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

Adam Altmejd Andrés Barrios-Fernández Marin Drlje Joshua Goodman

Michael Hurwitz Dejan Kovac Christine Mulhern Christopher Neilson Jonathan Smith

Latest Version : December 24, 2020

Abstract

This online appendix is organized in four sections. The first provides additional details about the higher education systems in Chile, Croatia, Sweden and the United States. It also explains how in this last setting we identify the hidden admission cutoffs. The second section discusses in detail our identification strategy and provides an in-depth description of the samples we use. The third section presents the robustness checks of the paper, and the fourth section additional results that either complement the analyses discussed in the main body of the paper, or extend them by exploring new outcomes or heterogeneity dimensions.

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

Contents

A Institutions: Further Details

This Section describes the higher education systems of Chile, Croatia, Sweden and the United States. We focus on the distinctive features of the admission systems that generate the discontinuities that we exploit in the paper to identify sibling spillovers. This Section also describes the procedure that we use to identify the subset of U.S. colleges using hidden test-score cutoffs in their admissions.

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).[1] 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).[2] Students take the PSU in December, at the end of the Chilean academic year, but they typically need to register before mid-August.[3] As of 2006, all public and voucher school graduates are eligible for a fee waiver that makes the PSU free for them.[4]

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

[1]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.

[2]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 mathematics sections and at least one of the other sections. Universities set the weights allocated to these instruments for selecting students in each program.

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

[4]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.

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. This system generates a sharp discontinuity in admission probabilities in each college-major combination with more applicants than vacancies.

Colleges that do not use the centralized system have their own admission processes in place.[5] Although they could use their own entrance exams, the PSU still plays an important role in the selection of their students, mostly due to the existence of strong financial incentives for both students and institutions.[6] 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; the only cost of rejecting the offer is losing it. This also makes it possible for some students originally rejected from a program to later receive an offer.

A.2 College Admission System in Croatia

In Croatia, there are 49 universities. Since 2010, all of which 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.[7] The national exam is taken in late June, approximately one month after the end of the Croatian academic year. How-

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

[6]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.

[7]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.

ever, students are required to submit a free-of-charge online registration form by mid-February.

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

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

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.[8]

There are two important differences between the Chilean and Croatian systems. First, all Croatian colleges use the centralized admission system. Second, rejecting an offer in the Croatian setting is more costly to students. If students do not accept the offer they receive the first time that they apply, they lose the tuition fee waiver offered by the government. This means that if students re-apply to college in the future, they will have to pay tuition fees.

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.[9]

[8] 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.

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

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, the same system is open to applications to full majors and shorter courses alike. To simplify, we will henceforth refer to all alternatives as *majors*. Moreover, applicants are ranked by different scores separately in a number of *admission groups*. Their best ranking is then used to determine their admission status.[10] Finally, the Swedish admission system has two rounds. Applicants who receive a first-round offer can choose to accept this offer or to participate in the second round of the application. Their scores and lists of preferences do not change between the two rounds, but the admission 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 the applicants above the second round admission cutoff receive 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 Sweden the jump in older siblings' admission and enrollment probabilities is smaller than in the other two countries. Applicants are free to reject their final offers. As in Chile, the only consequence of rejecting an offer is losing that place in college.

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.[11]

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.

[10] 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.

[11] 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.

Students can apply to majors starting in the fall or spring semester, with the application process occurring in the semester preceding the intended enrollment. In each application students may rank up to 20 alternatives.[12] 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 wait list for choices to which they have not yet been admitted. 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.[13]

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

A.4 College Admission System in the United States

In the U.S., 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.[14]. As in the case of the admission exams used in the other countries, the SAT has different sections and, in terms of application, it is common for colleges to consider students' "superscores"(**?**). 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.

[12]Students were only able to rank up to 12 alternatives until 2005.

[13]As in Croatia, we focus on first-round submissions. As many applicants stay on the wait 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.

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

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 colleges that 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 thresholds induce some students to retake the SAT until their scores meet the thresholds (**?**). 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 identify 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 for each college in each year.[15] 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:

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

[15] Our approach is similar to that used in **?**.

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 discontinuities 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 at the same threshold, given that most colleges that use minimum SAT scores in admissions are unlikely to change that policy from year to year and seeing a consistent threshold across years also reduces the chances of false positives. 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: Further Details

B.1 Definition of Estimation Samples

This section presents a more detailed description of the estimation samples that we use to estimate sibling spillovers on the choice of college-major, college and major in Chile, Croatia and Sweden.

B.1.1 College-Major Sample

As college-major combinations are unique, being above or below a cutoff always changes the college-major combination to which an older sibling is admitted to. Thus, this sample includes all individuals whose older siblings are within a given bandwidth from a target cutoff.

Let c_{cmt} be the cutoff for major m offered by college c. If the major m offered by college c is ranked before the major m' offered by college c' in student i's preference list, we write $(m, c) \succ (m', c')$.[16] Denoting the application score of individual i as a_{imc}, we can define marginal students in the college-major sample as those whose older siblings:

1. Listed major m offered in college c as a choice such that all majors preferred to m had a higher cutoff score than m (otherwise assignment to m is impossible):
   $\bar{c}_{mc} < c_{m'c'} \; \forall \; (m', c') \succ (m, c)$.
2. Had an application score sufficiently close to m's cutoff score to be within a given bandwidth bw around the cutoff:
   $|a_{imc} - \bar{c}_{mc}| \leq bw$.

Thus, this sample includes individuals whose older siblings were rejected from (c, m) $(a_{icm} < \bar{c}_{cm})$ and those whose older siblings scored just above the admission cutoff $(a_{icm} \geq \bar{c}_{cm})$. Note that the same applicant can narrowly miss several options that were highly ranked on her applications. This implies that the same individual may belong to more than one college-major marginal group.

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

B.1.2 College Sample

When investigating sibling spillovers on the choice of college, we use a sample similar to the one described in the previous section, but this time we add one extra restriction.

