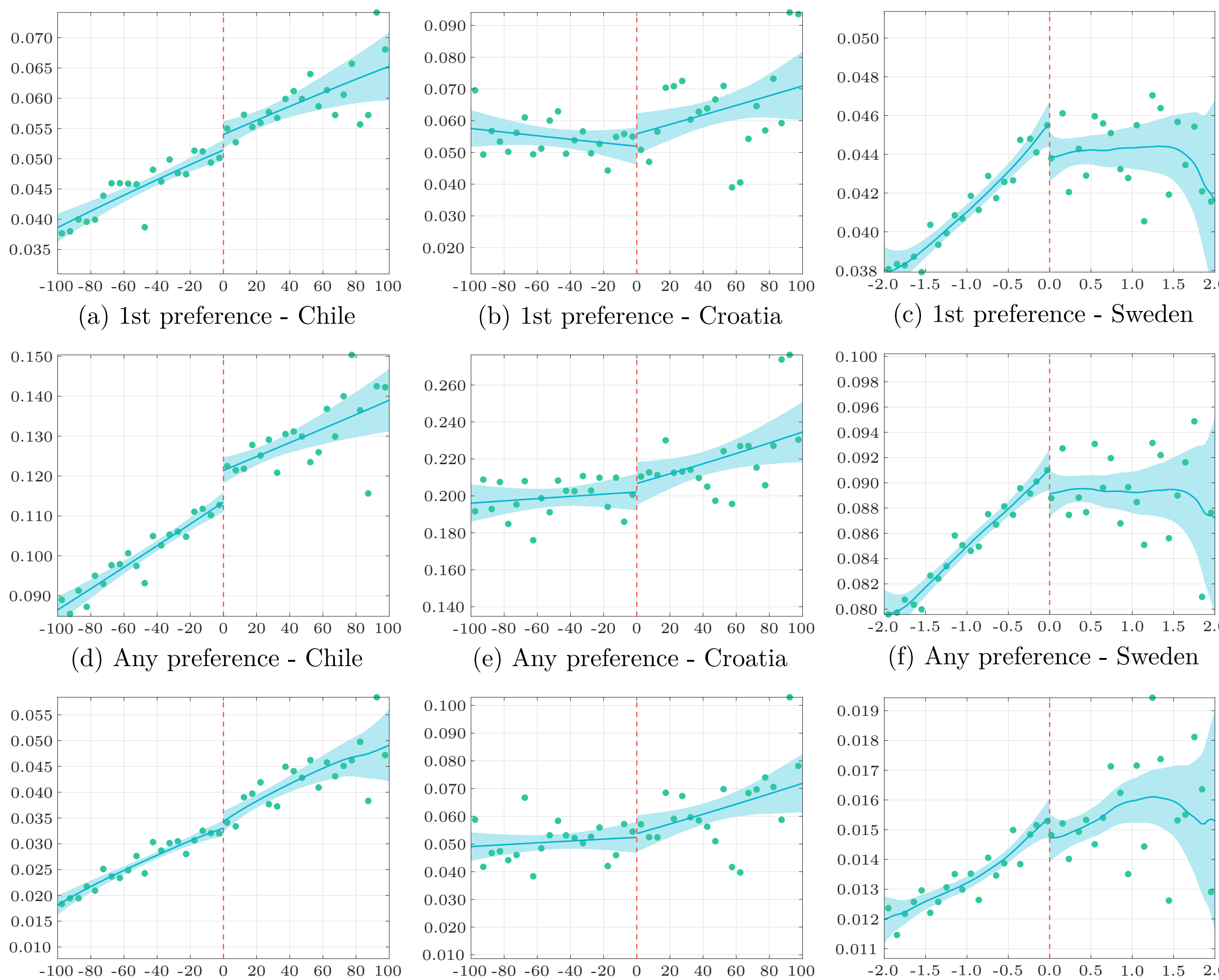


(a) 1st preference - Chile
(b) 1st preference - Croatia
(c) 1st preference - Sweden
(d) Any preference - Chile
(e) Any preference - Croatia
(f) Any preference - Sweden
(g) Enrolls - Chile
(h) Enrolls - Croatia
(i) Enrolls - Sweden

This figure illustrates the probabilities that younger siblings apply to and enroll in the target college of their older siblings in Chile, Croatia and Sweden. Figures (a), (d) and (e) illustrate the case of Chile, figures (b), (e) and (h) the case of Croatia, while figures (c), (f) and (i) the case of Sweden. Blue lines and the shadows in the back of them correspond to local polynomials of degree 1 and 95% confidence intervals. Green dots represent sample means of the dependent variable at different values of older sibling's admission score.

One hypothesis that may explain the big effects that we find on the choice of college is that they reflect at least in part geographic preferences. This would mean that individuals follow their older siblings to the city and not to the institution or major in which they enroll. To address this concern, we take advantage of the fact that in Chile there are three big cities —Santiago, Valparaíso and Concepción— that not only contain an important share of the population, but also multiple universities.[42].

Figure 4: Probabilities of Applying and Enrolling in Older Sibling's Target Field of Study



(a) 1st preference - Chile (b) 1st preference - Croatia (c) 1st preference - Sweden

(d) Any preference - Chile (e) Any preference - Croatia (f) Any preference - Sweden

(g) Enrolls - Chile (h) Enrolls - Croatia (i) Enrolls - Sweden

This figure illustrates the probability that younger siblings apply to and enroll in a program in the same field of study as the target program of their older siblings in Chile, Croatia and Sweden. Figures (a), (d) and (e) illustrate the case of Chile, figures (b), (e) and (h) the case of Croatia, while figures (c), (f) and (i) the case of Sweden. Blue lines and the shadows in the back of them correspond to local polynomials of degree 1 and 95% confidence intervals. Green dots represent sample means of the dependent variable at different values of older sibling's admission score.

Table 6 presents the results of an exercise in which we estimate the baseline specification on a sample of Chilean students from Santiago, Valparaíso and Concepción whose older siblings apply to institutions in their hometowns. If the effects documented in Table 5 were driven only by geographic preferences, we should not find siblings spillovers on the choice of college for this subsample. However, the coefficients that we obtain in this case are very similar to the main results previously discussed.

[42]In Santiago, there are campuses of 33 universities, in Valparaíso 11 and in Concepción 12

On the other hand, when investigating how the choice of field of study —defined by the three digit level code of the ISCED classification— is affected, we only find a marginally significant effect on younger siblings' applications in the case of Chile. In Croatia and Sweden none of the estimated coefficients is statistically significant (Table 7). Considering that the comparison of results across samples must be treated with caution, the results discussed so far suggest that individuals' major choice is only affected when younger siblings are likely to be admitted in their older siblings' specific major-college combination.

Since the choices of major and college seem to be the margins more affected by older siblings' higher education decisions, in the rest of the paper we will focus on these margins.[43]

5.4 Effects on Applications to Major and College by Gender:

This section explores if the responses in major and college choice documented in the previous sections vary depending on siblings' gender.[44]

The results of this section are summarized in Table 8. The first three columns look at differences on applications to majors, while the following three columns at differences in applications to colleges. To perform these analyses we expand the baseline specification by adding an interaction between the treatment and a dummy variable that indicates whether the gender of both siblings is the same. The main effect of the "same gender"dummy is also included as a control in all these specifications.

While columns (1) and (4) present results using the whole sample, the rest of the columns split the sample according to the gender of the older sibling. Thus, columns (2) and (5) look at pairs of siblings in which the older sibling is female, while columns (3) and (6) at pairs of siblings where the older sibling is male.

According to these results, older brothers are more likely to be followed to their specific major by males than by females. This difference is less clear when looking at older sisters. Apart from Sweden, where older sisters seem to generate stronger responses in their younger brothers, we find no significant differences in how male and female applicants respond to their major choice.

When looking instead at the college choice, we find no significant difference in how male and female applicants respond to the choices of their older brothers or sisters. Being of the same gender as younger siblings does not seem to increase the likelihood of being followed by them. However, in this case independently of their gender, younger siblings seem to be more responsive to older brothers than to older sisters.

Overall, the results discussed in this section indicate that males are more likely to apply to the same major and college of an older brother than of an older sister. However, their applications are

[43]Appendix C includes similar results for the field choice.

[44]The analyses presented in this section focus on applications to majors and colleges. Similar results for enrollment and for decisions related to the field of study are presented in Appendix Tables C1 and C2.

also affected by the higher education decisions of their older sisters. In the case of females, the pattern is less clear. They seem to be more responsive to what happens with their older sisters when choosing major, but the opposite is true when looking at applications to college.

5.5 Effects on Applications to Major and College by Differences in Age and in Academic Potential

In this section we investigate how the applications to major and college change depending on how close siblings are in terms of age and academic potential.[45] To investigate differential effects by age, we expand the baseline specification with an interaction between the treatment and a dummy variable indicating whether siblings were born 5 or more years apart. To investigate if the effects change depending on differences in academic potential, we proceed in a similar way by adding an interaction with the absolute difference in siblings' high school GPA.[46] In Croatia, we only observe high school GPA for students completing their secondary education before 2015; this explains the smaller sample used in this part of the analysis for Croatia.

Table 9 summarizes the results of this section. The first two columns look at the choice of major, while the last two at the choice of college. In Chile and Croatia the effects do not significantly decrease with the age difference between siblings. In the case of Sweden, the effects are stronger for siblings who are closer in age. However, even for those who are more than 5 years apart the effects are significant both statistically and economically.

The difference in siblings academic potential only seems to make a difference in Chile and Croatia (columns (2) and (4)). In Chile, a difference of 1σ in siblings' high school GPA score reduces the effect on applications to majors by 51.2% and on applications to colleges by 44.7%. In the case of Croatia, the estimates point in the same direction, but are less precisely estimated. A difference of 1σ in siblings' high school GPA decreases the effect on applications to majors by 44% and on applications to colleges by 15.9%. Finally, in Sweden we find no relevant differences in the effects on major and college choices depending on siblings' academic potential.

5.6 Effects on Application to College and Major by Older Siblings' Major Quality

This section studies how the effects documented in Section 5.2 change depending on the quality of the target major of the older sibling.[47] We measure quality in terms of admitted students' academic

[45] We present similar analyses for enrollment and for the choice of field of study in Appendix Tables C3 and. C4.

[46] Note that if younger siblings are still in high school when their older siblings apply to higher education, their high school GPA could be an outcome of the treatment. However, as shown in Section 5.8 "marginal enrollment"of older siblings in their target major does not seem to affect individuals' academic performance.

[47] Appendix Tables C5 and C6 present similar results for enrollment and for the choice of field of study respectively.

potential, first-year dropout rates and graduates' earnings.[48]

Student quality is the only variable in this section that we observe for the three countries. We define the quality of the students in a program in a given year using the average performance of admitted students in the college admission exams in Chile and Croatia, and as the average high school GPA of admitted students in Sweden. We are able to compute dropout rates and graduates earnings only for Chile and Sweden. We compute dropout rates for each major using individual level data provided by the Ministry of Education (Chile) and by the Council for Higher Education (Sweden). The data from Chile allow us to compute dropout rates for all college cohorts beginning in 2006;[49] in Sweden we observe dropout rates for the entire sample period. Variables measuring the labor market performance of former students in Chile are available at the major-college level. They are computed by the Ministry of Education with the support of the National Tax Authority.[50] In the case of Sweden, information on earnings comes from Statistics Sweden.

The main results of this section are summarized in Table 10. All variables, except for dropout rates, are standardized to facilitate the interpretation of the results. The first three columns of the table investigate heterogeneous effects on applications to majors, while the last three on applications to colleges.

When looking at heterogeneous effects on the major choice by the quality of the students admitted to that major, we only find a significant difference in Sweden. In this country, a difference of 1σ in the quality of the applicants admitted to the older sibling's major increases the younger sibling's applications to that major by 1.2 pp. Differences are more clear when looking instead at the college choice. In this case, an increase on the quality of the students admitted to the older siblings major increases younger siblings' applications to the college offering that major by 2.4 pp in Chile, 2.7 pp in Croatia and 3.6 pp in Sweden. [51]

Higher dropout rates seem to reduce younger siblings' applications to both the major and the college of the older sibling. However, this difference is only significant when looking at the college choice.

Finally, when looking at heterogeneity by graduates' labor market outcomes we find that younger

[48]We only observe earnings for Chile and Sweden. In the case of Chile, graduates average earnings are measured four years after graduation and reported by the Ministry of Education. We observe them only once for each major-college. This means that in our analysis this variable does not change over time. In the case of Sweden, we compute average earnings one year after graduation. We use as reference the cohort graduating the year in which older siblings apply to their target major.

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

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

[51]Note that since our sample only includes majors with positive number of individuals in the waiting list, our estimates are not valid for non-selective programs. This is particularly relevant in Chile, where the less selective institutions are not part of the sample at all.

siblings are more likely to apply to their older siblings' major when past graduates earnings are higher. A similar pattern arises when focusing on the college choice, but in this case the coefficients are unprecisely estimated.

Our results show that individuals do not follow their older siblings to all majors and colleges. The responses seem to be stronger when the quality of the major attended by the older sibling is higher.

Table 11 presents results of a similar exercise, but in which we study heterogeneous effects by the difference in the quality indexes of older siblings' target and counterfactual majors (counterfactual major is the major in which they would have been admitted in the event of being rejected from their target choice).[52] This forces us to restrict the sample to older siblings for whom it is possible to identify a counterfactual alternative. Therefore, those not admitted to any program are not part of this analysis. We find no heterogeneous effects by differences in any of the quality measures we use. In part, this could be due to the smaller sample size used for this exercise and to the fact that on average there is not a big difference between the quality of the target program and the quality of the next best option.

5.7 Effects on Application and Enrollment by Older Sibling's College Experience

This section investigates whether the effects on the choice of major and college depend on the experience of older siblings in their target major. Table 12 provides evidence consistent with the hypothesis that individuals learn from their older siblings' experience if a specific major or college would be a good match for them. Siblings are similar in many dimensions, and therefore if an older sibling has a negative experience in a specific major or college, their younger siblings may infer that applying and enrolling in that alternative is not necessarily good for them. In our data, the best available proxy for older siblings' experience in college is dropout. We are only able to compute dropout for Chile and Sweden, and therefore this section only presents results for these countries.

We add to the baseline specification an interaction between the treatment and a dummy that indicates whether the older sibling drops out from the major or college in which she first enrolls,[53] and the main effect of older siblings' dropout.[54] The results of this exercise should be interpreted with caution. Dropping out from college is not random, and although controlling by dropout helps to capture some of the differences that may exist between individuals who remain at and leave a particular college, there could still be differences that we are not able to control for.[55] In

[52] Appendix Tables C7 and C8 present results for major and college enrollment and for the choice of field of study.

[53] Note that the major in which older siblings enroll are not necessarily the ones to which they are admitted.

[54] We study dropout in the 4 years following enrollment. To be able to do this, we restrict the sample to sibling pairs in which the older sibling applies to college before 2011 in Chile and before 2012 in Sweden.

[55] In addition, note that with this specification we are comparing the effects found for admitted and rejected individuals who remain in the college in which they enroll, with the ones found when comparing admitted and rejected individuals who dropout from the college in which they enroll. In general, admitted and rejected individuals attend different majors.

addition, the dropout variable can only be built for older siblings who actually enroll in some major. Appendix Table B4 shows that in Chile and Sweden, marginal admission does not translate into relevant increases in older siblings' total enrollment. However, only focusing on applicants whose older siblings enroll in a program affects the composition of the sample used in this analysis.

Bearing these caveats in mind, the results of this exercise show that individuals whose older siblings dropout from their major or college are significantly less likely to follow them. Indeed, the effects documented in previous sections on both the choice of major and college virtually disappear if the older sibling drops out.

5.8 Effects on Academic Performance

In this section we study if the increase in the likelihood of applying and enrolling in the major attended by an older sibling could be driven by an improvement in younger siblings' academic performance. To study this we use the same fuzzy-RD strategy discussed in Section 4, but this time we look at younger siblings' high school GPA and at their scores in the admission exams. Since not all potential applicants take the admission exam, we replace missing values by zero. This means that when looking at effects on exams scores our estimates capture differences in performance, but also differences in the probability of taking the exam. The bandwidths used in this section are the same used in Section 5.2.

Table 13 summarizes these results. We show that, having an older sibling "marginally enrolling"in her target major does not seem to generate significant changes in younger siblings' high school performance or in their performance in the university admission exams.

These results hold for the three countries in our study, and suggest that the effects documented on the choice of program are not driven by an improvement in the academic performance of younger siblings.[56]

6 Discussion

The results presented in Section 5 show that the path followed by older siblings in higher education affects the major and college choice of their younger siblings. Although documenting the existence of sibling spillovers in the choice of college in three settings as different as Chile, Croatia and Sweden is interesting in itself, from a policy perspective it is also relevant to understand the mechanisms behind these responses. In the rest of this section, we discuss three broad classes of mechanisms

[56]We reach the same conclusion when investigating changes in academic performance in the Institution and Field samples. These results are presented in Appendix Tables C9 and C10. One reason why we may not detect changes in academic performance is that individuals may need some time after their older sibling's enrollment in order to respond. We explore this possibility in Appendix Table C11, but we find no significant effects even when looking at siblings born 5 or more years apart.

that could drive our results using a simple framework of discrete choice and utility maximization.

Let M_i be the set of majors m that form part of the alternatives to which individual i is considering to apply and $\vec{x_m}$ a vector of the attributes that characterize each major. Individuals have different preferences over these attributes and chose to apply to the major that maximizes their utility subject to a budget constraint B_i. P_m is the cost of enrolling in major m and it includes tuition fees, commuting costs and living costs.

$$\begin{aligned} &\underset{m \in M_i}{max} U_i(\vec{m}), \quad m = (x_{1m}, ..., x_{nm}) \\ &\text{s.t. } P_m \leq B_i \end{aligned}$$

With this simple framework in mind, a first way in which older siblings could affect the decision of applying and enrolling in a specific major or college is by affecting the costs of that option. For instance, by attending the same college as an older sibling, individuals might save in commuting and living costs. However, we find that the effects persist even among siblings who, due to age differences, are unlikely to attend college at the same time. This result, and the fact that the effects look very similar when we focus on a group of individuals whose older siblings apply to majors offered in their hometown, suggest that this convenience channel is not the main driver of our results. [57]

Alternatively, having an older sibling enrolling in a specific college could affect individuals' preferences. Preferences could change if individuals enjoy spending time with their older siblings or if they perceive them as role models and are inspired by them. Preferences could also be affected if siblings are competitive or if parental expectations are changed by the college choices of older siblings.

The persistence of the effects among siblings with large age differences suggests that our results are not driven by them enjoying each other's company. In addition, finding no heterogeneous effects by differences in the quality of target and counterfactual majors of older siblings and finding no effects on younger siblings' academic performance, suggests that there are no changes in individuals' aspirations. If this were the case, we would expect to see them exerting additional effort preparing for college, something that is not reflected in their applications, high school or college admission exam performance.

Joensen and Nielsen (2018) argue that the fact that their results are driven by brothers who are close in age and in academic performance is evidence in favor of competition being the main driver of

[57] In some settings, the admission systems give an advantage to siblings of current or former students. This however is not a concern in our case. In Chile, Croatia and Sweden universities use centralized admission systems that select students based only on their academic performance in high school and on a national level admission exam. Although in Chile some colleges offer discounts in tuition fees when many siblings simultaneously attend the same program, finding that the effect persists even when looking at siblings born 5 or more years apart makes this an unlikely driver of our results. In Croatia, students do not pay tuition fees if they accept the offer they receive the first time that they apply and in Sweden all institutions are free.

their results. As previously discussed, in our case the results persist even among siblings born more than 5 years apart, and among sisters and different-gender siblings, suggesting that if competition mostly arises between brothers close in age, it cannot be the main driver of our results.

The preferences of individuals could also be influenced by changes in their parents' expectations. However, we do not find heterogeneous effects based on differences in selectivity between target and counterfactual majors (i.e. the majors to which students would have enrolled in case of being rejected from their target option). We interpret this as evidence against the parental expectations channel. The intuition behind this argument is that if counterfactual majors are similarly selective, then having a child admitted to one or the other should not generate a gap in parental expectations.

Finally, older siblings' enrollment in a specific major-college could affect the choice set of their younger siblings by making some options more salient or by providing information about relevant attributes of the available options.[58] Considering the amount of major-college combinations from which applicants can chose, both hypothesis could play a relevant role. However, we find stronger effects when older siblings' majors are of higher quality, which goes against salience. If salience were the main driver of our results, we should see individuals following their older siblings independently of the quality of their majors. On the other hand, we show that the effects are driven by older siblings who enroll in majors that are better in terms of student quality, retention and graduates' labor market performance. This suggests that individuals learn about the quality of colleges from their older siblings. In addition, the difference found on the effects depending on older siblings' dropout suggest that the experience that they have in higher education matters, and that younger siblings are more likely to follow their older siblings when they have a good experience in college.

Even though the evidence discussed in this section does not allow us to perfectly distinguish the exact mechanisms behind our results, they suggest that information, particularly information about the college experience of someone close, might play a relevant role in college choices. Further research is required to investigate the precise information that individuals acquire through their close peers.

7 Conclusions

Despite the difference that a good college and major match can make on an individual's life, we know little about how the preferences and beliefs driving these choices are formed. The heterogeneity in colleges' and majors' characteristics, and the difficulty to observe some of their attributes make these decisions challenging. In this context, close relatives and other members of an individual's social network could significantly influence college related choices. However, causally identifying the effects of social interactions is notoriously challenging.

[58] Since in this framework a major is defined by its vector of attributes, any information that changes the perceived values of these attributes also modifies the choice set.

In this paper, we investigate how college application and enrollment decisions are affected by the higher education choices of older siblings. We study these sibling spillovers in Chile, Croatia and Sweden, where universities select students using centralized deferred acceptance systems that allocate students to majors and colleges only considering their declared preferences and academic performance. These admission systems create thousands of discontinuities that we exploit in a fuzzy Regression Discontinuity Design framework that allows us to overcome the main identification challenges that arise in the context of peers effects (i.e. correlated effects and the reflection problem).

