O Brother, Where Start Thou?
Sibling Spillovers on College and Major Choice in Four Countries[†]

Adam Altmejd Andrés Barrios-Fernández Marin Drlje Joshua Goodman

Michael Hurwitz Dejan Kovac Christine Mulhern Christopher Neilson Jonathan Smith

Latest Version : May 10, 2020

Abstract

Family and social networks are widely believed to influence important life decisions but identifying their causal effects is notoriously difficult. Using admissions thresholds that directly affect older but not younger siblings' college options, we present evidence from the United States, Chile, Sweden and Croatia that older siblings' college and major choices can significantly influence their younger siblings' college and major choices. On the extensive margin, an older sibling's enrollment in a better college increases a younger sibling's probability of enrolling in college at all, especially for families with low predicted probabilities of enrollment. On the intensive margin, an older sibling's choice of college or major increases the probability that a younger sibling applies to and enrolls in that same college or major. Spillovers in major choice are stronger when older siblings enroll and succeed in more selective and higher-earning majors. The observed spillovers are not well-explained by price, income, proximity or legacy effects, but are most consistent with older siblings transmitting otherwise unavailable information about the college experience and its potential returns. The importance of such personally salient information may partly explain persistent differences in college-going rates by geography, income, and other determinants of social networks.

Keywords: Sibling Effects, College and Major Choice, Peer and Social Network Effects
JEL codes: I21, I24.

[†]For granting us access to their administrative data, we thank the the College Board in the United States, the Ministries of Education of Chile and Croatia, the DEMRE, ASHE (AZVO), Riksarkivet, UHR and SCB. All errors are due to our siblings.

1 Introduction

Whether and where to enroll in college and, if so, what to major in is one of the most complex and consequential decisions a high school student will make (Altonji et al., 2012). These choices significantly impact future earnings and occupational trajectories(Hastings et al., 2013; Kirkebøen et al., 2016). We know little, however, about how individuals form the preferences and beliefs that underpin this complex decision (Altonji et al., 2016; Wiswall and Zafar, 2014). That college choices often vary substantially by income and geography has been frequently attributed to credit constraints (Belley and Lochner, 2007; Lochner and Monge-Naranjo, 2012), differences in teacher and school quality (Chetty et al., 2014; Deming et al., 2014), and spatial variation in college options (Hillman, 2016).[1]

Economists have paid less attention to social networks that may partly explain such income-based and geographic differences in college choices. If students' college choices are influenced by the choices of their peers, such spillovers would contribute to observed enrollment differences. High income students might be more likely than low income students to choose certain colleges and majors in part because of exposure to other high income peers who have made such choices. Communities with historically low college enrollment rates may struggle to improve such rates because current students have few peers with firsthand experience of higher quality colleges and majors. Such spillovers could partly explain persistent differences in college choices across families and communities of different socioeconomic status.

Suggestive evidence that social factors matter for college choices is scattered throughout recent papers. Hoxby and Turner (2013) note that high-achieving, low income students who fail to apply to any selective college are disproportionately from schools where such students "have only a negligible probability of meeting a schoolmate from an older cohort who herself attended a selective college." Dillon and Smith (2017) show that, controlling for a rich set of covariates, the share of a student's

[1] In the US, students from the highest income one percent of families are 77 times more likely to attend an Ivy League college than those from the bottom quintile (Chetty et al., 2017). Even among similarly low income students, enrollment rates vary substantially by geography. For those in the 25th percentile of the local parental income distribution, college enrollment rates range from less than 32 percent in the lowest-attending decile of commuting zones to over 55 percent in the highest decile commuting zones. See Online Appendix Figure VII, panel B from Chetty et al. (2014).

high school graduates enrolling in four-year college strongly predicts the quality of college chosen by that student, perhaps because they "have many role models to follow through the college choice process." Black et al. (2015) similarly find that Black and Hispanic students in Texas are more likely to apply to colleges that "recently enrolled any students from the same high school and recently graduated students from the same high school of the same race". They hypothesize that students have better information about a given campus when close peers have attended that campus, and that successful degree completion by such peers is a particularly important signal. Among high-achieving students in the UK, Campbell et al. (2019) find that high schools' college enrollment rates statistically account for about half of the socioeconomic gap in quality of college chosen.

Siblings are particularly important peers. When asked who most influences their thinking about postsecondary education, half of US high schoolers choose family members, compared to only four percent who choose friends (Oymak, 2018). Numerous studies suggest that siblings are particularly influential and even more so when parents lack college education. These include smaller scale qualitative studies showing that Black and Hispanic students report older siblings as important influences (Mwangi, 2015), as well as larger scale quantitative work showing that Black students' college enrollment rates are much higher when older siblings have enrolled earlier (Loury, 2004). Data on 1.6 million US sibling pairs show that one-fifth of younger siblings enroll in the same college as their older sibling and that younger siblings are 15–20 percentage points more likely to enroll in four-year or highly competitive colleges if their older siblings do so first (Goodman et al., 2015). Such correlations between sibling choices remain even after inclusion of extensive controls for potential confounds, such as siblings' academic achievement. Kaczynski (2011), Shahbazian (2018) and Hastings et al. (2016) present similar evidence from Chile and Sweden of strong correlations in siblings' college choices.

Such descriptive evidence does not, however, prove that older siblings' college choices causally influence the choices of their younger siblings. The empirical challenges of identifying such peer effects are well-known (Manski, 1993; Angrist, 2014). If two peers make similar college choices, two major issues arise in estimating peer effects. First, the "reflection problem" involves difficulty in distinguishing whether the first peer affects the second one or vice versa. We solve the reflection problem by estimating spillover effects from older to younger siblings, making the reasonable assumption that only the former's college choice can influence the latter's.[2] Second, the "common shock problem" arises from the fact that peers tend to share characteristics or environments that might be driving both of their choices. This problem is particularly acute when studying siblings, who share parents, homes, neighborhoods and other important potential determinants of college choice.

We solve the common shock problem by finding exogenous variation that affects the college choices available to older siblings but does not directly affect younger siblings through any channel other than their older siblings' choices. In particular, we exploit test-scored based college and major admissions thresholds that directly affect older siblings but, because younger siblings' scores are only weakly correlated with those of older siblings, do not directly affect younger siblings' own admissions outcomes. A regression discontinuity using older siblings' distance to these thresholds as a running variable allows us to estimate the causal impact of older siblings' college and major choices on those of their younger siblings. Identification comes from comparing the college choices of younger siblings whose otherwise identical older siblings were marginally above or below these college- or major-specific cutoffs and thus differed only in their college options. In the US, we use data on the universe of SAT-takers without siblings to identify colleges that use hidden admissions thresholds, the examine sibling pairs where the older sibling has applied to one of these "target" colleges. In Chile, Croatia and Sweden, we exploit centralized admissions systems that allocate students to colleges and majors in a way that creates sharp cutoffs in all oversubscribed programs.

Our first stage analysis shows that meeting these admissions thresholds improves the quality of older siblings' choices of colleges and majors. In the US, admissibility substantially increases older siblings' likelihood of enrolling in the target college and in any four-year college, as well as improving the graduation rate and peer quality of the institution they attend. In Chile, Sweden and Croatia, admissibility substantially increases older siblings' likelihood of enrolling in their target major, the combination of college and degree program to which they apply.

Such increases in older siblings' college enrollment rates and quality spill over onto younger siblings' college choices, both on the extensive margin (whether to enroll) and the intensive margin (where to enroll). In the US, older siblings' enrollment in the target college induces some younger siblings

[2] See Black et al. (2017) for a similar discussion of identifying peer effects among siblings.

to follow them to that same college and also induces a similar number of younger siblings to attend different four-year colleges. Many would not otherwise have attended college at all, so that younger siblings' college enrollment rates and quality of chosen colleges increase substantially. Importantly, spillover effects on college enrollment and quality are much larger for those whose family characteristics predict lower rates of enrollment in four-year colleges. In Chile, Sweden and Croatia, younger siblings are more likely to apply and enroll in a college or major if they have an older sibling who previously enrolled there.

We discuss three broad classes of mechanisms that could explain sibling spillovers on college and major choices. First, an older sibling going to college could affect the costs or benefits of continuing postsecondary education in general, and of attending the same college and major in particular. Second, the postsecondary trajectory of an older sibling could change the preferences younger siblings have over their options. Third, the choices and outcomes of older siblings could influence their younger siblings' choices by providing information unlikely to be transmitted through other means. Older siblings may report to their families information about different attributes of majors and colleges, but also about their overall experience in their target institutions. Younger siblings may place particularly high weight on their older siblings' college experiences, given that the educational success of a close family member may be more salient and predictive of one's own success than less personalized sources of information.

We present suggestive evidence that the provision of information, the third mechanism, is an important driver of our results. We show, for example, that these sibling spillovers exist even when the age difference between siblings makes it unlikely that they will be attending college at the same time. We show spillovers are stronger for younger siblings whose older siblings enroll in colleges and majors with higher scoring peers, lower dropout rates, and higher earning graduates. We also show that the effects are not present when the older sibling drops out of college. All of this is consistent with older siblings' college experiences changing younger siblings' perceptions of or information about the returns to college education. These patterns appear across four countries with fairly different income levels and rates of access to higher education, suggesting this information friction is not driven by any one country's specific institutional details.

These results contribute to two additional major strands of recent research. First, we provide some of the only evidence, and the only in the US context, of peer effects in college choices. Until recently, most of the voluminous peer effects literature exploited random or quasi-random assignment of classmates, schoolmates or roommates to study spillovers of peers' characteristics or risky behaviors onto students' own academic achievement or risky behaviors (Sacerdote, 2011). That literature rarely, if ever, focused on siblings as peers or considered college choices as either treatments or outcomes. Recent research has begun to provide evidence of spillovers between siblings in various behaviors, including: smoking and drinking (Altonji et al., 2017); military service (Bingley et al., 2019); and paternity leave usage (Dahl et al., 2014). The latter two papers argue that increased information, about the returns to military service and employers' reaction to leave-taking, are the most likely mechanism explaining sibling spillovers in these non-educational choices.

A handful of recent papers from outside the US suggest sibling spillovers in educational choices. Using distance to the nearest girls' school as an instrument, Qureshi (2018) shows that additional schooling for Pakistani eldest sisters induces younger brothers to pursue more schooling. Joensen and Nielsen (2018) use quasi-random variation in a school pilot scheme to show Danish older siblings' pursuit of advanced math and science coursework increases younger siblings' propensity to take such courses. Dustan (2018) uses randomness induced by Mexico City's high school assignment mechanism to show students prefer schools older siblings have attended.[3]

A second related literature, consistent with our argument that siblings can provide important information too costly or impossible to obtain otherwise, argues that "low-touch" interventions substantially underperform "high-touch" ones with respect to college choice. Multiple recent papers using nudge-style informational interventions at state or national scale fail to meaningfully impact college enrollment choices (Gurantz et al., 2002; Bird et al., 2019; Hyman, 2019). Researchers testing multiple treatments often find that information only interventions have little impact on students (Bettinger et al., 2012; Carrell and Sacerdote, 2017). Information on college-level earnings, released through the US federal government's College Scorecard website, does not affect the college application patterns of students from non-wealthy families or high schools (Hurwitz and Smith, 2018). One explanation consistent with these findings is that students, particularly low income

[3]Barrios-Fernández (2019) and Aguirre and Matta (2019), two recent working papers, investigate spillovers in higher education in Chile. The first uses a regression discontinuity design to investigate extensive margin spillovers from both close neighbors and siblings. The second paper shows evidence of siblings' spillovers in college choices. Both find results that are consistent with the results in this paper.

ones, may not view as salient the information provided by such interventions, information which is rarely personalized to their own circumstances.

Conversely, higher touch interventions with a more personalized flavor seem to make substantial differences in the college choices of low income students. Carrell and Sacerdote (2017), for example, show in-person mentoring improves college attendance, perhaps by substituting for costly or absent time investments of parents and teachers. Intensive after-school college counseling can improve the quality of colleges first generation students choose (Castleman and Goodman, 2018; Barr and Castleman, 2018). Traditional high school guidance counselors can make a substantial difference in the college choices of disadvantaged students, in part through the provision of information (Mulhern, 2019a). College choices can even be affected by personalized software showing students how their admission chances compare to prior students from their own high school with similar GPAs and SAT scores (Mulhern, 2019b). This body of work suggests students react powerfully to information from people close to them or sources that otherwise are highly personalized and thus salient. Siblings, and social networks more generally, may represent an important channel through which information frictions can be overcome.

In addition to potentially explaining persistent differences in college choices between people in different social networks, the sibling spillovers we document have two other potentially important implications. First, individuals respond to their older siblings' choices due to incomplete information, there is scope for improving the match between students and educational programs with policies that alleviate that friction. Second, educational interventions such as financial aid policy or affirmative action may have larger effects than those typically measured, if treated students' younger siblings are indirectly treated themselves. Multiplier effects that run through social networks may change the benefit-cost ratios of important education policies.

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

2 Institutions

This section describes the institutions and the college admission systems in Chile, Croatia, Sweden and the United States, emphasizing the features that generate the discontinuities we exploit to identify spillovers among siblings (Appendix A describes the institutions in more detail). As shown in Table I, the four countries have important differences in size, culture and in their stage of economic development. The absolute and relative size of the higher education systems is very different across countries. While in Chile and in the United States universities charge tuition fees, in Croatia students do not pay fees if they accept the offer they receive the first time they apply to college, and in Sweden higher education is free.

Chile, Croatia and Sweden use centralized admission systems that allocate applicants to majors and colleges only considering their declared preferences and their academic performance. In Chile and Croatia students compete for places based on a weighted average of their high school GPA and their scores in different sections of a university admission exam. In Sweden students can participate in two independent pools of applicants, one that only considers their high school GPA and a different one that only considers their performance on an admission exam. These systems generate sharp admissions cutoffs in oversubscribed programs. We exploit these discontinuities in older siblings' admission and enrollment in their target major-college to identify sibling spillovers.

Figure I illustrates how older siblings' admissions and enrollment change at the cutoff. The running variable corresponds to older siblings' application scores centered around their target major-college admission cutoff. In Chile and Croatia admission probability increases from 0 to 1 at the cutoff; in Sweden it increases from 0 to 0.6. In Sweden the application system has multiple rounds, and at the end of each round applicants can choose between accepting the offer that they have at that point or dropping it and waiting for the next round. The chart in Panel (c) of Figure I uses the final cutoffs and the final allocation of offers. Since not all applicants wait until the final round, some of them do not receive an offer even if their application scores were above the cutoff.

Figure I also shows that receiving an offer for a specific major-college increases the probability of enrolling there. However, in none of the three countries does admission translate one-to-one into enrollment. In Chile, only half of the universities—the most selective ones—participate in the centralized admission system, and rejecting an offer is not costly. Thus, some applicants reject their offers and some of the applicants on the waiting list can enroll in their target major-college. In Croatia, all universities use the centralized admission system and rejecting an offer is costly. Rejecting an offer forfeits the fee waiver, which means students will have to pay admission fees if they reapply. In this case, non-compliance arises because applicants submit their final list of preferences a couple of weeks after receiving their scores. During that period, they can modify their applications, and they observe their provisional admission outcomes each time. To avoid endogeneity concerns, we focus on the first set of preferences submitted by applicants immediately after they learn their scores. Finally, as in Chile, rejecting an offer in Sweden is not costly. This free disposal and the multiple rounds of the process explain the difference between admissions and enrollment.

In the United States, each college is free to set their own admissions criteria and there is no centralized admission system. However, most colleges take into account applicants' scores in a university admission exam. Although we do not observe the exact rules that each institution uses to select their students, we were able to identify a subset of 21 colleges that seem to be using sharp cutoffs as part of their admission processes.[4] These target colleges are largely public institutions (16 public, 5 private) with an average enrollment of over 10,000 full-time equivalent students, and they are located in eight different states on the East coast.

The median SAT threshold across years for these colleges ranges from 720 to 1060, with students widely distributed across these colleges and thresholds. Figure II illustrates how the probability of enrolling in any 4-year college and in the target college changes at the identified cutoffs (we do not observe admissions for the United States).

An important difference between the cutoffs generated by the centralized admission systems used in Chile, Croatia and Sweden and the cutoffs we identified in the United States is that, while the former do not generate relevant variation in total enrollment, the latter does. In Chile, Croatia and Sweden not all majors are oversubscribed; this means that even if older siblings are rejected from a specific college-major combination, they still have other enrollment options. These outside

[4] Appendix A explains in detail how we identified these colleges.

options, and the fact that it is easier to apply to multiple majors and colleges in these centralized admission systems, explain why the cutoffs in these settings do not create relevant extensive margin responses. This means that we will be able to study extensive margin spillovers only in the United States.

3 Data

We exploit administrative data provided by various public agencies in Chile, Croatia, Sweden and the United States. In these four countries, the main data sources are the agencies responsible for college admissions: DEMRE in Chile, NISpVU and ASHE (AZVO) in Croatia, UHR in Sweden and the College Board in the United States.

From the Chilean agency we obtained individual-level data on all students who registered to take the university admission exam (PSU) between 2004 and 2018. These datasets contain information on students' performance in high school and in the different sections of the college admission exam. They also contain student-level demographic and socioeconomic characteristics, information on applications, admissions to schools that use the centralized application system, and college enrollment. We complement this information with registers from the Ministry of Education and 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. We also observe some program and institution characteristics, including past students' performance in the labor market (i.e. annual earnings). Finally, using the registers from 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.

We were granted access to similar data for Croatia. These registers contain information on students' performance in high school and in the various sections of the college admission exam, and on applications to and enrollment in all Croatian colleges between 2012 and 2018. We also observe students' last names and their home addresses when they registered for the admission exam.

The Swedish application data consist of two parts. We combine applications data from the current system (2008–2016) and from an older system (1992–2005). While the modern system contains the universe of applications to higher education, colleges were not required to participate in centralized admissions before 2006. We do not observe applications to programs that don't participate in centralized admissions.[5] The data also contain information on students' high school GPA, their scores on the admission exam, and an individual identifier that allow us to observe individuals' higher education trajectories and their performance in the labor market.

In the United States we observe all students from the high school classes of 2004–2014 who took the PSAT, the SAT, or any Advanced Placement exam. We observe each student's name, home address and high school attended, as well as self-reported demographic information on gender, race, parental education and family income. We also observe scores each time the student takes the SAT, which allows us to construct the SAT "superscores" that are the index most commonly used in college admissions (Goodman et al., 2020). We observe all colleges to which students send their SAT scores, and by merging the College Board data and the National Student Clearinghouse (NSC), we are able to observe the specific college a student is enrolled in at any point in time.

3.1 Identifying Siblings

In Chile, students provide the national id number of their parents when they register for the university admission exam. Using this unique identifier we can match all siblings that correctly reported these numbers for at least one of their parents.[6] Although registering for the admission exam costs around USD 40, all the students graduating from subsidized high schools—93% of total high school enrollment—are eligible for a fee waiver that is automatically activated when they register for the exam.As a consequence of the fee waiver, even students who do not plan to apply to university usually register for the entrance exam.We complement this data with registers from the *Servicio de Registro Civil e Identificación* in which we observe all individuals born in 1992 or later and their mothers. Considering that in Chile very few individuals complete high school before

[5]Most of these programs had special admission groups and would have been excluded from our analysis for other reasons. The only large exception is Stockholm University, where admissions to some of the larger programs were managed locally for almost the entire period. Our results do not significantly change when we restrict the sample to the later period.

[6]79.4% of students report a valid national id number for at least one of their parents. 77.2% report the national id number of their mother.

they are 18 years old, these additional data let us increase the number of identified siblings for the cohorts completing their secondary education after 2010.

In Croatia and the United States we identify siblings through the home address and the surnames that students report when registering for the exam. We identify siblings as pairs of students from different high school classes whose last name and home address match. This approach should yield relatively few false positives, such as cousins living together. In Croatia, where individuals have two surnames, false positives are even less likely. This approach, however, likely generates many false negatives in which we mistakenly label individuals with siblings as only children. This can stem from two primary sources. First, we fail to identify siblings in families that change residential addresses between the years in which the siblings took the admission exams. Second, and likely a smaller concern, siblings may record their last names or home address differently.[7] Failing to identify siblings will have no impact on the internal validity of our subsequent estimates, but it does affect both sample size and the characteristics of the population we study. In order to have a sibling in this data, families must have at least two children in the high school classes of 2012–2018 in Croatia or of 2004–2014 in the United States, each of whom takes at least one College Board test and lives at the same home address the last time they take the test. We refer to anyone for whom we fail to identify a sibling as an "only child".

In Sweden, family connections and all the demographic and socioeconomic variables that we use are provided by Statistics Sweden. We observe the full set of sibling pairs regardless of whether they registered for an admission exam.

Because some families have more than two siblings, we use each family's oldest sibling to determine the "treatment" status of all younger siblings. The vast majority of siblings in our data appear in pairs, but some come from families where we identify three or more siblings. Families' demographic characteristics are assigned based on the oldest sibling's reports, for consistency across siblings and because treatment status is determined at the time of oldest sibling's college applications. We structure the data so that each observation is a younger sibling, whose characteristics and treatment status are assigned based on their oldest sibling. In addition, if older siblings have applied to college

[7] Our matching process also identifies twins as only children because they are in the same high school class. We do this in order to generate a set of siblings where influences clearly run from older to younger siblings. With twins, the direction of influence is unclear.

multiple times, we only take the first set of applications he or she submitted.

Using these data, we identify around $83,000$, $17,000$, and $300,000$ 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. In the United States we identify around $40,000$ pairs of siblings in which the oldest applied to at least one of the cutoff colleges. Table II presents summary statistics for these subsets of siblings and also for the full set of potential applicants.

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 we look at socioeconomic and academic variables. The individuals in the discontinuity 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. In Chile and Croatia, we observe that individuals with older siblings applying to college are more likely to have followed the academic track in high school. Finally, these individuals seem to have better academic performance in high school and in the college admission exam 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 or to a 4-year college in the past. On top of this, the institutions that use the centralized admission system in Chile are on average more selective. Thus, individuals with active applications to these colleges are usually better candidates than the average student in the population.

4 Empirical Strategy

Identifying siblings' effects is challenging. Since siblings share genetic characteristics and grow up under very similar circumstances, it is not surprising that their outcomes—including their higher education trajectories—are highly correlated. Thus, the first challenge is to distinguish these correlated effects from the effects generated by interactions among siblings. In addition, if siblings' choices simultaneously affect each other, this gives rise to what Manski (1993) described as the reflection problem. In our setting, since 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 jointly decide their higher education alternatives, so our empirical strategy needs to be robust to this sort of joint decision making.

We use two sources of variation to overcome these identification challenges. In Chile, Croatia and Sweden we exploit thousands of cutoffs generated by the deferred acceptance admission systems that universities in these countries use to select their students. For the United States we exploit the variation generated by hidden cutoffs that a subset of colleges seems to use in their admission processes (Appendix A.4 explains how we identify these colleges). Taking advantage of the admissions discontinuities created by these cutoffs, we use a Regression Discontinuity (RD) design to investigate sibling spillovers on the decision to enroll in any college (extensive margin spillovers) and on the choice of college and major (intensive margin spillovers).

Since individuals whose older siblings are marginally admitted to or rejected from a specific major or college are very similar, the RD allows us to rule out that the estimated effects are driven by differences in individual or family characteristics, eliminating concerns about correlated effects. Moreover, considering that the variation that we exploit in the major or 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 rule out concerns related to the reflection problem.

We only have relevant variation in older siblings' total college enrollment in the United States. In the other three countries, older siblings still have many other options available if they are rejected from an oversubscribed major, which makes the difference in total enrollment much smaller. Therefore, we will only investigate extensive margin spillovers in the United States. When looking at intensive margin spillovers, our sources of variation allow us to investigate how older siblings' admission to their target college affects the probability that their younger siblings will apply to or enroll in the same college. Finally, the variation that we have for Chile, Croatia and Sweden also allows us to study how older siblings' admission to their target major affects their younger siblings major choice.