We only want to keep in the sample individuals whose older siblings' target and next best college-major preferences are taught in different colleges. For them being below or above the admission threshold changes the college to which they are assigned to.

Thus, we define marginal students in the college sample as those whose older siblings meet restrictions 1 and 2, and:

3. Listed major m in college c as a choice such that majors not preferred to m in their application list are dictated by an institution different from c or if dictated by c 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 college-major would receive an offer to enroll in different major, but in the same target college.[17]

B.1.3 Major Sample

Finally, in order to investigate sibling spillovers in the choice of major, we follow the same logic used to define the two previous samples. In the "Major Sample"we want to keep older siblings for whom being below or above a college-major cutoff changes the major to which they are admitted to.

Thus, in order to be in this sample, apart from satisfying the first two restrictions discussed in Section B.1.1, older siblings need to:

3.B. list major m as a choice, such that options not preferred to m correspond to a major different

[17]In Appendix C we present additional results that investigate sibling spillovers on college choice in a modified version of this sample. In this alternative sample we only include individuals whose older siblings' target and next best options correspond to the same major, but are taught at different colleges (i.e. Economics at Princeton, and Economics at Boston University). The results are very similar to the ones we obtain using the College Sample.

from m (otherwise being above or below the cutoff would not generate variation in the major attended).

This means that we remove from this sample all older siblings whose target and next best option correspond to the same major.[18]

B.2 Identifying Assumptions

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 **??**, 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, and τ, and focus our analysis on a specific college-major 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 college-major 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 college-major 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

[18] In Section **??** we also present results that focus on individuals whose older siblings' target and next best college-major are taught in the same college. In this alternative sample, crossing the admission threshold changes the major, but not the college of the older sibling.

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 **?**, adapted to our setting:

1. Independence of the instrument:

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

2. Exclusion restriction:

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

3. First stage:

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

4. Monotonicity:

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

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

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

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

In addition to the usual monotonicity assumption that requires that admission to major u

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

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

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

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

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

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

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

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

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

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

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

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

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

Taken together, 3 and 4 imply:

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

□

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 for 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 college-major 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. In the U.S., where the cutoffs we exploit are hidden, we think violation of this assumption is just as unlikely. 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 the paper, 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. Because there are hundreds of these cutoffs, we pool them together in our analysis as 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 look for evidence of discontinuities around the admissions cutoff.

Figure C.II summarizes the result of this analysis for Chile, Croatia and Sweden. The figure 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 4 of the paper, 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 U.S.. 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 replicates this exercise for the U.S.. In this case, the coefficients also remain very stable when using a smaller bandwidth; when we increase it, the coefficients begin to drop, suggesting 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

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

First, in Figures C.VIII and C.VII we show that we observe an effect on younger siblings outcomes only at the real cutoff. This is not surprising since the placebo cutoffs that we use do not generate any change in older siblings' admissions. In the U.S. 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, we find no significant effects around placebo cutoffs that are far from the real cutoff. We do find some significant effects at points that are very close to the actual cutoff, but this is just the result of not observing the exact cutoffs and using instead estimates.

Second, in Figures C.X and C.XII we study if younger siblings' admission to their target college or major affect the application and enrollment decisions of their older siblings in Chile, Croatia and Sweden. Figure C.XI replicates this exercise for the U.S.. Since younger siblings apply to college after their older siblings, being marginally admitted or rejected from a major or college should not affect the outcomes of 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 small discontinuities at the cutoff, the size of the discontinuity is considerably smaller than the ones we document in the main body of the paper.

C.5 Alternative Specifications and Enrollment in Any College

We conclude this section by presenting results to alternative specifications.

Tables C.I and C.II summarize the results for the U.S.. The first table presents results of alternative specifications in which we control for additional covariates (column 2), include observations exactly at the cutoff (column 3), and compare the reduced form estimates that we obtain using our baseline specification with the ones that we obtain using instead the approach suggested by **?** to compute standard errors (columns 4 and 5). The second table presents results from specifications that control by a quadratic polynomial of the running variable (column 2), use a triangular kernel (column 3), and allow for different slopes of the running variable at each college's admission cutoff (column 4). Although we lose precision in some specifications, the size of the coefficients is very stable. The general picture that arises from these analyses is consistent with our main results and points to large sibling spillovers on both the decision to attend a four-year college and on the choice of college.

We present similar analyses for Chile, Croatia and Sweden distributed along multiple tables. First, Tables C.IV and C.III 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. In addition, in Tables C.VI and C.V we show that our results are robust to allowing the running variable to have cutoff-major specific slopes, and in Table C.VII we show that our main results are robust to control by covariates. Table C.VIII presents results from specifications in which we drop observations at the cutoff. Only the Swedish results change, with effect sizes decreasing to levels closer to the ones we find in the other countries. In Sweden, ties at the cutoff are much more frequent than in the other settings that we study. The donut specification thus removes many observations from the sample. Since these ties are broken by lottery, and we have no indication that admission at the cutoff could be manipulated, our main specifications also include these observations.

Since in the case of Chile, Croatia and Sweden we observe the full rank of individuals applications, in Table C.IX we present results from a specification in which we add two-way fixed effects that control for the target and next best option of older siblings. Thus, the identifying variation in these specifications only comes from individuals whose older siblings had the same target and next best option. It is comforting to see that the estimates we find here are very similar to the ones reported in the main body of the paper.

We finish this section going back to our baseline specification and estimating sibling spillovers on applications and enrollment in college, but on a new sample. In this sample we keep the major of the target and next best option fixed to ensure that the only difference at the margin is the college to which older siblings are allocated. The estimates that we obtain are once more very similar to the ones presented in the main body of the paper. Although the results for Croatia are less precise —the restrictions imposed to generate the new sample drastically reduced the number of observations in Croatia— the coefficients are similar in size to the ones discussed in the main body of the paper.