Despite the differences that exist between the three countries, we consistently find statistically and economically significant spillovers. In the three settings studied, we show that individuals are more likely to apply and enroll in the same major-college combination as their older siblings. In Chile, we document an increase of 2.8 pp (55%) in applications and 0.3 pp (30%) in enrollment; the same figures for Croatia are 3.4 pp (33%) and 1.4 pp (58%); and 3 pp (63.8%) and 0.4 pp (100%) for Sweden. These effects are stronger when individuals are more likely to be admitted in their older siblings' target major and persist even for individuals whose target and next best majors are offered by the same institution. This suggest that the spillovers we find in the specific major-college choice are not only driven by increased preferences for older siblings' colleges.

When looking at spillovers on the choice of college we find even bigger effects. Having older sibling enrolling in a particular institution increases the probability that their younger sibling applies there by between 8 pp and 15 pp and increases the likelihood of enrolling in that institution by 5 pp (50%) in Chile, 9 pp (30%) in Croatia and 6.4 pp in Sweden (188%). We find no significant spillovers on the field of study in any of the three countries. This and the results discussed in the previous paragraph suggest that the choice of field of study is only affected when individuals are likely to be admitted in their older siblings major-college combination.

We discuss three broad classes of mechanisms consistent with our results: a change in the costs, in the preferences or in the choice set of individuals. Firstly, attending the same college with a sibling could result in important savings (i.e. living or commuting costs). Alternatively, individuals could follow their siblings if, for instance, they enjoy spending time with them. Finally, individuals' choice sets could change as a consequence of salience or of information transmission.

We show that individuals only follow their older siblings to "high" quality colleges and that the experience that older siblings have in higher education makes an important difference in the observed response. We interpret these findings as suggestive evidence that information about the quality of colleges and majors and about the potential quality of the match for potential applicants is an important driver behind our results.

Our findings suggest that especially in contexts of incomplete information, policies that change the pool of students admitted to a specific college or major could have an indirect effect on their siblings and potentially on other members of their social networks. Our results also suggest that providing information about the experience that individuals would have in college, could improve their application and enrollment decisions.

Further research is needed to identify the type and accuracy of the information transmitted by siblings, and to find effective ways of closing the information gaps between applicants with different levels of exposure to college.

References

Aguirre, J. and J. J. Matta (2019). Walking in Your Footsteps: Sibling Spillovers in Higher Education Choices.

Altmejd, A. (2018). Education & Replication: Essays on the Determinants of College Choice and the Predictability of Lab Replications.

Altonji, J. G., E. Blom, and C. Meghir (2012). Heterogeneity in human capital investments: High school curriculum, college major, and careers. *Annu. Rev. Econ. 4*(1), 185–223.

Barrios-Fernandez, A. (2018). Should I Stay or Should I go? Neighbors' Effects on University Enrollment. *CEP Discussion Paper* (1653).

Barrios-Fernandez, A. (2019). Essays in economics of education.

Bettinger, E. P., B. T. Long, P. Oreopoulos, and L. Sanbonmatsu (2012, aug). The Role of Application Assistance and Information in College Decisions: Results from the H&R Block Fafsa Experiment. *The Quarterly Journal of Economics 127*(3), 1205–1242.

Björklund, A. and K. G. Salvanes (2011). Chapter 3 - Education and Family Background: Mechanisms and Policies. Volume 3 of *Handbook of the Economics of Education*, pp. 201 – 247. Elsevier.

Black, S. E., K. E. Cortes, and J. A. Lincove (2015, May). Academic Undermatching of High-Achieving Minority Students: Evidence from Race-Neutral and Holistic Admissions Policies. *American Economic Review 105*(5), 604–10.

Black, S. E. and P. J. Devereux (2011). Recent Developments in Intergenerational Mobility. In D. Card and O. Ashenfelter (Eds.), *Handbook of Labor Economics*, Volume 4, pp. 1487–1541.

Busso, M., T. Dinkelman, A. Claudia Martínez, and D. Romero (2017). The Effects of Financial Aid and Returns Information in Selective and Less Selective Schools: Experimental Evidence from Chile. *Labour Economics*.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2018). Regression Discontinuity Designs using Covariates. *Review of Economics and Statistics* (0).

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

Carrell, S. and B. Sacerdote (2017, July). Why do college-going interventions work? *American Economic Journal: Applied Economics 9*(3), 124–51.

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

Cattaneo, M. D., L. Keele, R. Titiunik, and G. Vazquez-Bare (2016). Interpreting regression discontinuity designs with multiple cutoffs. *The Journal of Politics 78*(4), 1229–1248.

Dillon, E. W. and J. A. Smith (2017). Determinants of the Match between Student Ability and College Quality. *Journal of Labor Economics 35*(1), 45–66.

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

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

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

French, R. and P. Oreopoulos (2017). Behavioral barriers transitioning to college.

Goodman, J., M. Hurwitz, C. Mulhern, and J. Smith (2019). O Brother, Where Start Thou? Sibling Spillovers in College Enrollment. *NBER Working Paper* (26502).

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

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

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

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

Hastings, J. S., C. A. Neilson, and S. D. Zimmerman (2013, July). Are some degrees worth more than others? evidence from college admission cutoffs in chile. Working Paper 19241, National Bureau of Economic Research.

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

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

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

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

Kaczynski, K. M. (2011). Exploring the influence of siblings and their relationships on the college choice process.

Kirkebøen, L. J., E. Leuven, and M. Mogstad (2016). Field of Study, Earnings, and Self-Selection. *131*(3), 1057–1111.

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

Long, B. T. (2004, aug). Does the Format of a Financial Aid Program Matter? The Effect of State In-Kind Tuition Subsidies. *Review of Economics and Statistics 86*(3), 767–782.

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

Oreopoulos, P. and U. Petronijevic (2013). Making college worth it: A review of the returns to higher education. *The Future of Children 23*(1), 41–65.

Pekkala Kerr, S., T. Pekkarinen, R. Uusitalo, et al. (2015). Post-secondary education and information on labor market prospects: A randomized field experiment.

Sacerdote, B. (2011). Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far? In *Handbook of the Economics of Education*, pp. 249–277.

Sacerdote, B. (2014, aug). Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward? *Annual Review of Economics 6*(1), 253–272.

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

Shahbazian, R. (2018). Under the influence of our older brother and sister: The association between sibling gender configuration and stem degrees.

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

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

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

Wiswall, M. and B. Zafar (2015). How Do College Students Respond to Public Information about Earnings? *Journal of Human Capital 9*(2), 117–169.

Table 3: Probability of Applying and Enrolling in Older Sibling's Target Major-College

	Applies 1st		**Applies**		**Enrolls**	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.008**	0.007*	0.028***	0.025***	0.003	0.002
	(0.003)	(0.003)	(0.005)	(0.006)	(0.002)	(0.003)
Reduced form	0.004**	0.003*	0.015***	0.012***	0.002	0.001
	(0.001)	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)
First stage	0.521***	0.488***	0.521***	0.488***	0.521***	0.488***
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
2SLS (Triangular kernel)	0.008*	0.008*	0.028***	0.028***	0.003	0.003
	(0.003)	(0.004)	(0.005)	(0.006)	(0.003)	(0.003)
Observations	136364	214840	136364	214840	136364	214840
Outcome mean	0.018	0.018	0.056	0.055	0.012	0.012
Bandwidth	20.000	35.000	20.000	35.000	20.000	35.000
F-statistics	13867.401	9520.717	13867.401	9520.717	13867.401	9520.717
			Panel B - Croatia			
2SLS	0.015***	0.014**	0.036***	0.038***	0.013**	0.015**
	(0.004)	(0.005)	(0.009)	(0.011)	(0.004)	(0.005)
Reduced form	0.012***	0.012**	0.030***	0.031***	0.011**	0.013**
	(0.004)	(0.004)	(0.007)	(0.009)	(0.003)	(0.004)
First stage	0.826***	0.820***	0.826***	0.820***	0.826***	0.820***
	(0.007)	(0.008)	(0.007)	(0.008)	(0.007)	(0.008)
2SLS (Triangular kernel)	0.014**	0.013*	0.040***	0.042***	0.014**	0.015**
	(0.005)	(0.006)	(0.009)	(0.011)	(0.004)	(0.005)
Observations	36757	48611	36757	48611	36757	48611
Outcome mean	0.029	0.029	0.129	0.130	0.024	0.024
Bandwidth	80.000	120.000	80.000	120.000	80.000	120.000
F-statistics	14512.301	10444.128	14512.301	10444.128	14512.301	10444.128
			Panel C - Sweden			
2SLS	0.020***	0.023***	0.029***	0.032***	0.004**	0.004**
	(0.003)	(0.003)	(0.005)	(0.006)	(0.001)	(0.002)
Reduced form	0.004***	0.005***	0.006***	0.007***	0.001**	0.001**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
First stage	0.217***	0.214***	0.217***	0.214***	0.217***	0.214***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
2SLS (Triangular kernel)	0.025***	0.027***	0.034***	0.035***	0.006***	0.006***
	(0.003)	(0.003)	(0.006)	(0.006)	(0.002)	(0.002)
Observations	730187	1034047	730187	1034047	730187	1034047
Outcome mean	0.011	0.010	0.047	0.046	0.004	0.003
Bandwidth	0.510	0.750	0.510	0.750	0.510	0.750
F-statistics	10817.599	8481.389	10817.599	8481.389	10817.599	8481.389

Notes: All the specifications in the table control for a linear or quadratic polynomial of older siblings' application score centered around target majors admission cutoff. Older siblings' application year, target major-year and younger siblings' birth year fixed effect are included as controls. 2SLS (Triangual Kernel) specifications use a triangular kernel to give more weight to observations close to the cutoff. Bandwidths were computed according to Calonico et al. (2014) for each outcome independently. The smallest one among the three is used for all the outcomes. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 4: Probability of Applying and Enrolling in Older Sibling's Target Major-College by Younger Siblings' Eligibility

	Major Sample			Major Sample Fixing College		
	Applies 1st (1)	Applies (2)	Enrolls (3)	Applies 1st (4)	Applies (5)	Enrolls (6)
			Panel A - Chile			
Older sibling enrolls	0.007**	0.024***	0.0004	0.002	0.010	-0.002
	(0.003)	(0.005)	(0.002)	(0.006)	(0.009)	(0.004)
Older sibling enrolls × Eligible = 1	0.004	0.019***	0.012***	0.010*	0.019*	0.014**
	(0.003)	(0.005)	(0.003)	(0.006)	(0.010)	(0.006)
Observations	136,364	136,364	136,364	39,343	39,343	39,343
Outcome mean	0.018	0.056	0.012	0.024	0.075	0.015
Bandwidth	20	20	20	20	20	20
F-statistics	6662.969	6662.969	6662.969	2794.937	2794.937	2794.937
			Panel B - Croatia			
Older sibling enrolls	0.009*	0.024**	-0.005	-0.004	-0.0004	-0.008
	(0.005)	(0.012)	(0.004)	(0.007)	(0.015)	(0.005)
Older sibling enrolls × Eligible = 1	0.011**	0.024**	0.029***	0.011*	0.035**	0.023***
	(0.005)	(0.011)	(0.004)	(0.006)	(0.014)	(0.005)
Observations	33,823	33,823	33,823	21,771	21,771	21,771
Outcome mean	0.031	0.141	0.026	0.032	0.150	0.027
Bandwidth	80	80	80	80	80	80
F-statistics	6770.281	6770.281	6770.281	4126.185	4126.185	4126.185
			Panel C - Sweden			
Older sibling enrolls	0.033***	0.046***	0.005**	0.008	-0.001	-0.005
	(0.005)	(0.010)	(0.003)	(0.012)	(0.022)	(0.007)
Older sibling enrolls × Eligible = 1	0.011**	0.010	0.014***	0.013	0.010	0.013*
	(0.004)	(0.009)	(0.003)	(0.011)	(0.019)	(0.007)
Observations	292,970	292,970	292,970	44367	44367	44367
Outcome mean	0.022	0.096	0.008	0.035	0.0133	0.014
Bandwidth	0.51	0.51	0.51	0.051	0.051	0.051
F-statistics	3270.581	3270.581	3270.581	830.621	830.621	830.621

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 3. In addition, they have an interaction between the treatment and a proxy of younger siblings' eligibility for their older siblings' target program. Columns (1) to (3) focus on the major sample, while columns (4) to (6) on the subset of individuals whose older siblings target and counterfactual major are offered by the same college. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 5: Probability of Applying and Enrolling in Older Sibling's Target College

	Applies 1st		Applies		Enrolls	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.072***	0.081***	0.101***	0.095***	0.044***	0.044***
	(0.012)	(0.011)	(0.015)	(0.014)	(0.010)	(0.009)
Reduced form	0.033***	0.038***	0.047***	0.045***	0.020***	0.020***
	(0.006)	(0.005)	(0.007)	(0.007)	(0.005)	(0.004)
First stage	0.466***	0.467***	0.466***	0.467***	0.466***	0.467***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
2SLS (Triangular Kernel)	0.080***	0.081***	0.103***	0.103***	0.051***	0.050***
	(0.013)	(0.013)	(0.017)	(0.016)	(0.011)	(0.010)
Observations	73331	152301	73331	152301	73331	152301
Outcome mean	0.161	0.157	0.302	0.292	0.101	0.097
Bandwidth	15.000	35.000	15.000	35.000	15.000	35.000
F-statistics	5441.604	5905.708	5441.604	5905.708	5441.604	5905.708
			Panel B - Croatia			
2SLS	0.075***	0.070**	0.109***	0.102***	0.084***	0.090***
	(0.019)	(0.023)	(0.019)	(0.024)	(0.018)	(0.023)
Reduced form	0.063***	0.058**	0.091***	0.085***	0.070***	0.075***
	(0.016)	(0.019)	(0.016)	(0.020)	(0.015)	(0.019)
First stage	0.835***	0.828***	0.835***	0.828***	0.835***	0.828***
	(0.010)	(0.013)	(0.010)	(0.013)	(0.010)	(0.013)
2SLS (Triangular Kernel)	0.086***	0.089***	0.105***	0.104***	0.092***	0.095***
	(0.020)	(0.024)	(0.021)	(0.025)	(0.020)	(0.024)
Observations	12950	17312	12950	17312	12950	17312
Outcome mean	0.321	0.322	0.555	0.559	0.287	0.287
Bandwidth	80.000	120.000	80.000	120.000	80.000	120.000
F-statistics	6459.562	4214.087	6459.562	4214.087	6459.562	4214.087
			Panel C - Sweden			
2SLS	0.149***	0.151***	0.153***	0.155***	0.064***	0.060***
	(0.009)	(0.009)	(0.013)	(0.013)	(0.006)	(0.006)
Reduced form	0.030***	0.030***	0.031***	0.031***	0.013***	0.012***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.001)	(0.001)
First stage	0.201***	0.198***	0.201***	0.198***	0.201***	0.198***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
2SLS (Triangular Kernel)	0.184***	0.169***	0.181***	0.169***	0.081***	0.071***
	(0.010)	(0.010)	(0.014)	(0.013)	(0.006)	(0.006)
Observations	443931	856200	443931	856200	443931	856200
Outcome mean	0.088	0.084	0.193	0.186	0.034	0.032
Bandwidth	0.370	0.730	0.370	0.730	0.370	0.730
F-statistics	6140.057	6084.386	6140.057	6084.386	6140.057	6084.386

Notes: All the specifications in the table control for a linear or quadratic polynomial of older siblings' application score centered around target majors admission cutoff. Older siblings' application year, target major-year and younger siblings' birth year fixed effect are included as controls. 2SLS (Triangual Kernel) specifications use a triangular kernel to give more weight to observations close to the cutoff. Bandwidths were computed according to Calonico et al. (2014) for each outcome independently. The smallest one among the three is used for all the outcomes. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 6: Probability of Applying and Enrolling in Older Sibling's Target College: Big Cities Sample

	Applies (1)	Enrolls (2)
2SLS	0.097^{***}	0.042^{**}
	(0.020)	(0.013)
Reduced form	0.053^{***}	0.023^{**}
	(0.011)	(0.007)
First stage	0.546^{***}	0.546^{***}
	(0.009)	(0.009)
Observations	32818	32818
Outcome mean	0.337	0.115
Bandwidth	15.000	15.000
F-statistics	3711.283	3711.283

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target college on younger siblings' probabilities of applying to and enrolling in the same college. The controls and bandwidths used in these specifications are the same described in Table 5. The sample only includes pairs of siblings who live in cities with at least 10 colleges and in which the older sibling target college is located in the same city. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 7: Probability of Applying and Enrolling in Older Sibling's Target Field of Study

	Applies 1st		**Applies**		**Enrolls**	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.011	0.011	0.023^{*}	0.021^{*}	0.001	-0.002
	(0.007)	(0.007)	(0.011)	(0.010)	(0.006)	(0.006)
Reduced form	0.005	0.005	0.010^{*}	0.009^{*}	0.000	-0.001
	(0.003)	(0.003)	(0.005)	(0.005)	(0.003)	(0.003)
First stage	0.442^{***}	0.442^{***}	0.442^{***}	0.442^{***}	0.442^{***}	0.442^{***}
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
2SLS (Triangular Kernel)	0.012	0.011	0.021	0.023^{*}	0.002	0.000
	(0.008)	(0.008)	(0.012)	(0.011)	(0.007)	(0.006)
Observations	74012	153713	74012	153713	74012	153713
Outcome mean	0.049	0.049	0.113	0.112	0.032	0.032
Bandwidth	15.000	35.000	15.000	35.000	15.000	35.000
F-statistics	4833.499	5187.871	4833.499	5187.871	4833.499	5187.871
			Panel B - Croatia			
2SLS	0.008	0.005	0.010	0.015	0.004	0.005
	(0.007)	(0.008)	(0.012)	(0.014)	(0.006)	(0.008)
Reduced form	0.007	0.004	0.008	0.012	0.003	0.004
	(0.005)	(0.007)	(0.009)	(0.012)	(0.005)	(0.006)
First stage	0.807^{***}	0.803^{***}	0.807^{***}	0.803^{***}	0.807^{***}	0.803^{***}
	(0.008)	(0.009)	(0.008)	(0.009)	(0.008)	(0.009)
2SLS (Triangular Kernel)	0.002	0.000	0.015	0.022	0.005	0.006
	(0.008)	(0.010)	(0.015)	(0.017)	(0.007)	(0.009)
Observations	31698	42421	31698	42421	31698	42421
Outcome mean	0.059	0.059	0.218	0.219	0.054	0.054
Bandwidth	80.000	120.000	80.000	120.000	80.000	120.000
F-statistics	10158.245	7440.903	10158.245	7440.903	10158.245	7440.903
			Panel C - Sweden			
2SLS	0.000	-0.004	-0.001	-0.009	0.000	-0.001
	(0.008)	(0.008)	(0.010)	(0.011)	(0.004)	(0.005)
Reduced form	0.000	-0.001	0.000	-0.002	0.000	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
First stage	0.201^{***}	0.199^{***}	0.201^{***}	0.199^{***}	0.201^{***}	0.199^{***}
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
2SLS (Triangular Kernel)	-0.004	-0.006	-0.012	-0.013	0.000	-0.001
	(0.008)	(0.008)	(0.011)	(0.011)	(0.005)	(0.005)
Observations	398036	624877	398036	624877	398036	624877
Outcome mean	0.040	0.039	0.087	0.085	0.014	0.013
Bandwidth	0.390	0.610	0.390	0.610	0.390	0.610
F-statistics	5103.422	4455.739	5103.422	4455.739	5103.422	4455.739

Notes: All the specifications in the table control for a linear or quadratic polynomial of older siblings' application score centered around target majors admission cutoff. Older siblings' application year, target major-year and younger siblings' birth year fixed effect are included as controls. 2SLS (Triangual Kernel) specifications use a triangular kernel to give more weight to observations close to the cutoff. Bandwidths were computed according to Calonico et al. (2014) for each outcome independently. The smallest one among the three is used for all the outcomes. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 8: Probability of Applying to Older Sibling's Target Major and Target College by Older Siblings' Gender

	Major			College		
	Older Siblings' Gender			Older Siblings' Gender		
	All	Female	Male	All	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
Older sibling enrolls	0.023***	0.023***	0.023**	0.094***	0.061**	0.124***
	(0.005)	(0.007)	(0.008)	(0.016)	(0.023)	(0.023)
Older sibling enrolls × Same gender	0.010**	0.001	0.019**	0.014	0.032	-0.001
	(0.004)	(0.005)	(0.006)	(0.012)	(0.017)	(0.017)
Observations	136364	73014	61982	73331	39129	32302
Outcome mean	0.056	0.051	0.062	0.012	0.010	0.014
Bandwidth	20.000	20.000	20.000	15.000	15.000	15.000
F-statistics	6933.231	3310.962	3530.694	2719.593	1278.857	1337.943
			Panel B - Croatia			
Older sibling enrolls	0.026**	0.031*	0.025	0.114***	0.098**	0.124***
	(0.009)	(0.013)	(0.015)	(0.022)	(0.031)	(0.033)
Older sibling enrolls × Same gender	0.023*	0.007	0.044**	-0.007	-0.027	0.001
	(0.009)	(0.012)	(0.016)	(0.020)	(0.027)	(0.032)
Observations	36757	22239	14203	12950	7545	5008
Outcome mean	0.129	0.123	0.141	0.555	0.552	0.556
Bandwidth	80.000	80.000	80.000	80.000	80.000	80.000
F-statistics	7220.184	3662.675	4025.070	3229.534	1651.529	1405.970
			Panel C - Sweden			
Older sibling enrolls	0.025***	0.036***	0.013	0.143***	0.154***	0.139***
	(0.006)	(0.008)	(0.009)	(0.014)	(0.019)	(0.024)
Older sibling enrolls × Same gender	0.008*	-0.019**	0.045***	0.011	-0.003	0.040*
	(0.004)	(0.006)	(0.007)	(0.011)	(0.014)	(0.019)
Observations	732025	438419	281549	444203	273981	160086
Outcome mean	0.047	0.042	0.057	0.193	0.183	0.211
Bandwidth	0.510	0.510	0.510	0.370	0.370	0.370
F-statistics	5419.139	2441.736	2717.178	3075.133	1484.510	1330.244