As discussed in Section 2 in none of the countries that we study admission translates one-to-one into enrollment. Thus, in order to study how older siblings' actual enrollment affects their younger siblings choices, we use a fuzzy RD in which we instrument older siblings' enrollment in a specific college or major with an indicator of admission to that college or major.

Our empirical design defines, for each college and major, the sample of older siblings marginally admitted to or marginally rejected from it, and then compares how this admissions decision affects younger siblings' choices. 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 or college.

In the United States the cutoffs that we observe affect admission probabilities to colleges. However, in the other three countries that we study the cutoffs define admission to specific major-college combinations. Since in these countries we observe all the applications submitted by the individuals through the centralized systems, we use this information to define two different samples for each country. These two samples allow us to investigate sibling spillovers on both the choice of college and the choice of major.

4.1 Major Sample

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

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 $(j, f, u) \succ (j', f', u')$.[8] Denoting the application

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

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 an application score sufficiently close to j's cutoff score to be within a given bandwidth bw around the cutoff:
   $|a_{ijfut} - c_{jfut}| \leq bw$.

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

4.2 College Sample

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

3. Listed major j in college u as a choice such that majors not preferred to j in their application list are dictated by an institution different from u or if dictated by u had cutoffs above their application scores (otherwise being above or below the cutoff would not generate variation in the college they attend).

This restriction removes from the sample older siblings who in case of being rejected from their target major-college combination would receive an offer to enroll in different major, but in the same target college. This makes our first stage stronger and improves the precision of our estimates.

4.3 Identifying Assumptions

As in any 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 this assumption plausible. In the United States the cutoffs that we exploit are not observed by applicants, which makes manipulation even less likely. However, to confirm that applicants are not manipulating their scores, we show that the distribution of the running variable in each setting is continuous around the cutoff (see Appendix C 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 C shows that this is indeed the case for a rich set of socioeconomic and demographic characteristics in all our settings.

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).[9] When focusing on intensive margin spillovers, in addition to the usual IV assumptions we also need to assume that receiving an offer for a specific major or college does not make enrollment in a different major or college more likely (Appendix B presents a detailed discussion of the identification assumptions.) Given the structure of the admission systems that we study, this additional assumption is not very demanding.[10]

[9]Independence, relevance, exclusion and monotonicity. In this setting, independence is satisfied around the cutoff. The existence of a first stage is shown in Figure I. The exclusion restriction implies that the only way older siblings' admission to a major or college affects younger siblings' outcomes is by increasing older siblings' enrollment in that major or college. Finally, the monotonicity assumption means that admission to a major or college weakly increases the probability of enrollment in that major or college (i.e. admission does not decrease the enrollment probability).

[10]In Chile—where not all colleges use centralized admissions—or in the United States—where each school runs its own admission system—this assumption could be violated if, for instance, other colleges 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 would define students' incentives based on admission cutoffs that they only observe ex-post or do not observe at all, we cannot completely rule out this possibility. In Croatia—where students lose their funding if they reject an offer—and Sweden—where there are no tuition fees—violations of this assumption seem unlikely.

A final consideration when interpreting our results on intensive margin spillovers relates to the findings of Barrios-Fernández (2019). According to this work, the probability of attending university increases with close peers' enrollment (i.e. neighbors and siblings). If marginal admission to majors or colleges translates into an increase in total enrollment, then our estimated results could simply reflect that individuals whose older siblings attend college are more likely to enroll in any college. In Appendix C we show that older siblings' marginal admission to their target majors does not generate a relevant difference in their own or in their younger siblings' total enrollment. This concern is more relevant in the United States, where we document a significant increase in total enrollment for both older and younger siblings. From this perspective, the variation that we have in the United States seem better suited to studying extensive margin responses. However, decomposing the extensive margin response among those following their older siblings to the same college and those going somewhere else is still interesting for understanding the drivers of the extensive margin responses.[11]

Appendix C presents multiple 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 generate a significant effect on any of the outcomes that we study.

4.4 Method

For all estimates, we pool observations from all "cutoff using"colleges or over-subscribed majors. We center older siblings' admission scores around the relevant admission cutoff. The following equation describes our baseline specification:[12]

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

[11]Note that the increase that we find in enrollment of younger siblings in the target college of their older siblings is well above the increase that we would observe if the increase in total enrollment were randomly allocated across the colleges chosen by similar individuals, but with no sibling effects. In addition, the effects we find have a similar magnitude in the US and in other countries.

[12]In the United States the variation is at the college level, so we can eliminate the major subscript. In addition, in this case the cutoffs are constant over time. Thus, the term $\mu_{ju\tau}$ is replaced by μ_u and μ_τ.

$y_{ijut\tau}$ is the outcome of interest for the younger sibling from sibling-pair i and cohort 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 for whether the older sibling from sibling-pair i had an admission score a_{ij} above the cutoff (c_{uj}) of major j offered by college u in year τ ($a_{iju\tau} \geq c_{uj\tau}$).

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

$\mu_{ju\tau}$ is a fixed effect for the older sibling's cohort and target college and major, and ε_{ijut} is an error term.

Our main results are based on local linear specifications in which we use a uniform kernel and control by the following linear function of the running variable:

$$f(a_{ijut\tau}; \gamma) = \gamma_0 * a_{ijut\tau} * 1[a_{iju\tau} < c_{uj\tau}] + \gamma_1 * a_{ijut\tau} * 1[a_{iju\tau} \geq c_{uj\tau}].$$

This specification allows the slope to change at the admission cutoff. In Appendix C we show that our results are robust to using a quadratic polynomial of $a_{iju\tau}$, a triangular kernel, or allowing the slope of the running variable to be different for each admission cutoff.

For the results from the United States we use a bandwidth of 90 SAT points. This is the median (and mean) Calonico et al. (2014) optimal bandwidth for the three main outcomes. When looking at intensive margin spillovers in Chile, Croatia and Sweden we investigate three outcomes: the probability of ranking the older sibling's target in first preference, the probability of ranking it in any preference, and the probability of enrolling in it. Depending on the margin of interest (i.e. college or major choice) we use one of the samples described in Section 4. For these analyses we also compute optimal bandwidths according to Calonico et al. (2014). We do this for each sample and outcome, but then we use a single bandwidth per sample: the smallest among the ones computed for the three outcomes under study.[13]

All of the specifications we use focus on individuals whose older siblings are near an admission cutoff.

[13] 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 C.IV, C.V and C.VI show that our estimates are robust to different bandwidth choices in all the settings that we investigate.

This means that our estimates represent the average effect of older siblings' marginal admission compared to the counterfactual of marginal rejection from a target major or college.[14]

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 ($admitted_{iju\tau}$). In Chile, Croatia and Sweden, 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, or if she has more than one younger sibling. For these countries, we cluster standard errors at the family level. In the United States we cluster standard errors at the older sibling's high school level since older siblings rarely appear in our sample multiple times.

To study heterogeneous effects, we augment the baseline specification by adding an interaction between older siblings' admission and the characteristic along which heterogeneous effects are being studied (i.e. $admitted_{iju\tau} \times x_{ijut\tau}$). We also use this interaction as an instrument for the interaction between older sibling's enrollment and $x_{ijut\tau}$. In these specifications we include $x_{ijut\tau}$ as a control.

5 Results

This section presents results on sibling spillovers. First, we show how older siblings' marginal admission to college increases younger siblings' four-year college enrollment in the United States (which is the only country where we have relevant variation in older siblings' total enrollment). Second, we show that older siblings' marginal admission to a college or major leads their younger siblings to apply to and enroll in the same college and major. Third, we show that older siblings' college choices do not influence the academic performance of their younger siblings. Finally, we show that younger siblings' responses vary depending on sibling and major characteristics. Sibling spillovers seem to be stronger when the older sibling has been successful in their chosen major or college.

[14]In Appendix Tables C.VII and C.VIII we present the results of an additional specification that also controls by target major $\times$ counterfactual major fixed effect. The estimates are very similar to the ones presented in the main section of the paper.

5.1 Sibling Effects on College Enrollment: Extensive Margin

First, we show that in the United States older siblings' enrollment in four-year colleges impact the types of institutions that younger siblings attend.

In the US, SAT admissions cutoffs generate variation in whether applicants enroll in any 4-year college.[15] Panel (A) in Figure II indicates that older siblings with SAT scores above the admission cutoff of their target college are 3 percentage points more likely to attend any four-year college. This increase is largely because these students are now more likely to attend the target (four-year) college than a two-year college. Panel (B) in Figure II indicates that older siblings with SAT scores above a target college's admission cutoff are 8.5 percentage points more likely to attend their target college than students with scores just below the threshold. Instrumenting the older siblings' enrollment in their target college with being above the admission threshold indicates that older siblings' enrollment in their target college increases their own probability of attending a 4-year college by 42 percentage points, and reduces 2-year college enrollment by 34 percentage points. Thus, only eight percent of the marginal older siblings would not have attended college if they had not crossed the threshold.

Panel (A) of Figure III indicates that older siblings' marginal admission to a target college increases younger siblings' enrollment in 4-year colleges. The IV estimate in column (1) of Table IV shows that applicants whose older siblings enroll in a target 4-year college are 22.9 percentage points more likely to enroll in a 4-year college than students whose older siblings just miss the admissions cutoff. Column (2) shows a small and insignificant decrease in 2-year college enrollment. This indicates that the older sibling's admission to her target college leads to some younger sibling movement from 2-year to 4-year colleges, as well as some increased enrollment among younger siblings who would not have enrolled in any type of college.

Columns (3) and (4) of Table IV also indicate that older siblings' admission to target colleges increases the quality of the colleges that younger siblings attend. Quality here is measured by the bachelor's degree completion rate and the standardized PSAT scores of students attending the

[15] Appendix C shows that older siblings' marginal admission to their target major-college only generates a small increase in their own total enrollment and does not generate significant differences in younger siblings' total enrollment in Chile, Croatia or Sweden.

institution.[16] Younger siblings whose older sibling attended the target college enroll in colleges with graduation rates 18 percentage points higher and peer quality 0.31 standard deviations higher than the ones they otherwise would have chosen.[17]

5.2 Sibling Effects on College Choice

Next, we show that an older sibling's enrollment in a target college increases the younger siblings' probability of enrolling in that same college. Our main outcome is an indicator for whether the younger sibling applies to or enrolls in their older sibling's target college. As before, the treatment is a dummy variable that indicates if the older sibling enrolls in his or her target college, and our instrument is crossing an admissions threshold.

Panel (B) of Figure II shows that older siblings' enrollment in their target college increases when they cross the admissions threshold. Figure III also indicates that younger siblings' enrollment choices are affected by their older siblings' proximity to the admissions threshold. Younger siblings are 1.4 pp more likely to enroll in the target college when the older sibling is above the admissions threshold. We can combine these first stage and reduced form results to obtain the IV estimates in columns (5) and (6) of Table V. These estimates indicate that individuals whose older siblings enroll in their target college are 27.0 pp more likely to apply to it and 17.2 pp more likely to enroll there than those whose older siblings were marginally rejected.

Although this evidence strongly suggests the existence of sibling spillovers on the intensive margin, we need to be careful when interpreting the magnitude of our estimates. Older siblings' enrollment in their target college also generates extensive margin spillovers. This means that at least part of the increase that we find on enrollment in older siblings' target college could be a mechanical

[16] We build a peers' quality measure following Smith and Stange (2016) and computing for each college the average standardized PSAT score of initial enrollees. This peer quality measure allows comparisons between two- and four-year institutions; the former do not require SAT scores and thus lack a peer quality measure in IPEDS. Students who do not enroll in college are assigned the mean PSAT score of all such students for this variable[1]. In addition, we build a second quality index using the NSC data to compute for each college the fraction of initial enrollees who eventually earn a B.A. from any college within six years. Unlike the IPEDS' graduation rate measures, this accounts for transfers between institutions and allows direction comparison of two- and four-year colleges. Students who do not enroll in college are assigned a zero for this variable.

[17] In Appendix D we present similar results for Chile, Croatia and Sweden. Given the nature of the variation that we exploit in these countries, in these settings we do not find significant changes in the quality of the programs attended by younger siblings.

consequence of the increase in the share of individuals going to 4-year colleges. However, the size of the effect (17.2 pp) makes it unlikely that it is only caused by extensive margin spillovers.[18]

In Chile, Croatia and Sweden this is not a concern. The variation that we exploit in these settings generates little variation in older siblings' total enrollment, and no significant variation in younger siblings' total enrollment. For these countries, apart from enrollment decisions, we also observe the preference rank submitted by younger siblings when they apply to college, so we investigate effects on three different outcomes: the probability of ranking older siblings' target college in the first slot, the probability of ranking it in any slot and finally the probability of enrolling in it. Our first stages, illustrated in Figure I, show that older siblings marginally admitted to their target major are significantly more likely to enroll there than those marginally rejected. Given that in this section our analyses focus on the "College Sample", this also generates variation in the college older siblings attend.

Figure IV shows that an older sibling's marginal admission into their target college also affects application and enrollment decision by their younger siblings. Younger siblings are more likely to rank older siblings' college in their first preference, in any preference and to enroll in it if the older sibling is admitted to that college. Table V summarizes these results and presents IV estimates for the effect of older siblings' actual enrollment in their target college on the outcomes of interest. According to these results, in Chile individuals are 6.7 pp more likely to rank their older siblings' target college first and 7.6 pp more likely to apply to it. They are also 3.8 pp more likely to enroll in that college. For Croatia, the same figures are 7.5 pp, 10.9 pp and 8.4 pp, and for Sweden they are 15 pp, 15.3 pp and 6.4 pp. Although the results on enrollment are slightly smaller than for the United States, this evidence points in the same direction. Having an older sibling enrolling in a particular college increases the likelihood of applying to and enrolling in it.

[18] On the left hand side of the admission cutoffs the share of individuals enrolling in the target college of their older sibling is 1.58% (0.006/0.38). On the right hand side it is 29.2% (0.178/0.609). If preferences were stable around the cutoff and older siblings did not affect preferences for specific colleges, we should find 1% (1.58% $\times$ 60.5%) of the younger siblings on that side enrolling in the target college of their older sibling. However, the increase we find is well above 0.4 percentage points.

5.3 Sibling Spillovers on Major Choice

This section discusses how older siblings' admission and enrollment in a specific major-college combination affects their younger siblings' probabilities of applying to and enrolling in that same program. To investigate changes along this margin we focus our attention in Chile, Croatia and Sweden, which are the only countries in which we have quasi-random variation in older siblings' major-college enrollment. The analyses that we present here closely follow the ones discussed in the previous section, but this time we use the "Major Sample"and we define the outcomes in terms of major-college instead of college.

The RD estimates illustrated in Figure V 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.

As discussed in Section 2, receiving an offer for a specific major does not translates one-to-one into enrollment in any of these countries. Thus, in order to estimate the effect of older siblings' enrollment on younger siblings' applications and enrollment decisions, we combine the reduced form results discussed in the previous paragraph with the first stages illustrated in Figure I, and obtain the fuzzy-RD estimates presented in Table V. Under the identification assumptions discussed in Section 4, these fuzzy-RD procedures provide consistent estimates for the effects of interest.

In Chile, having an older sibling "marginally enrolling"in a specific major increases the likelihood of applying to that major in the first preference by 1.2 pp, and in any preference by around 2.3 pp. This increase in applications also translates into an increase of around 0.6 pp in enrollment. The results for Croatia are very similar. Individuals are 1.4 pp more likely to apply to their older siblings' target major in the first preference, 3.4 pp more likely to apply to it in any preference, and 1.4 pp more likely to enroll in it. Finally, in Sweden, the likelihood of ranking an older sibling's target major in the first slot increases by around 2 pp, while the likelihood of ranking it in any position increases by around 3 pp. Enrollment in older siblings' major increases by roughly 0.4 pp.

Since applicants in these three settings receive their scores before submitting their applications, their rankings may depend on how likely they believe they are to be admitted in their older siblings' target major. In Table VI we present additional results that come from augmenting the baseline specification with an interaction between older siblings' marginal enrollment and a proxy of younger siblings' eligibility for their older sibling's target major.[19] According to the results presented in Table VI, younger siblings are more likely to apply to and enroll in their older siblings' target major if they are eligible for it.[20]

Despite the differences between Chile, Croatia and Sweden, the results in this section are consistent across countries. They indicate that, especially when younger siblings are likely to be admitted to their older siblings' target major-college, they are more likely to apply to and enroll in it.

5.4 Sibling Spillovers by Age Difference and Gender

This section explores whether the responses in major and college choice documented in the previous sections vary depending on siblings' age difference and gender.[21]

We summarize the results of this section in Table VII. Column (1) investigates differences by age, distance and siblings' gender on extensive margin spillovers (i.e. enrollment in any 4-year college); columns (2) to (5) focus instead on the probability that younger siblings apply to their older sibling's target college; and columns (6) to (8) focus on the probability that they apply to their older sibling's target major. When looking at heterogeneous effects by age difference in the United States, our results suggest that the effects on the decision to enroll in a 4-years college and also on the specific college chosen are stronger for siblings who were born five or more years apart. These results contrast with our findings for Chile, Croatia and Sweden. In these countries the probability of following an older sibling to her target college decreases with age. However, despite this decrease there is still a significant and meaningful effect even for siblings born more than 5 years apart. We find a similar pattern when looking at the choice of major. In this case the size of the effect also seems to decrease with age, but there is still a significant effect for siblings with big age differences.

[19] In Chile and Croatia the eligibility proxy is an indicator for whether the younger sibling's exam scores would let them gain admission to the older sibling's target major. In Sweden, since the scale of the GPA and the admission exam change during the period that we study, we use 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.

[20] In section 5.8, we show that older siblings' enrollment in 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 analysis.

[21] The analyses presented in this section focus on **applications** to majors and colleges. Similar results for **enrollment** are presented in in Appendix D.

When looking at heterogeneous effects by siblings' gender, we find that extensive margin responses seem to be stronger among siblings of opposite gender. However, we find no gender specific differences in the probability of applying to the target college of older siblings, and when looking at the probability of applying to the target major of older siblings we find that the responses are stronger among siblings of the same gender. As shown in Panels (C) and (D) of Table VII these gender differences are almost completely driven by what happens with older brothers. The probability of following an older brother to his target major-college is significantly larger for males than for females. On the other hand, older sisters are similarly followed by both male and female younger siblings.

5.5 Sibling Spillovers and College Quality

In this section we investigate whether sibling effects vary with the quality of older siblings' target college and major.[22] Our first measure of quality aims to capture peers' academic potential. We define the quality of the students in a program in a given year as the average performance of admitted students in the college admission exams in Chile, Croatia and the United States, and as the average high school GPA of admitted students in Sweden. Second, we compute first year dropout rates for each major and year using individual level data provided by the Ministry of Education in Chile and by the Council for Higher Education in Sweden. We do not observe enrollment after the first year in Croatia, or for the universe of students in the US, which prevents us from computing these figures in these countries.[23] We compute a third quality index based on college specific graduation rates. [24]

Finally, for Chile, Sweden and the United States we build an index based on graduates' earnings.

[22]Here we focus on younger siblings' **applications** to colleges and majors. Appendix D presents similar results for **enrollment**.

[23]The data from Chile allow us to compute dropout rates for all college cohorts beginning in 2006. The cohorts of older siblings applying to university in 2004 and 2005 are assigned the dropout rates observed for their target majors in 2006. Since some programs disappear from one year to the next, we do not have complete information for all the majors offered in 2004 and 2005. In Sweden we are able to compute first year dropout rates for the entire sample period.

[24]In Chile, this index is constant over time, and in the US and Sweden it is cohort specific. In Chile, the formal duration of undergraduate degrees ranges from 9 to 14 semesters. Thus, computing cohort specific graduation rates would force us to dramatically reduce our sample size. Thus, we compute graduation rates for the cohorts that we can observe for the formal duration of their major plus one additional year, and we assign this average graduation rate to all cohorts.

In Chile, graduates average earnings are measured four years after graduation and are reported by the Ministry of Education.[25] In Sweden, we compute average earnings one year after graduation using administrative registers managed by Statistics Sweden. We assign to each older sibling the average earnings of the cohort graduating from their target major the year in which they apply to college. For the United States, we use earnings measures reported by Chetty et al 2017. These date indicate the average earnings of a college's students approximately 10 years after students are expected to complete college. We standardize these earnings measures in our sample.

Table VIII summarizes our key results. We standardize all variables, except for dropout rates and graduation rates. The first column looks at heterogeneous effects on extensive margin spillovers, the next four columns focus on applications to older siblings' target colleges, and the last three columns focus on applications to older siblings' target majors.

When looking at extensive margin spillovers, although the differences are not significant, the responses seem to be stronger when the older sibling enrolls in colleges with better peers and higher graduation rates. When looking at the probability of following an older sibling to college, in Chile, Croatia and Sweden younger siblings are more likely to apply to their older sibling's target college when it has better peers, lower first year dropout rates and higher graduation rates. In the United States, although the results point in the opposite direction, none of differences is significant. Graduates' earnings, on the other hand, do not seem to affect the likelihood of following older siblings to college in any of the countries that we study.

When focusing on the major choice we find smaller and less precise differences in sibling effects by quality of the older sibling's target major. Only in Sweden does the likelihood of following older siblings to their target majors significantly increase with peers' quality. Although dropout and graduation rates have the expected sign, neither seem to make a significant difference. Finally, when looking at heterogeneity by graduates' earnings, we find that younger siblings are more likely to apply to their older sibling's target major when previous graduates from that major perform better in the labor market.

[25] Average earnings are computed by the Ministry of Education with the support of the National Tax Authority. These figures are only available for majors that were offered in 2018 and that had more than 4 cohorts of graduates. In addition, these statistics are only reported for majors with at least 10 graduates. In our analysis this variable does not change over time.

According to these results, individuals do not follow their older siblings everywhere. The responses, especially when looking at sibling effects on the choice of college, are stronger the older sibling attends a better institution.

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

This section shows that younger siblings are significantly less likely to follow their older sibling to a college or major if the older sibling drops out. Table IX shows that siblings' effects disappear if the older sibling drops out. This result is consistent with the hypothesis that individuals learn from their older siblings' experiences 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 to and enrolling in that alternative is not a good choice.

We are only able to compute dropout for Chile, Sweden and the United States, 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, and the main effect of older siblings' dropout.

The results of this exercise should be interpreted with caution because dropping out of college is not random. Although controlling for dropout helps us capture some of the differences that may exist between individuals who remain at and leave a particular college, there could still be differences we are unable to control for. In addition, we can only build the dropout variable for older siblings who actually enroll in some major. Appendix Table C.IX shows that in Chile and Sweden marginal admission does not translate into relevant increases in older siblings' total enrollment. Thus, in these countries we focus our analysis on individuals whose older siblings enroll in some college.In the United States, on the other hand, marginal admission increases older siblings' total enrollment and we include everyone in the estimation sample (i.e. individuals whose older siblings enroll and do not enroll in college). Since dropout can only be computed for older siblings who enroll, this specification does not control for its main effect.

Bearing these caveats in mind, the results of this exercise show that individuals whose older siblings drop out 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.

5.7 Effects on Application and Enrollment by Socioeconomic Status and Exposure to College

This section shows that, while in the United States sibling effects are significantly larger for individuals with low probabilities of enrolling in college, in Chile and Sweden the socioeconomic status of individuals does not seem to affect the size of the effects on college or on major choice. However, in Chile and Sweden the effects seem to be stronger for individuals less exposed to the target college of their older sibling. We compute exposure as the share of schoolmates completing high school one year before the younger sibling who enroll in the older sibling's target college.