C.6 Sibling Spillovers on College and College-Major Choice: Fixing Target and Next Best Option Major or College

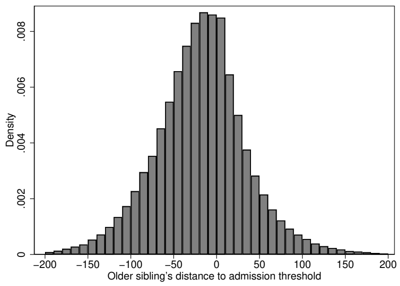
We start by expanding our study of sibling spillovers on college choice. In this Section we focus on individuals whose older siblings' target and next best options correspond to the same major, but are offered by different colleges. This means that crossing the threshold changes the college, but not the major to which older siblings are allocated. The results that we find—summarized in Table C.X— are very similar to the ones we document for college choice in the current section.In Croatia, the country for which we have the fewest number of observations, these estimates become less precise, but still they are similar in magnitude to the ones we present in the main body of the paper.

In order to investigate if sibling spillovers in the choice of major are only local—i.e. only affect preferences for the major in the same college of the older sibling— we build a new sample in which we only include individuals whose older sibling's target and next best option are offered by the same college (e.g. ranked first economics at Princeton and second sociology at Princeton). In the centralized admission systems used in Chile, Croatia and Sweden, individuals learn their scores before submitting their applications. This means that if after receiving their scores, they believe that it is unlikely to be admitted in the college-major of their older siblings, they might not even apply there. Thus, for this exercise we further restrict the sample to individuals who are likely to be admitted in their older siblings' target college-major if they apply.[19]

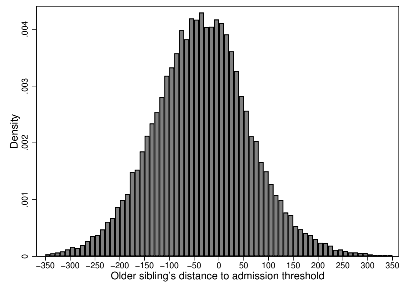
Table C.XII summarizes the results of this exercise. We find that when eligible for the older sibling's college-major choice, younger siblings' responses in terms of applications and enrollment are larger than the one we presented earlier in this Section. Most of the coefficients are significant only at the 10% level, but this lack of precision is a consequence of the reduced number of observations that we have in this new sample.

[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 college-major. In Sweden, the indicator is active whenever the younger sibling has a score above the cutoff in any admission group they are eligible for. In section **??**, we show that older siblings' enrollment in their target college-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.

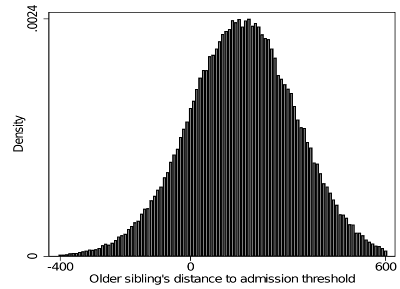
Figure C.I: Density of Older Siblings' Admission Exam and High School GPA at the Target College-Major Admission Cutoff