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major and college by siblings' gender. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Tables 3 and 5. Specifications also control by a dummy variable that indicates if the siblings are of the same gender. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 9: Probability of Applying in Older Sibling's Target Major and College by Siblings' Similarity

	Major		College	
	Δ Age > 5 (1)	Δ GPA (2)	Δ Age > 5 (3)	Δ GPA (4)
Panel A - Chile				
Older sibling enrolls	0.030***	0.056***	0.112***	0.170***
	(0.005)	(0.006)	(0.015)	(0.017)
Interaction	-0.004	-0.029***	-0.027*	-0.076***
	(0.004)	(0.002)	(0.012)	(0.007)
Observations	135777	133703	73030	71865
Outcome mean	0.056	0.057	0.302	0.308
Bandwidth	20.000	20.000	15.000	15.000
F-statistics	6904.432	6789.416	2710.198	2664.690
Panel B - Croatia				
Older sibling enrolls	0.039***	0.075**	0.109***	0.195***
	(0.009)	(0.025)	(0.020)	(0.052)
Interaction	-0.018	-0.033*	0.000	-0.031
	(0.013)	(0.014)	(0.026)	(0.032)
Observations	36756	8567	12950	2588
Outcome mean	0.129	0.160	0.555	0.609
Bandwidth	80.000	80.000	80.000	80.000
F-statistics	7225.706	1567.759	3230.667	648.627
Panel C - Sweden				
Older sibling enrolls	0.035***	0.032***	0.162***	0.179***
	(0.005)	(0.007)	(0.013)	(0.017)
Interaction	-0.015***	0.005	-0.030**	-0.002
	(0.004)	(0.003)	(0.011)	(0.008)
Observations	732025	591599	444203	359012
Outcome mean	0.047	0.055	0.193	0.222
Bandwidth	0.510	0.510	0.370	0.370
F-statistics	5255.957	4573.374	2975.652	2610.561

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major and college by siblings' similarity. Columns (1) and (3) investigate heterogeneous effects by age difference, while columns (2) and (4) by difference in high school GPA. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Tables 3 and 5. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 10: Probability of Applying in Older Sibling's Target Major and Target College by Quality

	Major			**College**		
	Admitted students quality (1)	Dropout (2)	Earnings (3)	Admitted students quality (4)	Dropout (5)	Earnings (6)
			Panel A - Chile			
Older sibling enrolls	0.021*	0.027***	0.026***	0.027	0.117***	0.099***
	(0.009)	(0.006)	(0.005)	(0.029)	(0.015)	(0.016)
Interaction	0.002	-0.004	0.007***	0.024***	-0.139*	0.010
	(0.002)	(0.029)	(0.002)	(0.006)	(0.069)	(0.006)
Observations	136364	121676	129847	73331	72642	69927
Outcome mean	0.056	0.057	0.057	0.302	0.302	0.304
Bandwidth	20.000	20.000	20.000	15.000	15.000	15.000
F-statistic	4914.155	5831.462	5732.572	1872.447	2459.612	2183.694
			Panel B - Croatia			
Older sibling enrolls	0.038			-0.010		
	(0.025)			(0.058)		
Interaction	-0.001			0.027*		
	(0.005)			(0.013)		
Observations	34510			10693		
Outcome mean	0.130			0.537		
Bandwidth	80.000			80.000		
F-statistic	6833.719			2598.965		
			Panel C - Sweden			
Older sibling enrolls	0.019**	0.015**	0.019***	0.120***	0.118***	0.110***
	(0.006)	(0.005)	(0.006)	(0.015)	(0.013)	(0.016)
Interaction	0.012***	-0.028	0.010***	0.036***	-0.126**	0.010
	(0.003)	(0.015)	(0.003)	(0.008)	(0.044)	(0.008)
Observations	732023	535714	358644	444203	320107	218552
Outcome mean	0.047	0.046	0.045	0.193	0.186	0.193
Bandwidth	0.510	0.510	0.510	0.370	0.367	0.367
F-statistic	4508.761	5465.470	2462.490	2577.150	2678.503	1380.629

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major or college by different quality measures of their target majors. Columns (1) and (4) investigate heterogeneous effects by the average quality of admitted students, columns (2) and (5) by first year dropout rates and columns (3) and (6) by graduates average earnings. Students' quality is measured by the average scores of admitted students in the admission exam. The measure of students quality and graduates average earnings are standardized. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Tables 3 and 5. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 11: Probability of Applying and Enrolling in Older Sibling's Target Major-College by Quality Difference respect Counterfactual Alternative

	Major			**College**		
	Δ Admitted students quality (1)	Δ Dropout (2)	Δ Earnings (3)	Δ Admitted students quality (4)	Δ Dropout (5)	Δ Earnings (6)
Panel A - Chile						
Older sibling enrolls	0.028***	0.028***	0.025***	0.108***	0.101***	0.103***
	(0.006)	(0.005)	(0.005)	(0.017)	(0.016)	(0.016)
Interaction	0.000	-0.003	0.006*	-0.005	-0.165	-0.013
	(0.005)	(0.037)	(0.003)	(0.015)	(0.105)	(0.021)
Observations	99652	90784	90082	45082	41229	40836
Outcome mean	0.062	0.062	0.062	0.319	0.322	0.323
Bandwidth	20.000	20.000	20.000	15.000	15.000	15.000
F-statistics	7674.012	7397.956	7219.418	3153.688	2959.387	2908.442
Panel B - Croatia						
Older sibling enrolls	0.034***			0.107***		
	(0.009)			(0.021)		
Interaction	-0.003			0.007		
	(0.005)			(0.010)		
Observations	34510			10693		
Mean y	0.130			0.537		
Bandwidth	80.000			80.000		
F-statistics	6854.732			2607.328		
Panel C - Sweden						
Older sibling enrolls	0.033***	0.017**		0.185***	0.116***	
	(0.006)	(0.006)		(0.015)	(0.014)	
Interaction	-0.015***	-0.002		-0.053***	-0.009	
	(0.003)	(0.002)		(0.010)	(0.007)	
Observations	472966	309934		262275	172027	
Mean y	0.054	0.053		0.200	0.196	
Bandwidth	0.510	0.510		0.367	0.367	
F-statistics	4439.812	4419.105		4439.812	4419.105	

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major and college by the gap between older siblings' target and counterfactual major in different quality measures. Columns (1) and (4) investigate heterogeneous effects by the difference in the average quality of admitted students, columns (2) and (5) by the difference in first year dropout rates and columns (3) and (6) by the difference in graduates average earnings. Students' quality is measured by the average scores of admitted students in the admission exam. The measure of students quality and graduates average earnings are standardized. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 3. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. In this table, the sample is restricted to older siblings with counterfactual programs in their application lists. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 12: Probability of Applying and Enrolling in Older Sibling's Target Major and Target College by Older Siblings' Dropout

	Chile		Sweden	
	Applies (1)	Enrolls (2)	Applies (3)	Enrolls (4)
	Panel A - Major			
Older sibling enrolls	0.024***	0.007*	0.046***	0.007***
	(0.008)	(0.004)	(0.008)	(0.002)
Older sibling enrolls × Older sibling drops-out	-0.024**	-0.005*	-0.037***	-0.005***
	(0.007)	(0.003)	(0.007)	(0.002)
Observations	49823	49823	732025	732025
Outcome mean	0.067	0.015	0.047	0.004
Bandwidth	20.000	20.000	0.510	0.510
F-statistics	4210.832	4210.832	3413.123	3413.123
	Panel B - College			
Older sibling enrolls	0.116***	0.044**	0.212***	0.088***
	(0.024)	(0.017)	(0.019)	(0.009)
Older sibling enrolls × Older sibling drops-out	-0.070**	-0.060***	-0.139***	-0.055***
	(0.023)	(0.015)	(0.017)	(0.008)
Observations	24753	24753	444203	444203
Outcome mean	0.348	0.126	0.193	0.034
Bandwidth	15.000	15.000	0.370	0.370
F-statistics	1516.263	1516.263	1945.998	1945.998

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major on younger siblings' probability of applying to and enrolling in that major. The specifications include the same controls and use the same bandwidths described in Tables 3 and 5. They also control for a dummy variable that indicates if older siblings dropout from the major in which they initially enroll. The samples used in these last columns only include individuals whose older siblings enroll in a major. In parenthesis, standard errors clustered at family level.*p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 13: Effect of Older Siblings' Enrollment in Target Major-College on Academic Performance (Major Sample)

	Takes admission exam (AE) (1)	Applies to college/higher ed. (2)	High School GPA (3)	Average Score AE (4)
		Panel A - Chile		
Older sibling enrolls	0.002	0.014	0.014	0.036
	(0.004)	(0.010)	(0.025)	(0.024)
Observations	136,364	136,364	136,364	136,364
Outcome mean	0.957	0.583	-0.105	0.256
Bandwidth	20.000	20.000	20.000	20.000
F-statistic	13867.401	13867.401	13867.401	13867.401
		Panel B - Croatia		
Older sibling enrolls	-0.013		-0.120	-0.102
	(0.017)		(0.127)	0.085
Observations	12,443		12,443	12,443
Outcome mean	0.825		-1.298	-0.834
Bandwidth	80.000		80.000	80.000
F-statistic	4498.481		4498.481	4498.481
		Panel C - Sweden		
Older sibling enrolls	-0.056***	-0.034**	0.007	0.032
	(0.012)	(0.011)	(0.025)	(0.035)
Observations	732,025	732,025	613,294	344,442
Outcome mean	0.484	0.577	0.219	0.051
Bandwidth	0.510	0.510	0.510	0.510
F-statistic	10838.800	10838.800	9529.889	6498.021

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major on younger siblings' probability of taking the admission exam and applying to college (columns 1 and 2), and on different measures of academic performance: high school GPA (column 3), reading and math sections of the admission exam (columns 4 and 5) and average performance on the admission exam (column 6). While in Chile and Croatia we only observe applications to college degrees, in Sweden we also observe applications to other higher education programs. These analyses focus on the Major Sample. This means that in this case, marginal admission or rejection from their target major, changes the major, but not necessarily the college or field in which older siblings are admitted. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 5. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

A Identification Strategy: Further Discussion

This section discusses the assumptions under which our identification strategy provides us with a consistent estimator of the effects of interest. As discussed in Section 4.4, a fuzzy RD can be thought as an IV. In what follows, and for ease of notation, we drop time and individual indices t, i, τ and focus our analysis on a specific major-college u. Following this notation, the treatment in which we are interested is:

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

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

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

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

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

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

1. Independence of the instrument:

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

2. Exclusion restriction:

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

3. First stage:

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

4. Monotonicity:

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

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

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

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

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

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

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

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

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

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

$$\frac{\hat{E}[Y_u|adm_u = 1] - \hat{E}[Y_u|adm_u = 0]}{\hat{E}[O_u|adm_u = 1] - \hat{E}[O_u|adm_u = 0]} = \frac{\hat{E}[Y_u(1) - Y_u(0)]}{\hat{E}[O_u(1) - O_u(0)]}. \tag{2}$$

The numerator of equation 2, after applying law of iterated expectations, becomes:

$$\hat{E}[Y_u(1) - Y_u(0)] = \tag{3}$$

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

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

Similarly, by virtue of assumption 4.1., the denominator of equation 2 becomes:

$$\hat{E}[O_u(1)-O_u(0)] = P(O_u(1)=1, O_u(0)=0). \tag{4}$$

Taken together, 3 and 4 imply:

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

□

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

B Robustness Checks

This section investigates if the identification assumptions of our empirical strategy are satisfied. We start by investigating if there is any evidence of manipulation of the running variable. Next, we check if other variables that could affect individuals' application and enrollment decisions present jumps at the cutoff and if the results are robust to different bandwidths. We continue by performing two types of placebo exercises. In the first, we study if similar effects arise when looking at placebo cutoffs (i.e. cutoffs that do not affect older siblings' admission). In the second, we analyze if similar effects arise when looking at the effect of the younger sibling enrollment on older siblings decisions. We then investigate if our conclusions change when allowing the slope of the running variable to vary by major-college and year and when re-weighting the observations around each cutoff by the number of applicants around them (i.e. to make all the cutoffs that we are pooling together equally relevant in the estimation). Finally, we end this section by showing that there are no extensive margin responses (i.e. increases in total enrollment) that could explain our findings.

B.1 Manipulation of the Running Variable

A first condition for the validity of our RD estimates is that individuals should not be able to manipulate their older siblings' application scores around the admission cutoff. The structures of the admission systems in Chile, Croatia and Sweden make the violation of this assumption unlikely. However, to confirm this we study whether the distribution of the running variable (i.e. older sibling's application score centered around the relevant cutoff) is continuous at the cutoff. We do this by implementing the test suggested by Cattaneo et al. (2018), the results of which are presented in Figure B1. As expected, we do not detect discontinuities in the distribution of the running variable at the cutoff for any of the three countries.[59] In Sweden, Figure B1 only focuses on the distribution of the high school GPA. As discussed in Section 2, the admission exam is voluntary in Sweden, and institutions select their students using two independent pools that consider either the applicants' high school GPA or the applicants' scores in the admission exam. Considering that the distribution of admission exam scores is coarser, to investigate manipulation of these scores we present histograms of these variables in Figures B2. In Sweden, the admission exam was changed in 2013. Thus, Appendix Figure B2 presents two histograms. One for the distribution before and one for the distribution after the change.

Strictly speaking, the density of the running variable needs to be continuous around each admission cutoff. In our analysis, we pool them together because there are thousands of cutoffs in our samples and studying them independently would be impractical.

[59] The density tests illustrated in Figure B1 omit observations exactly at the cutoff. This explains the pattern of confidence intervals close the cutoff. We omit observations exactly at 0 because pooling together multiple cutoffs mechanically generates an excess of mass at that point.

B.2 Discontinuities in Potential Confounders

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

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

Figure B3 summarizes this result. It plots the estimated discontinuities at the cutoff and their 95% confidence intervals. To estimate these discontinuities we control for a linear polynomial of the running variable and allow for the slope to change at the cutoff. Using the same bandwidths reported for linear specifications in Section 5, we find no statistically significant jump at the cutoff for any of the potential confounders being investigated.

The only exception is the age at which individuals apply to higher education in Sweden. In this case, we find that individuals with older siblings marginally admitted to their target major in the past are older than those with older sibling marginally rejected. However, this difference is very small. They are less than 14.6 days older.

B.3 Different Bandwidths

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

Figures B4, B5 and B6 show how the estimated coefficients change when reducing the bandwidth used in the estimations. Although the standard errors increase as the sample size is reduced, the coefficients remain stable.

B.4 Placebo Exercises

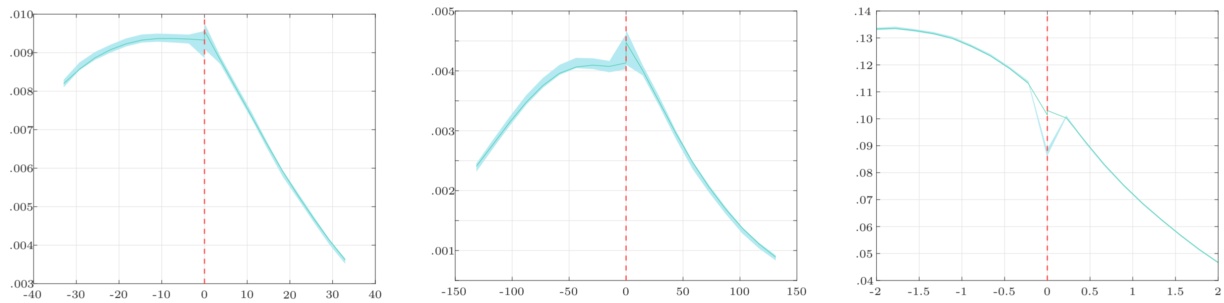
This setting allows us to perform two types of placebo exercises. First, in Figures B10, B11 and B12 we study if younger siblings' enrollment affect the application decisions of their older siblings. Since younger siblings apply to college after their older siblings, being marginally admitted or rejected from a major or college should not affect what happens with older siblings. These figures show that this is indeed the case. In addition, in Figures B7, B8 and B9 we show that only at the real cutoff we observe a discontinuity on younger siblings outcomes This is not surprising since these fake cutoffs do not generate any increase in older siblings' admission.

B.5 Alternative Specifications and Total Enrollment

Figures B13, B14 and B15 and Tables B1, B2, B3, B5, B6 and B7 study how robust our estimates are to the degree of the local polynomial used, to re-weighting the observations by the inverse of the total number of applicants in the proximity of each cutoff and to allowing the running variable to have different slopes for each cutoff-major. In addition, Tables B8, B9 and B10 present results in which target $\times$ counterfactual major fixed effects are used. The results are robust to these changes, and although the magnitude of the coefficients is smaller when re-weighting the observations, the general picture remains unchanged.

Finally, Table B4 investigates if the marginal admission of older siblings translates into an increase in total enrollment (i.e. enrollment in any college in the system) for them or for their younger siblings. We do not find evidence of extensive margin responses in any of the countries studied. Thus, our findings are not driven by a general increase on younger siblings enrollment. In terms of older siblings' enrollment, we observe a small increase in total enrollment in Chile relative to Croatia. This is not surprising because the group of universities studied in Chile is more selective than the ones we study in Croatia. This means that in Chile, older siblings still have many available colleges in case of rejection.

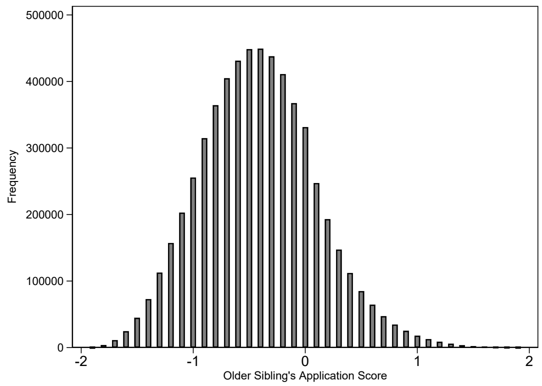
Figure B1: Density of Older Siblings' Application Scores at the Target Major-College Admission Cutoff



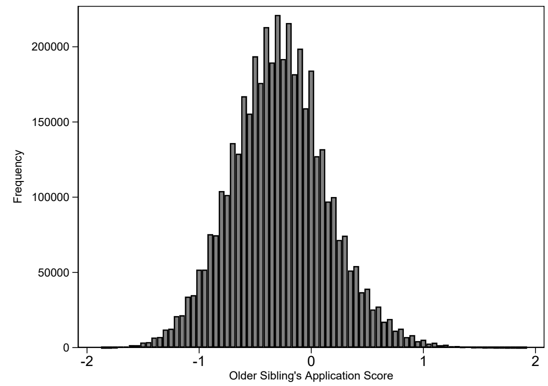
(a) 1st preference - Chile (b) 1st preference - Croatia (c) 1st preference - Sweden

This figure illustrates the density of older siblings' application scores around the cutoff. Figure (a) illustrates this density for Chile, figure (b) for Croatia and figure (c) for Sweden. In Sweden, students can apply to college using their high school GPA or their score in an admission exam (SAT score). In this figure we consider only the students who applied with GPA score, since it is dense enough to be understood as a continuous variable. In the appendix Figure B2, we present the distribution of SAT scores as well. Green lines represent local quadratic polynomials and the blue shadows 95% confidence intervals. In all cases, triangular kernels are used. Bandwidths are estimated according to Cattaneo et al. (2018). The p-values associated to the null hypothesis of no discontinuity at the cutoff are 0.379 , 0.725 and 0.250 respectively.