Panels (B) and (C) of Table IV summarize the results for the United States. Panel (B) indicates that older siblings' college enrollment has the largest impact on the enrollment of younger siblings who are "uncertain college goers". Uncertain college goers are defined as students whose probabilities of attending a four-year college, based on observable characteristics, are in the bottom third of our sample. We use the sample of "only children"and sociodemographic characteristics we observe in the college board data to predict the relationship between observable characteristics and the likelihood of enrolling in a four-year college. The second row of results in Table IV shows that the spillover effects on four-year college enrollment are considerably larger for the third of students in the uncertain college-goers group. Among this group of students, having an older sibling enroll in a target college increases their four-year college enrollment by 52 percentage points. These students also significantly alter the quality, price and location of the institutions they attend based on the enrollment of their older sibling. Panel (C) indicates that older siblings have little impact on the enrollment of younger siblings who are probable college goers. As in the case of extensive margin spillovers, the probability of enrolling in the older sibling's target college increases more for uncertain college goers.

In Chile and Sweden, Table X shows that the socioeconomic status of individuals does not generate a significant difference in the effects we find on the probabilities of applying to and enrolling in older siblings' target colleges and majors. It is worth highlighting that these results are not directly comparable with the ones we find for the United States. As discussed in Section 3, the individuals we observe in Chile and Sweden are positively selected. They come from families where at least one child was in the margin of being admitted to a selective program. This positive selections means that we observe very few uncertain college-goers in these countries .

Table XI presents the results of a similar exercise, but this time we focus on the exposure that younger siblings have to the target college of their older siblings. Although our estimates are not always precise, these estimates suggest that sibling effects are stronger when the younger sibling has less exposure to the college of the older sibling.

5.8 Effects on Academic Performance

In this section we study whether the increase in the likelihood of applying to and enrolling in the major attended by an older sibling could be driven by an improvement in younger siblings' academic performance. We use the same fuzzy-RD strategy discussed in Section 4, but this time we look at different measures of younger siblings' academic performance. Since not all potential applicants take the admission exam, we replace missing values with zero. When looking at effects on exam scores our estimates capture differences in performance, but also differences in the probability of taking the exam. We use the same bandwidths from the previous sections.

Table XII summarizes these results. We show that having an older sibling enrolling in her target college in the United States or in her target major in the other three countries does not generate significant changes in younger siblings' high school performance, or in their performance in the college admission exams. We find no significant increase in the probability of taking the admission exam or of submitting applications to college. In Chile, Croatia and Sweden we study applications using a dummy variable for whether younger siblings submit at least one application to college; in the United States we use the total number of applications submitted. In Sweden, where students do not need to take the admission exam to apply, we find a decrease in the share of younger siblings taking it, and we also find a small reduction in the share of younger siblings applying to college.

On balance, we do not find any evidence that the sibling effects on the choice of college and major are driven by an improvement in the academic performance of younger siblings.

6 Discussion

Our results in Section 5 show that older siblings' higher education trajectories influence the choice of college and major of their younger siblings.

We discuss three broad mechanisms through which older siblings can affect younger siblings' choices. The first possibility is that the older sibling's educational trajectory can affect the costs and benefits of attending college, and possibly the value proposition of a specific institution or major. Another possibility is that the older siblings' outcomes affect the utility younger siblings derive from higher education. A third broad category of mechanisms is that the older sibling's experiences can provide information that would otherwise not be available. We explore each of these potential mechanisms in detail.

On the extensive margin, having an older sibling attending her target college could affect the family's budget constraint. If the family spent a large share of its budget on the older sibling's education, this could even decrease college attendance among younger siblings. However, our results from the United States indicate that older siblings' enrollment increases younger siblings' enrollment in any college and also in 4-year colleges. These results suggest that the additional costs faced by families in which one of their children enrolls in college do not make it less likely that their younger siblings will apply to and enroll in college. In general, having an older sibling attending a particular college campus may affect the costs faced by younger siblings. This would be the case if, for instance, siblings attending the same college could save on commuting and living costs. Costs could also be affected if having an older sibling enrolled in college increases the amount of financial aid available for younger siblings, or if the same college would offer them a tuition discount. In the four countries that we study, sibling spillovers persist even among siblings who, due to age differences, are unlikely to attend college at the same time. These price effects therefore seem unlikely to explain a large portion of the observed spillovers, especially since in two out of the four settings that we study universities do not charge tuition fees.

Although not directly related to prices, another way sibling spillovers could arise in this setting is if colleges offer family members some type of advantage in the admission process. In the United States, such legacy effects are common because some colleges give admissions preferences to students whose family members have previously enrolled. Hurwitz (2011) noted that this practice if frequent among colleges that seek to increase the likelihood of donations among family members. However, legacy effects are unlikely to explain the spillovers that we find. First, the target colleges that we identified in the United States are largely public, non-flagship institutions, whereas legacy admissions are concentrated in more prestigious colleges. Second, in Chile, Croatia and Sweden colleges select their students only based on their previous academic performance. Thus, legacy effects do not play any role in these countries.

Alternatively, having an older sibling enrolled in a specific college or major could affect individuals' preferences. Preferences could change if younger siblings experience utility gains from being close to their older siblings. This would be the case if they enjoy the company of their older sibling, but also if they think that their older sibling could support them with their studies or make their college experience easier. Preferences could also be affected if individuals perceive their older siblings as role models and get inspired by them, if they are competitive, or if parental pressure changes as a consequence of older sibling enrollment.

The persistence of the effects among siblings with large age differences suggests that our results are not driven by siblings enjoying each other's company or by the benefits that may arise from attending the same campus simultaneously. In the United States younger siblings' four-year college enrollment rose by twice as much as enrollment in their older siblings' target college, further suggesting that this sibling proximity channel is not the main driver of our results.

On the other hand, finding no effects on younger siblings' academic performance suggests that there are no changes in individuals' aspirations or on the pressure that parents put on them to apply to and enroll in college. If this were the case, we would expect to see younger siblings exerting additional effort in preparation for college. We do not see evidence of additional effort in younger siblings' applications, high school performance or admission exams scores. 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 that competition is driving their results. As previously discussed, in our case the results persist even among siblings with important age differences, and also among opposite gender siblings, suggesting that competition is not the main driver of our results.

Finally, older siblings' enrollment in specific colleges and majors 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. Considering the amount of college and major options from which applicants can chose, both hypotheses could play a role. An older sibling's enrollment at a particular college campus may generate information for parents or a younger sibling that would otherwise be costly or impossible to obtain. Information from older siblings may change younger siblings' perceptions of the potential returns to attending high quality colleges, which would be consistent with our observations of younger siblings applying to and enrolling in higher quality colleges than they would otherwise. Younger siblings may place particularly high weight on their older siblings' college experiences, given that the educational success of a close family member may be more salient and predictive of one's own success than less personalized sources of information. We find stronger sibling spillovers when older siblings' colleges are higher quality, a result that goes against a pure salience story. If salience were the main driver of our results, we should see individuals following their older siblings independently of the quality of their majors and colleges. On the other hand, we show that the effects are driven by older siblings who enroll in colleges that are better in terms of student quality, retention, and graduates' labor market performance. These results and the improvement in the quality of colleges attended by younger siblings in the United States suggests that individuals learn from their siblings about college quality. In addition, the fact that sibling effects vanish if the older sibling drops out suggests that the experience older siblings have in higher education matters, and that younger siblings are more likely to follow their older siblings when the older sibling has a good experience in higher education.

Even though the evidence discussed in this section does not allow us to perfectly distinguish the exact mechanisms behind our results, it suggests that information, particularly information about the college experience of someone close, might play a relevant role in college related choices. Further research is required to learn about the exact information transmitted by close peers.

7 Conclusion

Despite the importance of choosing a good college and major for high school students, we know little about how students form the preferences and beliefs that drive these choices. These decisions are complex: colleges and majors are heterogeneous, and it is often difficult to observe some of the important features of each program. In this context, close relatives and other members of an individual's social network can significantly influence college choices. However, it is notoriously challenged to causally identify the effects of social interactions.

In this paper, we provide some of the first causal evidence of peer effects in college choice by investigating how college application and enrollment decisions are affected by the higher education trajectories of older siblings. We study these sibling spillovers in Chile, Croatia, Sweden and the United States. Our empirical strategy takes advantage of admission cutoffs that generate discontinuities in older siblings' enrollment in their target colleges and majors. In the first three countries the admission cutoffs are generated by the centralized admission systems that universities use to select their students. In the United States the cutoffs come from a group of colleges that use a threshold rule as part of their admission process. We exploit this variation in a fuzzy Regression Discontinuity Design framework that allows us to overcome the main identification challenges that arise in the context of peer effects (i.e. correlated effects and the reflection problem).

Despite the differences that exist between these four countries, we consistently find statistically and economically significant sibling effects. In the United States we find that an older sibling enrolling in her target college increase her younger siblings' enrollment in the target college, and also their enrollment in other colleges. Similarly, in Chile, Croatia and Sweden younger siblings are more likely to apply to and enroll in a specific college and major if they have an older sibling who went there before. These sibling spillovers persist even when siblings' age difference makes it unlikely that they will attend college at the same time.

Finding similar results in four settings as different as the ones that we investigate not only helps to address concerns about the external validity of our results, but also to study the mechanisms behind them. We discuss three broad classes of mechanisms consistent with our main results: changes in the family budget constraint, changes in preferences, and changes in the choice set of individuals. Our results show that individuals enroll in better colleges when an older sibling has gone to college before them, and that they are more likely to follow their older siblings to "high" quality colleges. In addition, we find that the experience of older siblings in higher education makes a difference on their younger siblings' responses. We argue that an older sibling's enrollment in a high quality college can provide families with information about postsecondary education that would otherwise be difficult or impossible to obtain. In this sense, an older sibling's college choice is a particularly high touch intervention, providing prolonged exposure to another person's college experience.

Our findings suggest that, especially in settings with incomplete information, policies that change the pool of students admitted to a specific college or major could have an indirect effect on these students' 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.

If students' college choices are deeply affected by the college experiences of people in their social networks, such social factors may partly explain persistent differences in college trajectories by income, race and geography. Future research might investigate the role played by college choices of other socially close individuals, such as parents, friends or neighbors. Such social factors have been relatively understudied by economists and might shed further light on the origins of inequalities in postsecondary outcomes.

References

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

Altonji, J. G., P. Arcidiacono, and A. Maurel (2016). The analysis of field choice in college and graduate school: Determinants and wage effects. In *Handbook of the Economics of Education*, Volume 5, pp. 305–396. Elsevier.

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

Altonji, J. G., S. Cattan, and I. Ware (2017). Identifying sibling influence on teenage substance use. *Journal of Human Resources 52*(1), 1–47.

Angrist, J. (2014). The perils of peer effects. *Labour Economics 30*(C), 98–108.

Barr, A. and B. Castleman (2018). An engine of economic opportunity: Intensive advising, college success, and social mobility. . *Working Paper*.

Barrios-Fernández, A. (2019). Should I Stay or Should I Go? Neighbors' Effects on University Enrollment. CEP Discussion Papers, Centre for Economic Performance, LSE.

Belley, P. and L. Lochner (2007). The Changing Role of Family Income and Ability in Determining Educational Achievement. *Journal of Human Capital 1*(1), 37–89.

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.

Bingley, P., L. P., and S. Lyk-Jensen (2019). Brothers in arms: Spillovers from a draft lottery. *Journal of Human Resources.*

Bird, K. A., B. Castleman, J. T. Denning, J. Goodman, C. Lamberton, and K. O. Rosinger (2019). Nudging at scale: Experimental evidence from FAFSA Completion Campaigns. *NBER Working Papers* (2615).

Black, S. E., S. Breining, D. N. Figlio, J. Guryan, K. Karbownik, H. S. Nielsen, J. Roth, and M. Simonsen (2017, January). Sibling spillovers. Working Paper 23062, National Bureau of Economic Research.

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.

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

Campbell, S., L. Macmillan, R. Murphy, and G. Wyness (2019). Inequalities in student to course match: Evidence from linked administrative data. Cep discussion papers, Centre for Economic Performance, LSE.

Carrell, S. and B. Sacerdote (2017, July). Why Do College-Going Interventions Work? *American Economic Journal: Applied Economics 9*(3), 124–51.

Castleman, B. and J. Goodman (2018). Intensive College Counseling and the Enrollment and Persistence of low-income Students. . *Education Finance and Policy 1*(13), 19–41.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014, September). Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review 104*(9), 2633–79.

Chetty, R., J. N. Friedman, E. Saez, N. Turner, and D. Yagan (2017, July). Mobility Report Cards: The Role of Colleges in Intergenerational Mobility. NBER Working Papers 23618, National Bureau of Economic Research, Inc.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014, 09). Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States *. *The Quarterly Journal of Economics 129*(4), 1553–1623.

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

Deming, D. J., J. S. Hastings, T. J. Kane, and D. O. Staiger (2014, March). School Choice, School Quality, and Postsecondary Attainment. *American Economic Review 104*(3), 991–1013.

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.

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.

Gurantz, O., J. Howell, M. Hurwitz, C. Larson, M. Pender, and B. White (2002). Realizing your college potential? Impacts of College Board's RYCP Campaign on Postsecondary Enrollment. *Technical Report Annenberg Institute. Brown University. 43*(40).

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.

Hillman, N. W. (2016). Geography of College Opportunity: The Case of Education Deserts. *American Educational Research Journal 53*(4), 987–1021.

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).

Hurwitz, M. (2011). The impact of legacy status on undergraduate admissions at elite colleges and universities. *Economics of Education Review 30*(3), 480–492.

Hurwitz, M. and J. Smith (2018). Student responsiveness to earnings data in the College Scorecard. *Economic Inquiry 2*(40), 1220–1243.

Hyman, J. (2019). Can light-touch college-going interventions make a difference? Evidence from a statewide experiment in Michigan. *Journal of Policy Analysis and Management.*.

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. *The Quarterly Journal of Economics 131*(3), 1057–1111.

Lochner, L. and A. Monge-Naranjo (2012). Credit Constraints in Education. *Annual Review of Economics 4*(1), 225–256.

Loury, L. D. (2004). Siblings and gender differences in african-american college attendance. *Economics of Education Review 23*(3), 213 – 219.

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

Mulhern, C. (2019a). Beyond teachers: Estimating individual guidance counselors' effects on educational attainment. *Working Paper*.

Mulhern, C. (2019b). Changing college choices with personalized information at scale: Evidence on Naviance. *Working Paper*.

Mwangi, C. A. G. (2015). (Re)Examining the role of family and community in college access and choice: A metasynthesis. *The Review of Higher Education 39*(1), 123–151.

Oymak, C. (2018). High school students' views on who influences their thinking about education and careers. Stats in Brief, National Center for Education Statistics.

Qureshi, J. A. (2018). Additional returns to investing in girls' education: Impact on younger sibling human capital. *The Economic Journal 616*(128), 3285–3319.

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.

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

Wiswall, M. and B. Zafar (2014). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies 82*(2), 791–824.

Table I: Institutional Characteristics of Investigated Countries

	Chile (1)	Croatia (2)	Sweden (3)	US (4)
		A. Countries Characteristics		
Population	17,969,353	4,203,604	9,799,186	320,742,673
Area (km^2)	756,700	56,590	447,430	9,834,000
GDP per Capita	$22,688,01	$23,008.21	$48,436.98	$56,803.47
GDP Growth (2000-2015)	285.60%	227.47%	185.25%	156.34%
GINI Index	47.7	31.1	29.2	41.5
Human Development Index	0.84	0.827	0.929	0.917
Adults w/ Postsecondary Ed.	15.17%	18.30%	34.56%	39.95%
Main Religious Affiliation	Christian (80.1%)	Christian (93.2%)	Christian (62.9%)	Christian (73.9%)
Official Language	Spanish	Croatian	Swedish	English
		B. University System Characteristics		
Colleges	33/60	49/49	35/36	21/3004
Majors	1,423	564	2,421	
Tuition Fees	Yes	Yes	No	Yes
Funding	Student loans and scholarships	Fee waiver when accepting offer[1].	NA	Student loans and scholarships

Notes: The statistics presented in Panel A come from the World Bank (`https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD`), from the United Nations (`http://hdr.undp.org/en/data`) and from the World Population Review (`https://worldpopulationreview.com/countries/religion-by-country/` websites. All the statistics reported in the table correspond to the values observed in 2015. The only exceptions are the share of adults with complete postsecondary education —which we observe in 2011— and religious affiliation . 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. In the United States we only report the total number of colleges in 2015. 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.

[1] 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.

Table II: Summary Statistics

	Chile		Croatia		Sweden		US	
	Younger Siblings Sample	Whole Sample	Younger Siblings Sample	Whole Sample	Younger Siblings Sample	Whole Sample	Younger Siblings Sample	Whole Sample
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Demographic characteristics								
Female	0.522	0.525	0.563	0.567	0.579	0.595	0.530	. 0.533
	(0.500)	(0.499)	(0.496)	(0.495)	(0.493)	(0.490)	(0.500)	(0.500)
Age when applying	19.028	20.059	18.880	19.158	20.589	20.872		
	(1.021)	(3.380)	(0.654)	(0.963)	(2.374)	(2.562)		
Household size[1]	4.632	4.322	2.790	1.925	3.086	2.946	2.250	1.288
	(1.737)	(1.830)	(1.243)	(1.198)	(1.142)	(1.186)	(0.510)	(0.611)
Race: White							0.570	0.543
							(0.490)	(0.497)
B. Socioeconomic characteristics								
High income[2]	0.373	0.113			0.349	0.339	0.19	0.15
	(0.484)	(0.316)			(0.477)	(0.473)	(0.39)	(0.36)
Mid income[2]	0.387	0.286			0.262	0.290	0.27	0.21
	(0.487)	(0.452)			(0.440)	(0.454)	(0.44)	(0.41)
Low income[2]	0.240	0.478			0.389	0.371	0.16	0.23
	(0.427)	(0.500)			(0.488)	(0.483)	(0.41)	(0.42)
Parental ed: 4-year college[3]	0.434	0.207			0.562	0.517	0.650	0.595
	(0.496)	(0.405)			(0.496)	(0.500)	(0.480)	(0.489)
C. Academic characteristics								
High school track: academic[4]	0.905	0.582	0.439	0.416				
	(0.294)	(0.493)	(0.496)	(0.496)				
Takes admission test	0.995	0.864	0.865	0.835	0.679	0.628	0.850	0.963
	(0.072)	(0.343)	(0.342)	(0.372)	(0.467)	(0.483)	(0.360)	(0.191)
High school GPA score	-0.147	-0.757	268.373	265.298	0.673	0.432		
	(1.035)	(1.069)	(65.766)	(66.600)	(0.766)	(0.773)		
Admission test avg. score	-0.322	-0.534	312.800	286.247	0.281	-0.061	987.19	1026.095
	(2.128)	(1.722)	(102.568)	(112.787)	(0.991)	(1.000)	(171.87)	(212.640)
Applicants	140,043	3,889,550	16,721	199,475	301,967	3,822,188	44,191	14,432,122

Notes: The table present summary statistics for Chile, Croatia, Sweden and the United States. Columns (1), (3), (5) and (7) describe individuals in the siblings discontinuity samples used in this paper, while columns (2), (4), (6) and (8) describe all potential applicants. While in Chile, Croatia and the United States "potential applicants"include all students who register for the admission exam, in Sweden the term refers to all students applying to higher education.

[1] In Croatia and in the United States *Household Size* refers only to the number of children in the 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 monthly 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). In the US, low income refers to students from families earning less than $50,000 USD per year. Middle income refers to families with $50,000-10,000 and high income refers to families with incomes above $100,000. In the US, incomes are self-reported by the students and are missing for many students.

[3] While in Chile and Sweden, parental education refers to the maximum level of education reached by any of the applicants' parents, in the United States it refers to the education of the mother.

[4] 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).

Table III: First Stage: Older Siblings' College and Major Enrollment

	Enroll in a 4-year College	**Target College Enrollment**				**Target Major Enrollment**		
	US	CHI	CRO	SWE	US	CHI	CRO	SWE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Older Sibling Admitted = 1	0.030***	0.484***	0.835***	0.201***	0.083***	0.536***	0.826***	0.217***
	(0.011)	(0.006)	(0.010)	(0.003)	(0.007)	(0.004)	(0.007)	(0.002)
Observations	44190	86521	12950	443931	44190	170886	36757	730187
Bandwidth	93.000	12.500	80.000	0.370	93.000	18.000	80.000	0. 510
Counterfactual mean	0.693	0.113	0.002	0.039	0.140	0.118	0.002	0.032

Notes: All the specifications in the table control for a linear polynomial of older siblings' application score centered around target college or target majors admission cutoffs. Older siblings' application year, target major-year and younger siblings' birth year fixed effect are included as controls. Bandwidths corresponds to the ones used for younger siblings' outcome and were computed according to Calonico et al. (2014). In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table IV: Sibling Effects on Total College Enrollment and College Quality in the US

	College type		College quality		Price, location	
	4-year college (1)	2-year college (2)	B.A. completion rate (3)	Peer quality (Z-score) (4)	Net price (000s) (5)	50+ miles from home (6)
			Panel A - All Students			
Older Sibling Enrolls	0.229*	-0.004	0.179**	0.314**	2.287	0.136
	(0.133)	(0.111)	(0.080)	(0.150)	(2.352)	(0.127)
Counterfactual mean	0.38	0.20	0.31	-0.20	8.71	0.22
			Panel B - Uncertain college-goers			
Older Sibling Enrolls	0.522**	0.073	0.466***	0.687***	10.801***	0.459**
	(0.240)	(0.205)	(0.145)	(0.255)	(4.040)	(0.217)
Counterfactual mean	0.09	0.08	0.01	-0.67	0.09	-0.06
			Panel C - Probable college-goers			
Older Sibling Enrolls	0.050	-0.046	0.011	0.095	-2.754	-0.076
	(0.161)	(0.139)	(0.099)	(0.188)	(3.084)	(0.164)
Counterfactual mean	0.57	0.26	0.49	0.09	13.95	0.42

Notes: Heteroskedasticity robust standard errors clustered by oldest sibling's high school are in parentheses (*p-value<0.1 **p-value<0.05 ***p-value<0.01). Each coefficient is an instrumental variables estimate of the impact of an older sibling's enrollment in the target college on younger siblings' college choices, using admissibility as an instrument. Each estimate comes from a local linear regression with a bandwidth of 93 SAT points, a donut hole specification that excludes observations on the threshold, and fixed effects for each combination of older sibling's cohort, younger sibling's cohort, and older sibling's target college. The first row includes all students, while the second and third rows divide the sample into those in the bottom third and top two-thirds of the distribution of predicted four-year college enrollment. College quality is measured by the fraction of students starting at that college who complete a B.A. anywhere within six years (column 3) and the mean standardized PSAT score of students at that college (column 4). Also listed below each coefficient is the predicted value of the outcome for control compliers.