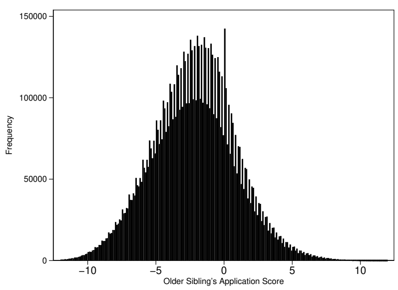
(a) Chile



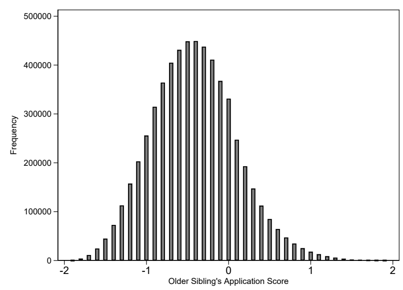
(b) Croatia



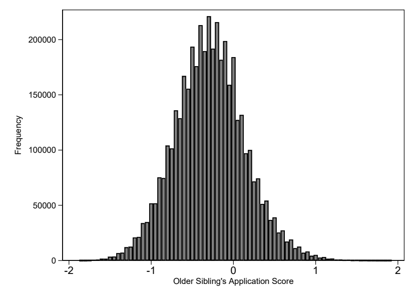
(c) United States



(d) High School GPA - Sweden



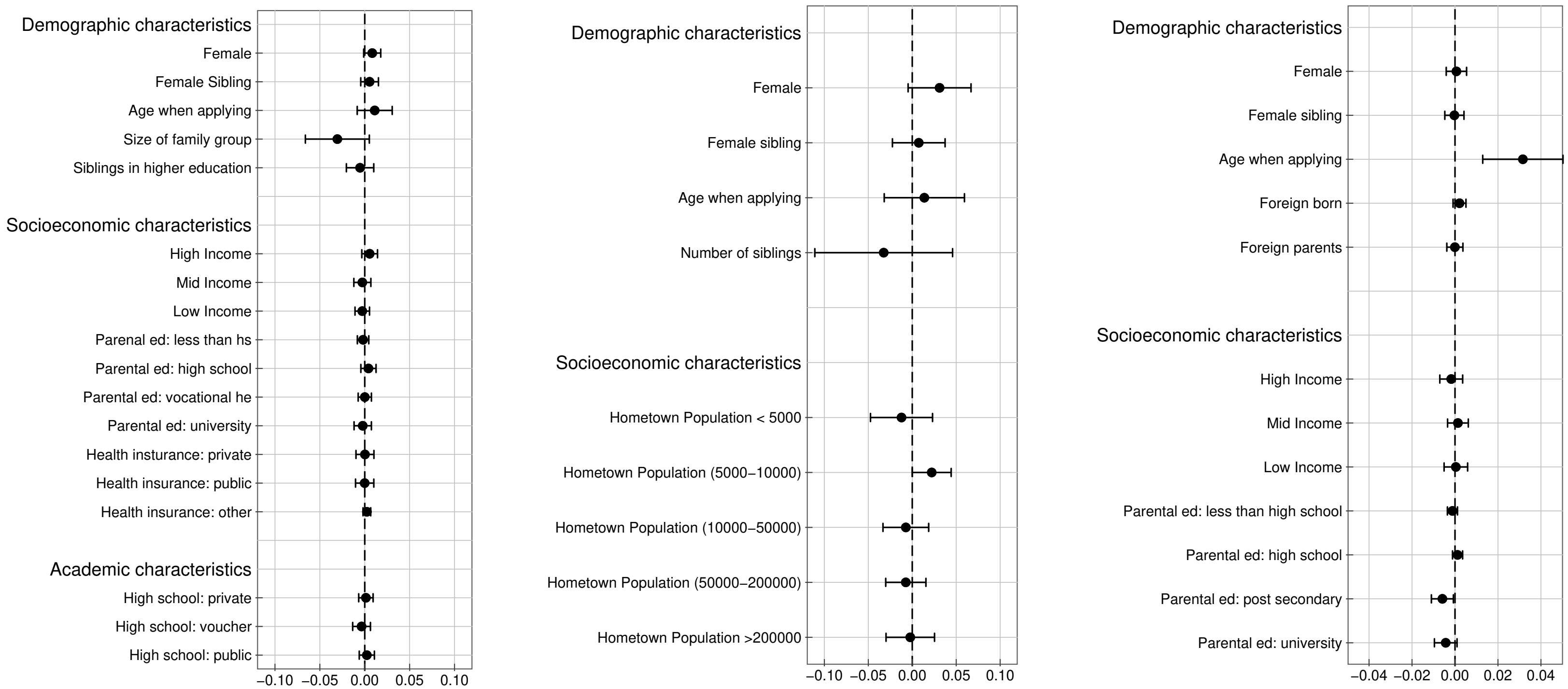
(e) Exam Before 2013 - Sweden



(f) Exam After 2013 - Sweden

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 (e) 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 (f) presents the distribution of scores after 2013.

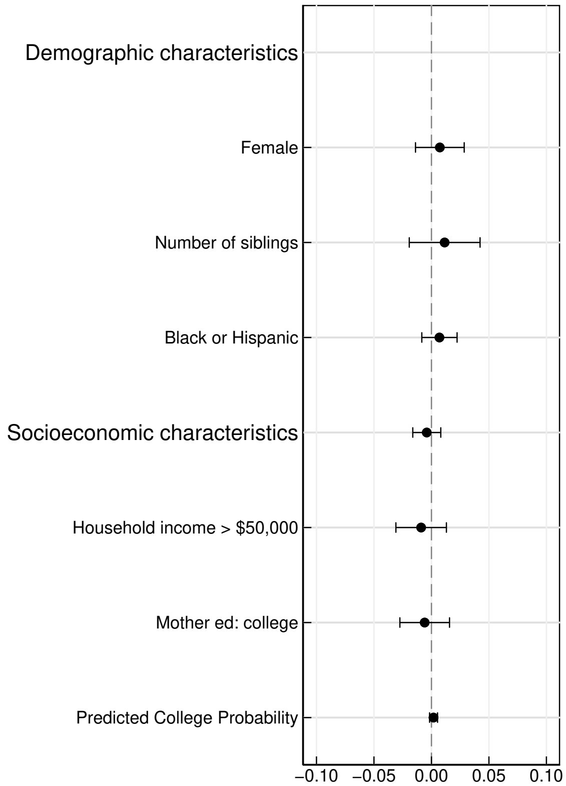
Figure C.II: Discontinuities in other Covariates at the Cutoff



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

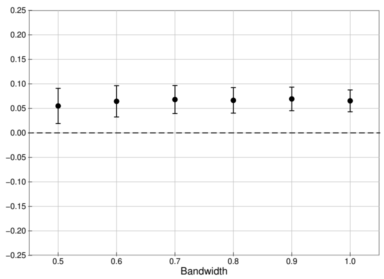
This figure illustrates the estimated jumps at the cutoff for a vector of socioeconomic and demographic characteristics. These estimates come from parametric specifications that control for a linear polynomial of the running variable. As the main specifications, these also include major-college-year 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 represent 95% confidence intervals.

Figure C.III: Discontinuities in other Covariates at the Cutoff (United States)

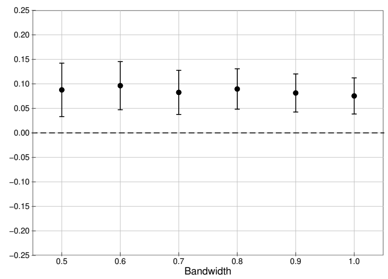


This figure illustrates how demographic and socioeconomic characteristics vary at the admissions cutoff in the United States. The range of the running variable corresponds to the bandwidth used in our main specifications. The points represent the estimated coefficient , while the lines represent 95%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

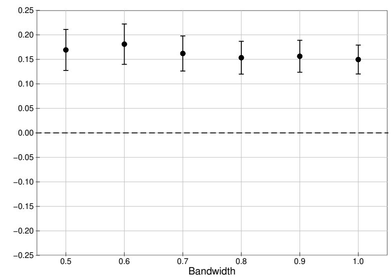
Figure C.IV: Probabilities of Applying and Enrolling in Older Sibling's Target College - Different Bandwidths