Figure B2: Density of Older Siblings' SAT Application Scores at the Target Major Admission Cutoff (Sweden)



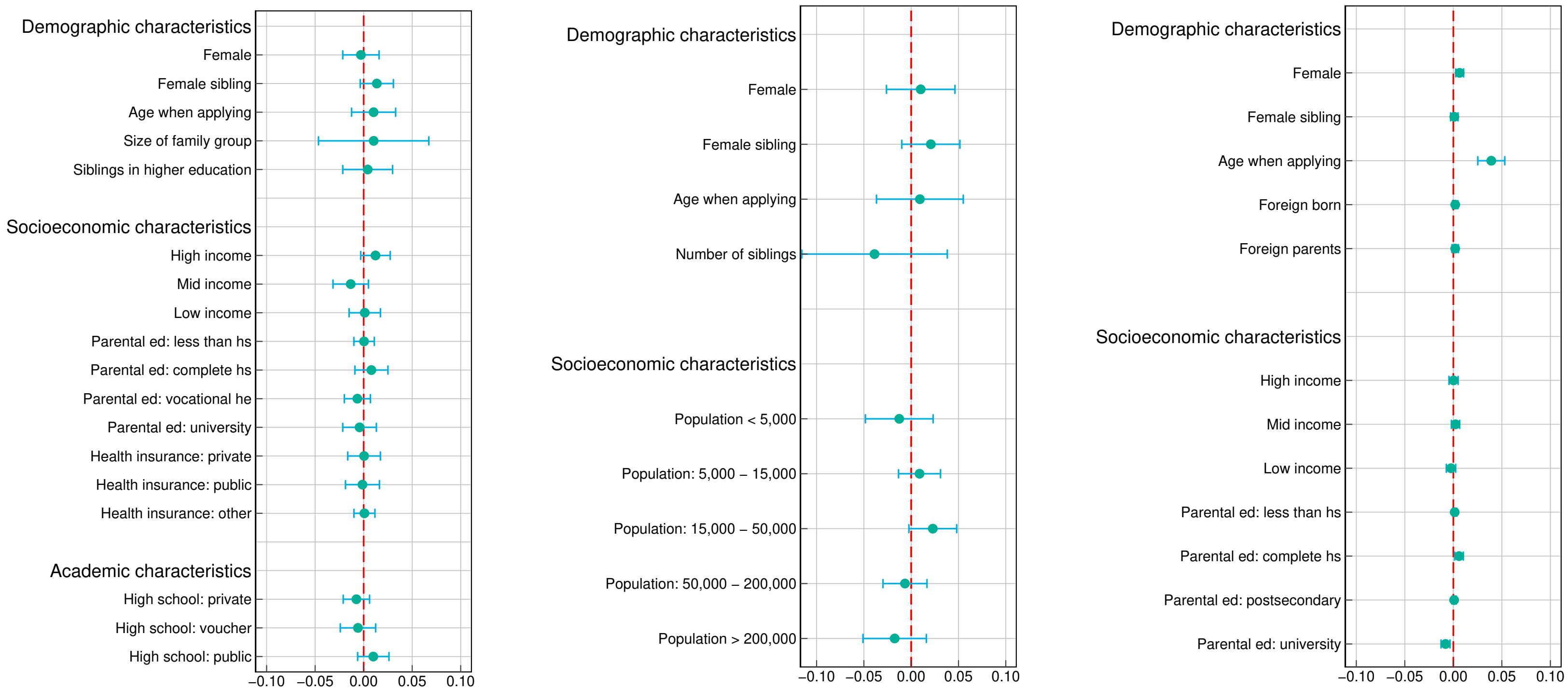
(a) Before 2013



(b) After 2013

These histograms illustrate distributions of older siblings' SAT application scores centred around admission cutoffs in Sweden. The left panel of the figure corresponds to applicants who took the admission exam before 2013 (including 2013). In 2013 there was a structural change in the measurement of SAT scores. The right panel corresponds to applicants who took the admission exam after 2013.

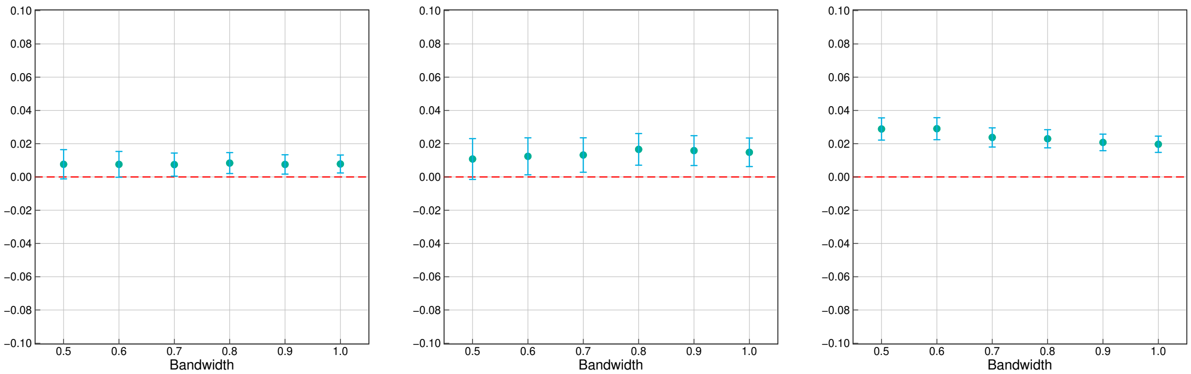
Figure B3: Disconitnuities in other Covariates at the Cutoff



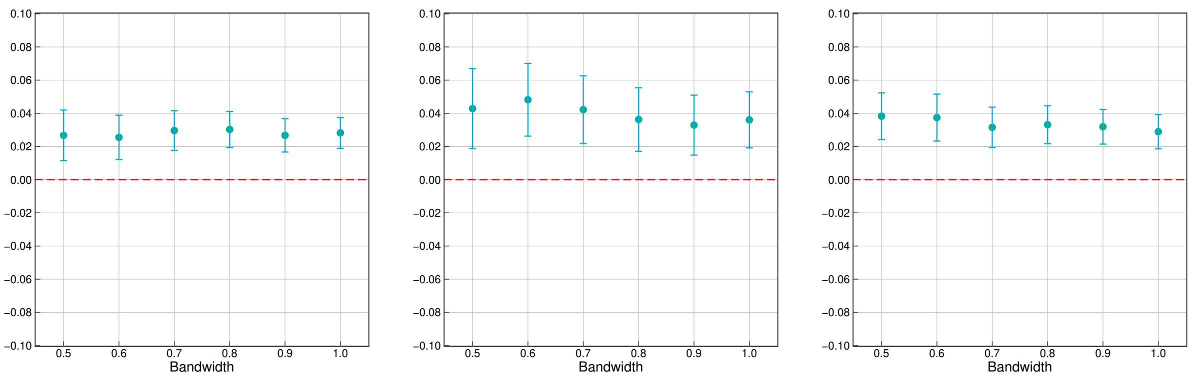
(a) Chile (b) Croatia (c) Sweden

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

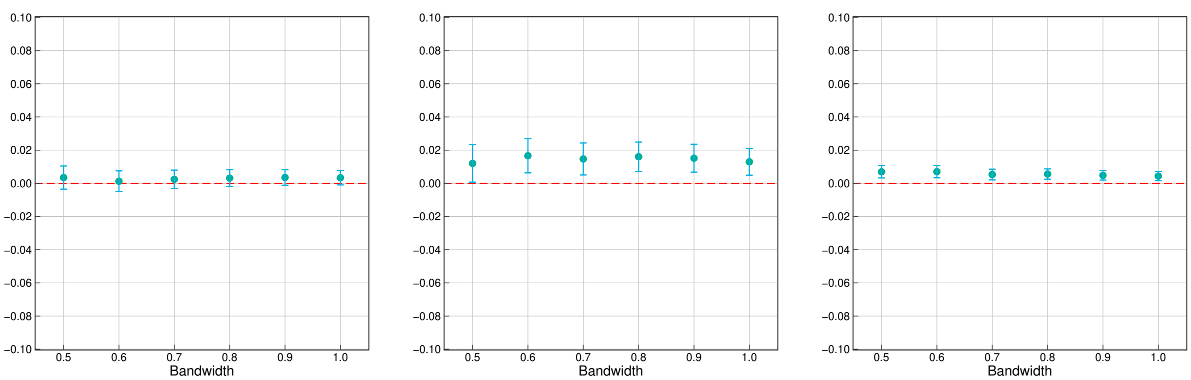
Figure B4: Probabilities of Applying and Enrolling in Older Sibling's Target Major-College - Different Bandwidths



(a) 1st preference - Chile (b) 1st preference - Croatia (c) 1st preference - Sweden



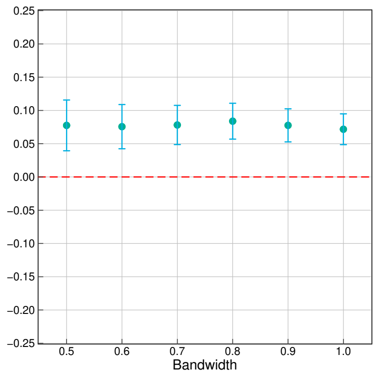
(d) Any preference - Chile (e) Any preference - Croatia (f) Any preference - Sweden



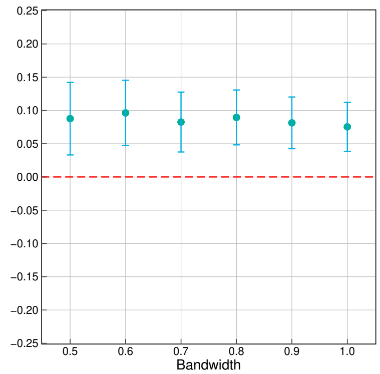
(g) Enrolls - Chile (h) Enrolls - Croatia (i) Enrolls - Sweden

This figure illustrates how being admitted to a specific program changes younger siblings' probabilities of applying and enrolling in the same major. The x-axis corresponds to different bandwidths used to build these figures, chosen as multiples of the optimal bandwidths computed following Calonico et al. (2014). Blue points illustrate estimated effect, and the blue bars denote the 95% confidence intervals. Figures (a), (d) and (g) illustrate the case of Chile, figures (b), (e) and (h) the case of Croatia, while figures (c), (f) and (i) the case of Sweden. The coefficients and their confidence intervals come from parametric specifications that control for a linear polynomial of the running variable.

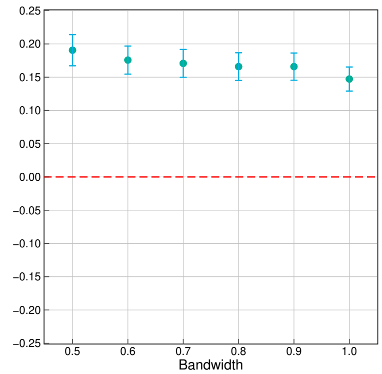
Figure B5: Probabilities of Applying and Enrolling in Older Sibling's Target College - Different Bandwidths



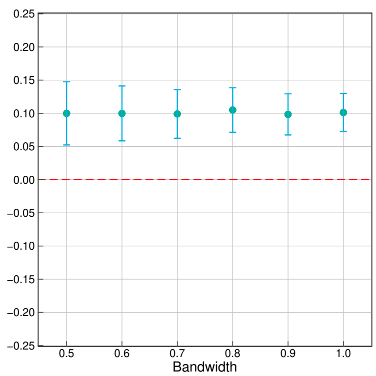
(a) 1st preference - Chile



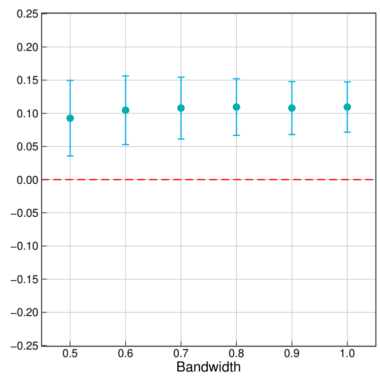
(b) 1st preference - Croatia



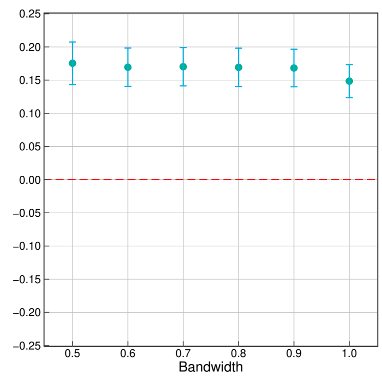
(c) 1st preference - Sweden



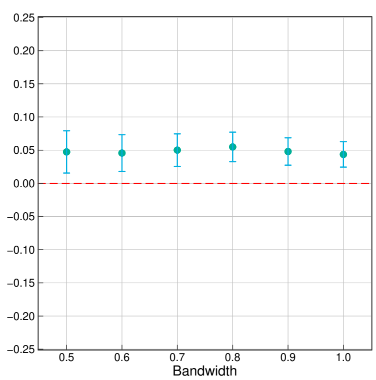
(d) Any preference - Chile



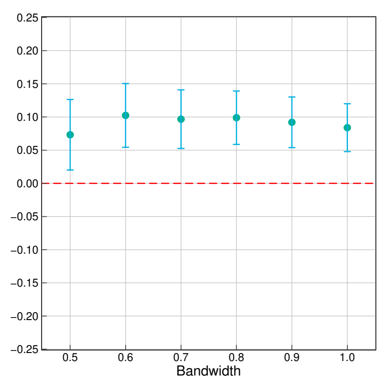
(e) Any preference - Croatia



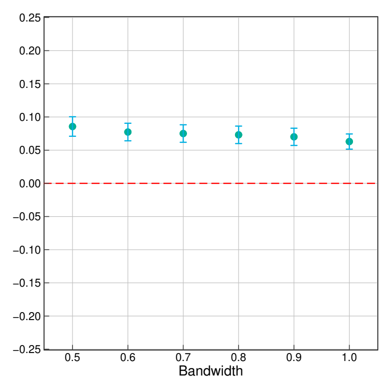
(f) Any preference - Sweden



(g) Enrolls - Chile



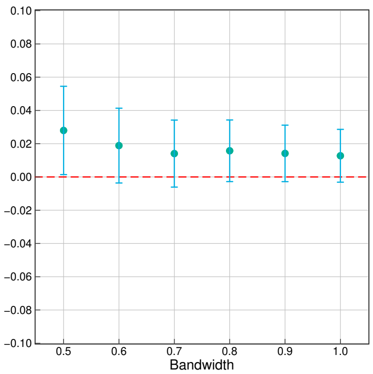
(h) Enrolls - Croatia



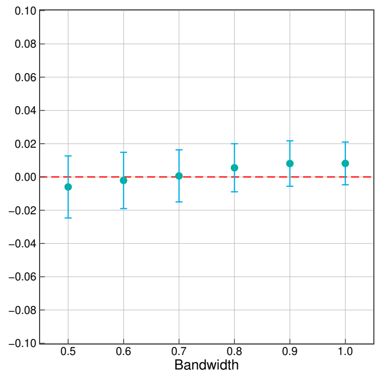
(i) Enrolls - Sweden

This figure illustrates how being admitted to a specific institution changes younger siblings' probabilities of applying and enrolling in the same college. The x-axis corresponds to different bandwidths used to build these figures, chosen as multiples of the optimal bandwidths computed following Calonico et al. (2014). Blue points illustrate estimated effect, and the blue bars denote the 95% confidence intervals. Figures (a), (d) and (g) illustrate the case of Chile, figures (b), (e) and (h) the case of Croatia, while figures (c), (f) and (i) the case of Sweden. The coefficients and their confidence intervals come from parametric specifications that control for a linear polynomial of the running variable.

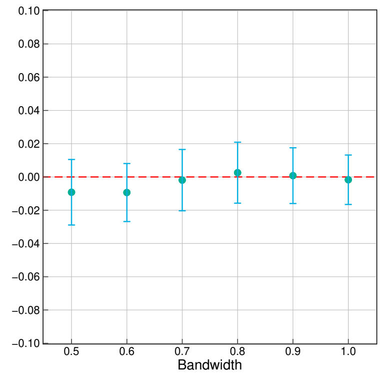
Figure B6: Probabilities of Applying and Enrolling in Older Sibling's Target Field of Study - Different Bandwidths



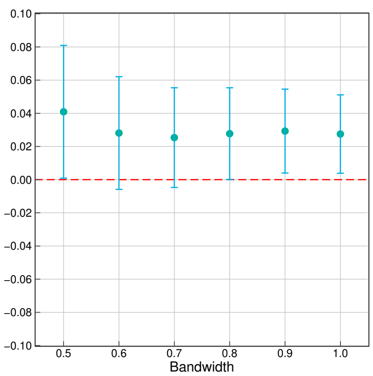
(a) 1st preference - Chile



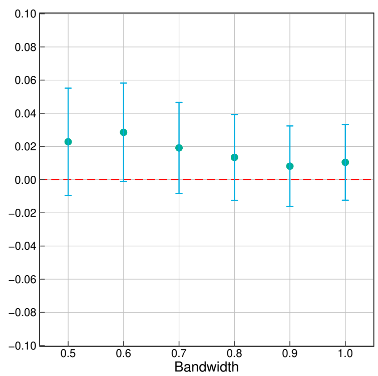
(b) 1st preference - Croatia



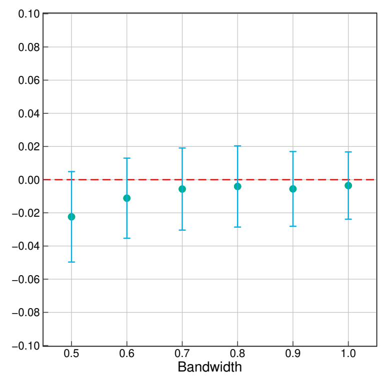
(c) 1st preference - Sweden



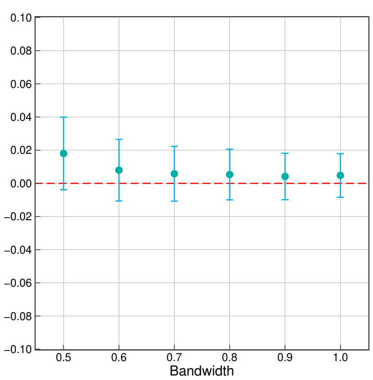
(d) Any preference - Chile



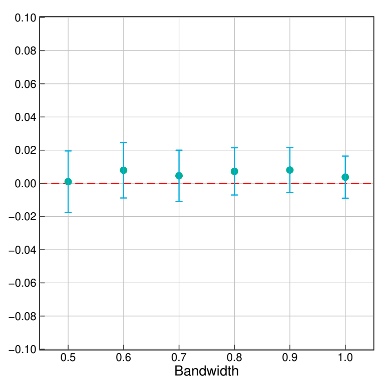
(e) Any preference - Croatia



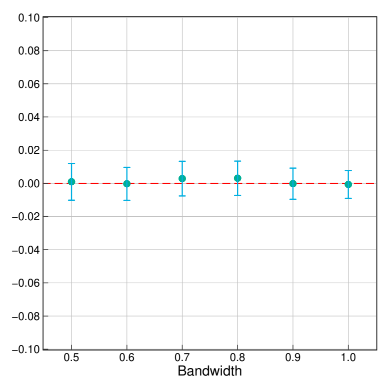
(f) Any preference - Sweden



(g) Enrolls - Chile



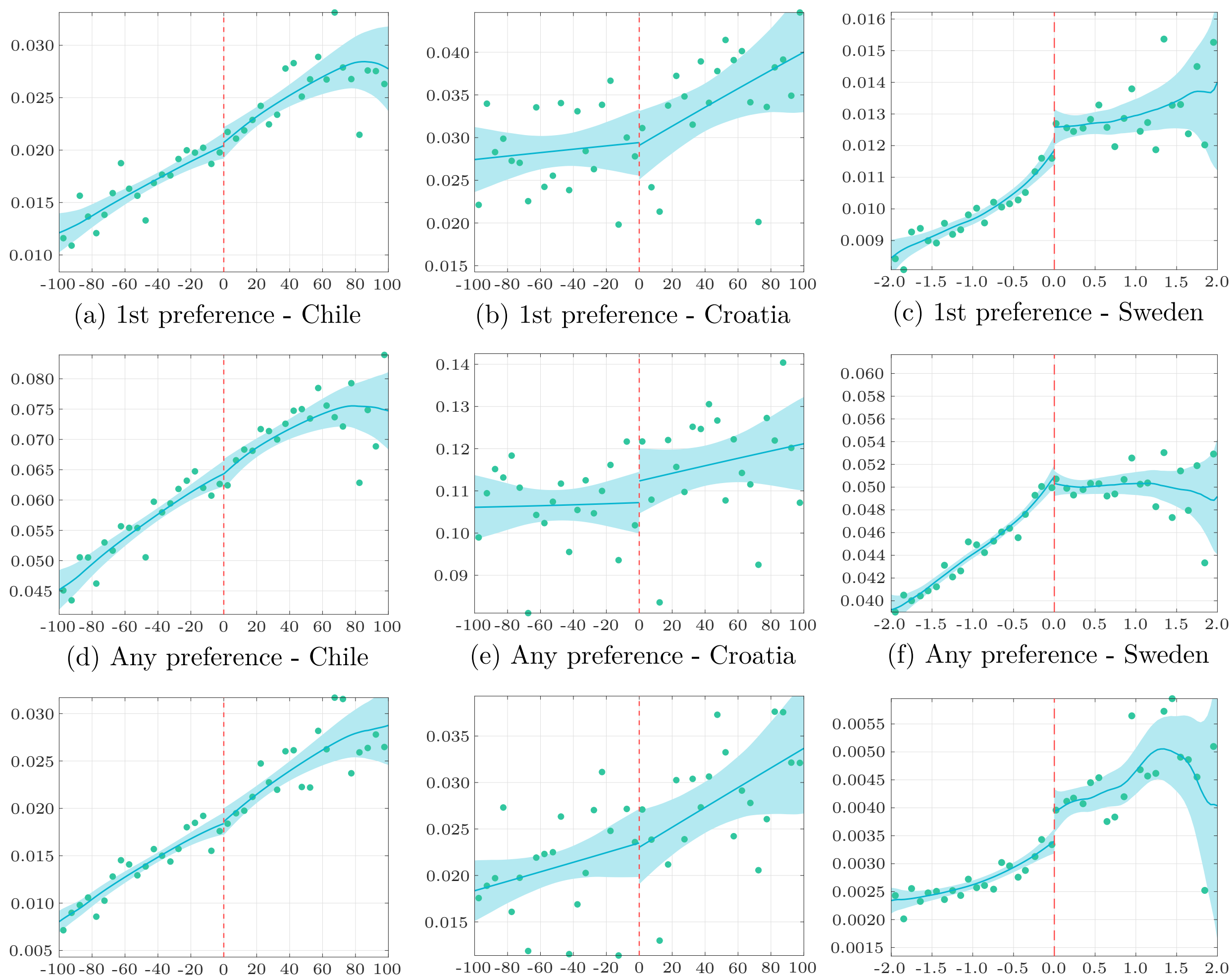
(h) Enrolls - Croatia



(i) Enrolls - Sweden

This figure illustrates how being admitted to a major in a specific field of study changes younger siblings' probabilities of applying and enrolling in a major in the same field. The x-axis corresponds to different bandwidths used to build these figures, chosen as multiples of the optimal bandwidths computed following Calonico et al. (2014). Blue points illustrate estimated effect, and the blue bars denote the 95% confidence intervals. Figures (a), (d) and (g) illustrate the case of Chile, figures (b), (e) and (h) the case of Croatia, while figures (c), (f) and (i) the case of Sweden. The coefficients and their confidence intervals come from parametric specifications that control for a linear polynomial of the running variable. Standard errors are clustered at the family level.