Table V: Sibling Effects on Applications to and Enrollment in Older Sibling's Target College and Major

	College Choice			Major Choice		
	Applies 1st	Applies	Enrolls	Applies 1st	Applies	Enrolls
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A - Chile						
2SLS	0.067***	0.076***	0.038***	0.012***	0.023***	0.006***
	(0.012)	(0.014)	(0.011)	(0.003)	(0.005)	(0.002)
Reduced form	0.033***	0.037***	0.018***	0.006**	0.012***	0.003***
	(0.006)	(0.007)	(0.005)	(0.001)	(0.003)	(0.001)
Observations	86521	86521	86521	170886	170886	170886
Counterfactual mean	0.222	0.447	0.132	0.019	. 0.064	0.012
Bandwidth	12.500	12.500	12.500	18.000	18.000	18.000
Kleibergen-Paap Wald F statistic	5576.25	5576.25	5576.25	14765.19	14765.19	14765.19
Panel B - Croatia						
2SLS	0.075***	0.109***	0.084***	0.015***	0.036***	0.013**
	(0.019)	(0.019)	(0.018)	(0.004)	(0.009)	(0.004)
Reduced form	0.063***	0.091***	0.070***	0.012***	0.030***	0.011**
	(0.016)	(0.016)	(0.015)	(0.004)	(0.007)	(0.003)
Observations	12950	12950	12950	36757	36757	36757
Counterfactual mean	0.293	0.523	0.253	0.022	0.111	0.017
Bandwidth	80.000	80.000	80.000	80.000	80.000	80.000
Kleibergen-Paap Wald F statistic	6459.562	6459.562	6459.562	14512.301	14512.301	14512.301
Panel C - Sweden						
2SLS	0.149***	0.153***	0.064***	0.020***	0.029***	0.004**
	(0.009)	(0.013)	(0.006)	(0.003)	(0.005)	(0.001)
Reduced form	0.030***	0.031***	0.013***	0.004***	0.006***	0.001**
	(0.002)	(0.003	(0.001)	(0.001)	(0.001)	(0.000)
Observations	443931	443931	443931	730187	730187	730187
Counterfactual mean	0.088	0.193	0.034	0.011	0.047	0.004
Bandwidth	0.370	0.370	0.370	0.510	0.510	0.510
Kleibergen-Paap Wald F statistic	6140.057	6140.057	6140.057	10817.599	10817.599	10817.599
Panel D - United States						
2SLS		0.270**	0.172***			
		(0.105)	(0.053)			
Reduced form		0.022**	0.014***			
		(0.009)	(0.005)			
Observations		44190	44190			
Counterfactual mean		0.106	0.006			
Bandwidth		93	93			
Kleibergen-Paap Wald F statistic		129.730	129.730			

Notes: All the specifications in the table control for a linear 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. 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 VI: Sibling Effects on Applications to and Enrollment in Older Sibling's Target Major-College by Younger Siblings' Eligibility

	Applies 1st (1)	Applies (2)	Enrolls (3)
	Panel A - Chile		
Older sibling enrolls	0.012***	0.017***	0.003
	(0.003)	(0.005)	(0.002)
Older sibling enrolls × Eligible = 1	0.001	0.028***	0.015***
	(0.003)	(0.006)	(0.003)
Observations	161,123	161,123	161,123
Kleibergen-Paap Wald F statistic	6767.580	6767.580	6767.580
	Panel B - Croatia		
Older sibling enrolls	0.009*	0.024**	-0.005
	(0.005)	(0.012)	(0.004)
Older sibling enrolls × Eligible = 1	0.011**	0.024**	0.029***
	(0.005)	(0.011)	(0.004)
Observations	33,823	33,823	33,823
Kleibergen-Paap Wald F statistic	6770.281	6770.281	6770.281
	Panel C - Sweden		
Older sibling enrolls	0.033***	0.046***	0.005**
	(0.005)	(0.010)	(0.003)
Older sibling enrolls × Eligible = 1	0.011**	0.010	0.014***
	(0.004)	(0.009)	(0.003)
Observations	292,970	292,970	292,970
Kleibergen-Paap Wald F statistic	3270.581	3270.581	3270.581

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table V. In addition, they include an interaction between the treatment and a proxy of younger siblings' eligibility for their older siblings' target major. The eligibility proxy is also included in these specifications as a control. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table VII: Sibling Effects on College and Major Choice by Age Difference and Gender

	Enroll in a 4-year College	College Choice				Major Choice		
	US	CHI	CRO	SWE	US	CHI	CRO	SWE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Age Difference $\geq$ 5								
Older Sibling Enrolls = 1	0.216	0.092***	0.109***	0.162***	0.262**	0.025***	0.039***	0.035***
	(0.132)	(0.015)	(0.020)	(0.013)	(0.104)	(0.005)	(0.009)	(0.005)
Interaction	0.130	-0.035***	0.000	-0.030**	0.076	-0.004	-0.018	-0.015***
	(0.141)	(0.011)	(0.026)	(0.011)	(0.101)	(0.004)	(0.013)	(0.004)
Observations	44190	86364	12950	444203	44190	170570	36756	732025
Kleibergen-Paap Wald F statistic	64.970	2767.580	3230.667	2975.652	64.970	7330.470	7225.706	5255.957
Panel B: Same Gender = 1 - Whole Sample								
Older Sibling Enrolls = 1	0.3064**	0.070***	0.114***	0.143***	0.294***	0.017***	0.026**	0.025***
	(0.138)	(0.016)	(0.022)	(0.014)	(0.108)	(0.005)	(0.009)	(0.006)
Interaction	-0.149**	0.011	-0.007	0.011	-0.049	0.011***	0.023*	0.008*
	(0.072)	(0.012)	(0.020)	(0.011)	(0.054)	(0.004)	(0.009)	(0.004)
Observations	44190	86521	12950	444203	44190	170886	36757	732025
Kleibergen-Paap Wald F statistic	64.780	2788.470	3229.534	3075.133	64.780	7383.02	7220.184	5419.139
Panel C: Same Gender = 1 - Older Brothers								
Older Sibling Enrolls = 1		0.078***	0.124***	0.139***		0.016***	0.025	0.013
		(0.022)	(0.033)	(0.024)		(0.007)	(0.015)	(0.009)
Interaction		0.001	0.001	0.040*		0.020***	0.044**	0.045***
		(0.017)	(0.032)	(0.019)		(0.006)	(0.016)	(0.007)
Observations		39919	5008	160086		81072	14203	281549
Kleibergen-Paap Wald F statistic		1435.860	1405.970	1330.244		4150.72	4025.070	2717.178
Panel D: Same Gender = 1 - Older Sisters								
Older Sibling Enrolls = 1		0.079***	0.098**	0.154***		0.018**	0.031*	0.036***
		(0.024)	(0.031)	(0.019)		(0.007)	(0.013)	(0.008)
Interaction		-0.004	-0.027	-0.003		-0.000	0.007	-0.019**
		(0.018)	(0.027)	(0.014)		(0.006)	(0.012)	(0.006)
Observations		44222	7545	273981		87895	22239	438419
Kleibergen-Paap Wald F statistic		1223.530	1651.529	1484.510		7383.02	3662.675	2441.736

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table V. In addition, they include an interaction between the treatment and a dummy variable that indicates if siblings are 5 or more years apart (Panel A) or between the treatment and a dummy variable that indicates if siblings are of the same gender (Panel B). Panel C and Panel D do something similar but while Panel C focus only on siblings pairs in which the older one is male, Panel D looks at cases in which the older one is female. The dummy variables are also included in the specifications as controls. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table VIII: Sibling Effects on College and Major Choice by Older Sibling's Target Option Quality

	Enrollment in any 4-year College	College Choice				Major Choice		
	US	CHI	CRO	SWE	US	CHI	CRO	SWE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Peers' Quality (Standardize performance in SAT or HS)								
Older Sibling Enrolls = 1	-0.586	0.063***	-0.010	0.120***	0.319	0.021*	0.038	0.019**
	(0.426)	(0.018)	(0.058)	(0.015)	(0.310)	(0.009)	(0.025)	(0.006)
Interaction	2.138***	0.016**	0.027*	0.036***	-0.128	0.002	-0.001	0.012***
	(0.743)	(0.007)	(0.013)	(0.008)	(0.541)	(0.002)	(0.005)	(0.003)
Observations	44190	86521	10693	444203	44190	136364	34510	732023
Kleibergen-Paap Wald F statistic	17.194	1856.76	2598.965	2577.150	17.194	4914.155	6833.719	4508.761
Panel B: First Year Dropout Rate								
Older Sibling Enrolls = 1		0.106***		0.118***		0.029***		0.015**
		(0.016)		(0.013)		(0.009)		(0.005)
Interaction		-0.248***		-0.126**		-0.035		-0.028
		(0.073)		(0.044)		(0.040)		(0.015)
Observations		84076		320107		167804		535714
Kleibergen-Paap Wald F statistic		2652.770		2678.503		7387.420		5465.470
Panel C: Graduation Rate								
Older Sibling Enrolls = 1	-0.297	-0.054		0.065***	0.054	0.015		0.017**
	(0.208)	(0.039)		(0.018)	(0.070)	(0.013)		(0.006)
Interaction	0.448	0.300***		0.100***	-0.077	0.014		0.016**
	(0.305)	(0.078)		(0.024)	(0.010)	(0.026)		(0.008)
Observations	44190	85697		314434	44190	169557		509583
Kleibergen-Paap Wald F statistic	1.269	2579.88		2844.24	1.269	6697.58		5421.94
Panel D: Graduates' Earnings (Standardized annual earnings)								
Older Sibling Enrolls = 1	0.229*	0.075***		0.110***	0.270**	0.025***		0.019***
	(0.133)	(0.014)		(0.016)	(0.105)	(0.005)		(0.006)
Interaction	0.004	0.002		0.010	-0.013*	0.002**		0.010***
	(0.007)	(0.003)		(0.008)	(0.007)	(0.001)		(0.003)
Observations	44190	81112		218552	44190	160627		358644
Kleibergen-Paap Wald F statistic	129.787	2121.960		1380.629	129.787	5764.60		2462.490

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table V. In addition, they include an interaction between the treatment and an index of older siblings' major or college quality: peers' academic performance (Panel A), first year dropout rates (Panel B), graduation rates (Panel C) and graduates' earnings (Panel D). The quality index is also included in the specifications as a control. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table IX: Sibling Effects on College and Major Choice by Older Sibling's Dropout

	Enrollment in any 4-year College	College Choice			Major Choice	
	US	CHI	SWE	US	CHI	SWE
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A - Applications						
Older Sibling Enrolls = 1	0.362***	0.128***	0.212***	0.317***	0.033***	0.046***
	(0.103)	(0.016)	(0.019)	(0.085)	(0.006)	(0.008)
Interaction	-0.511***	-0.135***	-0.139***	-0.119**	-0.036***	-0.037***
	(0.072)	(0.014)	(0.017)	(0.060)	(0.004)	(0.007)
Observations	44190	64247	444203	44190	130917	732025
Kleibergen-Paap Wald F statistic	66.658	2936.770	1945.998		8078.940	3413.123
Panel B - Enrollment						
Older Sibling Enrolls = 1	0.487***	0.075***	0.088***	0.179***	0.010***	0.007***
	(0.112)	(0.012)	(0.009)	(0.044)	(0.003)	(0.002)
Interaction	-0.648***	-0.092***	-0.055***	-0.017	-0.011***	-0.005***
	(0.077)	(0.010)	(0.008)	(0.030)	(0.002)	(0.002)
Observations	44190	64247	444203	44190	130917	732025
Kleibergen-Paap Wald F statistic		2936.770	1945.998		8078.940	3413.123

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table V. In addition, they include an interaction between the treatment and a dummy variable that takes value 1 if the older sibling drops out. This dummy variable is also included in the specifications as a control. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table X: Sibling Effects on College and Major Choice by Socioeconomic Status (Bottom $\leq$ 40% of Income Distribution)

	College				**Major**			
	Applies		**Enrolls**		**Applies**		**Enrolls**	
	CHI	SWE	CHI	SWE	CHI	SWE	CHI	SWE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Older Sibling Enrolls = 1	0.081***	0.158***	0.036***	0.056***	0.024***	0.034***	0.004	0.003***
	(0.201)	(0.013)	(0.015)	(0.006)	(0.006)	(0.005)	(0.003)	(0.001)
Interaction	-0.008	0.014	0.002	0.007	-0.002	0.004	0.003	0.003***
	(0.016)	(0.012)	(0.011)	(0.005)	(0.005)	(0.005)	(0.002)	(0.001)
Observations	86521	472037	86521	472037	170886	749483	170886	749483
Kleibergen-Paap Wald F statistic	4921.048	3673.514	4921.048	3673.514	7225.228	6244.625	7225.228	6244.625

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table V. In addition, they include an interaction between the treatment and a dummy variable that takes value 1 if siblings belong to the bottom 40% of the income distribution. The main effect of the interaction is also included in the specifications as a control. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table XI: Sibling Effects on College Choice by Exposure to Older Sibling's Target College

	Applies		**Enrolls**	
	CHI	SWE	CHI	SWE
	(1)	(2)	(3)	(4)
Older Sibling Enrolls = 1	0.099***	0.143***	0.054***	0.035***
	(0.018)	(0.014)	(0.012)	(0.006)
Interaction	-0.317***	-0.448***	-0.192	0.042
	(0.138)	(0.064)	(0.137)	(0.046)
Observations	84911	390877	84911	390877
Avg. exposure in the sample	0.075	0.247	0.075	0.038
Kleibergen-Paap Wald F statistic	2775.363	3281.852	2775.363	3281.852

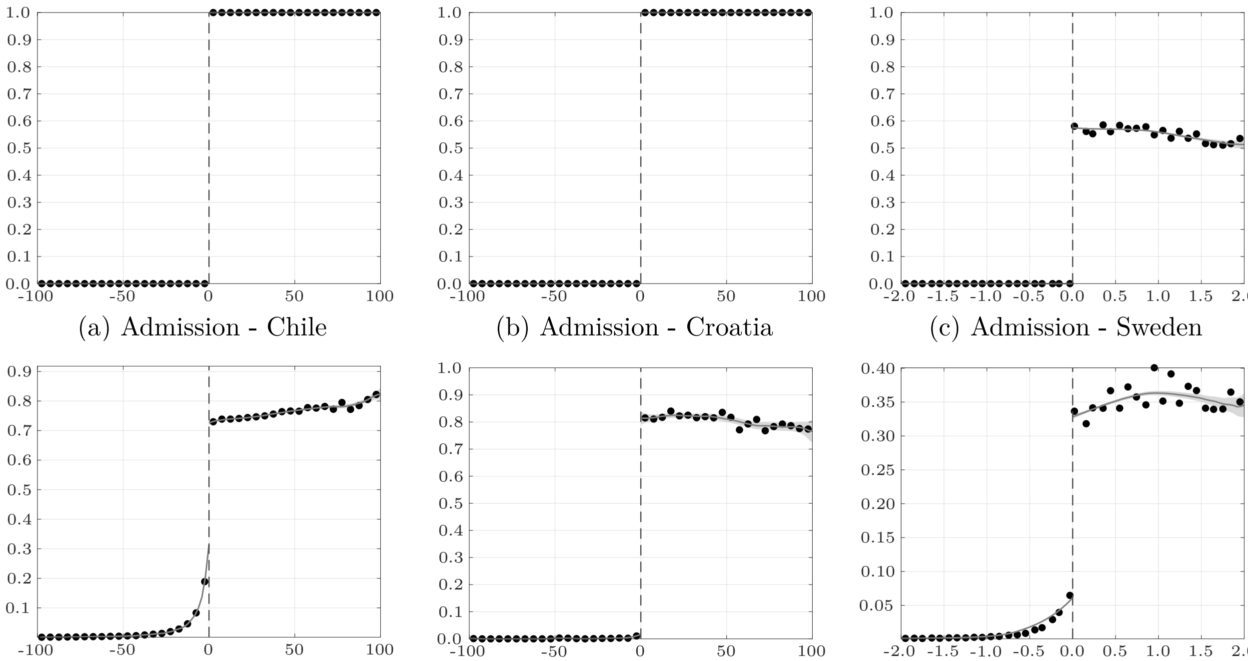
Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table V. In addition, they include an interaction between the treatment and the share of individuals from the younger sibling high school going to the older sibling's target college one year before the younger sibling completes high school. The main effect is also included in the specifications as a control. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table XII: Sibling Effects in Target Major-College on Academic Performance

	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	0.001	-0.002	-0.009	-0.011
	(0.001)	(0.005)	(0.019)	(0.011)
Observations	170,886	170,886	170,886	170,886
Counterfactual mean	0.995	0.930	-0.170	-0.240
F-statistic	14765.190	14765.190	14765.190	14765.190
	Panel B - Croatia			
Older sibling enrolls	-0.013	-0.008	-0.043	-0.054
	(0.017)	(0.009)	(0.045)	(0.045)
Observations	12,443	36,757	12,443	12,443
Counterfactual mean	0.810	0.866	-0.030	-0.035
Kleibergen-Paap Wald F statistic	4498.481	14512.30	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
Counterfactual mean	0.484	0.577	0.219	0.051
Kleibergen-Paap Wald F statistic	10838.800	10838.800	9529.889	6498.021
	Panel D - United States			
Older sibling enrolls	0.073	0.159		46.9
	(0.096)	(0.125)		(43.0)
Observations	44,190	44,190		37,554
Counterfactual mean	0.830	0.545		951 .000
Kleibergen-Paap Wald F statistic	129.730	129.730		120.758

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major (Chile, Croatia and Sweden) or college (United States) 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), and average performance on the admission exam (column 4). For the United States column (2) looks at differences in the total number of applications submitted instead of at the probability of submitting at least 1. While in Chile, Croatia and the United States we only observe applications to college, in Sweden we also observe applications to other higher education institutions. The analyses for Chile, Croatia and Sweden focus on the Major Sample. This means that in this case, marginal admission or rejection from their target major, changes the major-college combination but not necessarily 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 Tables C.IV and C.IV. In parenthesis, standard errors clustered at family level in Chile, Croatia and Sweden and at the older sibling's high school in the United States. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

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



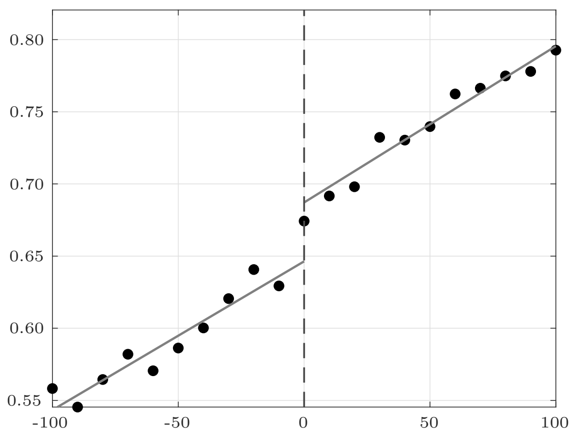
(a) Admission - Chile (b) Admission - Croatia (c) Admission - Sweden

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

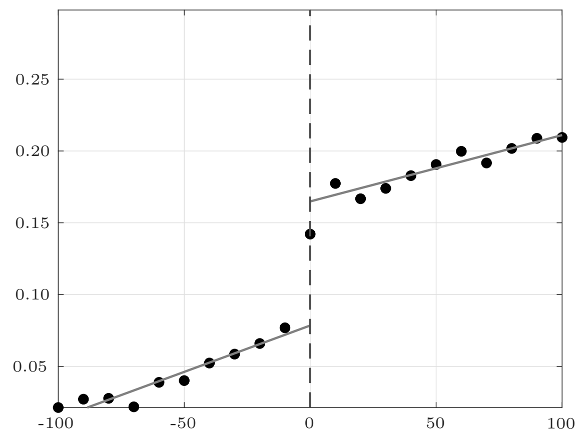
---

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. Gray lines and the shadows in the back of them represent local linear polynomials and 95% confidence intervals. Black dots represent sample means of the dependent variable at different values of older siblings' own application score.

Figure II: Older Siblings' Enrollment Probability in 4-year Colleges and in Target College at the Admission Cutoff (First Stage)



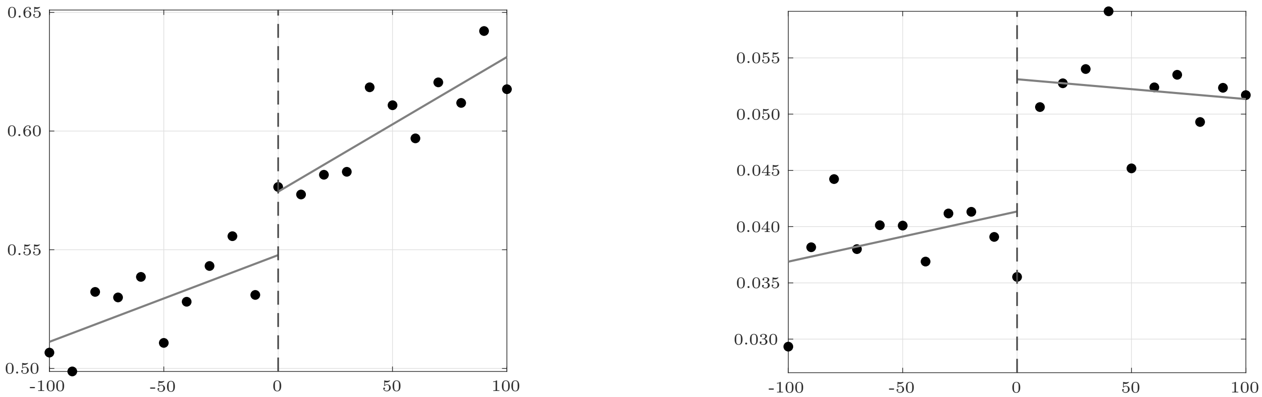
(a) 4-year Colleges



(b) Target College

This figure illustrates older siblings' enrollment probability in any 4-year college and in their target college around the admission cutoffs in the United States. Panel (a) illustrates the change in enrollment in any 4-year college and panel (b) the change in enrollment in the target college. Gray lines represent local linear polynomials. Black dots represent sample means of the dependent variable at different values of older siblings' SAT score.

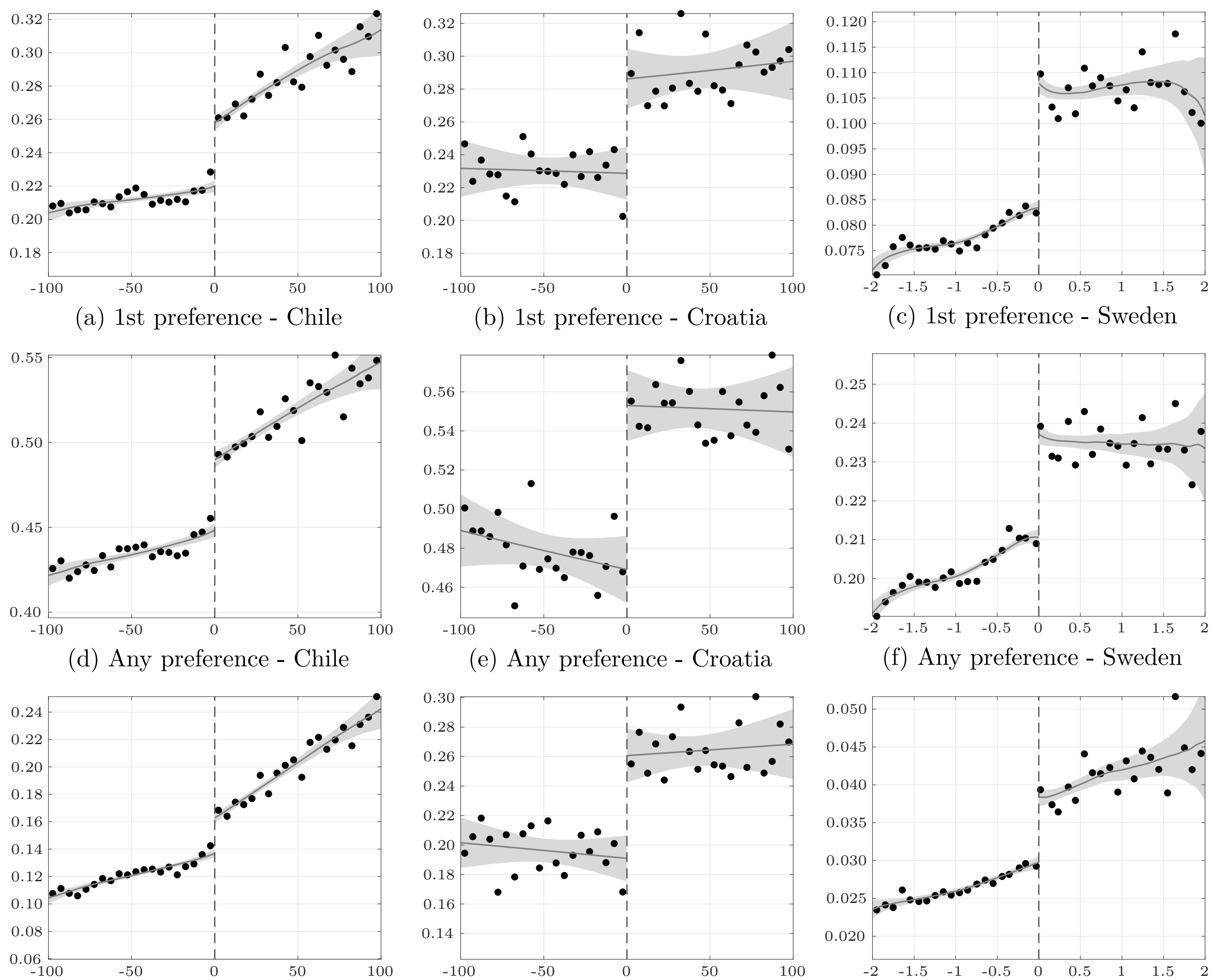
Figure III: Probabilities of Enrolling in any 4-year College and in the Older Sibling's Target College



(a) 4-year Colleges (b) Target College

This figure illustrates the probability that younger siblings enroll in any 4-year college and in target college of their older siblings in the United States. Panel (a) focus on 4-year colleges while panel (b) on the target college of the older sibling. Gray lines correspond to local polynomials of degree 1. Black dots represent sample means of the dependent variable at different values of older sibling's admission score.