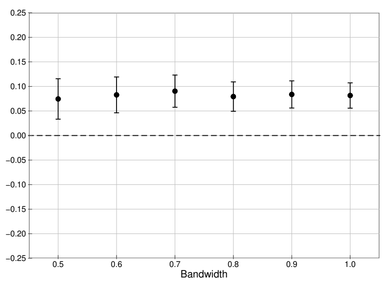
(a) 1st preference - Chile



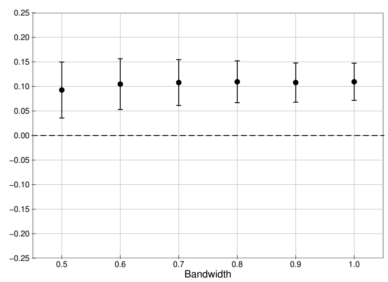
(b) 1st preference - Croatia



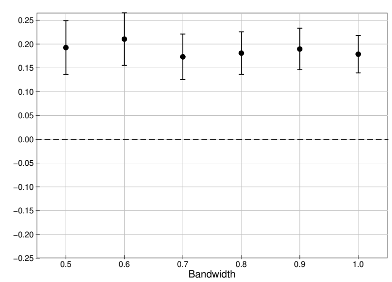
(c) 1st preference - Sweden



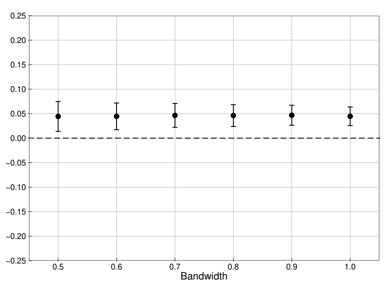
(d) Any preference - Chile



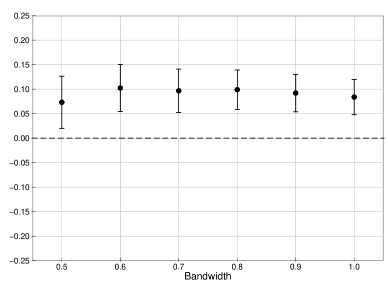
(e) Any preference - Croatia



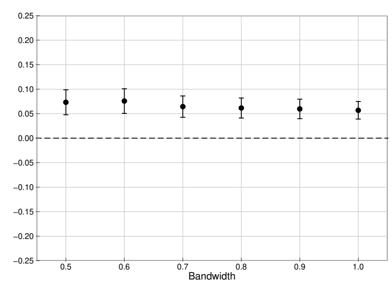
(f) Any preference - Sweden



(g) Enrolls - Chile



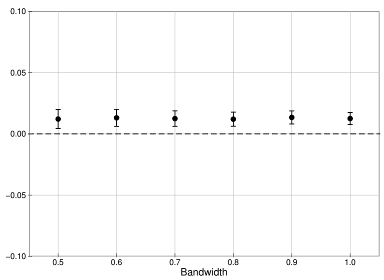
(h) Enrolls - Croatia



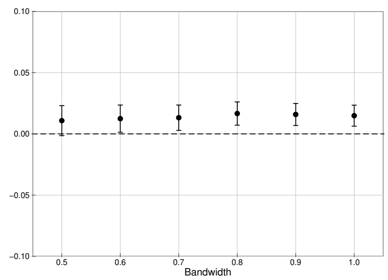
(i) Enrolls - Sweden

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 **?**. 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.

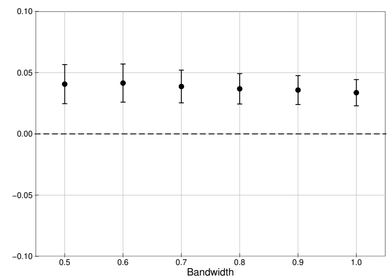
Figure C.V: Probabilities of Applying and Enrolling in Older Sibling's Target Major-College - Different Bandwidths