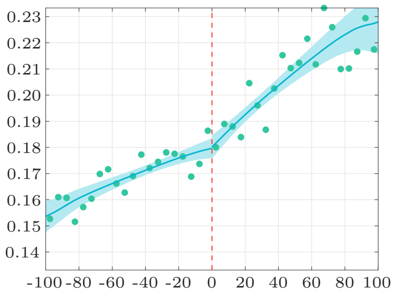
Figure B7: Placebo - Probabilities of Applying and Enrolling in Younger Sibling's Target Major-College



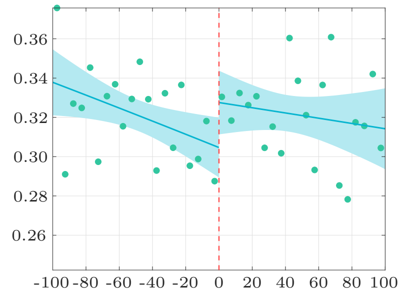
(a) 1st preference - Chile
(b) 1st preference - Croatia
(c) 1st preference - Sweden
(d) Any preference - Chile
(e) Any preference - Croatia
(f) Any preference - Sweden
(g) Enrolls - Chile
(h) Enrolls - Croatia
(i) Enrolls - Sweden

This figure illustrates a placebo exercise that investigates if younger siblings marginal admission to a specific major-college affects the college-major to which older siblings apply to and enroll in. Blue lines and the shadows in the back of them correspond to local polynomials of degree 1 and 95% confidence intervals. Green dots represent sample means of the dependent variable for different values of the running variable.

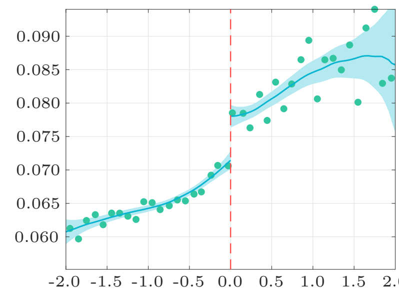
Figure B8: Placebo - Probabilities of Applying and Enrolling in Younger Sibling's Target College



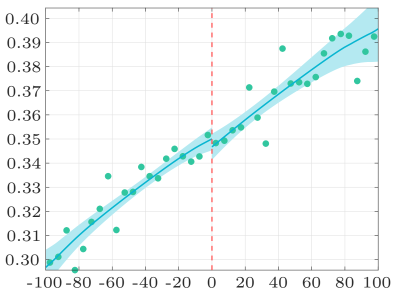
(a) 1st preference - Chile



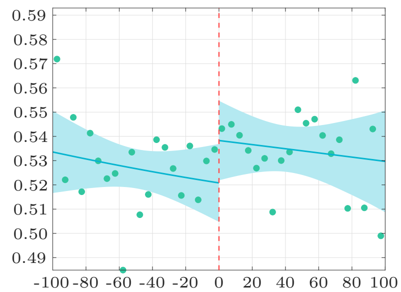
(b) 1st preference - Croatia



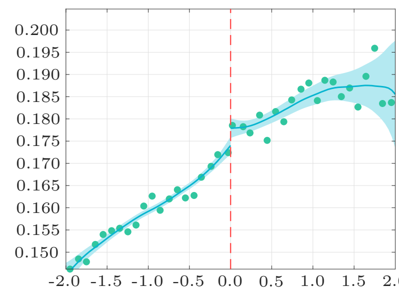
(c) 1st preference - Sweden



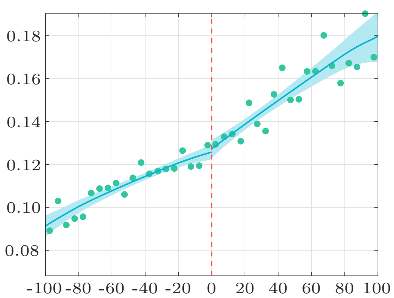
(d) Any preference - Chile



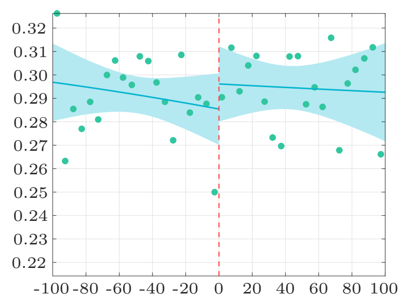
(e) Any preference - Croatia



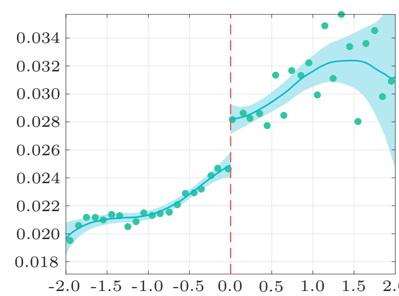
(f) Any preference - Sweden



(g) Enrolls - Chile



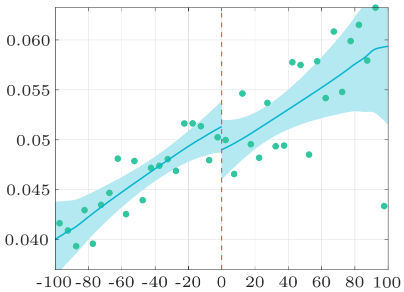
(h) Enrolls - Croatia



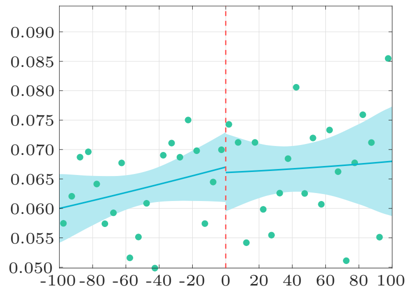
(i) Enrolls - Sweden

This figure illustrates a placebo exercise that investigates if younger siblings marginal admission to a college affects the institution to which older siblings apply to and enroll in. Blue lines and the shadows in the back of them correspond to local polynomials of degree 1 and 95% confidence intervals. Green dots represent sample means of the dependent variable for different values of the running variable.

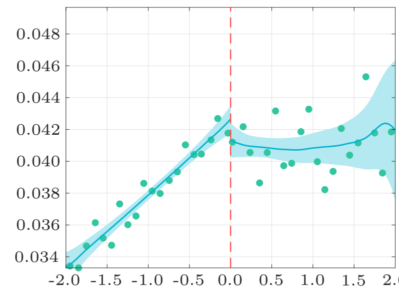
Figure B9: Placebo - Probabilities of Applying and Enrolling in Younger Sibling's Target Field of Study



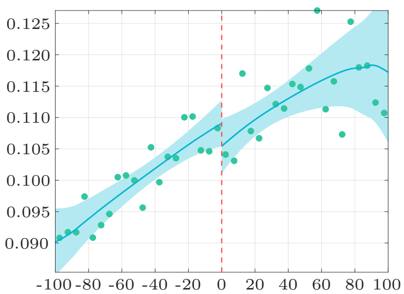
(a) 1st preference - Chile



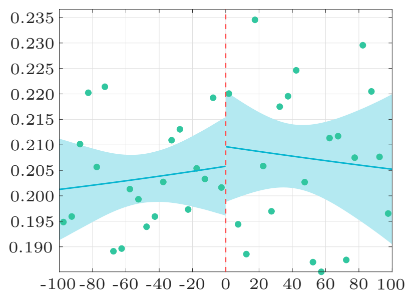
(b) 1st preference - Croatia



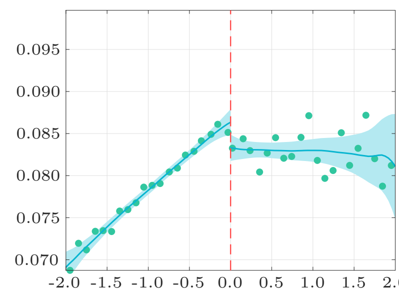
(c) 1st preference - Sweden



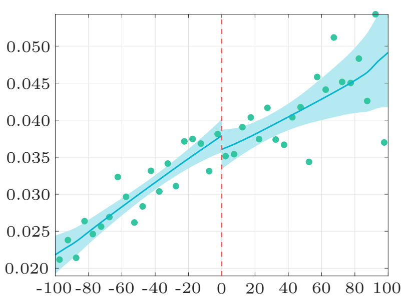
(d) Any preference - Chile



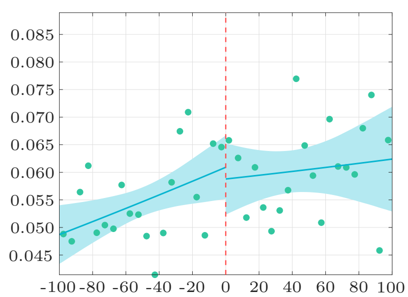
(e) Any preference - Croatia



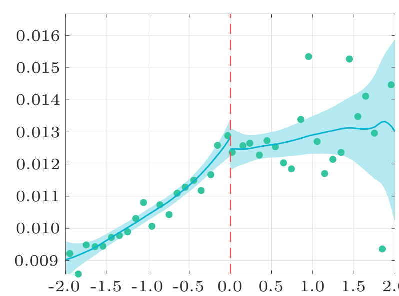
(f) Any preference - Sweden



(g) Enrolls - Chile



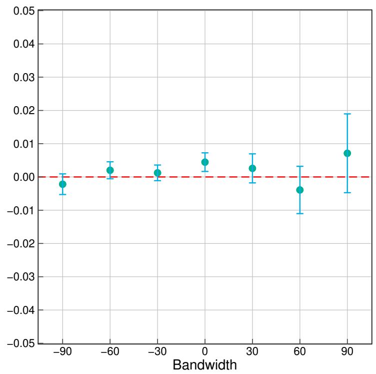
(h) Enrolls - Croatia



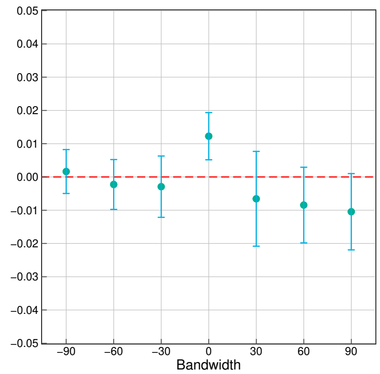
(i) Enrolls - Sweden

This figure illustrates a placebo exercise that investigates if younger siblings marginal admission to a major in a specific field of study affects the field of study to which older siblings apply to and enroll in. Blue lines and the shadows in the back of them correspond to local polynomials of degree 1 and 95% confidence intervals. Green dots represent sample means of the dependent variable for different values of the running variable.

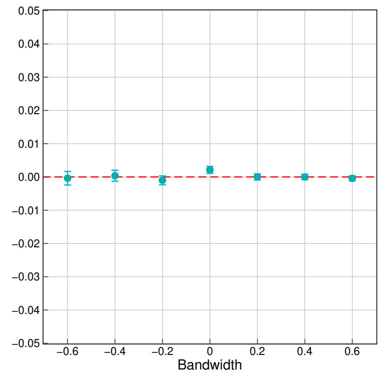
Figure B10: Placebo Cutoffs - Probabilities of Applying and Enrolling in Older Sibling's Target Major-College



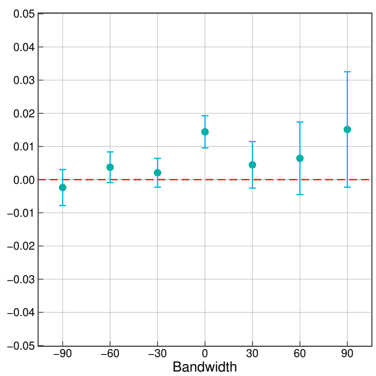
(a) 1st preference - Chile



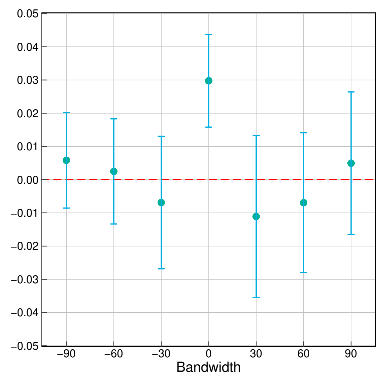
(b) 1st preference - Croatia



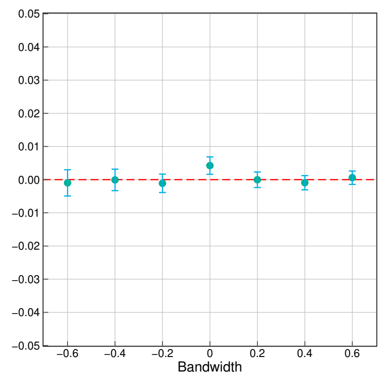
(c) 1st preference - Sweden



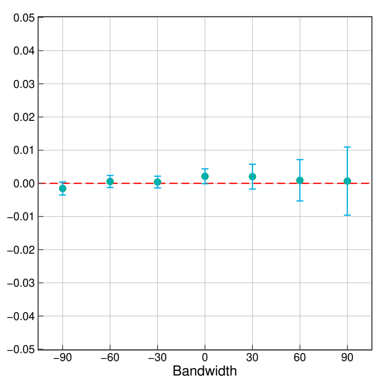
(d) Any preference - Chile



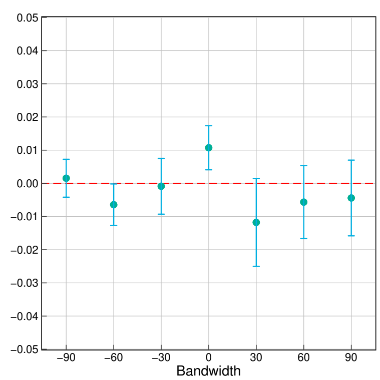
(e) Any preference - Croatia



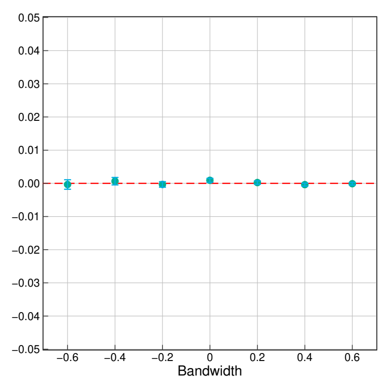
(f) Any preference - Sweden



(g) Enrolls - Chile



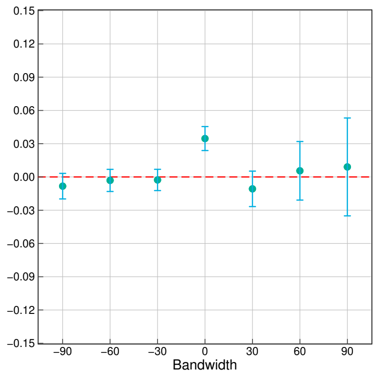
(h) Enrolls - Croatia



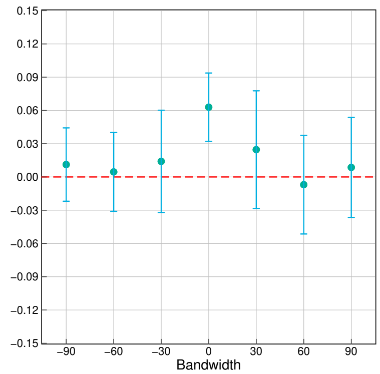
(i) Enrolls - Sweden

This figure illustrates the results of a placebo exercise that investigates if effects similar to the ones documented in figure 2 arise at different values of the running variable. Therefore, the x-axis corresponds to different (hypothetical) values of cutoffs - 0 corresponds to the actual cutoff used in the main body of the paper. The other values correspond to points where older siblings' probability of being admitted to their target major is continuous. Blue points illustrate estimated effect, and the blue bars denote the 95% confidence intervals. Figures (a), (d) and (g) illustrate the case of Chile, figures (b), (e) and (h) the case of Croatia, while figures (c), (f) and (i) the case of Sweden.

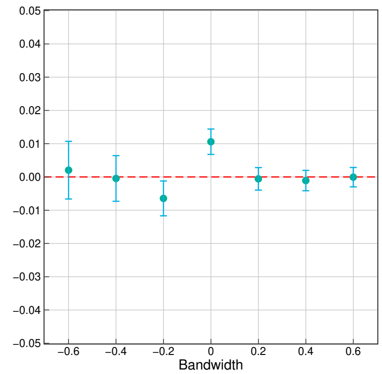
Figure B11: Placebo Cutoffs - Probabilities of Applying and Enrolling in Older Sibling's Target College



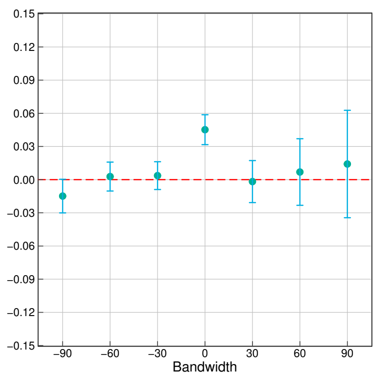
(a) 1st preference - Chile



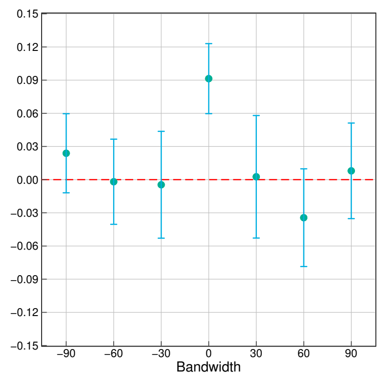
(b) 1st preference - Croatia



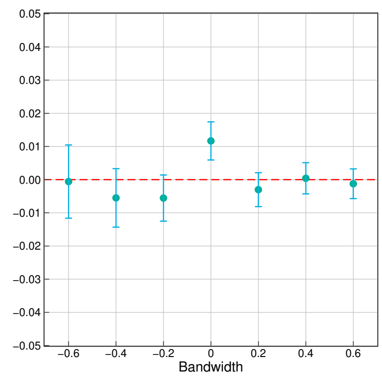
(c) 1st preference - Sweden



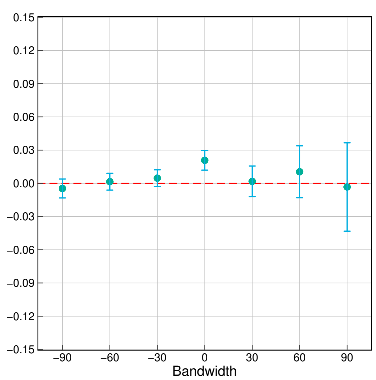
(d) Any preference - Chile



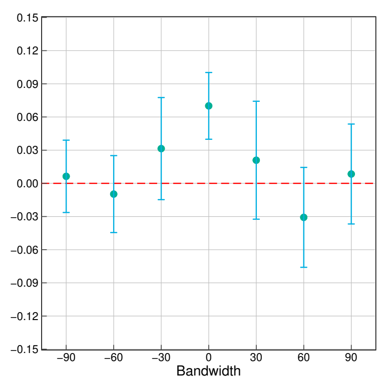
(e) Any preference - Croatia



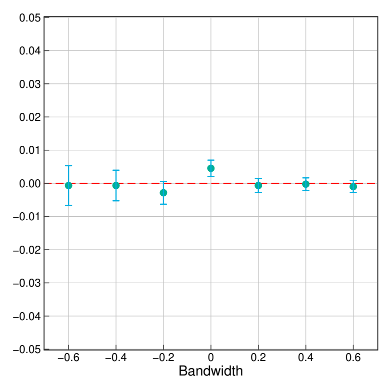
(f) Any preference - Sweden



(g) Enrolls - Chile



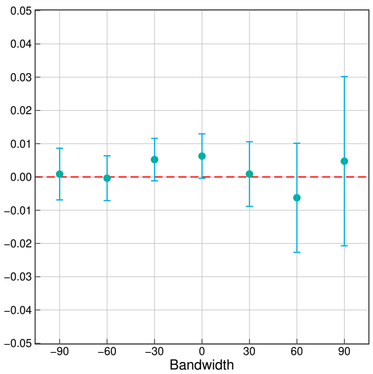
(h) Enrolls - Croatia



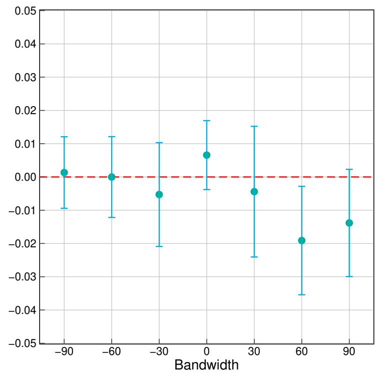
(i) Enrolls - Sweden

This figure illustrates the results of a placebo exercise that investigates if effects similar to the ones documented in figure 3 arise at different values of the running variable. Therefore, the x-axis corresponds to different (hypothetical) values of cutoffs - 0 corresponds to the actual cutoff used in the main body of the paper. The other values correspond to points where older siblings' probability of being admitted to their target majors is continuous. Blue points illustrate estimated effect, and the blue bars denote the 95% confidence intervals. Figures (a), (d) and (g) illustrate the case of Chile, figures (b), (e) and (h) the case of Croatia, while figures (c), (f) and (i) the case of Sweden.