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



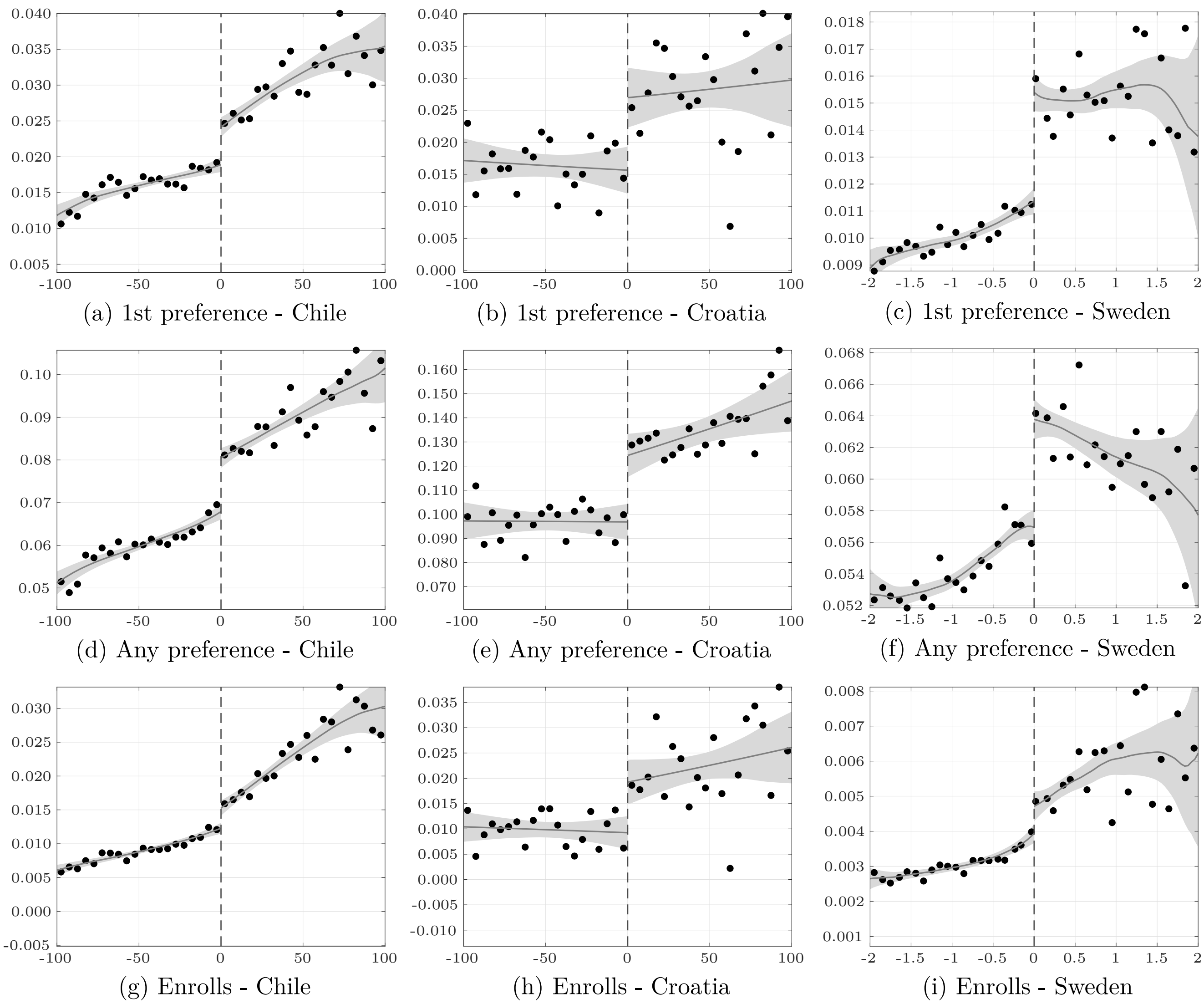
(a) 1st preference - Chile (b) 1st preference - Croatia (c) 1st preference - Sweden

(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 (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. Gray lines and the shadows in the back of them correspond to local polynomials of degree 1 and 95% confidence intervals. Black dots represent sample means of the dependent variable at different values of older sibling's admission score.

Figure V: 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. Gray lines and the shadows in the back of them correspond to local polynomials of degree 1 and 95% confidence intervals. Black dots represent sample means of the dependent variable at different values of older sibling's admission score.

A Institutions: Further Details

A.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).[26] 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).[27] Students take the PSU in December, at the end of the Chilean academic year, but they typically need to register before mid-August.[28] As of 2006, all public and voucher school graduates are eligible for a fee waiver that makes the PSU free for them.[29]

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 I, 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.[30] Although

[26]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.

[27]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.

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

[29]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.

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

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.[31] 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 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 I illustrates how the admission to a major translates into enrollment.

A.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.[32] 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.

[31] 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.

[32] 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.

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 I).

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 I).[33]

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.

A.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.[34]

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

[33] 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.

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

ranking is then used to determine their admission status.[35] 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 I).

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.[36]

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 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

[35] 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.

[36] 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.

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.[37]

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.

A.4 College Admission System in the United States

In the United States, each college is free to set their own admission criteria and there is no centralized admission system in place. However, when selecting students the majority of the colleges take into account applicants' scores in a university admission exam (i.e. PSAT, SAT, or Advanced Placement exams).

During the period that we study the SAT was offered seven times a year and could be taken as often as the college application timeline allowed.[38]. As in the case of the admission exams used in the other countries, the SAT has different sections and in terms of application is common for colleges to consider students' "superscores"(Goodman et al., 2020). The "superscores"are the sum of a student's maximum math and maximum critical reading scores, regardless of whether those scores occurred on the same attempt. In order to apply to college students need to submit their SAT scores and any other application material requested by the institutions in which they are interested.

Since colleges are free to consider other variables to select their students, this admission system does not necessarily generate sharp admission cutoffs. Thus, we use our data to detect which colleges admit students in part on the basis of minimum SAT thresholds not known to applicants. Many colleges use minimum SAT scores as one criterion for determining admissions decisions, so that meeting or exceeding a college's threshold typically increases a student's probability of being admitted to that college. We focus on thresholds hidden from applicants because publicly known

[37] 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.

[38] Retakes cost roughly $40, with low income students eligible for fee waivers for up to two attempts

thresholds induce some students to retake the SAT until their scores meet the thresholds (Goodman et al., 2017). Such behavior creates endogenous sorting around the threshold that invalidates the regression discontinuity design. Conversely, students can not react endogenously to cutoffs about which they are unaware.

We search for such thresholds using the only child sample, which is independent of the sibling sample that we use to estimate spillover effects. This avoids the potentially spurious findings that might be generated by searching for thresholds using the same observations and outcomes used to estimate treatment effects. For each college and year, we find all only children who sent their SAT scores to that college, generating an indicator for a student enrolling in that college within one year of graduating high school. We then search for discontinuities by SAT score in a given college's enrollment rate among its applicants. We limit our search to the 526 colleges that received SAT scores from at least 1,000 students each year, in order to minimize the possibility of false positives arising from small samples.

To search for discontinuities, we estimate local linear regression discontinuity models at each SAT score that might represent a potential threshold.[39] We define the set of potential thresholds for each college as the set of SAT scores in the 5th to 50th percentiles of the applicant distribution for the specified college and year. Colleges are unlikely to set minimum thresholds lower or higher in their applicant distributions. For every potential threshold T and all applicants i to college c in year y, we run regressions of the form:

$$Enrolled_{icy} = \beta_0 + \beta_1 1(SAT_i \geq T_{cy}) + \beta_2(SAT_i - T_{cy}) + \beta_3 1(SAT_i \geq T_{cy}) \times (SAT_i - T_{cy}) + \varepsilon_{icy} \quad (2)$$

We define the running variable using students' SAT "superscores", the most frequently used form of scores considered by college admissions offices. To minimize false positives driven by specification error, we use a bandwidth of 60 SAT points, within which enrollment graphs look generally linear.

The coefficient of interest β_1 estimates the magnitude of any potential discontinuity in enrollment rates at the given threshold T. To further limit potential false positives, we consider as disconti-

[39] Our approach is similar to that used in Andrews et al. (2017).

nuities only those instances where discontinuities in enrollment rates exceed five percentage points and where we reject the null hypothesis of no discontinuity with $p > 0.0001$. Finally, we discard any colleges where thresholds are detected in fewer than five years, given that most colleges that use minimum SAT scores in admissions are unlikely to change that policy from year to year. We also discard a small number of colleges for which we find evidence from admissions websites that the detected thresholds are publicly known.

This procedure yields 21 threshold-using colleges, which we refer to as "target" colleges both for brevity and because of older siblings' interest in attending these institutions. These target colleges are largely public institutions (16 public, 5 private) with an average enrollment of over 10,000 full-time equivalent students, and are located in eight different East coast states. The median SAT threshold across years for these colleges ranges from 720 to 1060, with students relatively widely distributed across these colleges and thresholds. These target colleges' average graduation rate is 63 percent and the average PSAT z-score of their students is 0.27. They have average net prices of $12,500, making them $4,000 less expensive per year than the average college attended by students in our full sample.

B 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 4.3, a fuzzy RD can be thought as an IV. In what follows, and for ease of notation, we drop time and individual indices t, i, τ and focus our analysis on a specific 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:

$$\frac{\hat{E}[Y_u|adm_u = 1] - \hat{E}[Y_u|adm_u = 0]}{\hat{E}[O_u|adm_u = 1] - \hat{E}[O_u|adm_u = 0]} = \frac{\hat{E}[Y_u(1) - Y_u(0)]}{\hat{E}[O_u(1) - O_u(0)]}. \tag{3}$$

The numerator of equation 3, after applying law of iterated expectations, becomes:

$$\hat{E}[Y_u(1) - Y_u(0)] = \tag{4}$$

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

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 3 becomes:

$$\hat{E}[O_u(1) - O_u(0)] = P(O_u(1) = 1, O_u(0) = 0). \tag{5}$$

Taken together, 4 and 5 imply:

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

$\square$

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

C Robustness Checks

This section investigates if the identification assumptions of our empirical strategy are satisfied. We start by checking if there is any evidence of manipulation of the running variables. 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 using a second order polynomial of the running variable, when using a triangular kernel and when allowing the slope of the running variable to vary by major-college and year. Finally, we end this section by showing that there are no extensive margin responses of younger siblings (i.e. increases in total enrollment) in Chile, Croatia and Sweden that could explain our findings.

C.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. Something similar occurs in the United States, where the cutoffs that we exploit are hidden. 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. As discussed in Section 2, in Sweden the admission exam is voluntary and institutions select their students using either their high school GPA or their scores in the admission exam. Both of these measures are not fully continuous and in addition, the admission exam suffered some transformations in 2013. Therefore, to investigate manipulation of these scores we present independent histograms for each one of these variables. Figure C.I illustrates the density of the relevant running variables for all the countries that we study. These histograms do not show any evidence of manipulation.

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 hundreds of them in our samples and studying them independently would be impractical.

C.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 that we observe in the 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 C.II summarizes the result of this analysis for Chile, Croatia and Sweden. It plots the estimated discontinuities at the cutoff and their 95% confidence intervals. To estimate these discontinuities at the cutoff we use the same specification described in the main body of the paper. This means that we control for a linear polynomial of the running variable and allow 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.

Figure C.III presents similar results to the United States. Here instead of presenting the estimated jump at the cutoff we illustrate how the variable on the y-axis evolves with the running variable. None of the potential confounders studied in this figure seem to jump at the cutoff.

C.3 Different Bandwidths

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

Figures C.IV and C.V show how the estimated coefficients change when reducing the bandwidth used in the estimations for Chile, Croatia and Sweden. Although the standard errors increase as the sample size gets smaller, the coefficients remain stable. Figure C.VI does the same exercise for the outcomes that we investigate in the United States. In this case, the coefficients remain also very stable when using smaller bandwidth; when we increase it, the coefficients begin to drop what suggests a non-linear relationship between the running variable and the outcomes outside the 100 SAT points window used in our analyses.

C.4 Placebo Exercises

This setting allows us to perform two types of placebo exercises.

First, in Figures C.VIII and C.VII we show that only at the real cutoff we observe a discontinuity on younger siblings outcomes This is not surprising since these fake cutoffs do not generate any increase in older siblings' admission. In the case of the United States we do not perfectly observe the actual cutoffs; instead, we estimate them from the data. Figure C.IX present results for an exercise similar to the one we just discussed. As before there do not seem to be discontinuities around placebo cutoffs. Finding some discontinuities at points that are very close to the actual cutoffs is just a consequence of not observing them and using instead estimates.

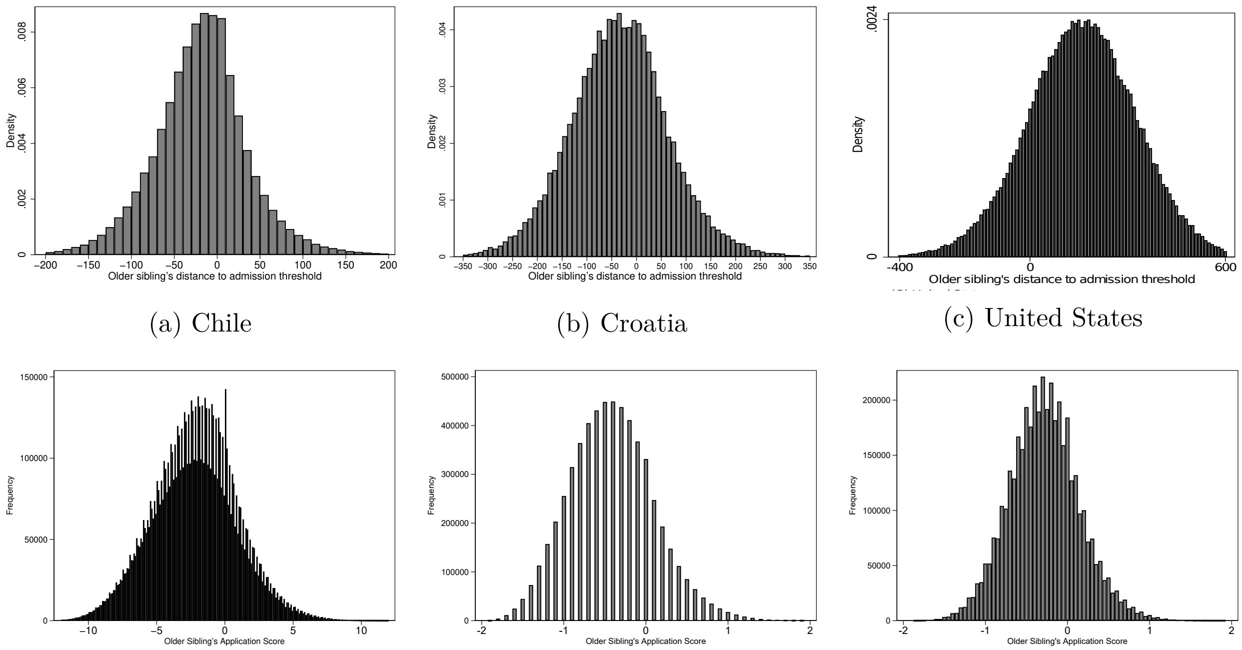
Second, in Figures C.X and C.XI we study if younger siblings' admission to their target college or major affect the application decisions of their older siblings in Chile, Croatia and Sweden. In Figure C.XII we do something similar for the United States. Since younger siblings apply to college after their older siblings, being marginally admitted or rejected from a major or college should not affect what happens with their older siblings. These figures show that this is indeed the case. Even though when looking at the placebo on college choice in Sweden we find some discontinuities at the cutoff, their size is considerably smaller than the ones documented in the main body of the paper.

C.5 Alternative Specifications and Total Enrollment

Tables C.III and C.IV show that our results are robust to using a second degree polynomial of the running variable and also to use a triangular instead of a uniform kernel. We complement these analyses by illustrating in Figures C.XIII and C.XIV the reduced forms of our outcomes when using a second degree polynomial of the running variable. In addition, in Tables C.V and C.VI we show that our results are robust to allowing the running variable to have cutoff-major specific slopes. In Tables C.VII and C.VIII we present results of specifications in which older siblings' target $\times$ counterfactual major fixed effects are used.

Finally, Table C.IX 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 in Chile, Croatia and Sweden. We do not find evidence of younger siblings extensive margin responses in any of these countries. Thus, our findings do not seem to be driven by a general increase on younger siblings enrollment.

Figure C.I: Density of Older Siblings' Admission Exam and High School GPA at the Target College-Major Admission Cutoff

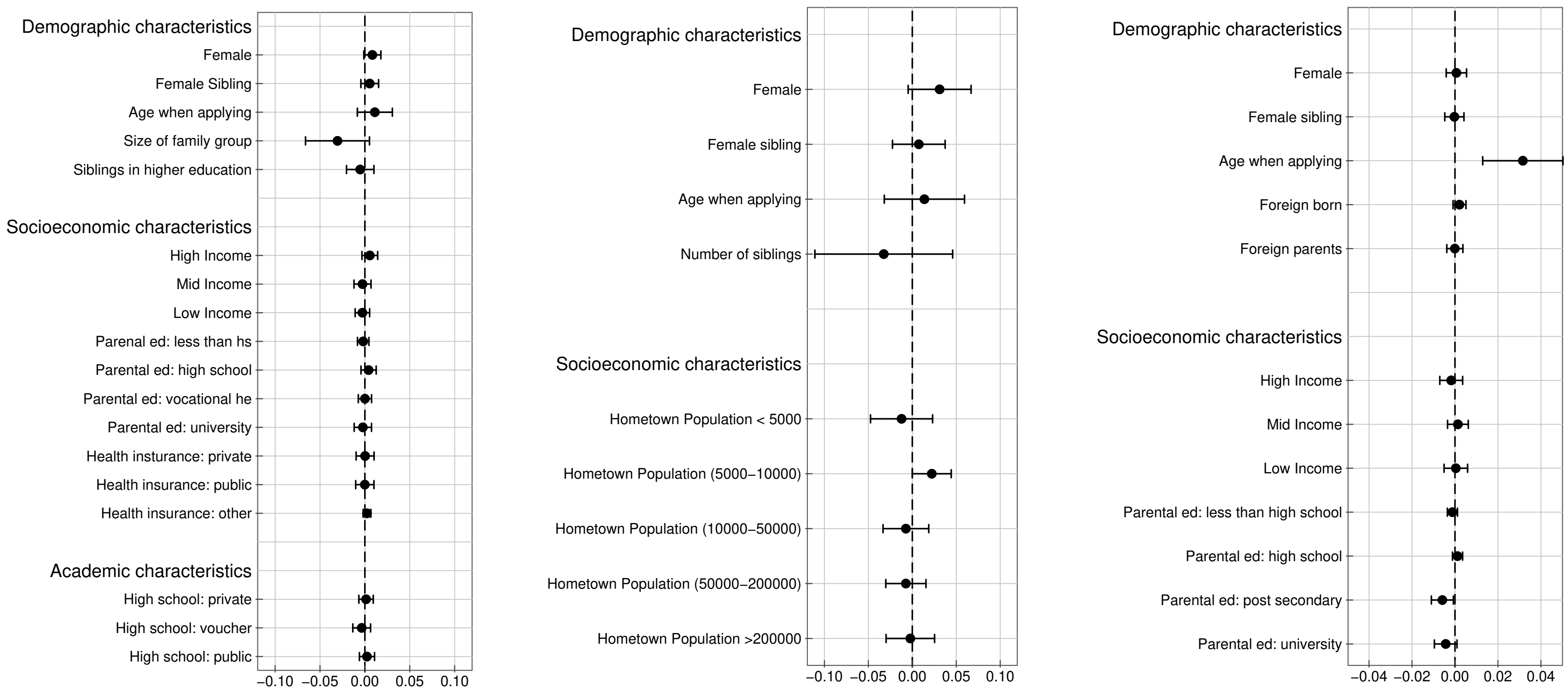


(a) Chile (b) Croatia (c) United States

(d) High School GPA - Sweden (e) Exam Before 2013 - Sweden (f) Exam After 2013 - Swedem

These histograms illustrate distributions of older siblings' admission exam and high school GPA around admission cutoffs for Chile, Croatia, Sweden and the United States. Panels (a), (b) and (c) illustrate the distribution of admission exam scores in Chile, Croatia and the United States respectively. Panel (d) illustrates the distribution of high school GPA in Sweden and panel (b) corresponds to the distribution of admission exam scores until 2013 in Sweden. In 2013 there was a structural change in the admission exam, including its scale . Panel (c) presents the distribution of scores after 2013.