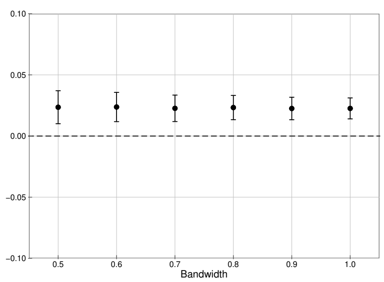
(a) 1st preference - Chile



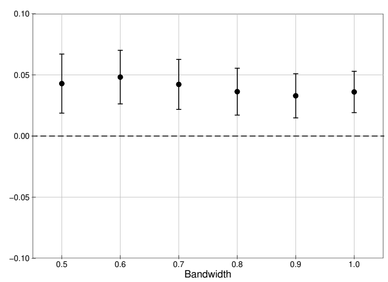
(b) 1st preference - Croatia



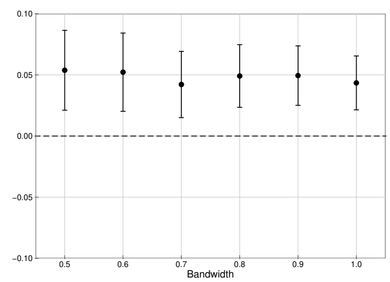
(c) 1st preference - Sweden



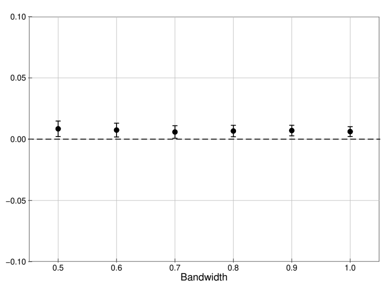
(d) Any preference - Chile



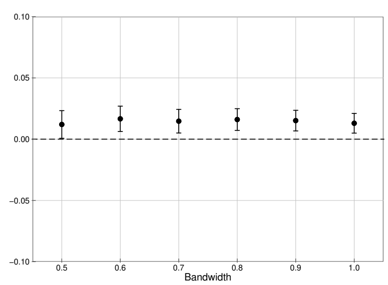
(e) Any preference - Croatia



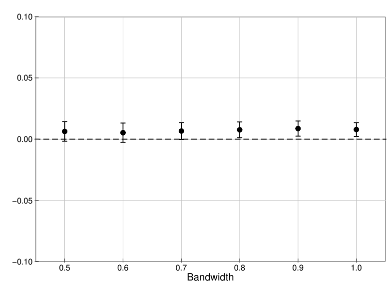
(f) Any preference - Sweden



(g) Enrolls - Chile



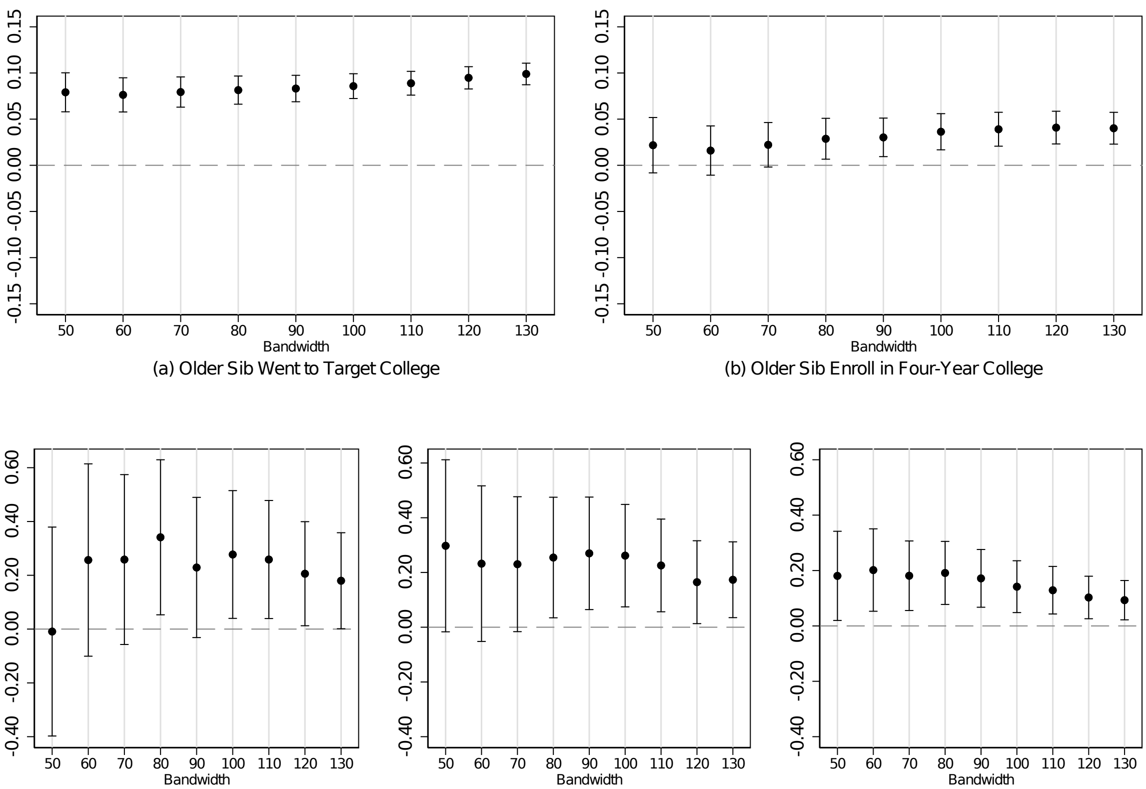
(h) Enrolls - Croatia



(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 **?**. 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-year College and in Older Sibling's Target College - Different Bandwidths (United States)



(a) Older Sib Went to Target College

(b) Older Sib Enroll in Four-Year College

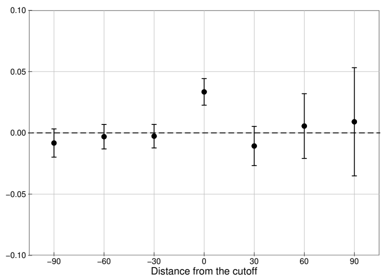
(c)Younger Sib Enroll in Four-Year College

(d) Younger Sib Apply to Target College

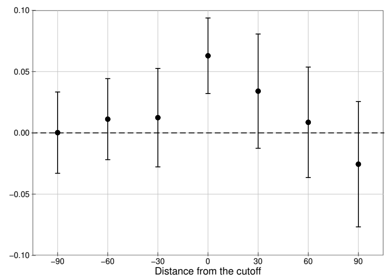
(e) Younger Sib Enroll in Target College

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.

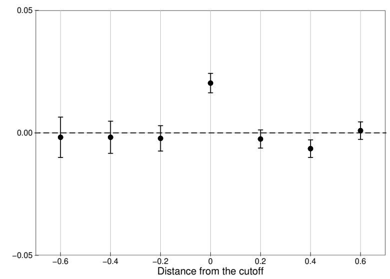
Figure C.VII: Placebo Cutoffs - Probabilities of Applying and Enrolling in Older Sibling's Target College