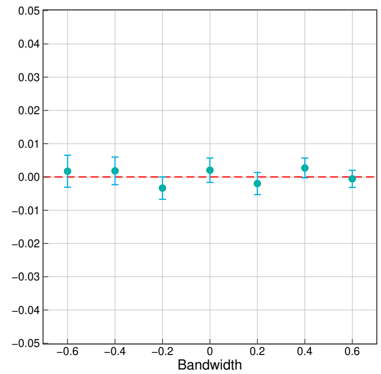
Figure B12: Placebo Cutoffs - Probabilities of Applying and Enrolling in Older Sibling's Target Field of Study



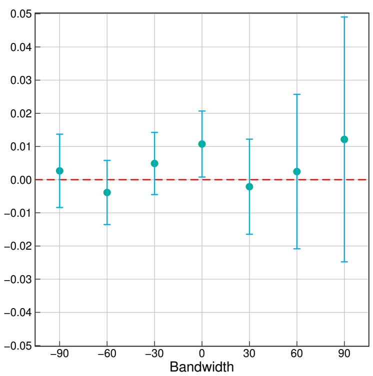
(a) 1st preference - Chile



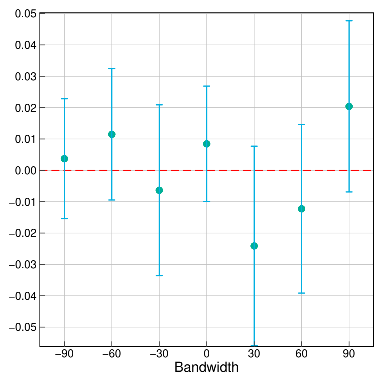
(b) 1st preference - Croatia



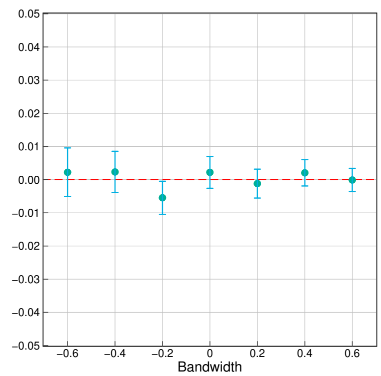
(c) 1st preference - Sweden



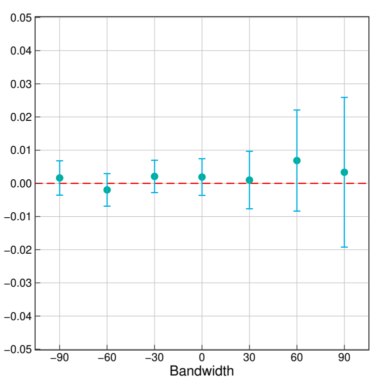
(d) Any preference - Chile



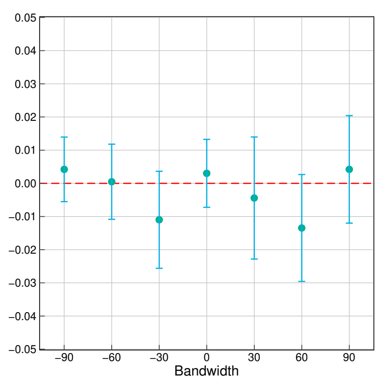
(e) Any preference - Croatia



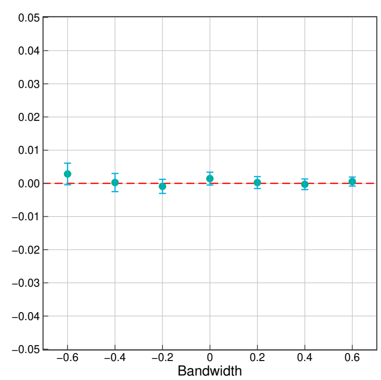
(f) Any preference - Sweden



(g) Enrolls - Chile



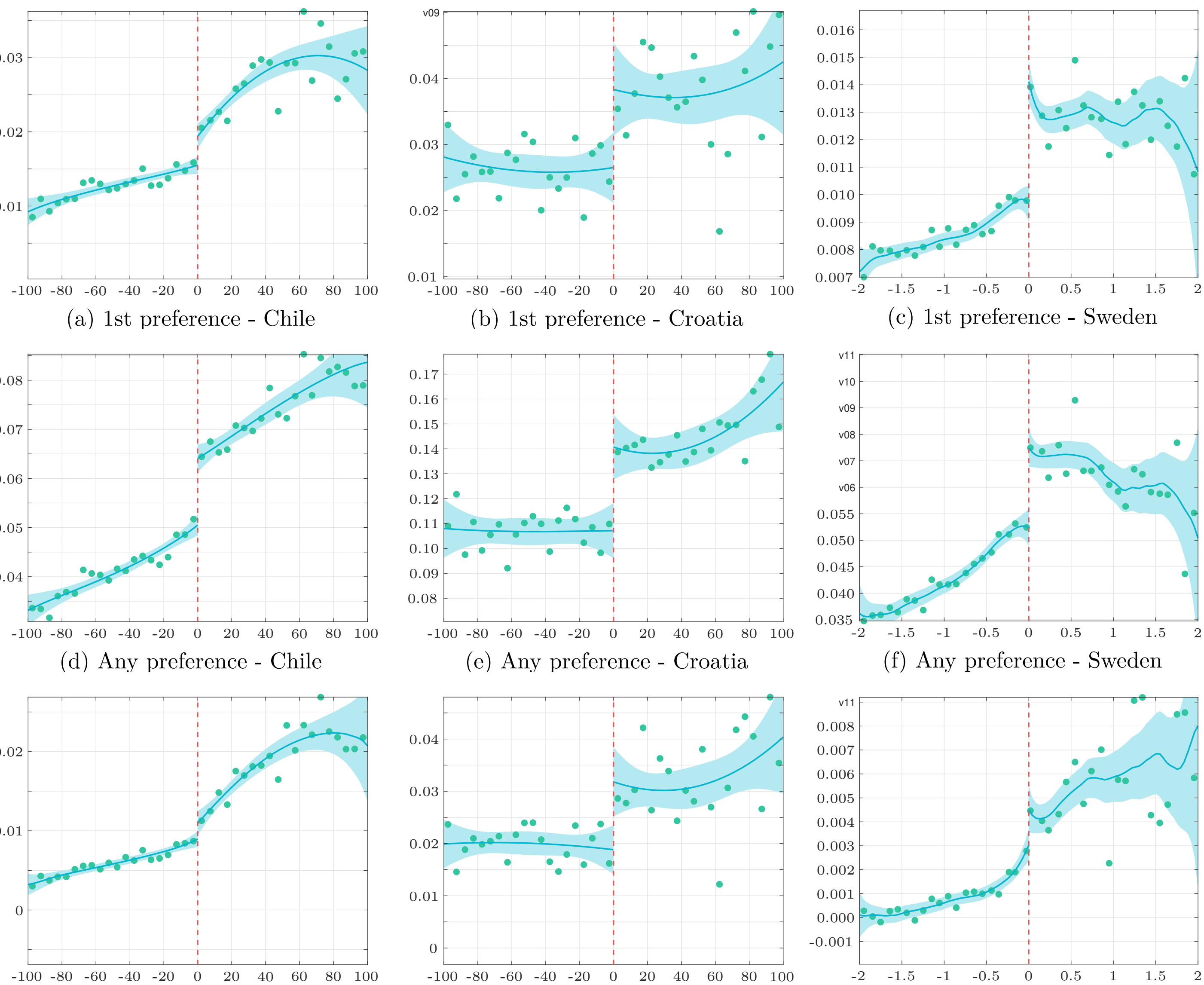
(h) Enrolls - Croatia



(i) Enrolls - Sweden

This figure illustrates the results of a placebo exercise that investigates if effects similar to the ones documented in figure 4 arise at different values of the running variable. Therefore, the x-axis corresponds to different (hypothetical) values of cutoffs - 0 corresponds to the actual cutoff used in the main body of the paper. The other values correspond to points where older siblings' probability of being admitted to their target major is continuous. Blue points illustrate estimated effect, and the blue bars denote the 95% confidence intervals. Figures (a), (d) and (g) illustrate the case of Chile, figures (b), (e) and (h) the case of Croatia, while figures (c), (f) and (i) the case of Sweden.

Figure B13: Probabilities of Applying and Enrolling in Older Sibling's Target Major-College (Polynomial of degree 2)



(a) 1st preference - Chile

(b) 1st preference - Croatia

(c) 1st preference - Sweden

(d) Any preference - Chile

(e) Any preference - Croatia

(f) Any preference - Sweden

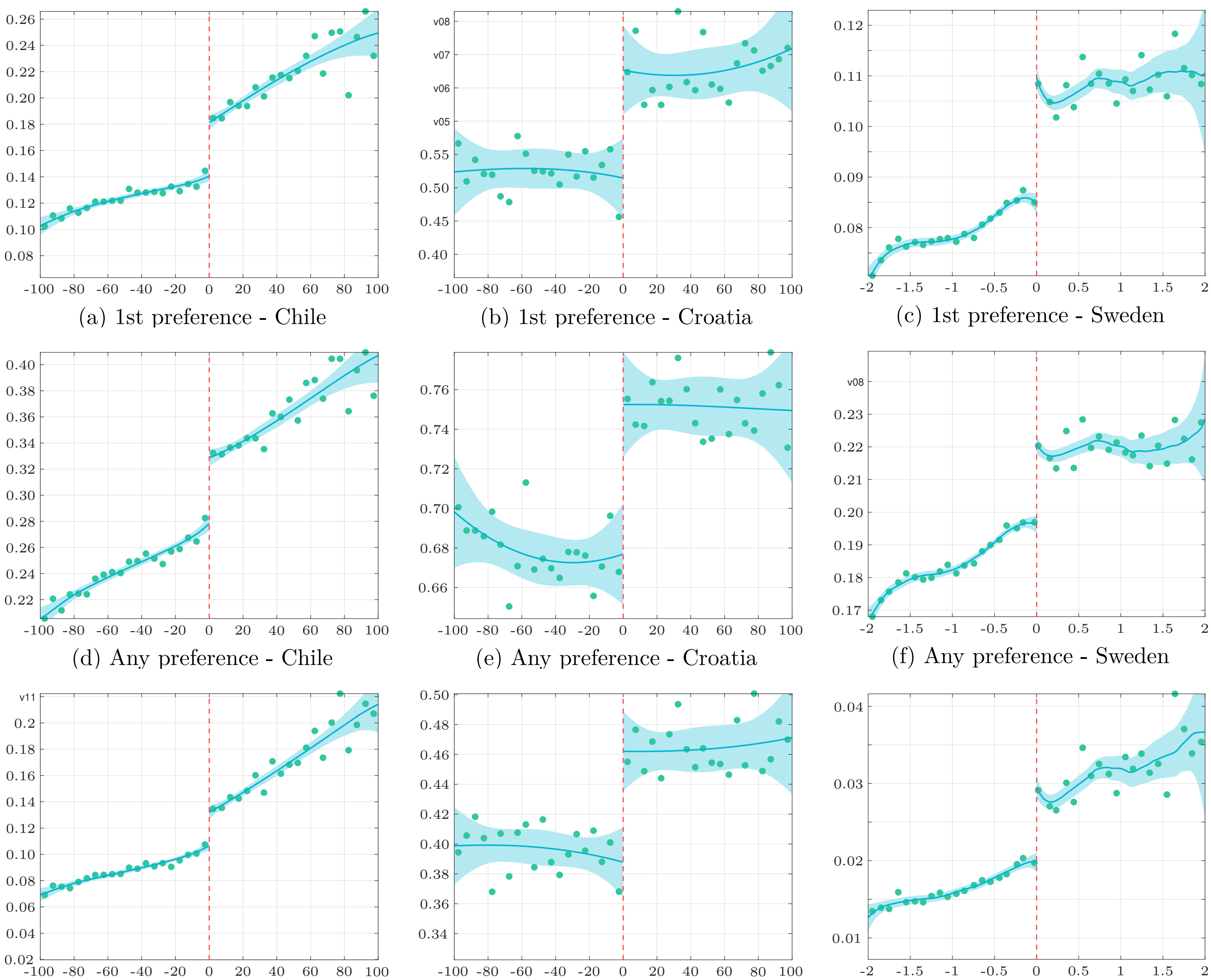
(g) Enrolls - Chile

(h) Enrolls - Croatia

(i) Enrolls - Sweden

This figure illustrates the probabilities that younger siblings apply to and enroll in the target major-college combination of their older siblings in Chile, Croatia and Sweden.Figures (a), (d) and (g) illustrate the case of Chile, figures (b), (e) and (h) the case of Croatia, while figures (c), (f) and (i) the case of Sweden. Blue lines and the shadows in the back of them correspond to local polynomials of degree 1 and 95% confidence intervals. In all cases triangular kernels are used. The bandwidths used to build these figures correspond to optimal bandwidths computed following Calonico et al. (2014) for estimating the discontinuities at the cutoff. Green dots represent sample means of the dependent variable at different values of the older sibling's admission score.

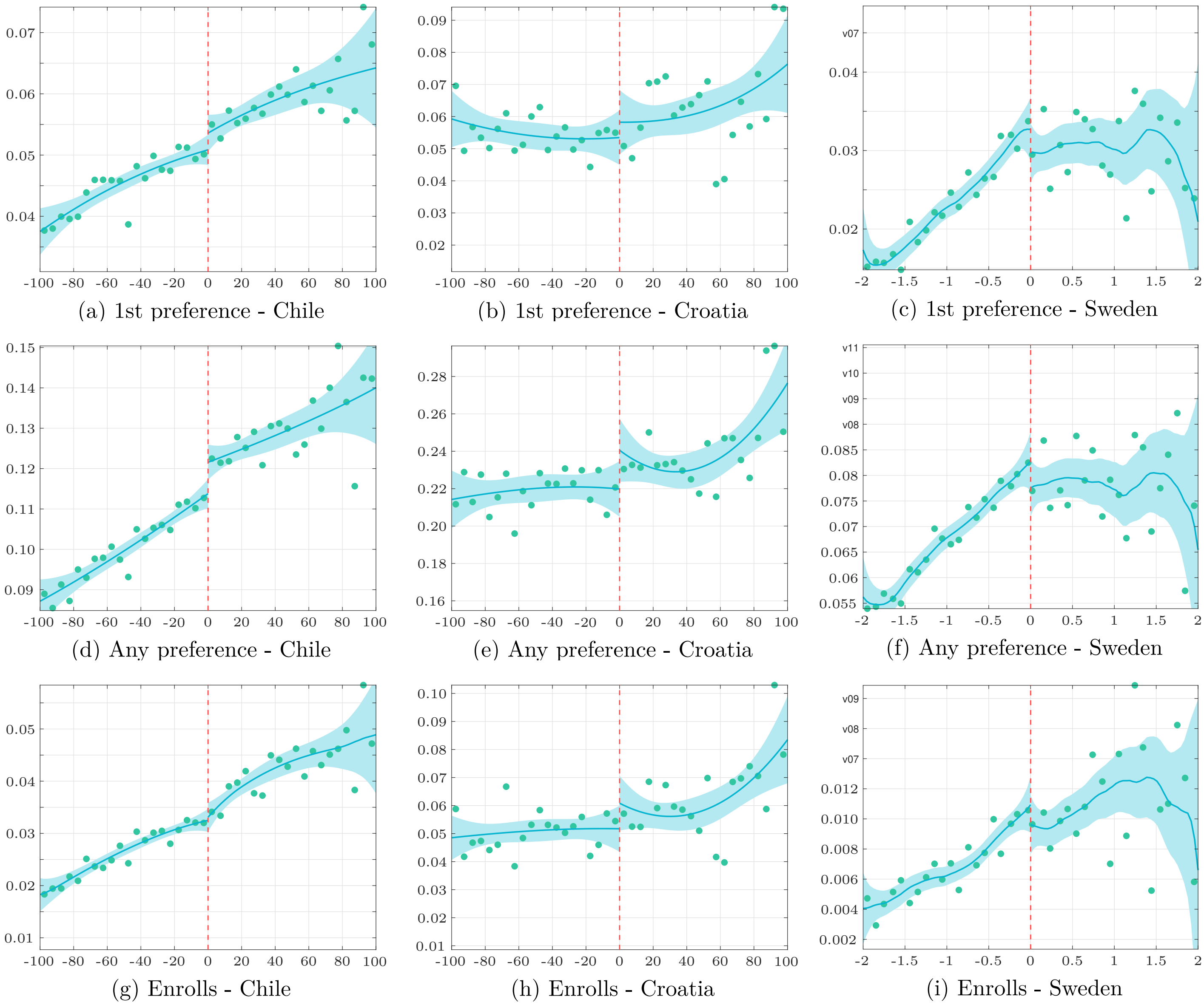
Figure B14: Probabilities of Applying and Enrolling in Older Sibling's Target College (Polynomial of degree 2)



(a) 1st preference - Chile
(b) 1st preference - Croatia
(c) 1st preference - Sweden
(d) Any preference - Chile
(e) Any preference - Croatia
(f) Any preference - Sweden
(g) Enrolls - Chile
(h) Enrolls - Croatia
(i) Enrolls - Sweden

This figure illustrates the probabilities that younger siblings apply to and enroll in the target college of their older siblings in Chile, Croatia and Sweden. Figures (a), (d) and (g) illustrate the case of Chile, figures (b), (e) and (h) the case of Croatia, while figures (c), (f) and (i) the case of Sweden. Blue lines and the shadows in the back of them correspond to local polynomials of degree 2 and 95% confidence intervals. In all cases triangular kernels are used. The bandwidths used to build these figures correspond to optimal bandwidths computed following Calonico et al. (2014) for estimating the discontinuities at the cutoff. Green dots represent sample means of the dependent variable at different values of the older sibling's admission score.

Figure B15: Probabilities of Applying and Enrolling in Older Sibling's Target Field of Study (Polynomial of degree 2)



This figure illustrates the probabilities that younger siblings apply to and enroll in a program in the same field of study as the target program of their older siblings in Chile, Croatia and Sweden. Figures (a), (d) and (e) illustrate the case of Chile, figures (b), (e) and (h) the case of Croatia, while figures (c), (f) and (i) the case of Sweden. Blue lines and the shadows in the back of them correspond to local polynomials of degree 2 and 95% confidence intervals. In all cases, triangular kernels are used. The bandwidths used to build these figures correspond to optimal bandwidths computed following Calonico et al. (2014) for estimating the discontinuities at the cutoff. Green dots represent sample means of the dependent variable at different values of the older sibling's admission score.

Table B1: Probability of Applying and Enrolling in Older Sibling's Target Major - Reweighting

	Applies 1st		**Applies**		**Enrolls**	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.003	0.003	0.024***	0.016	0.001	0.002
	(0.003)	(0.004)	(0.007)	(0.008)	(0.003)	(0.004)
Reduced form	0.001	0.001	0.011***	0.007	0.000	0.001
	(0.002)	(0.002)	(0.003)	(0.004)	(0.001)	(0.002)
Observations	136364	214840	136364	214840	136364	214840
Outcome mean	0.014	0.014	0.050	0.049	0.011	0.011
Bandwidth	20.000	35.000	20.000	35.000	20.000	35.000
F-statistics	5791.853	3479.052	5791.853	3479.052	5791.853	3479.052
			Panel B - Croatia			
2SLS	0.019***	0.020***	0.026**	0.021	0.012**	0.013*
	(0.005)	(0.006)	(0.009)	(0.011)	(0.005)	(0.006)
Reduced form	0.015***	0.016***	0.021**	0.017	0.010**	0.011*
	(0.004)	(0.005)	(0.007)	(0.009)	(0.004)	(0.005)
Observations	36757	48611	36757	48611	36757	48611
Outcome mean	0.020	0.020	0.093	0.094	0.017	0.018
Bandwidth	80.000	120.000	80.000	120.000	80.000	120.000
F-statistics	8076.129	5369.296	8076.129	5369.296	8076.129	5369.296
			Panel C - Sweden			
2SLS	0.007**	0.010***	0.012*	0.012*	0.000	0.001
	(0.002)	(0.003)	(0.005)	(0.006)	(0.002)	(0.002)
Reduced form	0.002**	0.002***	0.003*	0.003*	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Observations	732025	1033985	732025	1033985	732025	1033985
Outcome mean	0.007	0.007	0.033	0.032	0.003	0.003
Bandwidth	0.510	0.750	0.510	0.750	0.510	0.750
F-statistics	7710.134	5944.291	7710.134	5944.291	7710.134	5944.291

Notes: All the specifications in the table control for a linear or quadratic polynomial of older siblings' application score centered around target majors admission cutoff. Observations are re-weighted by the inverse of the number of observations around the cutoff in each major-year. Older siblings' application year, target cutoff-year and younger siblings' birth year fixed effect are included as controls. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table B2: Probability of Applying and Enrolling in Older Sibling's Target College - Reweighting

	Applies 1st		**Applies**		**Enrolls**	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.061***	0.067***	0.082***	0.067**	0.030*	0.043**
	(0.016)	(0.018)	(0.020)	(0.022)	(0.014)	(0.015)
Reduced form	0.025***	0.027***	0.033***	0.027**	0.012*	0.017**
	(0.007)	(0.007)	(0.008)	(0.009)	(0.006)	(0.006)
Observations	73331	152301	73331	152301	73331	152301
Outcome mean	0.157	0.155	0.292	0.286	0.102	0.099
Bandwidth	15.000	35.000	15.000	35.000	15.000	35.000
F-statistics	2576.800	2319.288	2576.800	2319.288	2576.800	2319.288
			Panel B - Croatia			
2SLS	0.090***	0.085**	0.102***	0.095**	0.087***	0.113***
	(0.024)	(0.030)	(0.024)	(0.030)	(0.024)	(0.030)
Reduced form	0.074***	0.070**	0.084***	0.078**	0.071***	0.093***
	(0.020)	(0.025)	(0.020)	(0.025)	(0.019)	(0.025)
Observations	12950	17312	12950	17312	12950	17312
Outcome mean	0.344	0.347	0.582	0.587	0.307	0.307
Bandwidth	80.000	120.000	80.000	120.000	80.000	120.000
F-statistics	3981.458	2474.691	3981.458	2474.691	3981.458	2474.691
			Panel C - Sweden			
2SLS	0.095***	0.085***	0.097***	0.089***	0.034***	0.032***
	(0.010)	(0.010)	(0.013)	(0.014)	(0.006)	(0.007)
Reduced form	0.022***	0.020***	0.022***	0.021***	0.008***	0.008***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.001)	(0.002)
Observations	444203	856457	444203	856457	444203	856457
Outcome mean	0.081	0.077	0.167	0.158	0.033	0.032
Bandwidth	0.370	0.730	0.370	0.730	0.370	0.730
F-statistics	4819.332	4601.144	4819.332	4601.144	4819.332	4601.144