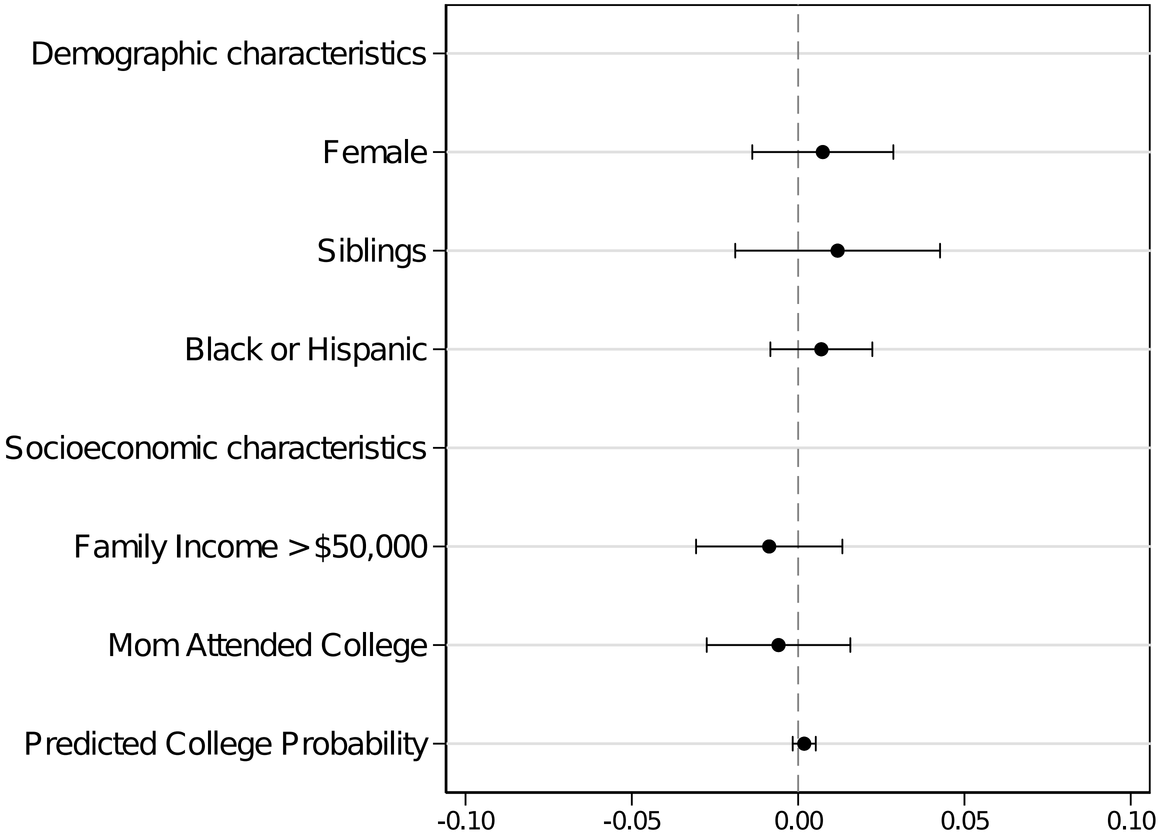
Figure C.II: Disconitnuities in other Covariates at the Cutoff



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

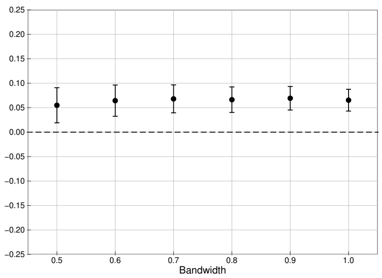
This figure illustrates the estimated jumps at the cutoff for a vector of socioeconomic and demographic characteristics. These estimates come from parametric specifications that control for a linear polynomial of the running variable. As the main specifications, these also include major-college-year 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 C.III: Disconitnuities in other Covariates at the Cutoff (United States)

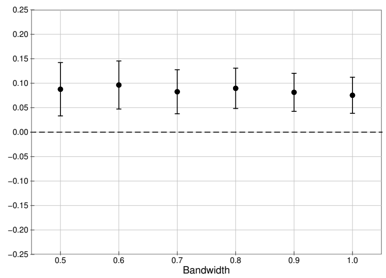


This figure illustrates how demographic and socioeconomic characteristics vary at the admissions cutoff in the United States. The range of the running variable corresponds to the bandwidth used in our main specifications. The points represent the estimated coefficient , while the lines represent 95% confidence intervals.

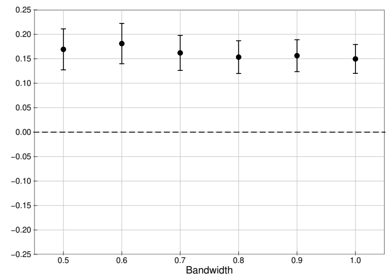
Figure C.IV: Probabilities of Applying and Enrolling in Older Sibling's Target College - Different Bandwidths



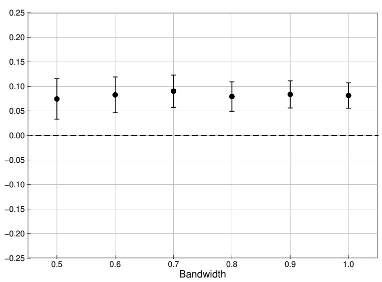
(a) 1st preference - Chile



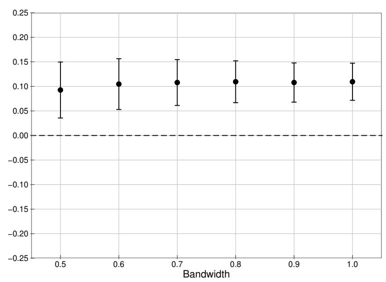
(b) 1st preference - Croatia



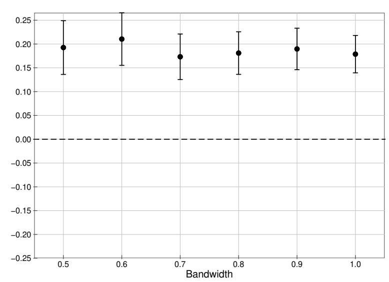
(c) 1st preference - Sweden



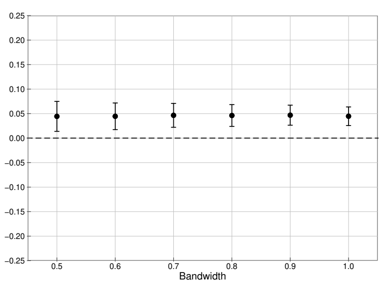
(d) Any preference - Chile



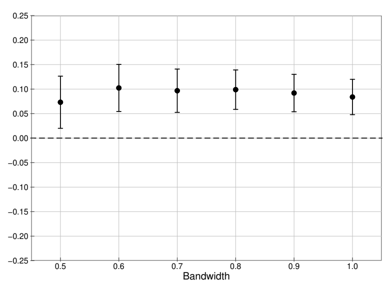
(e) Any preference - Croatia



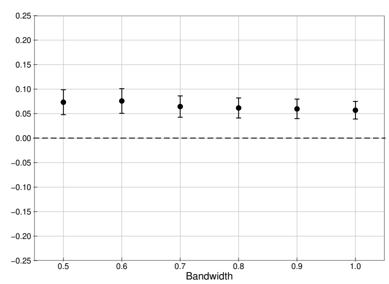
(f) Any preference - Sweden



(g) Enrolls - Chile



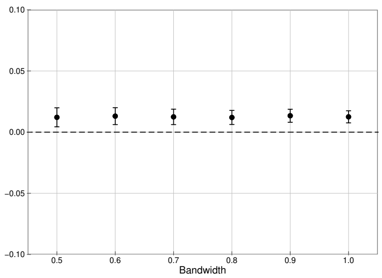
(h) Enrolls - Croatia



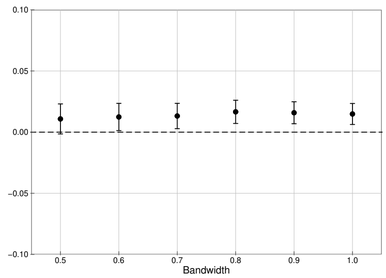
(i) Enrolls - Sweden

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). The points illustrate the estimated effect, and the lines 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 specifications that control for a linear polynomial of the running variable.

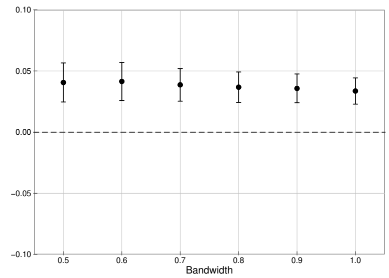
Figure C.V: Probabilities of Applying and Enrolling in Older Sibling's Target Major-College - Different Bandwidths



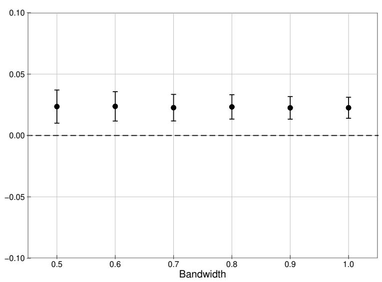
(a) 1st preference - Chile



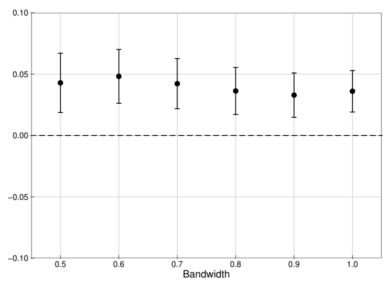
(b) 1st preference - Croatia



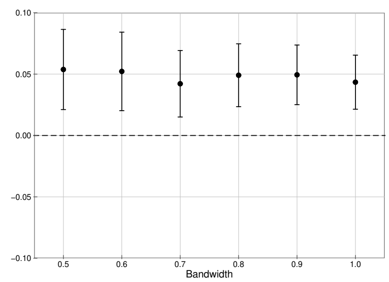
(c) 1st preference - Sweden



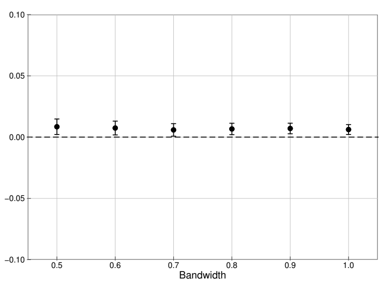
(d) Any preference - Chile



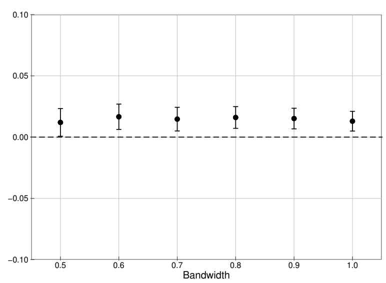
(e) Any preference - Croatia



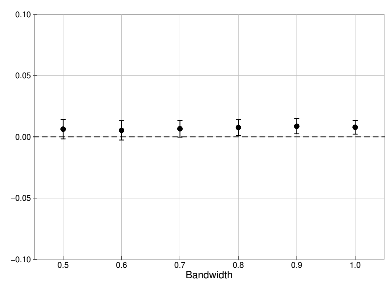
(f) Any preference - Sweden



(g) Enrolls - Chile



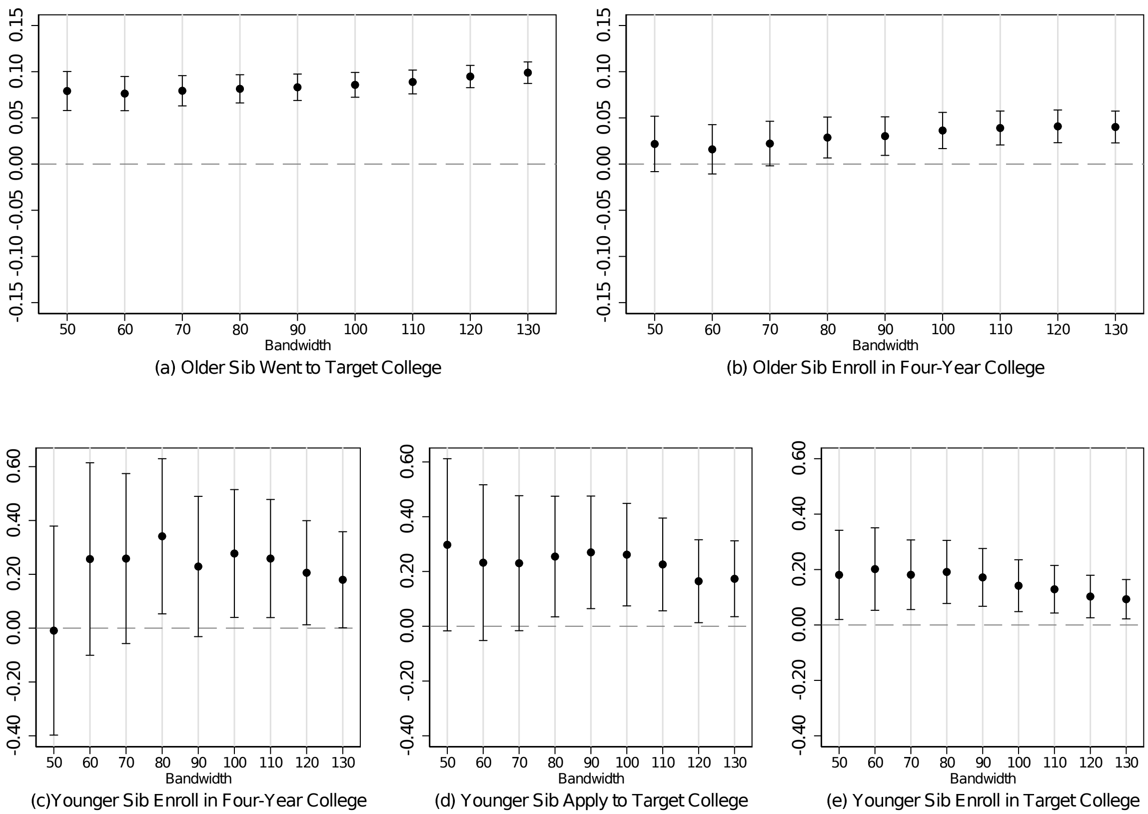
(h) Enrolls - Croatia



(i) Enrolls - Sweden

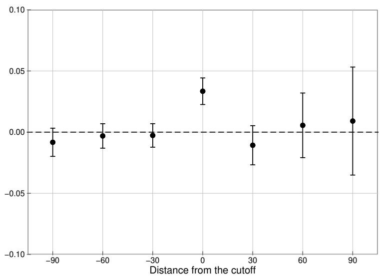
This figure illustrates how being admitted to a specific program changes younger siblings' probabilities of applying and enrolling in the same major. The x-axis corresponds to different bandwidths used to build these figures, chosen as multiples of the optimal bandwidths computed following Calonico et al. (2014). The points illustrate the estimated effect, and the lines 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 specifications that control for a linear polynomial of the running variable.

Figure C.VI: Probabilities of Enrolling in any 4-years College and in Older Sibling's Target College - Different Bandwidths (United States)

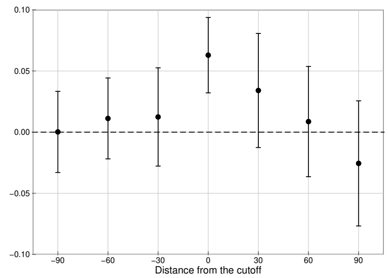


This figure illustrates how an older sibling's marginal enrollment in her target college changes a younger sibling's probability of enrolling in any 4-year college and in the older sibling's target college The x-axis corresponds to different bandwidths used to build these figures. The dots represent the estimated effect, and the lines denote the 95% confidence intervals. The coefficients and their confidence intervals come from specifications that control for a linear polynomial of the running variable.

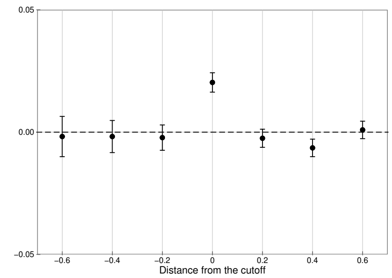
Figure C.VII: Placebo Cutoffs - Probabilities of Applying and Enrolling in Older Sibling's Target College



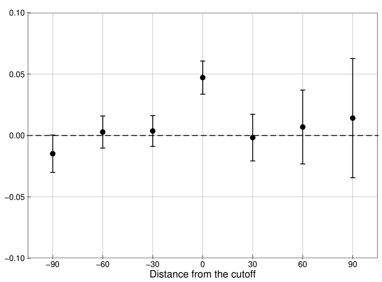
(a) 1st preference - Chile



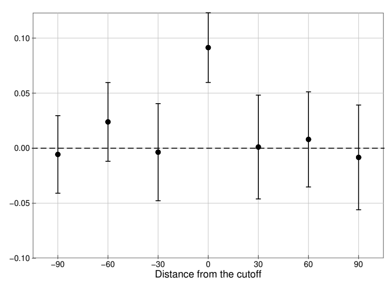
(b) 1st preference - Croatia



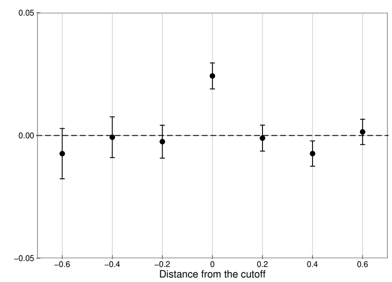
(c) 1st preference - Sweden



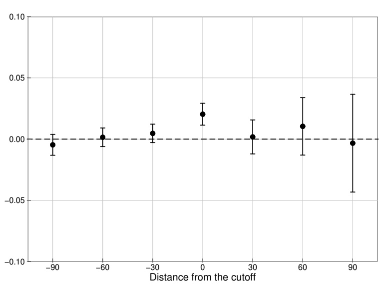
(d) Any preference - Chile



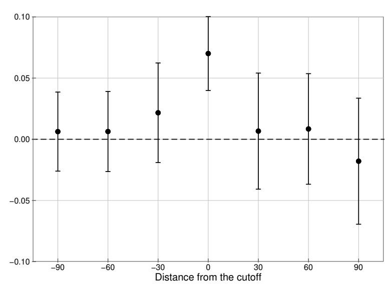
(e) Any preference - Croatia



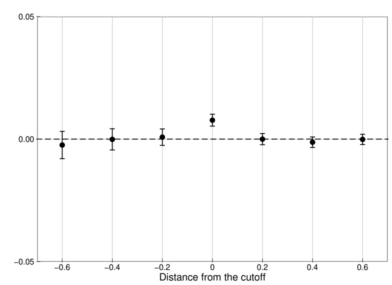
(f) Any preference - Sweden



(g) Enrolls - Chile



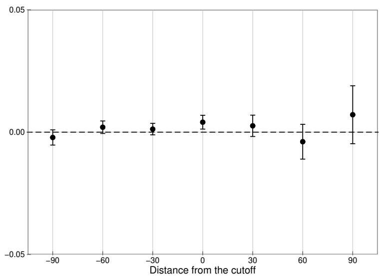
(h) Enrolls - Croatia



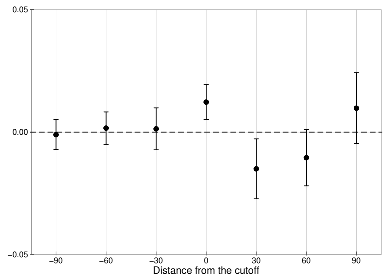
(i) Enrolls - Sweden

This figure illustrates the results of a placebo exercise that investigates if effects similar to the ones documented in figure IV 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. Black points illustrate estimated effect, and the lines 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.

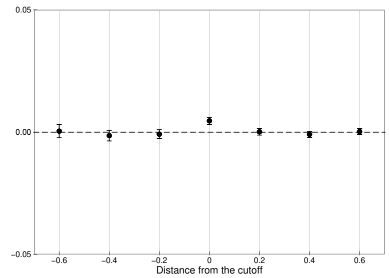
Figure C.VIII: Placebo Cutoffs - Probabilities of Applying and Enrolling in Older Sibling's Target Major-College



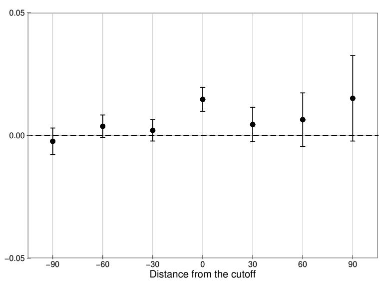
(a) 1st preference - Chile



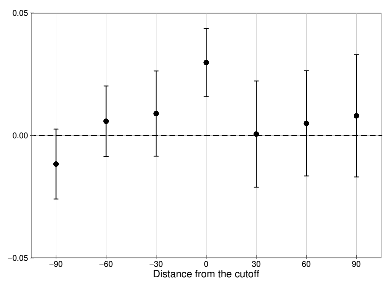
(b) 1st preference - Croatia



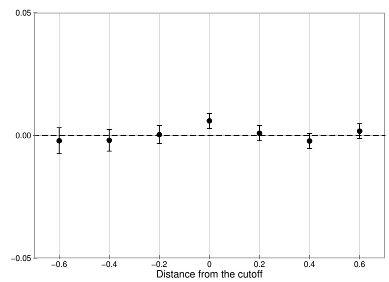
(c) 1st preference - Sweden



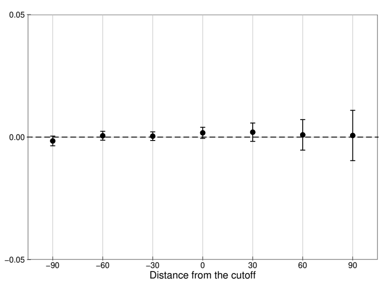
(d) Any preference - Chile



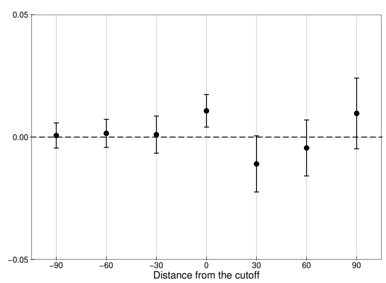
(e) Any preference - Croatia



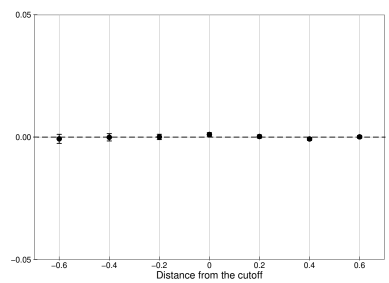
(f) Any preference - Sweden



(g) Enrolls - Chile



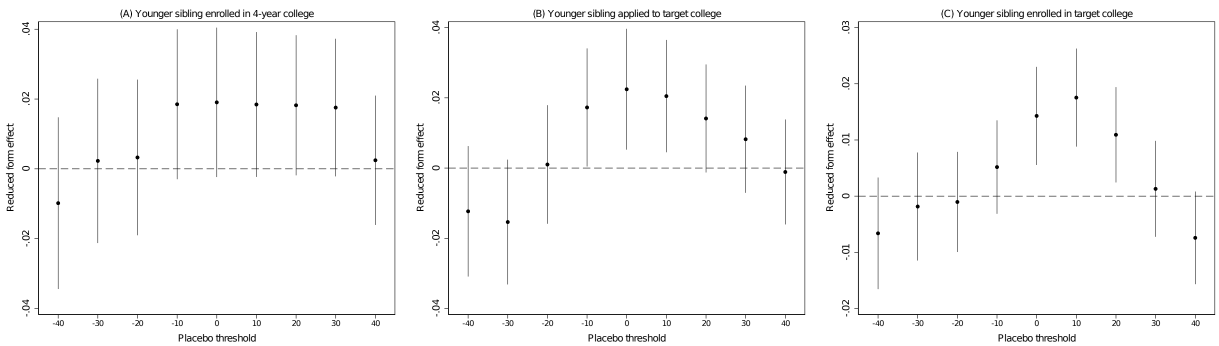
(h) Enrolls - Croatia



(i) Enrolls - Sweden

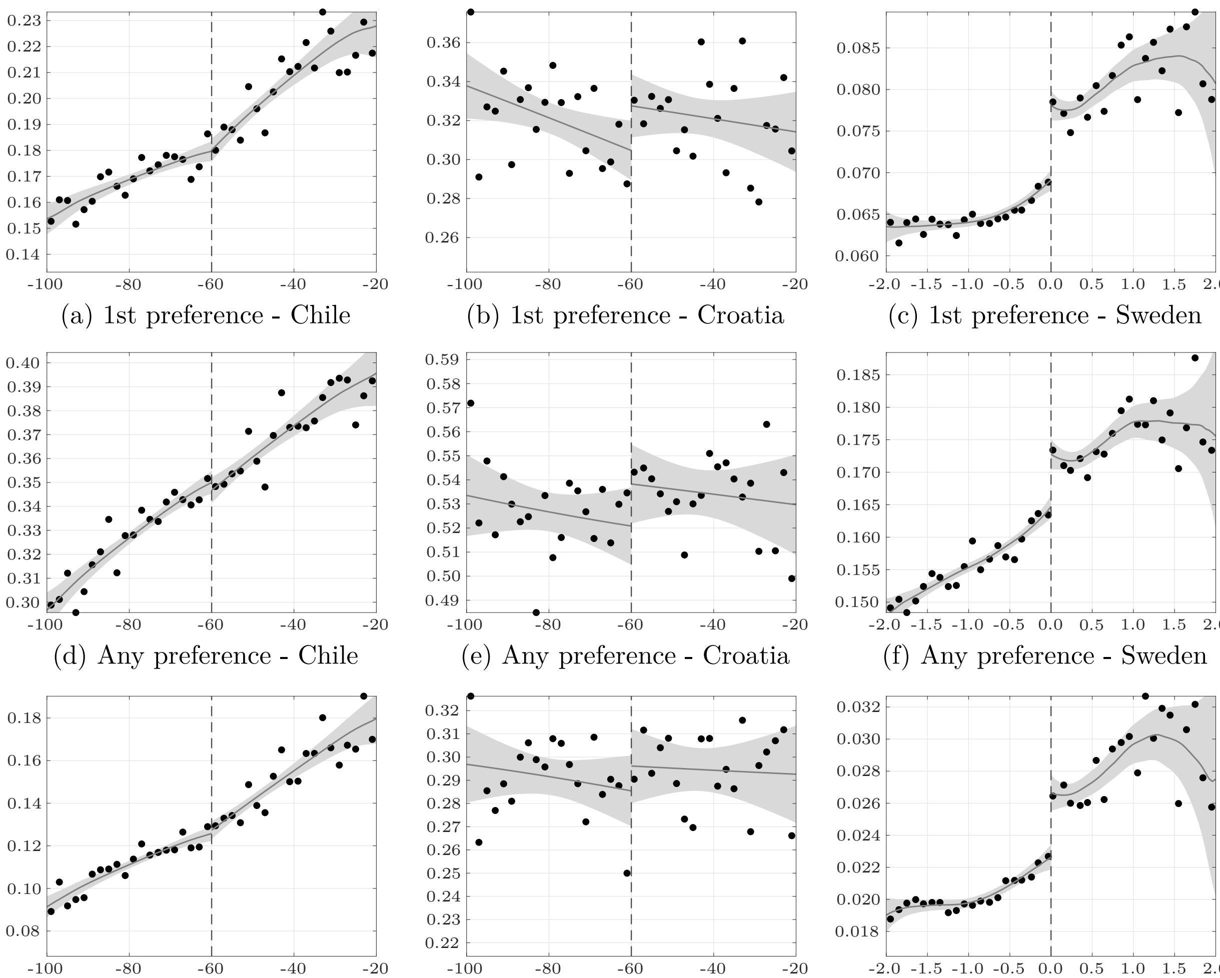
This figure illustrates the results of a placebo exercise that investigates if effects similar to the ones documented in figure V arise at different values of the running variable. Therefore, the x-axis corresponds to different (hypothetical) values of cutoffs - 0 corresponds to the actual cutoff used 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. Black points illustrate estimated effect, and the lines 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 C.IX: Placebo Cutoffs - Probability of Enrolling in any 4-year College and Applying or Enrolling in Older Sibling's Target College (United States)



This figure illustrates the results of a placebo exercise that investigates if effects similar to the ones documented in the main body of the paper arise at different values of the running variable. Therefore, the x-axis corresponds to different (hypothetical) values of cutoffs and 0 corresponds to the actual cutoff. The other values correspond to points where older siblings' probability of being admitted to their target major is continuous. The black dots represent the estimated effect, and the lines denote the 95% confidence intervals.