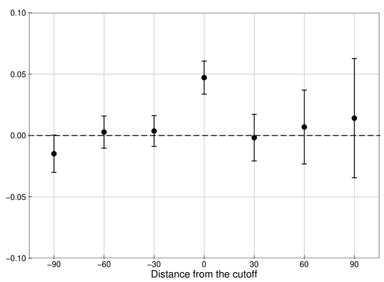
(a) 1st preference - Chile



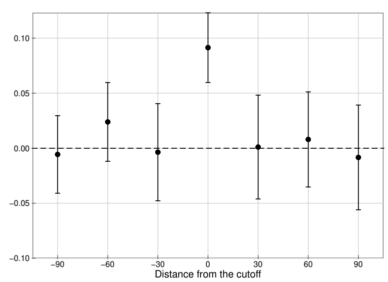
(b) 1st preference - Croatia



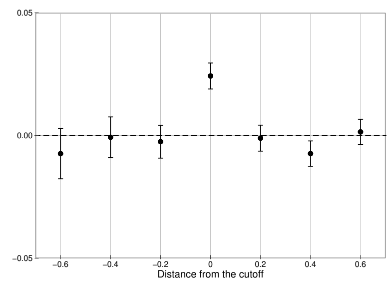
(c) 1st preference - Sweden



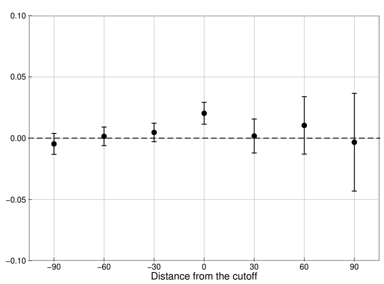
(d) Any preference - Chile



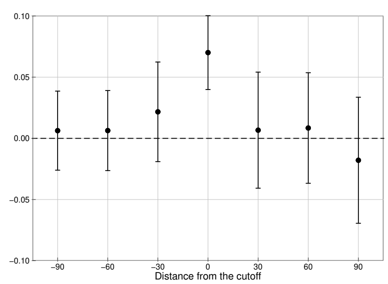
(e) Any preference - Croatia



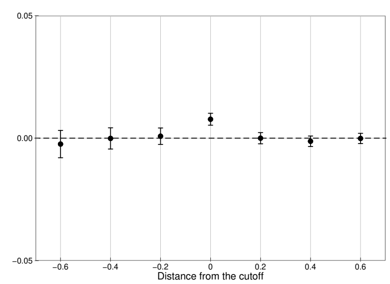
(f) Any preference - Sweden



(g) Enrolls - Chile



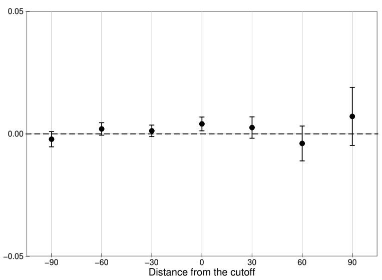
(h) Enrolls - Croatia



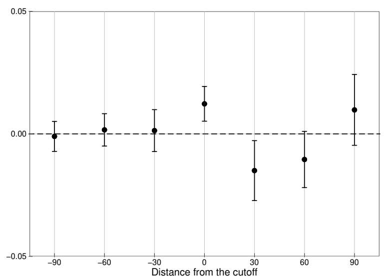
(i) Enrolls - Sweden

This figure illustrates the results of a placebo exercise that investigates if effects similar to the ones documented in figure **??** 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.

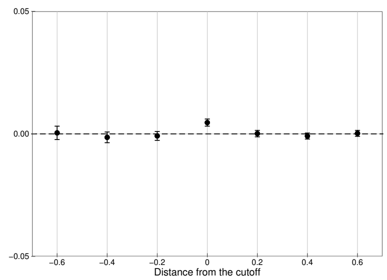
Figure C.VIII: Placebo Cutoffs - Probabilities of Applying and Enrolling in Older Sibling's Target Major-College