Notes: All the specifications in the table control for a linear or quadratic polynomial of older siblings' application score centered around target majors admission cutoff. Observations are re-weighted by the inverse of the number of observations around the cutoff in each major-year. Older siblings' application year, target cutoff-year and younger siblings' birth year fixed effect are included as controls. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table B3: Probability of Applying and Enrolling in Older Sibling's Target Field - Reweighting

	Applies 1st		**Applies**		**Enrolls**	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.011	0.008	0.016	0.025	0.006	0.001
	(0.010)	(0.011)	(0.014)	(0.015)	(0.009)	(0.010)
Reduced form	0.004	0.003	0.006	0.010	0.002	0.001
	(0.004)	(0.004)	(0.006)	(0.006)	(0.003)	(0.004)
Observations	74012	153713	74012	153713	74012	153713
Outcome mean	0.051	0.051	0.113	0.114	0.035	0.036
Bandwidth	15.000	35.000	15.000	35.000	15.000	35.000
F-statistics	2655.255	2310.756	2655.255	2310.756	2655.255	2310.756
			Panel B - Croatia			
2SLS	0.023^{**}	0.027^{*}	0.027	0.035	0.007	0.008
	(0.008)	(0.011)	(0.015)	(0.019)	(0.008)	(0.010)
Reduced form	0.018^{**}	0.021^{*}	0.021	0.028	0.006	0.006
	(0.007)	(0.008)	(0.012)	(0.015)	(0.007)	(0.008)
Observations	31698	42421	31698	42421	31698	42421
Outcome mean	0.051	0.052	0.198	0.198	0.048	0.048
Bandwidth	80.000	120.000	80.000	120.000	80.000	120.000
F-statistics	6215.082	4240.732	6215.082	4240.732	6215.082	4240.732
			Panel C - Sweden			
2SLS	-0.014^{*}	-0.015^{*}	-0.020^{*}	-0.018	-0.003	-0.002
	(0.006)	(0.007)	(0.009)	(0.010)	(0.004)	(0.004)
Reduced form	-0.003^{*}	-0.004^{*}	-0.005^{*}	-0.004	-0.001	0.000
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Observations	398220	625535	398220	625535	398220	625535
Outcome mean	0.030	0.028	0.067	0.065	0.011	0.011
Bandwidth	0.390	0.610	0.390	0.610	0.390	0.610
F-statistics	4402.932	3898.206	4402.932	3898.206	4402.932	3898.206

Notes: All the specifications in the table control for a linear or quadratic polynomial of older siblings' application score centered around target majors admission cutoff. Observations are re-weighted by the inverse of the number of observations around the cutoff in each major-year. Older siblings' application year, target cutoff-year and younger siblings' birth year fixed effect are included as controls. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table B4: Probability of Enrolling in any College Depending on Older Siblings' Admission to Target Major-College

	Younger siblings (1)	(2)	**Older siblings** (3)	(4)
	Panel A - Chile			
Older sibling admitted to target major = 1	-0.002	-0.004	0.017***	0.019***
	(0.006)	(0.006)	(0.004)	(0.004)
Observations	101955	206940	69170	139469
Outcome mean	0.529	0.526	0.929	0.916
Bandwidth	15.000	35.000	15.000	35.000
	Panel B - Croatia			
Older sibling admitted to target major = 1	-0.003	0.000	0.123***	0.131***
	(0.007)	(0.008)	(0.007)	(0.008)
Observations	36757	48611	36757	48611
Outcome mean	0.90	0.90	0.88	0.85
Bandwidth	80	120	80	120
	Panel C - Sweden			
Older sibling admitted to target major = 1	0.004	0.003	0.046***	0.039***
	(0.004)	(0.003)	(0.003)	(0.004)
Observations	239690	387184	431007	704370
Outcome mean	0.342	0.338	0.326	0.292
Bandwidth	0.550	1.040	0.550	1.040

Notes: The table presents estimates for the effect of older siblings' marginal admission in their target major on their own and on their younger siblings' probability of enrolling in any institution of the system. The specifications controls for a linear or quadratic local polynomial of older siblings' application score centered around their target major admission cutoff. While older siblings' application year fixed effects are used in all specifications, younger siblings' birth year fixed effects are only used in columns (1) and (2). The slope of the running variable is allowed to change at the cutoff. In addition, target major-year fixed effects are included in all specifications. In the case of Chile, we observe enrollment for all the colleges of the system from 2007 onwards. Thus, the sample is adjusted accordingly. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table B5: Probability of Applying and Enrolling in Older Sibling's Target Major-College - Different Slope for each Admission Cutoff

	Applies 1st		**Applies**		**Enrolls**	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.010**	0.009*	0.029***	0.027***	0.003	0.000
	(0.003)	(0.004)	(0.005)	(0.007)	(0.002)	(0.003)
Reduced form	0.005**	0.004*	0.016***	0.014***	0.001	0.000
	(0.002)	(0.002)	(0.003)	(0.003)	(0.001)	(0.002)
Observations	136364	214840	136364	214840	136364	214840
Outcome mean	0.018	0.018	0.056	0.055	0.012	0.012
Bandwidth	20.000	35.000	20.000	35.000	20.000	35.000
F-statistics	12251.360	7965.265	12251.360	7965.265	12251.360	7965.265
			Panel B - Croatia			
2SLS	0.016**	0.016*	0.044***	0.051***	0.014**	0.017**
	(0.005)	(0.007)	(0.010)	(0.013)	(0.005)	(0.006)
Reduced form	0.013**	0.013*	0.036***	0.042***	0.012**	0.014**
	(0.004)	(0.006)	(0.008)	(0.011)	(0.004)	(0.005)
Observations	36757	48611	36757	48611	36757	48611
Outcome mean	0.029	0.029	0.129	0.130	0.024	0.024
Bandwidth	80.000	120.000	80.000	120.000	80.000	120.000
F-statistics	12626.492	7917.659	12626.492	7917.659	12626.492	7917.659
			Panel C - Sweden			
2SLS	0.024***	0.036***	0.034***	0.047***	0.007***	0.010***
	(0.003)	(0.005)	(0.007)	(0.009)	(0.002)	(0.003)
Reduced form	0.005***	0.007***	0.007***	0.009***	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.000)	(0.001)
Observations	718979	1020696	718979	1020696	718979	1020696
Outcome mean	0.011	0.010	0.048	0.047	0.004	0.003
Bandwidth	0.510	0.750	0.510	0.750	0.510	0.750
F-statistics	6882.985	3855.300	6882.985	3855.300	6882.985	3855.300

Notes: All the specifications in the table control for a linear or quadratic polynomial of older siblings' application score centered around target majors admission cutoff. The slope of the running variable is allowed to change at the cutoff and for each target major-year. Older siblings' application year, target cutoff-year and younger siblings' birth year fixed effect are included as controls. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table B6: Probability of Applying and Enrolling in Older Sibling's Target College - Different Slope for each Admission Cutoff

	Applies 1st		Applies		Enrolls	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.076***	0.075***	0.106***	0.092***	0.048***	0.040***
	(0.014)	(0.014)	(0.018)	(0.017)	(0.012)	(0.011)
Reduced form	0.037***	0.037***	0.052***	0.045***	0.023***	0.020***
	(0.007)	(0.007)	(0.009)	(0.009)	(0.006)	(0.006)
Observations	73331	152301	73331	152301	73331	152301
Outcome mean	0.161	0.157	0.302	0.292	0.101	0.097
Bandwidth	15.000	35.000	15.000	35.000	15.000	35.000
F-statistics	4228.409	4390.981	4228.396	4390.993	4228.409	4390.978
			Panel B - Croatia			
2SLS	0.080**	0.081*	0.107***	0.115**	0.085***	0.096**
	(0.024)	(0.037)	(0.025)	(0.038)	(0.023)	(0.036)
Reduced form	0.068***	0.067*	0.090***	0.096**	0.072***	0.080**
	(0.020)	(0.031)	(0.021)	(0.031)	(0.020)	(0.030)
Observations	12950	17312	12950	17312	12950	17312
Outcome mean	0.321	0.322	0.555	0.559	0.287	0.287
Bandwidth	80.000	120.000	80.000	120.000	80.000	120.000
F-statistics	4398.579	1945.206	4398.579	1945.206	4398.579	1945.206
			Panel C - Sweden			
2SLS	0.193***	0.227***	0.186***	0.217***	0.086***	0.102***
	(0.014)	(0.016)	(0.019)	(0.021)	(0.009)	(0.010)
Reduced form	0.036***	0.041***	0.035***	0.039***	0.016***	0.018***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)
Observations	432924	843955	432924	843955	432924	843955
Outcome mean	0.088	0.084	0.193	0.187	0.034	0.032
Bandwidth	0.370	0.730	0.370	0.730	0.370	0.730
F-statistics	2985.240	2446.559	2985.240	2446.559	2985.240	2446.559

Notes: All the specifications in the table control for a linear or quadratic polynomial of older siblings' application score centered around target majors admission cutoff. The slope of the running variable is allowed to change at the cutoff and for each target major-year. Older siblings' application year, target cutoff-year and younger siblings' birth year fixed effect are included as controls. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table B7: Probability of Applying and Enrolling in Older Sibling's Target Field - Different Slope for each Admission Cutoff

	Applies 1st		**Applies**		**Enrolls**	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.011	0.007	0.016	0.014	0.000	-0.007
	(0.009)	(0.009)	(0.013)	(0.013)	(0.007)	(0.007)
Reduced form	0.005	0.005	0.010*	0.009*	0.000	-0.001
	(0.003)	(0.003)	(0.005)	(0.005)	(0.003)	(0.003)
Observations	74012	153713	74012	153713	74012	153713
Outcome mean	0.049	0.049	0.113	0.112	0.032	0.032
Bandwidth	15.000	35.000	15.000	35.000	15.000	35.000
F-statistics	3612.147	3682.283	3612.147	3682.307	3612.147	3682.307
			Panel B - Croatia			
2SLS	0.004	-0.005	0.012	0.011	0.006	0.002
	(0.007)	(0.010)	(0.013)	(0.018)	(0.007)	(0.010)
Reduced form	0.004	-0.004	0.010	0.009	0.005	0.001
	(0.006)	(0.008)	(0.011)	(0.014)	(0.006)	(0.008)
Observations	31698	42421	31698	42421	31698	42421
Outcome mean	0.059	0.059	0.218	0.219	0.054	0.054
Bandwidth	80.000	120.000	80.000	120.000	80.000	120.000
F-statistics	8616.156	5280.547	8616.156	5280.521	8616.156	5280.547
			Panel C - Sweden			
2SLS	-0.000	-0.000	-0.004	-0.011	0.002	-0.000
	(0.011)	(0.015)	(0.016)	(0.020)	(0.006)	(0.008)
Reduced form	-0.000	-0.000	-0.001	-0.002	0.000	-0.000
	(0.002)	(0.003)	(0.003)	(0.004)	(0.001)	(0.001)
Observations	386777	612955	386777	612955	386777	612955
Outcome mean	0.041	0.039	0.087	0.086	0.014	0.014
Bandwidth	0.390	0.610	0.390	0.610	0.390	0.610
F-statistics	2261.735	1424.370	2261.735	1424.370	2261.735	1424.370

Notes: All the specifications in the table control for a linear or quadratic polynomial of older siblings' application score centered around target majors admission cutoff. The slope of the running variable is allowed to change at the cutoff and for each target major-year. Older siblings' application year, target cutoff-year and younger siblings' birth year fixed effect are included as controls. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table B8: Probability of Applying and Enrolling in Older Sibling's Target Major-College - Target $\times$ Counterfactual Major Fixed Effects

	Applies 1st		**Applies**		**Enrolls**	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.012***	0.013***	0.029***	0.026***	0.003	0.001
	(0.004)	(0.005)	(0.007)	(0.008)	(0.003)	(0.004)
Reduced form	0.006***	0.006***	0.015***	0.013***	0.002	0.001
	(0.002)	(0.002)	(0.004)	(0.004)	(0.002)	(0.002)
Observations	92821	154561	92821	154561	92821	154561
Outcome mean	0.019	0.020	0.058	0.057	0.013	0.013
Bandwidth	20.000	35.000	20.000	35.000	20.000	35.000
F-statistics	7232.029	5490.28	7232.029	5490.28	7232.029	5490.28
			Panel B - Croatia			
2SLS	0.012	0.010	0.038***	0.40**	0.011	0.015
	(0.008)	(0.009)	(0.014)	(0.017)	(0.007)	(0.008)
Reduced form	0.010	0.009	0.033***	0.035**	0.010	0.013
	(0.006)	(0.008)	(0.012)	(0.014)	(0.006)	(0.007)
Observations	23076	32230	23076	32230	23076	32230
Outcome mean	0.033	0.032	0.144	0.143	0.027	0.027
Bandwidth	80.000	120.000	80.000	120.000	80.000	120.000
F-statistics	10630.120	7653.077	10630.120	7653.077	10630.120	7653.077
			Panel C - Sweden			
2SLS	0.017***	0.020***	0.026***	0.029***	0.006***	0.008***
	(0.002)	(0.002)	(0.004)	(0.003)	(0.001)	(0.001)
Reduced form	0.004***	0.005***	0.006***	0.007***	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.0003)	(0.0003)
Observations	567548	818146	567548	818146	567548	818146
Outcome mean	0.011	0.010	0.047	0.046	0.004	0.003
Bandwidth	0.510	0.745	0.510	0.745	0.510	0.745
F-statistics	14168.46	18488.9	14168.46	18488.9	14168.46	18488.9

Notes: All the specifications in the table control for a linear or quadratic polynomial of older siblings' application score centered around target majors admission cutoff. The slope of the running variable is allowed to change at the cutoff. Older siblings' application year, target $\times$ counterfactual cutoff-year and younger siblings' birth year fixed effect are included as controls. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table B9: Probability of Applying and Enrolling in Older Sibling's Target College - Target × Counterfactual Major Fixed Effects

	Applies 1st		**Applies**		**Enrolls**	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.067***	0.086***	0.106***	0.0110***	0.043***	0.039***
	(0.017)	(0.016)	(0.018)	(0.019)	(0.014)	(0.013)
Reduced form	0.030***	0.038***	0.047***	0.049***	0.019***	0.017***
	(0.008)	(0.007)	(0.009)	(0.009)	(0.006)	(0.005)
Observations	50076	111993	50076	111993	50076	111993
Outcome mean	0.173	0.167	0.313	0.301	0.108	0.102
Bandwidth	15.000	35.000	15.000	35.000	15.000	35.000
F-statistics	2790.058	3442.876	2790.058	3442.876	2790.058	3442.876
			Panel B - Croatia			
2SLS	0.053	0.042	0.106***	0.092**	0.078**	0.068*
	(0.033)	(0.039)	(0.032)	(0.037)	(0.033)	(0.038)
Reduced form	0.047	0.037	0.094***	0.081**	0.069***	0.060*
	(0.030)	(0.034)	(0.028)	(0.033)	(0.029)	(0.034)
Observations	6743	9596	6743	9596	6743	9596
Outcome mean	0.355	0.352	0.588	0.592	0.319	0.318
Bandwidth	80.000	120.000	80.000	120.000	80.000	120.000
F-statistics	2517.738	3540.023	2517.738	3540.023	2517.738	3540.023
			Panel C - Sweden			
2SLS	0.134***	0.141***	0.133***	0.142***	0.056***	0.061***
	(0.008)	(0.006)	(0.011)	(0.007)	(0.005)	(0.004)
Reduced form	0.029***	0.034***	0.028***	0.034***	0.012***	0.015***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)
Observations	353602	697976	353602	697976	353602	697976
Outcome mean	0.089	0.085	0.193	0.186	0.035	0.033
Bandwidth	0.367	0.733	0.367	0.733	0.367	0.733
F-statistics	7604.52	15313.80	7604.52	15313.80	7604.52	15313.80

Notes: All the specifications in the table control for a linear or quadratic polynomial of older siblings' application score centered around target majors admission cutoff. The slope of the running variable is allowed to change at the cutoff. Older siblings' application year, target × counterfactual cutoff-year and younger siblings' birth year fixed effect are included as controls. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table B10: Probability of Applying and Enrolling in Older Sibling's Target Field - Target × Counterfactual Major Fixed Effects

	Applies 1st		**Applies**		**Enrolls**	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.014	0.015	0.021	0.023	-0.001	-0.008
	(0.012)	(0.011)	(0.017)	(0.015)	(0.009)	(0.008)
Reduced form	0.005	0.005	0.010*	0.009*	0.000	-0.001
	(0.003)	(0.003)	(0.005)	(0.005)	(0.003)	(0.003)
Observations	47027	107632	47027	107632	47027	107632
Outcome mean	0.051	0.051	0.114	0.112	0.033	0.033
Bandwidth	15.000	35.000	15.000	35.000	15.000	35.000
F-statistics	1944.226	2482.383	1944.226	2482.383	1944.226	2482.383
			Panel B - Croatia			
2SLS	-0.010	-0.017	-0.005	-0.001	-0.007	-0.007
	(0.012)	(0.014)	(0.019)	(0.023)	(0.011)	(0.013)
Reduced form	-0.009	-0.014	-0.004	-0.001	-0.006	-0.005
	(0.010)	(0.011)	(0.016)	(0.019)	(0.009)	(0.011)
Observations	18862	26932	18862	26932	18862	26932
Outcome mean	0.064	0.064	0.229	0.229	0.057	0.057
Bandwidth	80.000	120.000	80.000	120.000	80.000	120.000
F-statistics	6159.354	4672.655	6159.354	4672.655	6159.354	4672.655
			Panel C - Sweden			
2SLS	-0.0002	0.004	0.003	0.002	0.002	0.001
	(0.006)	(0.004)	(0.008)	(0.006)	(0.003)	(0.003)
Reduced form	-0.000	-0.000	-0.001	-0.002	0.000	-0.000
	(0.002)	(0.003)	(0.003)	(0.004)	(0.001)	(0.001)
Observations	310122	495991	310122	495991	310122	495991
Outcome mean	0.040	0.039	0.086	0.084	0.013	0.013
Bandwidth	0.389	0.606	0.389	0.606	0.389	0.606
F-statistics	6632.403	11502.85	6632.403	11502.85	6632.403	11502.85

Notes: All the specifications in the table control for a linear or quadratic polynomial of older siblings' application score centered around target majors admission cutoff. The slope of the running variable is allowed to change at the cutoff. Older siblings' application year, target-counterfactual cutoff and younger siblings' birth year fixed effect are included as controls. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

C Additional Results

The heterogeneity analyses presented in the main body of the paper focus on applications to major and college. This appendix presents similar results looking at heterogeneous effects in major and college enrollment, as well as in applications to and enrollment in fields of study. The results that we find in terms of major and college enrollment follow a similar pattern to the ones we find when focusing on applications. Something similar happens with the results we obtain when looking instead at the choice of field of study. However, since average effects on the choice of field of study (i.e. applications and enrollment) are smaller, few of the interactions we document are significant. As in the case of the major and college choices, when looking at the field of study our results suggest that males are more likely to follow older brothers than sisters, and that for females the gender of the older sibling seems less relevant. Effects also seem stronger for siblings who are closer in age and in academic potential. We find no significant differences on applications or enrollment in older siblings' field of study depending on the quality of older siblings' target major.

Finally, we investigate changes in younger siblings' academic performance by the age difference they have with their older siblings in the three samples that we use in this project (i.e. major, college and field). These results are consistent with the ones presented in the main body of the paper and provide additional evidence that the effects we find in major and college enrollment are not driven by an improvement of individuals' academic performance.