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



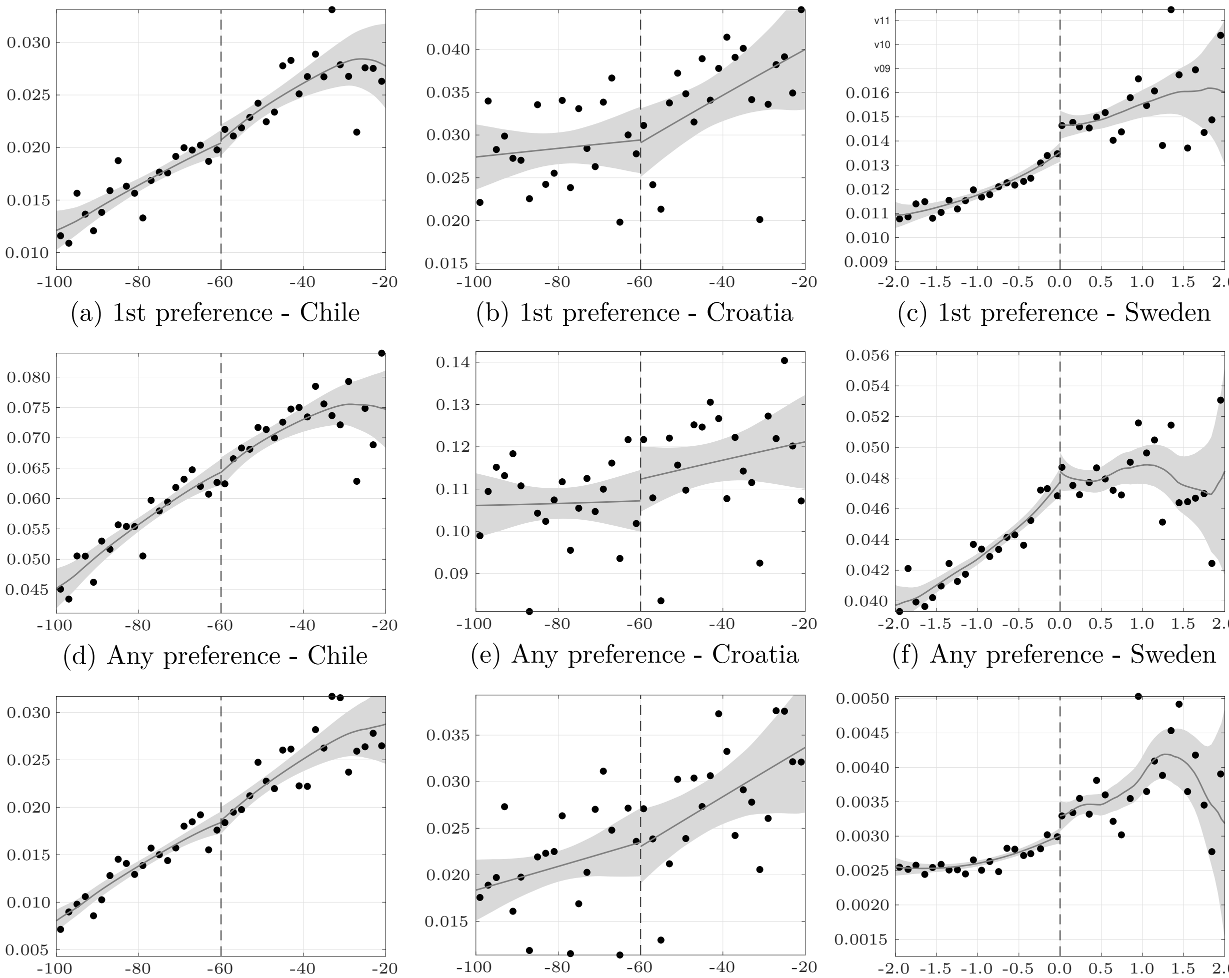
(a) 1st preference - Chile (b) 1st preference - Croatia (c) 1st preference - Sweden

(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. Gray lines and the shadows in the back of them correspond to local polynomials of degree 1 and 95% confidence intervals. Black dots represent sample means of the dependent variable for different values of the running variable.

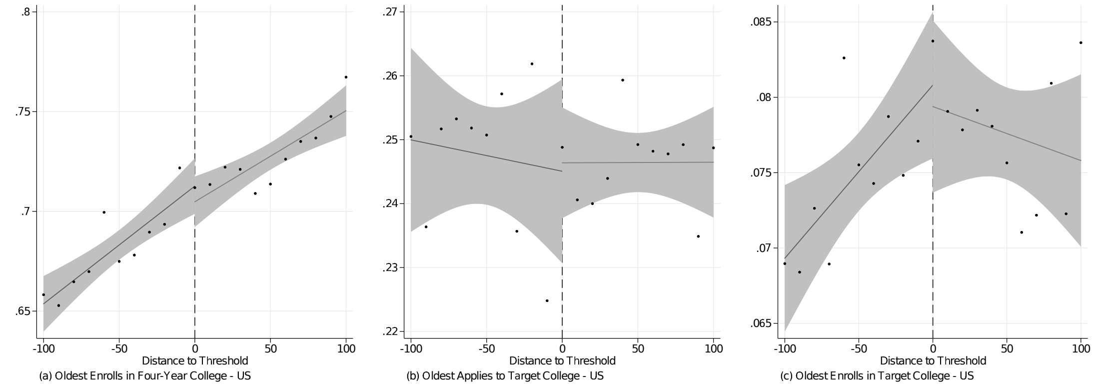
Figure C.XI: Placebo - Probabilities of Applying and Enrolling in Younger Sibling's Target Major-College



(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

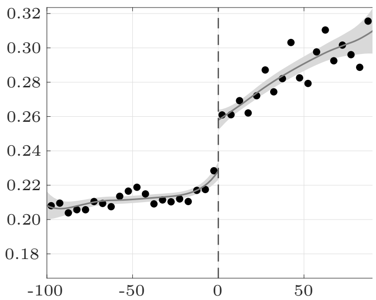
This figure illustrates a placebo exercise that investigates if younger siblings marginal admission to a specific major-college affects the college-major to which older siblings apply to and enroll in. Gray lines and the shadows in the back of them correspond to local polynomials of degree 1 and 95%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

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

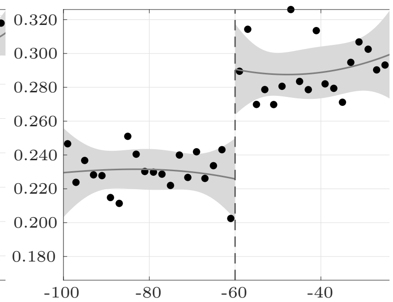


This figure illustrates a placebo exercise that investigates if younger siblings marginal admission to their target college affects the college choices of their older siblings. Gray lines and the shadows in the back of them correspond to local polynomials of degree 1 and 95% confidence intervals. Black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

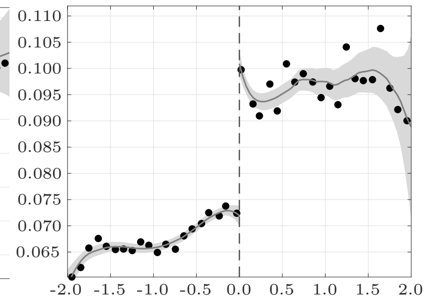
Figure C.XIII: Probabilities of Applying and Enrolling in Older Sibling's Target College (Polynomial of degree 2)



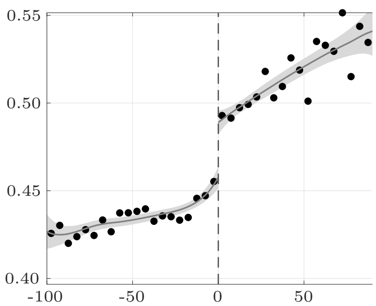
(a) 1st preference - Chile



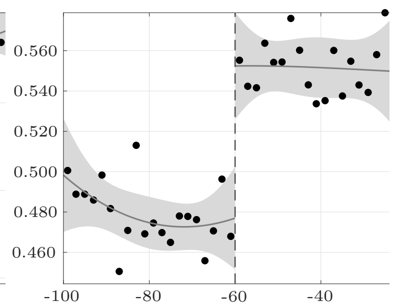
(b) 1st preference - Croatia



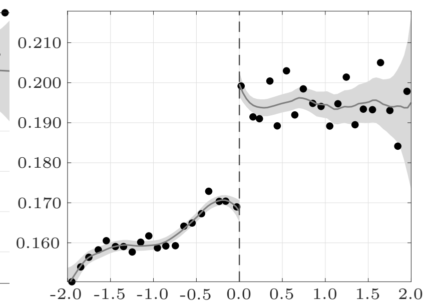
(c) 1st preference - Sweden



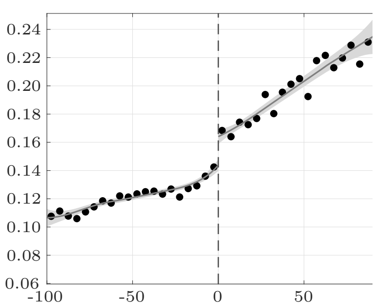
(d) Any preference - Chile



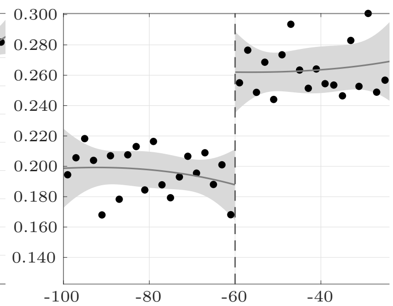
(e) Any preference - Croatia



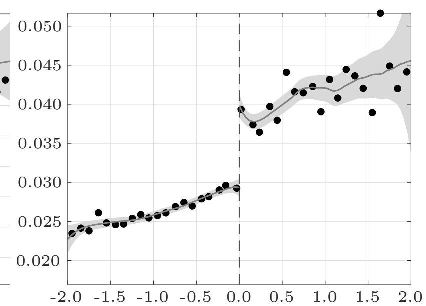
(f) Any preference - Sweden



(g) Enrolls - Chile



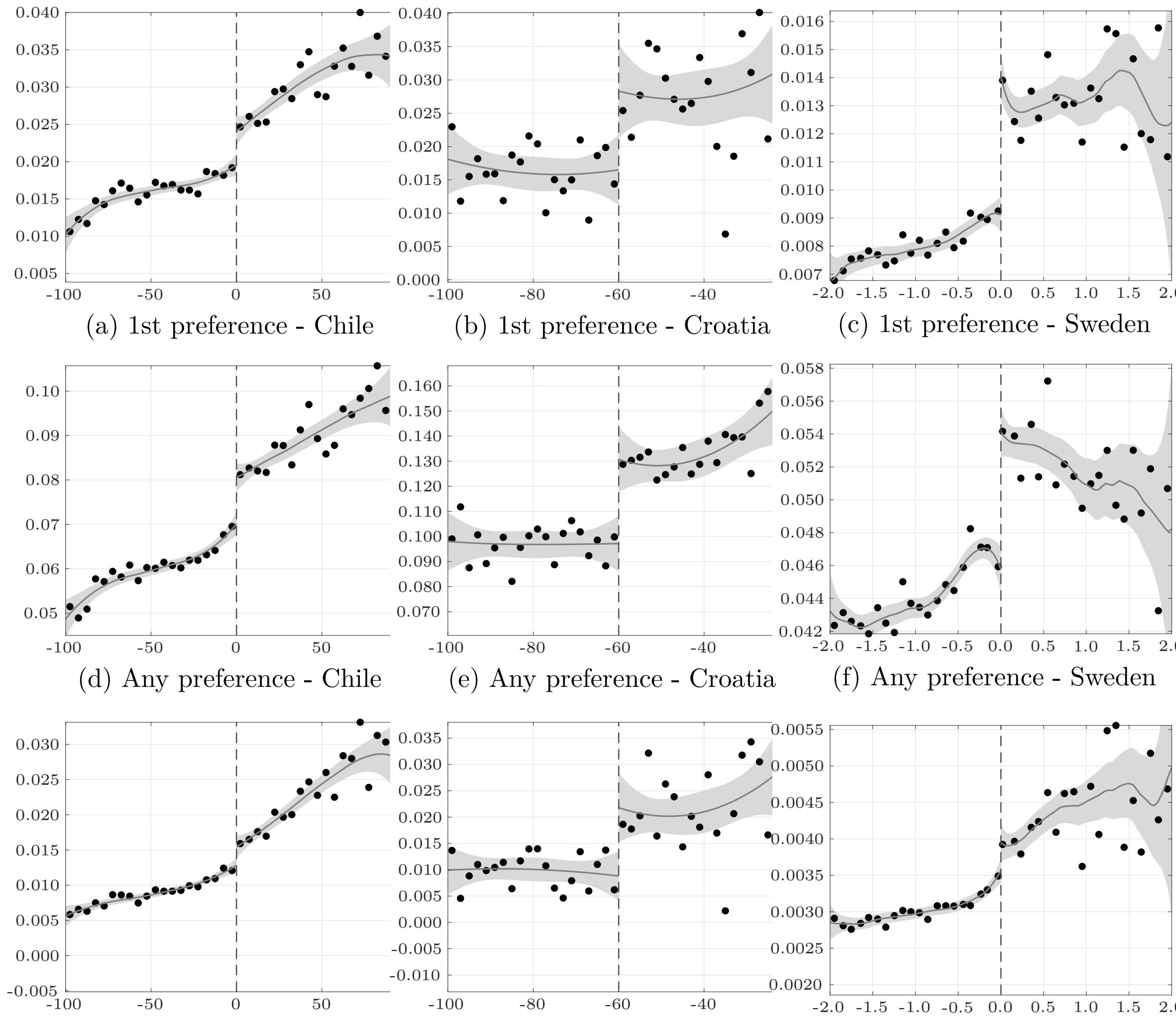
(h) Enrolls - Croatia



(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. Gray lines and the shadows in the back of them correspond to local polynomials of degree 2 and 95% confidence intervals. Black dots represent sample means of the dependent variable at different values of the older sibling's admission score.

Figure C.XIV: 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. Gray lines and the shadows in the back of them correspond to local polynomials of degree 1 and 95% confidence intervals. Black dots represent sample means of the dependent variable at different values of the older sibling's admission score.

Table C.I: Characteristics of Older Siblings' Chosen Colleges

	College type		College quality		Price, location	
	4-year college (1)	2-year college (2)	B.A. completion rate (3)	Peer quality (Z-score) (4)	Net price (000s) (5)	50+ miles from home (6)
All students	0.364***	-0.281**	0.238***	0.439***	2.693	0.300**
	(0.125)	(0.115)	(0.069)	(0.114)	(2.172)	(0.137)
Counterfactual outcome	0.64	0.28	0.43	-0.06	10.92	0.26
Uncertain college-goers	0.369	-0.458**	0.321***	0.631***	1.894	0.310
	(0.227)	(0.217)	(0.121)	(0.206)	(3.511)	(0.235)
Counterfactual outcome	0.63	0.46	0.31	-0.38	11.54	0.24
Probable college-goers	0.275*	-0.122	0.149*	0.296**	2.106	0.248
	(0.152)	(0.137)	(0.085)	(0.140)	(2.849)	(0.171)
Counterfactual outcome	0.73	0.12	0.54	0.15	11.62	0.31

Notes: Heteroskedasticity robust standard errors clustered by oldest sibling's high school are in parentheses (* p<.10 ** p<.05 *** p<.01). Each coefficient is an instrumental variables estimate of the impact of an older sibling's enrollment in the target college on their own college choices, using admissibility as an instrument. Each estimate comes from a local linear regression with a bandwidth of 100 SAT points, a donut hole specification that excludes observations on the threshold, and fixed effects for each combination of older sibling's cohort, younger sibling's cohort, and older sibling's target college. The first row includes all students, while the second and third rows divide the sample into those in the bottom third and top two-thirds of the distribution of predicted four-year college enrollment. College quality is measured by the fraction of students starting at that college who complete a B.A. anywhere within six years (column 3) and the mean standardized PSAT score of students at that college (column 4). Also listed below each coefficient is the predicted value of the outcome for control compliers.

Table C.II: Robustness Checks of Younger Siblings' College Choices

	Baseline specification (1)	Including covariates (2)	Eliminating donut hole (3)	Quadratic Polynomial (4)	Triangular Kernel (5)	Different Slope at each Cutoff (6)
Panel A - All students						
Enrolled in 4-year college	0.229*	0.185	0.117	0.248	0.229	0.283
	(0.133)	(0.126)	(0.164)	(0.245)	(0.152)	(0.214)
Applied to target college	0.270**	0.251**	0.349***	0.157	0.227*	0.202
	(0.105)	(0.104)	(0.129)	(0.191)	(0.119)	(0.148)
Enrolled in target college	0.172***	0.167***	0.278***	0.175*	0.171***	0.177**
	(0.053)	(0.053)	(0.073)	(0.100)	(0.061)	(0.087)
Panel B - Uncertain college-goers						
Enrolled in 4-year college	0.522**	0.531**	0.562*	0.293	0.403	0.546
	(0.240)	(0.236)	(0.313)	(0.419)	(0.251)	(0.324)
Applied to target college	0.263	0.241	0.460*	-0.126	0.112	0.015
	(0.181)	(0.178)	(0.235)	(0.326)	(0.193)	(0.242)
Enrolled in target college	0.262***	0.262***	0.427***	0.339*	0.273***	0.217**
	(0.098)	(0.098)	(0.145)	(0.186)	(0.104)	(0.112)

Notes: Heteroskedasticity robust standard errors clustered by oldest sibling's high school are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$). Each coefficient is an instrumental variables estimate of the impact of an older sibling's enrollment in the target college on younger siblings' college choices, using admissibility as an instrument. Each estimate comes from a local linear regression that includes fixed effects for each combination of older sibling's cohort, younger sibling's cohort, and the target college to which the older sibling applied. Column 1 uses a bandwidth of 90 SAT points and a donut hole specification that exclude observations on the threshold itself. Column 2 adds to that regression covariates, including gender, race, income and parental education. Column 3 eliminates the donut hole, including observations on the threshold itself. Column 4 uses a bandwidth of 90 SAT points, a donut hole specification and it controls for a quadratic polynomial of the older sibling's distance to the threshold. Column 5 uses a bandwidth of 90 SAT points, a donut hole and triangular kernel specification. Column 6 allows for the bandwidth and slope at each admission cutoff to vary. Panel A includes all students, while panel B includes those in the bottom third of the distribution of predicted four-year college enrollment.

Table C.III: Sibling Effects on Applications to and Enrollment in Older Sibling's Target College

	Applies 1st		**Applies**		**Enrolls**	
	P1	P2	P1	P2	P1	P2
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.067***	0.060***	0.076***	0.068***	0.038***	0.031**
	(0.012)	(0.015)	(0.014)	(0.017)	(0.011)	(0.013)
Reduced form	0.033***	0.027***	0.037***	0.031***	0.018***	0.014**
	(0.006)	(0.007)	(0.007)	(0.008)	(0.005)	(0.006)
First stage	0.484***	0.455***	0.484***	0.455***	0.484***	0.455***
	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)
2SLS (Triangular Kernel)	0.069***	0.067***	0.079***	0.075***	0.042***	0.038***
	(0.014)	(0.016)	(0.016)	(0.019)	(0.012)	(0.010)
Observations	86521	136868	86521	136868	86521	136868
Counterfactual mean	0.225	0.222	0.450	0.446	0.136	0.132
Bandwidth	12.500	20.500	12.500	20.500	12.500	20.500
Kleibergen-Paap Wald F statistic	5576.25	3750.78	5576.25	3750.78	5576.25	3750.78
			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
Counterfactual mean	0.293	0.295	0.523	0.529	0.253	0.255
Bandwidth	80.000	120.000	80.000	120.000	80.000	120.000
Kleibergen-Paap Wald F statistic	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
Counterfactual 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
Kleibergen-Paap Wald F statistic	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 C.IV: Sibling Effects on Applications and Enrollment in Older Sibling's Target Major-College

	Applies 1st		**Applies**		**Enrolls**	
	P1	P2	P1	P2	P1	P2
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.012***	0.014***	0.023***	0.024***	0.006***	0.007**
	(0.003)	(0.004)	(0.005)	(0.006)	(0.002)	(0.003)
Reduced form	0.006***	0.007***	0.012***	0.012***	0.003***	0.003***
	(0.001)	(0.002)	(0.003)	(0.003)	(0.001)	(0.001)
First stage	0.536***	0.501***	0.536***	0.501***	0.536***	0.501***
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
2SLS (Triangular kernel)	0.012***	0.013***	0.024***	0.026***	0.006***	0.007***
	(0.003)	(0.004)	(0.005)	(0.006)	(0.003)	(0.003)
Observations	170886	247412	170886	247412	170886	247412
Counterfactual mean	0.020	0.019	0.066	0.065	0.012	0.012
Bandwidth	18.000	27.500	18.000	27.500	18.000	27.500
Kleibergen-Paap Wald F statistic	14765.19	8835.99	14765.19	8835.99	14765.19	8835.99
			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
Counterfactual mean	0.022	0.021	0.111	0.111	0.017	0.016
Bandwidth	80.000	120.000	80.000	120.000	80.000	120.000
Kleibergen-Paap Wald F statistic	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
Counterfactual 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
Kleibergen-Paap Wald F statistic	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 C.V: Sibling Effects on Applications and Enrollment in Older Sibling's Target College - Different Slope for each Admission Cutoff

	Applies 1st		**Applies**		**Enrolls**	
	P1	P2	P1	P2	P1	P2
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.060***	0.056***	0.082***	0.090***	0.054***	0.052***
	(0.015)	(0.020)	(0.018)	(0.023)	(0.013)	(0.017)
Reduced form	0.030***	0.027***	0.041***	0.043***	0.027***	0.025***
	(0.008)	(0.010)	(0.009)	(0.011)	(0.006)	(0.008)
Observations	86521	136868	86521	136868	86521	136868
Counterfactual outcome mean	0.222	0.218	0.447	0.441	0.132	0.127
Bandwidth	12.500	20.500	12.500	20.500	12.500	20.500
Kleibergen-Paap Wald F statistic	3948.401	2421.742	3948.401	2421.742	3948.401	2421.742
			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
Counterfactual 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
Kleibergen-Paap Wald F statistic	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
Counterfactual 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
Kleibergen-Paap Wald F statistic	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 C.VI: Sibling Effects on Applications and Enrollment in Older Sibling's Target Major-College - Different Slope for each Admission Cutoff

	Applies 1st		Applies		Enrolls	
	P1	P2	P1	P2	P1	P2
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.013***	0.015***	0.025***	0.025***	0.007***	0.007*
	(0.003)	(0.004)	(0.005)	(0.007)	(0.003)	(0.003)
Reduced form	0.007***	0.008***	0.014***	0.013***	0.004***	0.003*
	(0.002)	(0.002)	(0.003)	(0.004)	(0.001)	(0.002)
Observations	170886	247412	170886	247412	170886	247412
Counterfactual mean	0.019	0.018	0.065	0.063	0.012	0.011
Bandwidth	18.000	27.500	18.000	27.500	18.000	27.500
Kleibergen-Paap Wald F statistic	12905.771	7216.201	12905.771	7216.201	12905.771	7216.201
			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
Counterfactual 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
Kleibergen-Paap Wald F statistic	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
Counterfactual 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
Kleibergen-Paap Wald F statistic	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 C.VII: Sibling Effects on Applications and Enrollment in Older Sibling's Target College - Target × Counterfactual Major Fixed Effects

	Applies 1st		**Applies**		**Enrolls**	
	P1	P2	P1	P2	P1	P2
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.041**	0.041*	0.056***	0.052**	0.034**	0.030*
	(0.018)	(0.021)	(0.021)	(0.024)	(0.016)	(0.018)
Reduced form	0.019**	0.018*	0.026***	0.023**	0.016**	0.013*
	(0.009)	(0.009)	(0.010)	(0.011)	(0.007)	(0.008)
Observations	64886	106436	64886	106436	64886	106436
Counterfactual mean	0.230	0.230	0.460	0.450	0.140	0.130
Bandwidth	12.500	20.500	12.500	20.500	12.500	20.500
Kleibergen-Paap Wald F statistic	2639.50	2035.28	2639.50	2035.28	2639.50	2035.28
			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
Counterfactual 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
Kleibergen-Paap Wald F statistic	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
Counterfactual 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
Kleibergen-Paap Wald F statistic	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 C.VIII: Sibling Effects on Applications and Enrollment in Older Sibling's Target Major-College - Target × Counterfactual Major Fixed Effects

	Applies 1st		**Applies**		**Enrolls**	
	P1	P2	P1	P2	P1	P2
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.018***	0.017***	0.029***	0.029***	0.006*	0.006
	(0.004)	(0.005)	(0.007)	(0.009)	(0.003)	(0.004)
Reduced form	0.009***	0.009***	0.015***	0.015***	0.003*	0.003
	(0.002)	(0.003)	(0.004)	(0.004)	(0.002)	(0.002)
Observations	128112	191980	128112	191980	128112	191980
Counterfactual mean	0.020	0.020	0.070	0.060	0.010	0.010
Bandwidth	12.500	20.500	12.500	20.500	12.500	20.500
Kleibergen-Paap Wald F statistic	7526.410	5003.480	7526.410	5003.480	7526.410	5003.48
			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
Counterfactual 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
Kleibergen-Paap Wald F statistic	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
Counterfactual 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
Kleibergen-Paap Wald F statistic	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 × 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 C.IX: Probability of Enrolling in any College Depending on Older Siblings' Admission to Target Major-College

	Younger siblings		Older siblings	
	(1)	(2)	(3)	(4)
	Panel A - Chile			
Older sibling admitted to target major = 1	-0.010*	-0.009	0.044***	0.044***
	(0.005)	(0.007)	(0.005)	(0.007)
Observations	114424	136355	78655	93826
Counterfactual outcome mean	0.760	0.757	0.804	0.799
Bandwidth	12.000	14.500	12.000	14.500
	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
Counterfactual 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
Counterfactual 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.

D Additional Results

D.1 Heterogeneous Effects on College and Major Enrollment

The heterogeneity analyses presented in the main body of the paper focus on applications to colleges and majors. In this section, we presents similar results but looking instead at college and major enrollment. As shown in Tables D1 and D2 the results that we find in this case follow a similar pattern to the ones we find when focusing only on applications.

D.2 Sibling Spillover on the Quality of Younger Siblings' Choices

In this section we investigate if the quality of the younger siblings' first preference in their application list changes when an older sibling marginally enrolls in her target major. We measure quality using the same indexes that we defined in the main body of the paper. As shown in Table D3 we find no significant change in the quality of the preferred option of younger siblings. This differs with the results discussed for the United States in the main body of the paper. This is in part explained by the fact that in Chile, Croatia and Sweden we find no relevant extensive margin responses in either older or younger siblings. Thus, if target and counterfactual programs of older siblings do not greatly differ in quality, it is not surprising finding no significant differences in the quality of the programs that younger siblings rank in first place. In the case of the United States, marginal admission generates a much bigger change in older siblings' higher education trajectories,

D.3 Sibling Spillovers on Academic Performance

In the main body of the paper we find that marginal admission of older siblings into their target college or major does not generate significant improvements on their academic performance in high school or in the admission exam. In this section we further investigate sibling spillovers on academic performance by checking if some responses arise depending on the age difference between siblings. Table D4 summarizes these results. We find no increases on younger siblings academic performance no matter of the age difference that they have with their older sibling.

D.4 Sibling Spillover on College Choices and Location Preferences

One hypothesis that may explain the big effects that we find on the choice of college is that they reflect at least in part a change in 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 investigate this possibility, 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.[40].

Table D5 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 C.III were driven only by geographic preferences, we should not find sibling spillovers on the choice of college for this sub-sample. However, the coefficients that we obtain in this case are very similar to the main results discussed in the main body of the paper.

D.5 Sibling Spillovers on the Field of Study

The data that we have available for Chile, Croatia and Sweden also makes it possible to investigate sibling spillovers on the choice of the field of study. We define field of study by the three digit level code of the ISCED classification. In order to investigate sibling spillovers in this dimension we follow the same logic discussed in Section 4 and define a "Field of Study Sample". In order to be in this sample, apart from satisfying the first two assumptions discussed in the main body of the paper, older siblings need to:

3.B. list 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 older siblings for whom the marginal admission or rejection from their target major changes the field of study to which they are allocated. Using this

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

sample we study sibling spillovers on applications to and enrollment in fields of study. However, as shown in Figure D1 and Table D6 we only find a marginal significant increase in applications to the older siblings' field in the case of Chile. In Croatia and Sweden none of the estimated coefficients is statistically significant. Considering that the comparison of results across samples must be treated with caution, our results suggest that individuals' major choice is only affected when younger siblings are likely to be admitted in their older siblings' specific major-college combination.

Table D1: Sibling Effects on College and Major Enrollment by Age Difference and Gender

	Enrollment in any 4-year College	College Enrollment				Major Enrollment		
	US	CHI	CRO	SWE	US	CHI	CRO	SWE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Age Difference ≥ 5								
Older Sibling Enrolls = 1	0.216	0.052****	0.089***	0.067***	0.163***	0.006***	0.013**	0.006***
	(0.0132)	(0.011)	(0.019)	(0.006)	(0.053)	(0.002)	(0.004)	(0.001)
Interaction	0.130	-0.029***	-0.029	-0.010	0.087	-0.001	0.001	-0.003***
	(0.141)	(0.009)	(0.026)	(0.005)	(0.057)	(0.002)	(0.006)	(0.001)
Observations	44190	86364	12950	444203	44190	170570	36756	732025
Kleibergen-Paap Wald F statistic	64.780	2767.580	3230.667	2975.652	64.780	7330.470	7225.706	5255.957
Panel B: Same Gender = 1								
Older Sibling Enrolls = 1	0.306**	0.033***	0.065***	0.056***	0.184***	0.002	0.007	0.002
	(0.138)	(0.011)	(0.021)	(0.006)	(0.056)	(0.002)	(0.009)	(0.001)
Interaction	-0.149***	0.010	0.037*	0.014**	-0.024	0.007***	0.013**	0.006***
	(0.072)	(0.009)	(0.019)	(0.005)	(0.030)	(0.002)	(0.004)	(0.001)
Observations	44190	86521	12950	444203	44190	170886	36757	732025
Kleibergen-Paap Wald F statistic	64.970	2788.47	3229.534	3075.133	64.970	7383.020	7220.184	5419.139
Panel C: Same Gender = 1 - Older Brothers								
Older Sibling Enrolls = 1		0.041**	0.066	0.059***		-0.001	0.008	0.002
		(0.016)	(0.034)	(0.011)		(0.003)	(0.007)	(0.003)
Interaction		0.014	0.014	0.015		0.015***	0.031***	0.009***
		(0.012)	(0.031)	(0.009)		(0.003)	(0.008)	(0.002)
Observations		39919	5008	160086		81072	14203	281549
Kleibergen-Paap Wald F statistic		1435.860	1405.970	1330.244		4150.072	4025.070	2717.178
Panel D: Same Gender = 1 - Older Sisters								
Older Sibling Enrolls = 1		0.034*	0.044	0.061***		0.006*	0.006	0.001
		(0.018)	(0.029)	(0.009)		(0.003)	(0.006)	(0.002)
Interaction		0.008	0.046	0.013		-0.002	0.004	0.003*
		(0.013)	(0.026)	(0.007)		(0.003)	(0.005)	(0.001)
Observations		44222	7545	273981		87895	22239	438419
Kleibergen-Paap Wald F statistic		1223.530	1651.529	1484.510		3096.64	3662.675	2441.736

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table V. In addition, they include an interaction between the treatment and a dummy variable that indicates if siblings are 5 or more years apart (Panel A) or between the treatment and a dummy variable that indicates if siblings are of the same gender (Panel B). Panel C and Panel D do something similar but while Panel C focus only on siblings pairs in which the older one is male, Panel D looks at cases in which the older one is female. The dummy variables are also included in the specifications as controls. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D2: Sibling Effects on College and Major Enrollment by Older Sibling's Target Choice Quality

	Enrollment in any 4-year College	**College Choice**				**Major Choice**		
	US	CHI	CRO	SWE	US	CHI	CRO	SWE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Peers' Quality (Standardize performance in SAT or in HS)								
Older Sibling Enrolls = 1	-0.586	0.026**	-0.024	0.043***	0.316**	0.004	0.021	0.000
	(0.426)	(0.013)	(0.012)	(0.007)	0.161	(0.003)	(0.058)	(0.002)
Interaction	2.138***	0.015***	0.029*	0.026***	-0.378	0.002**	-0.002	0.005***
	(0.743)	(0.005)	(0.012)	(0.004)	(0.279)	(0.001)	(0.003)	(0.001)
Observations	44190	86521	10693	444203	44190	170886	34510	732023
Kleibergen-Paap Wald F statistic	17.194	1856.76	2598.965	2577.150	17.194	5123.72	6833.719	4508.761
Panel B: First Year Dropout Rate								
Older Sibling Enrolls = 1		0.057***		0.059***		0.013***		0.005**
		(0.012)		(0.007)		(0.004)		(0.002)
Interaction		-0.155***		-0.079***		-0.037*		-0.006
		(0.054)		(0.023)		(0.019)		(0.005)
Observations		84076		320107		167804		535714
Kleibergen-Paap Wald F statistic		2652.77		2678.503		7387.42		5465.479
Panel C: Graduation Rate								
Older Sibling Enrolls = 1	-0.297	-0.028		0.017**	0.0577	0.003		-0.003
	(0.208)	(0.029)		(0.009)	(0.047)	(0.006)		(0.002)
Interaction	0.448	0.152***		0.049***	-0.084	0.006		0.012***
	(0.305)	(0.059)		(0.011)	(0.070)	(0.013)		(0.002)
Observations	44190	85697		314434	44190	169557		509583
Kleibergen-Paap Wald F statistic	1.269	2796.37		2844.24	1.269	6697.58		5421.94
Panel D: Graduates' Earnings (Standardized annual earnings)								
Older Sibling Enrolls = 1	0.229*	0.042***		0.053***	0.172***	0.009***		0.002
	(0.133)	(0.010)		(0.008)	(0.053)	(0.002)		(0.002)
Interaction	0.004	0.003		0.008*	-0.004	0.001***		0.003**
	(0.007)	(0.002)		(0.004)	(0.004)	(0.0004)		(0.001)
Observations	44190	81112		218552	44190	160627		358644
Kleibergen-Paap Wald F statistic	129.787	2121.96		1380.629	129.787	5764.600		2462.490

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table V. In addition, they include an interaction between the treatment and an index of older siblings' major or college quality: peers' academic performance (Panel A), first year dropout rates (Panel B), graduation rates (Panel C) and graduates' earnings (Panel D). The quality index is also included in the specifications as a control. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D3: Sibling Spillovers on Quality of Younger Siblings Major-College

	Peers' Quality (Standardized SAT or HS performance)	1st Year Dropout Rates	Graduation Rates	Standardized Earnings
	(1)	(2)	(3)	(4)
Panel A - Chile				
Older sibling enrolls	0.033	-0.003	0.008***	0.012
	(0.031)	(0.003)	(0.002)	(0.012)
Observations	86519	75627	83060	78782
Kleibergen-Paap Wald F statistic	5575.82	5062.37	5416.03	5090.04
Panel B - Croatia				
Older sibling enrolls	-0.030			
	(0.024)			
Observations	33096			
Kleibergen-Paap Wald F statistic	13708.41			
Panel C - Sweden				
Older sibling enrolls	-0.052**	0.002	-0.014	-0.018*
	(0.022)	(0.005)	(0.011)	(0.011)
Observations	480386	335103	267106	181681
Kleibergen-Paap Wald F statistic	9444.93	7212.60	5078.18	2406.57

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major-college on the quality of the top choice of younger siblings. The measure of students quality and graduates' average earnings are standardized. These specifications use the same set of controls and bandwidths used in Table V. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D4: Sibling Effects on Academic Performance by Age Difference

	Major Sample		College Sample	
	High School GPA	Average Score AE	High School GPA	Average Score AE
	(1)	(2)	(3)	(4)
Panel A - Chile				
Older sibling enrolls	-0.004	-0.005	0.010	0.014
	(0.0218)	(0.014)	(0.035)	(0.022)
Older sibling enrolls × 2 < Δ Age ≤ 4	0.004	-0.007	0.017	-0.011
	(0.019)	(0.012)	(0.031)	(0.019)
Older sibling enrolls × 4 < Δ Age	-0.010	-0.009	-0.024	-0.010
	(0.018)	(0.011)	(0.029)	(0.018)
Observations	170886	170886	86521	86521
Kleibergen-Paap Wald F statistic	4889.680	4889.680	1843.23	1843.23
Panel B - Croatia				
Older sibling enrolls	-0.146	-0.133	-0.327	-0.302*
	(0.139)	(0.093)	(0.239)	(0.157)
Older sibling enrolls × 2 < Δ Age ≤ 4	0.066	0.093	0.007	0.097
	(0.170)	(0.111)	(0.202)	(0.134)
Older sibling enrolls × 4 < Δ Age	0.211	0.125	-0.235	0.280
	(0.568)	(0.392)	(0.590)	(0.402)
Observations	12,433	12,443	4,170	4,170
Counterfactual mean	-1.300	-0.834	-1.313	-0.909
Bandwidth	80.000	80.000	80.000	80.000
Kleibergen-Paap Wald F statistic	1461.978	1461.978	659.829	659.829
Panel C - Sweden				
Older sibling enrolls	0.288	0.015	0.015	0.080
	(0.027)	(0.038)	(0.038)	(0.055)
Older sibling enrolls × × 2 < Δ Age ≤ 4	0.010	0.070**	0.007	0.106
	(0.024)	(0.035)	(0.038)	(0.055)
Older sibling enrolls × × 4 < Δ Age	-0.057**	-0.017	-0.008	-0.006
	(0.024)	(0.036)	(0.037)	(0.055)
Observations	613,294	344,442	372,578	206,613
Counterfactual mean	0.219	0.051	0.232	0.055
Bandwidth	0.51	0.51	0.367	0.367
Kleibergen-Paap Wald F statistic	3070.585	2086.53	1747.338	1177.487
Panel D - United States				
Older sibling enrolls × Δ Age ≤ 2				36.286
				(27.521)
Older sibling enrolls × Δ Age ≤ 4				49.694
				(39.036)
Older sibling enrolls × Δ Age ≤ 10				-10.110
				(54.228)
Observations				37554
Counterfactual mean				950.926
Bandwidth				93
Kleibergen-Paap Wald F statistic				39.448

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major or college 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 V. 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.

Table D5: Sibling Effects on Applications to and Enrollment in Older Sibling's Target College: Cities with Multiple Colleges

	Applies 1st (1)	Applies (2)	Enrolls (3)
2SLS	0.073***	0.082***	0.065***
	(0.017)	(0.019)	(0.015)
Reduced form	0.041***	0.045***	0.036***
	(0.010)	(0.011)	(0.008)
First stage	0.556***	0.556***	0.556***
	(0.010)	(0.010)	(0.010)
Observations	37279	37279	37279
Counterfactual mean	0.25	0.50	0.15
Bandwidth	12.500	12.500	12.500
Kleibergen-Paap Wald F statistic	3353.800	3353.800	3353.800

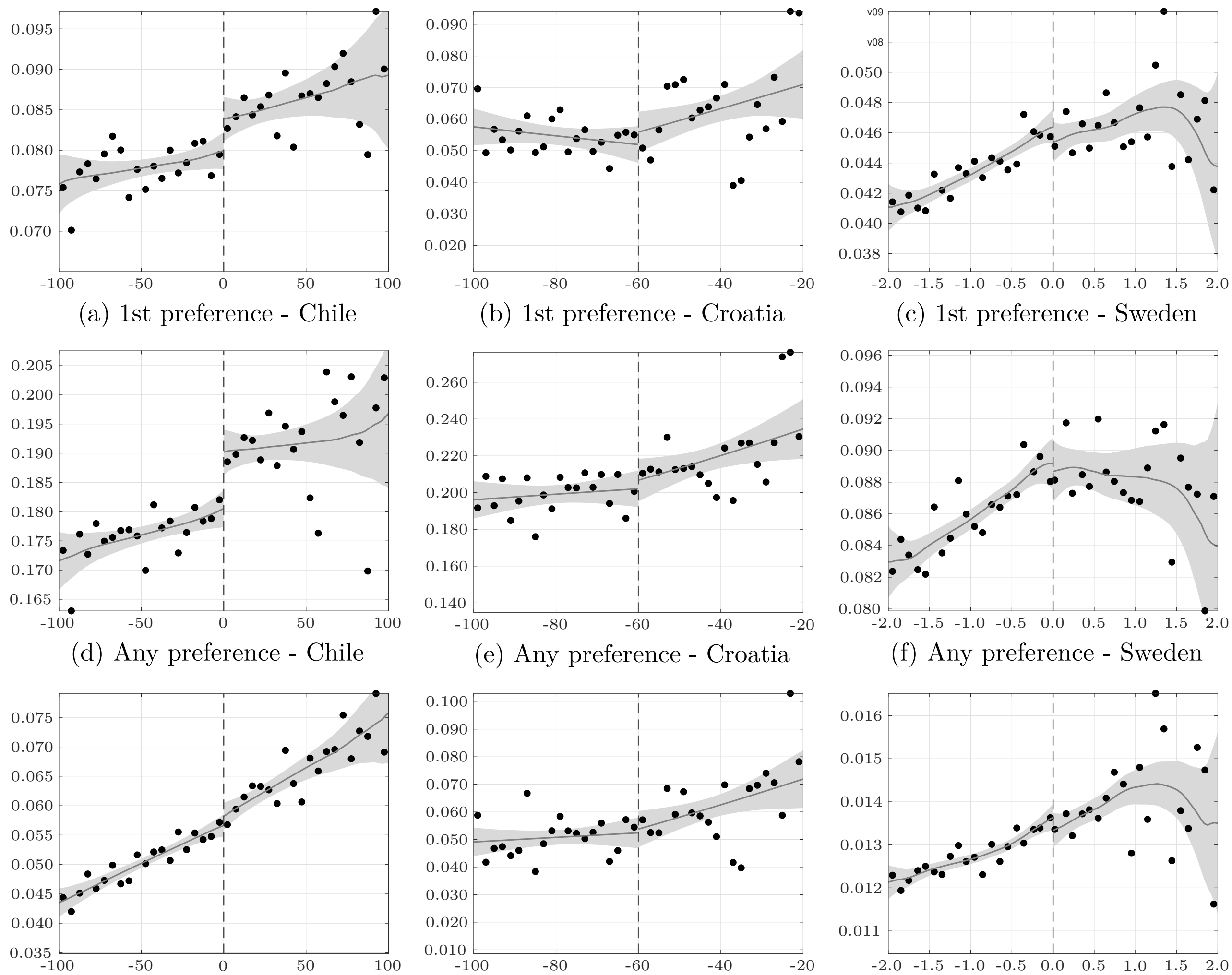
Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target college on younger siblings' probabilities of applying to and enrolling in the same college. The controls and bandwidths used in these specifications are the same described in Table V. The sample 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 D6: Sibling Effects on Applications to and Enrollment in Older Sibling's Target Field of Study

	Applies 1st		**Applies**		**Enrolls**	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
2SLS	0.012	0.009	0.017*	0.014	-0.001	-0.004
	(0.007)	(0.009)	(0.010)	(0.012)	(0.006)	(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)
First stage	0.478***	0.449***	0.478***	0.449***	0.478***	0.449***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
2SLS (Triangular Kernel)	0.008	0.008	0.015	0.013	0.000	-0.001
	(0.008)	(0.008)	(0.011)	(0.013)	(0.006)	(0.008)
Observations	106085	162122	106085	162122	106085	162122
Counterfactual mean	0.079	0.079	0.179	0.178	0.054	0.051
Bandwidth	16.000	25.500	16.000	25.500	16.000	25.500
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
Counterfactual 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
Counterfactual 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.

Figure D1: 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. Gray lines and the shadows in the back of them correspond to local polynomials of degree 1 and 95% confidence intervals. Black dots represent sample means of the dependent variable at different values of older sibling’s admission score.