Table C1: Probability of Enrolling in Older Sibling's Target Major and Target College by Older Siblings' Gender

	Major			College		
	Older Siblings' Gender			Older Siblings' Gender		
	All	Female	Male	All	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A - Chile						
Older sibling enrolls	0.001	0.001	0.001	0.037***	0.027	0.042**
	(0.002)	(0.003)	(0.004)	(0.010)	(0.015)	(0.015)
Older sibling enrolls × Same gender	0.005**	0.000	0.011***	0.013	0.015	0.020
	(0.002)	(0.002)	(0.003)	(0.008)	(0.011)	(0.012)
Observations	136364	73014	61982	73331	39129	32302
Outcome mean	0.012	0.010	0.014	0.101	0.102	0.099
Bandwidth	20.000	20.000	20.000	15.000	15.000	15.000
F-statistics	6933.231	3310.962	3530.694	2719.593	1278.857	1337.943
Panel B - Croatia						
Older sibling enrolls	0.007	0.006	0.008	0.065**	0.044	0.066
	(0.004)	(0.006)	(0.007)	(0.021)	(0.029)	(0.034)
Older sibling enrolls × Same gender	0.013**	0.004	0.031***	0.037	0.046	0.014
	0.004)	(0.005)	(0.008)	(0.019)	(0.026)	(0.031)
Observations	36757	22239	14203	12950	7545	5008
Outcome mean	0.024	0.022	0.029	0.287	0.284	0.290
Bandwidth	80.000	80.000	80.000	80.000	80.000	80.000
F-statistics	7220.184	3662.675	4025.070	3229.534	1651.529	1405.970
Panel C - Sweden						
Older sibling enrolls	0.002	0.001	0.002	0.056***	0.061***	0.059***
	(0.001)	(0.002)	(0.003)	(0.006)	(0.009)	(0.011)
Older sibling enrolls × Same gender	0.006***	0.003*	0.009***	0.014**	0.013	0.015
	(0.001)	(0.001)	(0.002)	(0.005)	(0.007)	(0.009)
Observations	732025	438419	281549	444203	273981	160086
Outcome mean	0.004	0.003	0.005	0.034	0.032	0.038
Bandwidth	0.510	0.510	0.510	0.370	0.370	0.370
F-statistics	5419.139	2441.736	2717.178	3075.133	1484.510	1330.244

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major and college by siblings' gender. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Tables 3 and 5. Specifications also control by a dummy variable that indicates if the siblings are of the same gender. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C2: Probability of Applying and Enrolling in Older Sibling's Target Field of Study by Older Siblings' Gender

	Older Siblings' Gender					
	All	Female	Male	All	Female	Male
		Applies			Enrolls	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
Older sibling enrolls	0.014	0.020	0.010	-0.002	-0.002	0.001
	(0.011)	(0.015)	(0.017)	(0.006)	(0.008)	(0.010)
Older sibling enrolls × Same gender	0.019*	0.002	0.033*	0.006	0.003	0.009
	(0.008)	(0.011)	(0.013)	(0.005)	(0.006)	(0.008)
Observations	74012	40123	31964	74012	40123	31964
Outcome mean	0.113	0.103	0.124	0.032	0.026	0.039
Bandwidth	15.000	15.000	15.000	15.000	15.000	15.000
F-statistics	2416.376	1201.441	1111.501	2416.376	1201.441	1111.501
			Panel B - Croatia			
Older sibling enrolls	0.012	0.020	0.004	0.003	0.007	0.002
	(0.015)	(0.017)	(0.020)	(0.008)	(0.010)	(0.012)
Older sibling enrolls × Same gender	0.009	-0.019	0.040	-0.001	-0.011	0.018
	(0.015)	(0.017)	(0.022)	(0.008)	(0.009)	(0.012)
Observations	31698	19269	12085	31698	19269	12085
Outcome mean	0.218	0.206	0.238	0.054	0.049	0.062
Bandwidth	80.000	80.000	80.000	80.000	80.000	80.000
F-statistics	5027.422	2501.951	2815.384	5027.422	2501.951	2815.384
			Panel C - Sweden			
Older sibling enrolls	0.001	0.033*	-0.032	-0.002	0.004	-0.007
	(0.011)	(0.016)	(0.018)	(0.004)	(0.006)	(0.008)
Older sibling enrolls × Same gender	-0.010	-0.056***	0.052***	0.003	-0.007	0.016**
	(0.009)	(0.012)	(0.014)	(0.004)	(0.005)	(0.006)
Observations	398220	240016	148034	398220	240016	148034
Outcome mean	0.087	0.077	0.104	0.014	0.012	0.017
Bandwidth	0.390	0.390	0.390	0.390	0.390	0.390
F-statistics	2558.556	1064.952	1253.694	2558.556	1064.952	1253.694

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target field of study by siblings' gender. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 5. Specifications also control by a dummy variable that indicates if the siblings are of the same gender. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C3: Probability of Enrolling in Older Sibling's Target Major and TargetCollege by Siblings' Similarity

	Major		**College**	
	Δ Age > 5 (1)	Δ GPA (2)	Δ Age > 5 (3)	Δ GPA (4)
		Panel A - Chile		
Older sibling enrolls	0.002	0.012***	0.047***	0.091***
	(0.002)	(0.003)	(0.010)	(0.012)
Interaction	0.003	-0.010***	-0.007	-0.052***
	(0.002)	(0.001)	(0.008)	(0.005)
Observations	135777	133703	73030	71865
Outcome mean	0.012	0.012	0.101	0.103
Bandwidth	20.000	20.000	15.000	15.000
F-statistics	6904.432	6789.416	2710.198	2664.690
		Panel B - Croatia		
Older sibling enrolls	0.013**	0.053***	0.089***	0.189***
	(0.004)	(0.012)	(0.019)	(0.055)
Interaction	0.001	-0.028***	-0.029	-0.040
	(0.006)	(0.007)	(0.026)	(0.032)
Observations	36756	8567	12950	2588
Outcome mean	0.024	0.030	0.287	0.338
Bandwidth	80.000	80.000	80.000	80.000
F-statistics	7225.706	1567.759	3230.667	648.627
		Panel C - Sweden		
Older sibling enrolls	0.035***	0.032***	0.067***	0.087***
	(0.005)	(0.007)	(0.006)	(0.008)
Interaction	-0.015***	0.005	-0.010	-0.017***
	(0.004)	(0.003)	(0.005)	(0.003)
Observations	732025	591599	444203	359012
Outcome mean	0.047	0.055	0.034	0.039
Bandwidth	0.510	0.510	0.370	0.370
F-statistics	5255.957	4573.374	2975.652	2610.561

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major and college by siblings' similarity. Columns (1) and (3) investigate heterogeneous effects by age difference, while columns (2) and (4) by difference in high school GPA. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Tables 3 and 5. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C4: Probability of Applying and Enrolling in Older Sibling's Target Field of Study by Siblings' Similarity

	Applies		**Enrolls**	
	Δ Age > 5 (1)	Δ GPA (2)	Δ Age > 5 (3)	Δ GPA (4)
		Panel A - Chile		
Older sibling enrolls	0.024*	0.047***	0.002	0.008
	(0.011)	(0.013)	(0.006)	(0.007)
Interaction	-0.006	-0.025***	-0.002	-0.007*
	(0.008)	(0.005)	(0.005)	(0.003)
Observations	73665	72463	73665	72463
Outcome mean	0.113	0.115	0.032	0.033
Bandwidth	15.000	15.000	15.000	15.000
F-statistics	2411.227	2363.090	2411.227	2363.090
		Panel B - Croatia		
Older sibling enrolls	0.021	-0.019	0.002	0.017
	(0.014)	(0.044)	(0.008)	(0.021)
Interaction	-0.034	-0.014	-0.001	-0.024
	(0.020)	(0.026)	(0.011)	(0.013)
Observations	31697	7167	31697	7167
Outcome mean	0.218	0.251	0.054	0.061
Bandwidth	80.000	80.000	80.000	80.000
F-statistics	5058.433	1063.448	5058.433	1063.448
		Panel C - Sweden		
Older sibling enrolls	0.002	-0.023	0.001	-0.001
	(0.011)	(0.014)	(0.004)	(0.006)
Interaction	-0.012	0.033***	-0.004	0.000
	(0.009)	(0.006)	(0.004)	(0.003)
Observations	398220	320212	398220	320212
Outcome mean	0.087	0.101	0.014	0.016
Bandwidth	0.390	0.390	0.390	0.390
F-statistics	2482.598	2129.958	2482.598	2129.958

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target field of study by siblings' similarity. Columns (1) and (3) investigate heterogeneous effects by age difference, while columns (2) and (4) by difference in high school GPA. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 5. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C5: Probability of Enrolling in Older Sibling's Target Major and College by Quality

	Major			College		
	Admitted students quality (1)	Dropout (2)	Earnings (3)	Admitted students quality (4)	Dropout (5)	Earnings (6)
Panel A - Chile						
Older sibling enrolls	-0.006	0.004	0.003	-0.017	0.057***	0.040***
	(0.004)	(0.003)	(0.002)	(0.019)	(0.010)	(0.010)
Interaction	0.003**	-0.006	0.002*	0.020***	-0.112*	0.011**
	(0.001)	(0.014)	(0.001)	(0.004)	(0.046)	(0.004)
Observations	136364	121676	129847	73331	72642	69927
Outcome mean	0.012	0.012	0.012	0.101	0.101	0.102
Bandwidth	20.000	20.000	20.000	15.000	15.000	15.000
F-statistic	4914.155	5831.462	5732.572	1872.447	2459.612	2183.694
Panel B - Croatia						
Older sibling enrolls	0.021			-0.024		
	(0.058)			(0.012)		
Interaction	-0.002			0.029*		
	(0.003)			(0.012)		
Observations	34510			10693		
Outcome mean	0.024			0.268		
Bandwidth	80.000			80.000		
F-statistic	6833.719			2598.965		
Panel C - Sweden						
Older sibling enrolls	0.000	0.005**	0.002	0.043***	0.059***	0.053***
	(0.002)	(0.002)	(0.002)	(0.007)	(0.007)	(0.008)
Interaction	0.005***	-0.006	0.003**	0.026***	-0.079***	0.008*
	(0.001)	(0.005)	(0.001)	(0.004)	(0.023)	(0.004)
Observations	732023	535714	358644	444203	320107	218552
Outcome mean	0.004	0.004	0.004	0.034	0.036	0.038
Bandwidth	0.510	0.510	0.510	0.370	0.367	0.367
F-statistic	4508.761	5465.479	2462.490	2577.150	2678.503	1380.629

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major and college by different quality measures of their target majors. Columns (1) and (4) investigate heterogeneous effects by the average quality of admitted students, columns (2) and (5) by first year dropout rates and columns (3) and (6) by graduates average earnings. Students' quality is measured by the average scores of admitted students in the admission exam. The measure of students quality and graduates average earnings are standardized. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Tables 3 and 5. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C6: Probability of Applying and Enrolling in Older Sibling's Target Field of Study by Quality

	Applies			Enrolls		
	Admitted students quality (1)	Dropout (2)	Earnings (3)	Admitted students quality (4)	Dropout (5)	Earnings (6)
Panel A - Chile						
Older sibling enrolls	0.031	0.015	0.024*	0.005	0.000	0.002
	(0.020)	(0.012)	(0.011)	(0.011)	(0.007)	(0.006)
Interaction	-0.003	0.061	-0.003	-0.002	0.012	-0.004
	(0.005)	(0.048)	(0.005)	(0.003)	(0.026)	(0.003)
Observations	74012	72888	69487	74012	72888	69487
Outcome mean	0.113	0.113	0.115	0.032	0.032	0.033
Bandwidth	15.000	15.000	15.000	15.000	15.000	15.000
F-statistic	1824.898	2308.953	1953.139	1824.898	2308.953	1953.139
Panel B - Croatia						
Older sibling enrolls	-0.007			0.001		
	(0.035)			(0.020)		
Interaction	0.003			0.000		
	(0.007)			(0.004)		
Observations	29466			29466		
Outcome mean	0.218			0.053		
Bandwidth	80.000			80.000		
F-statistic	4664.494			4664.494		
Panel C - Sweden						
Older sibling enrolls	-0.008	0.011	-0.001	-0.002	0.001	-0.002
	(0.012)	(0.011)	(0.013)	(0.005)	(0.005)	(0.006)
Interaction	0.006	-0.077**	-0.001	0.001	-0.018	0.002
	(0.006)	(0.029)	(0.006)	(0.003)	(0.013)	(0.003)
Observations	398220	283534	190647	398220	283534	190647
Outcome mean	0.087	0.083	0.085	0.014	0.015	0.016
Bandwidth	0.389	0.389	0.389	0.389	0.389	0.389
F-statistic	2206.902	2408.936	1064.776	2206.902	2408.936	1064.776

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target field by different quality measures of their target programs. Columns (1) and (4) investigate heterogeneous effects by the average quality of admitted students, columns (2) and (5) by first year dropout rates and columns (3) and (6) by graduates average earnings. Students' quality is measured by the average scores of admitted students in the admission exam. The measure of students quality and graduates average earnings are standardized. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 7. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C7: Probability of Enrolling in Older Sibling's Target Major and College by Quality Difference respect to Counterfactual Alternative

	Major			College		
	Δ Admitted students quality (1)	Δ Dropout (2)	Δ Earnings (3)	Δ Admitted students quality (4)	Δ Dropout (5)	Δ Earnings (6)
Panel A - Chile						
Older sibling enrolls	0.005	0.006*	0.005	0.044***	0.042***	0.042***
	(0.003)	(0.002)	(0.002)	(0.011)	(0.011)	(0.011)
Interaction	-0.001	0.017	0.000	-0.002	-0.120	-0.016
	(0.002)	(0.016))	(0.001)	(0.010)	(0.066)	(0.013)
Observations	99652	90784	90082	45082	41229	40836
Outcome mean	.013	0.013	0.013	0.105	0.106	0.106
Bandwidth	20.000	20.000	20.000	15.000	15.000	15.000
F-statistics	7674.012	7397.956	7219.418	3153.688	2959.387	2908.442
Panel B - Croatia						
Older sibling enrolls	0.013**			0.101***		
	(0.004)			(0.020)		
Interaction	0.002			0.007		
	(0.002)			(0.010)		
Observations	34510			10693		
Outcome mean	0.024			0.268		
Bandwidth	80.000			80.000		
F-statistics	6854.732			2607.328		
Panel C - Sweden						
Older sibling enrolls	0.006***	0.004**		0.071***	0.049***	
	(0.002)	(0.002)		(0.007)	(0.007)	
Interaction	-0.002	0.000		-0.016***	-0.005	
	(0.001)	(0.001)		(0.005)	(0.004)	
Observations	472966	309934		262275	172027	
Outcome mean 0.004	0.005		0.032	0.036		
Bandwidth	0.510	0.510		0.367	0.367	
F-statistics	4439.812	4419.105		2282.347	2063.087	

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major and college by the gap between older siblings' target and counterfactual major in different quality measures. Columns (1) and (4) investigate heterogeneous effects by the difference in the average quality of admitted students, columns (2) and (5) by the difference in first year dropout rates and columns (3) and (6) by the difference in graduates average earnings. Students quality is measured by the average scores of admitted students in the admission exam. The measure of students quality and graduates average earnings are standardized. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Tables 3 and 5. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. In this table, the sample is restricted to older siblings with counterfactual programs in their application lists. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C8: Probability of Applying and Enrolling in Older Sibling's Target Field of Study by Difference in Quality respect Counterfactual Alternative

	Applies			**Enrolls**		
	Δ Admitted students quality (1)	Δ Dropout (2)	Δ Earnings (3)	Δ Admitted students quality (4)	Δ Dropout (5)	Δ Earnings (6)
			Panel A - Chile			
Older sibling enrolls	0.012	0.013	0.012	-0.002	-0.006	-0.006
	(0.013)	(0.012)	(0.012)	(0.007)	(0.007)	(0.007)
Interaction	0.006	0.022	0.001	0.000	0.059	-0.001
	(0.012)	(0.077)	(0.005)	(0.006)	(0.040)	(0.003)
Observations	45591	40142	39660	45591	40142	39660
Outcome mean	0.122	0.124	0.125	0.034	0.035	0.035
Bandwidth	15.000	15.000	15.000	15.000	15.000	15.000
F-statistics	2608.326	2397.713	2325.023	2608.326	2397.713	2325.023
			Panel B - Croatia			
Older sibling enrolls	0.005			0.000		
	(0.012)			(0.007)		
Interaction	0.010			0.005		
	(0.006)			(0.004)		
Observations	29466			29466		
Outcome mean	0.218			0.053		
Bandwidth	80.000			80.000		
F-statistics	4707.803			4707.803		
			Panel C - Sweden			
Older sibling enrolls	0.012	-0.006		0.006	0.003	
	(0.014)	(0.013)		(0.006)	(0.006)	
Interaction	-0.023***	-0.005		-0.004	0.001	
	(0.007)	(0.004)		(0.003)	(0.002)	
Observations	207042	126204		207042	126204	
Outcome mean	0.094	0.090		0.015	0.016	
Bandwidth	0.390	0.390		0.390	0.390	
F-statistics	1746.185	1454.422		1746.185	1454.422	

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target field of study by the gap between older siblings' target and counterfactual program in different quality measures. Columns (1) and (4) investigate heterogeneous effects by the difference in average quality of admitted students, columns (2) and (5) by the difference in first year dropout rates and columns (3) and (6) by the difference in graduates average earnings. Students' quality is measured by the average scores of admitted students in the admission exam. The measure of students quality and graduates average earnings are standardized. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 5. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. In this table, the sample is restricted to older siblings with counterfactual programs in their application lists. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C9: Effect of Older Siblings' Enrollment in Target Program on Academic Performance (College Sample)

	Takes admission exam (AE) (1)	Applies to college/higher ed. (2)	High School GPA (3)	Average Score AE (4)
		Panel A - Chile		
Older sibling enrolls	0.000	0.028	0.026	0.021
	(0.006)	(0.016)	(0.039)	(0.038)
Observations	73,741	73,741	73,741	73,741
Outcome mean	0.957	0.580	-0.103	0.272
Bandwidth	15.000	15.000	15.000	15.000
F-statistic	5446.004	5446.004	5446.004	5446.004
		Panel B - Croatia		
Older sibling enrolls	-0.023		-0.329	-0.027*
	(0.031)		(0.228)	(0.150)
Observations	4,170		4,170	4,170
Outcome mean	0.824		-1.313	-0.909
Bandwidth	80.000		80.000	80.000
F-statistic	2008.201		2008.201	2008.201
		Panel C - Sweden		
Older sibling enrolls	-0.064***	-0.043**	0.009	0.113*
	(0.016)	(0.015)	(0.034)	(0.049)
Observations	444,203	444,203	372,578	206,613
Outcome mean	0.484	0.584	0.232	0.055
Bandwidth	0.367	0.367	0.367	0.367
F-statistic	6151.602	6151.602	5451.560	3681.775

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major on younger siblings' probability of taking the admission exam and applying to college (columns 1 and 2), and on different measures of academic performance: high school GPA (column 3), reading and math sections of the admission exam (columns 4 and 5) and average performance on the admission exam (column 6). While in Chile and Croatia we only observe applications to college degrees, in Sweden we also observe applications to other higher education programs. These analyses focus on the College Sample. This means that in this case, marginal admission or rejection from their target major, changes the college in which older siblings are admitted. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 5. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C10: Effect of Older Siblings' Enrollment in Target Program on Academic Performance (Field of Study Sample)

	Takes admission exam (AE)	Applies to university/higher ed.	High School GPA	Average Score AE
	(1)	(2)	(3)	(4)
		Panel A - Chile		
Older sibling enrolls	0.003	0.004	-0.027	0.024
	(0.007)	(0.017)	(0.041)	(0.040)
Observations	74,012	74,012	74,012	74,012
Outcome mean	0.955	0.567	-0.149	0.200
Bandwidth	15.000	15.000	15.000	15.000
F-statistic	4833.498	4833.498	4833.498	4833.498
		Panel B - Croatia		
Older sibling enrolls	-0.004		-0.051	-0.043
	(0.020)		(0.146)	(-0.099)
Observations	10,719		10,719	10,719
Outcome mean	0.822		-1.328	-0.851
Bandwidth	80.000		80.000	80.000
F-statistic	3147.714		3147.714	3147.714
		Panel C - Sweden		
Older sibling enrolls	-0.074***	-0.055***	-0.014	0.052
	(0.018)	(0.017)	(0.038)	(0.053)
Observations	398,220	398,220	331,901	182,819
Outcome mean	0.481	0.577	0.226	0.058
Bandwidth	0.389	0.389	0.389	0.389
F-statistic	5116.605	5116.605	4430.987	3023.592

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target field on younger siblings' probability of taking the admission exam and applying to university (columns 1 and 2), and on different measures of academic performance: high school GPA (column 3), reading and math sections of the admission exam (columns 4 and 5) and average performance on the admission exam (column 6). While in Chile and Croatia we only observe applications to university degrees, in Sweden we also observe applications to other higher education programs. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 7. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C11: Effect of Older Siblings' Enrollment in Target Major-College on Academic Performance by Age Difference

	Major Sample		College Sample		Field Sample	
	High School GPA	Average Score AE	High School GPA	Average Score AE	High School GPA	Average Score AE
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
Older sibling enrolls	0.011	0.034	-0.017	0.039	-0.088*	0.026
	(0.029)	(0.028)	(0.042)	(0.041)	(0.052)	(0.051)
Δ Age ≤ 2	-0.014	-0.004	0.048**	-0.010	0.051*	-0.013
	(0.025)	(0.024)	(0.024)	(0.022)	(0.028)	(0.027)
$2 < \Delta$ Age ≤ 2	0.022	0.006	0.072**	-0.049	0.089***	-0.005
	(0.024)	(0.006)	(0.028)	(0.028)	(0.032)	(0.032)
Observations	136364	136364	73,741	73,741	62,011	62,011
Outcome mean	-0.105	0.256	-0.103	0.272	-0.165	0.195
Bandwidth	20.000	20.000	15.000	15.000	15.000	15.000
F-statistics	4614.009	4614.009	1812.148	1812.148	1184.061	1184.061
			Panel B - Croatia			
Older sibling enrolls	-0.146	-0.133	-0.327	-0.302*	-0.145	-0.114
	(0.139)	(0.093)	(0.239)	(0.157)	(0.157)	(0.106)
Δ Age ≤ 2	0.066	0.093	0.007	0.097	0.285*	0.207**
	(0.170)	(0.111)	(0.202)	(0.134)	(0.152)	(0.102)
$2 < \Delta$ Age ≤ 2	0.211	0.125	-0.235	0.280	0.032	0.233
	(0.568)	(0.392)	(0.590)	(0.402)	(0.422)	(0.295)
Observations	12,433	12,443	4,170	4,170	10,719	10,719
Outcome mean	-1.300	-0.834	-1.313	-0.909	-1.328	-0.851
Bandwidth	80.000	80.000	80.000	80.000	80.000	80.000
F-statistics	1461.978	1461.978	659.829	659.829	1022.964	1022.964
			Panel C - Sweden			
Older sibling enrolls	0.288	0.015	0.015	0.080	-0.015	0.027
	(0.027)	(0.038)	(0.038)	(0.055)	(0.041)	(0.058)
Δ Age ≤ 2	0.010	0.070**	0.007	0.106	0.059	0.068
	(0.024)	(0.035)	(0.038)	(0.055)	(0.038)	(0.055)
$2 < \Delta$ Age ≤ 2	-0.057**	-0.017	-0.008	-0.006	-0.030	0.006
	(0.024)	(0.036)	(0.037)	(0.055)	(0.038)	(0.056)
Observations	613,294	344,442	372,578	206,613	331,901	182,819
Outcome mean	0.219	0.051	0.232	0.055	0.226	0.058
Bandwidth	0.51	0.51	0.367	0.367	0.389	0.389
F-statistics	3070.585	2086.53	1747.338	1177.487	1441.458	969.494

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major on high school GPA (column 1) and on average performance on the admission exam (column 2). The effect is allowed to vary with age difference between siblings. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 3. Age difference between siblings is added as control. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.