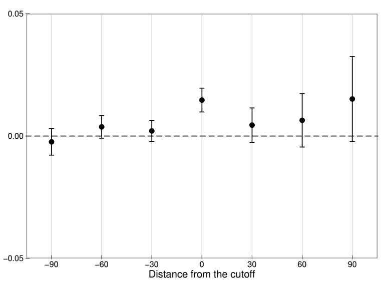
(a) 1st preference - Chile



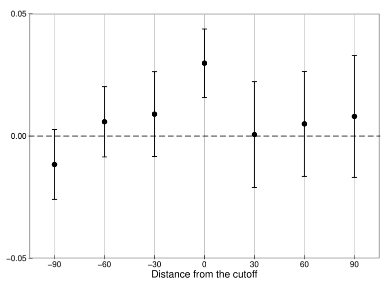
(b) 1st preference - Croatia



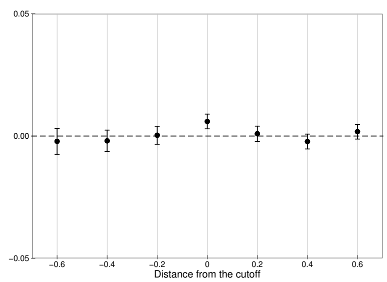
(c) 1st preference - Sweden



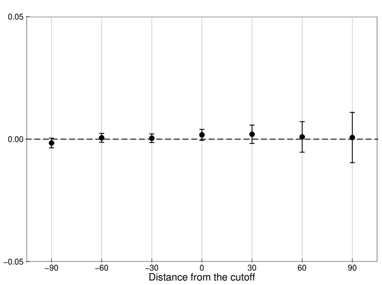
(d) Any preference - Chile



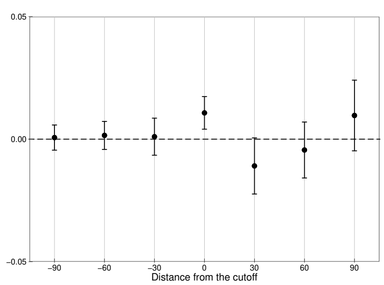
(e) Any preference - Croatia



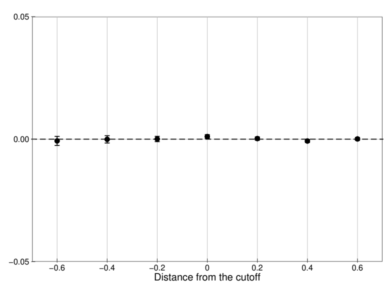
(f) Any preference - Sweden



(g) Enrolls - Chile



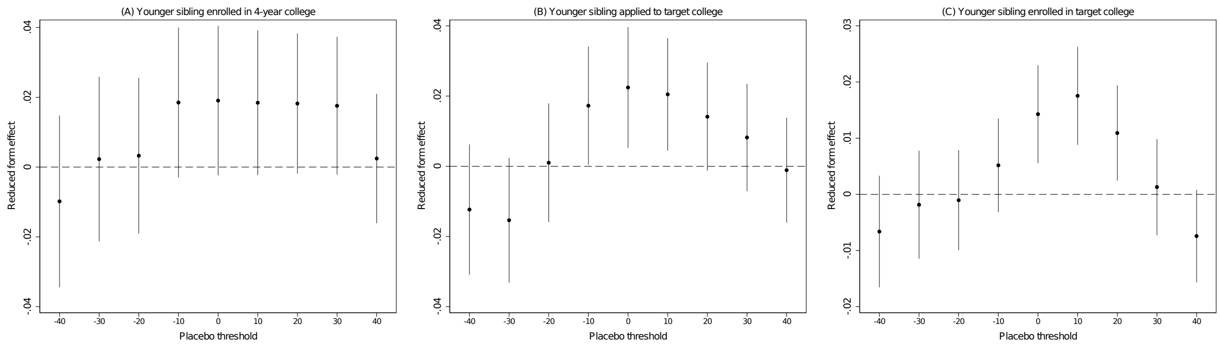
(h) Enrolls - Croatia



(i) Enrolls - Sweden

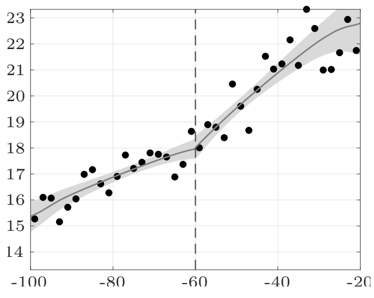
This figure illustrates the results of a placebo exercise that investigates if effects similar to the ones documented in figure **??** arise at different values of the running variable. Therefore, the x-axis corresponds to different (hypothetical) values of cutoffs - 0 corresponds to the actual cutoff 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)

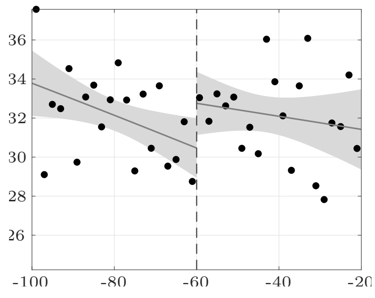


This figure illustrates the results of a placebo exercise that investigates if effects similar to the ones documented in the main body of the paper arise at different values of the running variable. Therefore, the x-axis corresponds to different (hypothetical) values of cutoffs and 0 corresponds to the 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.

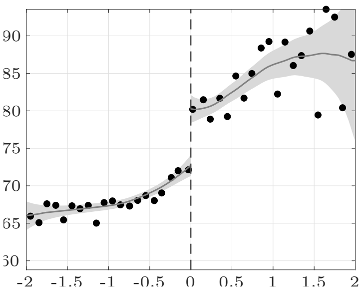
Figure C.X: Placebo - Probabilities of Applying and Enrolling in Younger Sibling's Target College



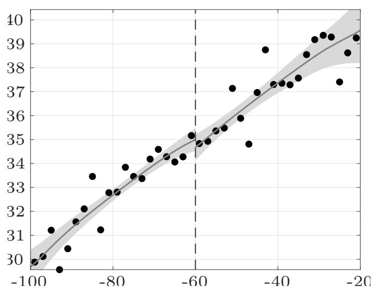
(a) 1st preference - Chile



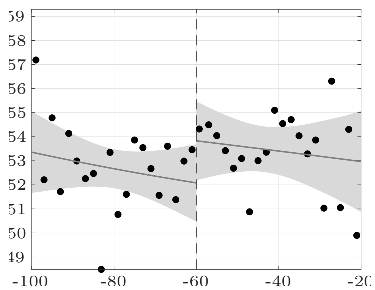
(b) 1st preference - Croatia



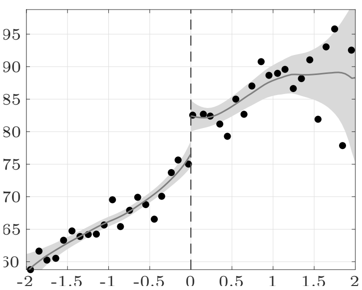
(c) 1st preference - Sweden



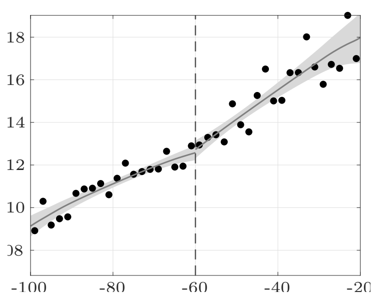
(d) Any preference - Chile



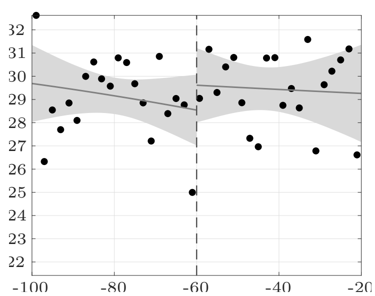
(e) Any preference - Croatia



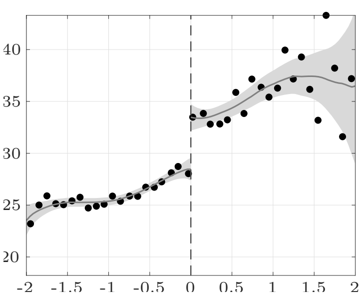
(f) Any preference - Sweden



(g) Enrolls - Chile



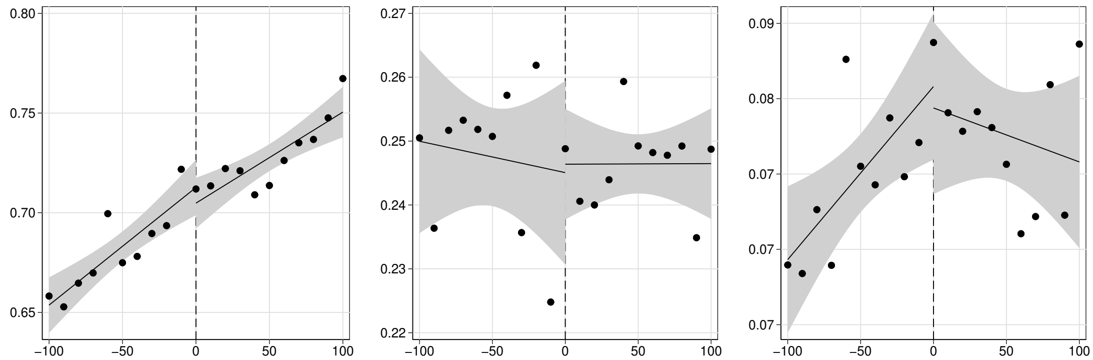
(h) Enrolls - Croatia



(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.

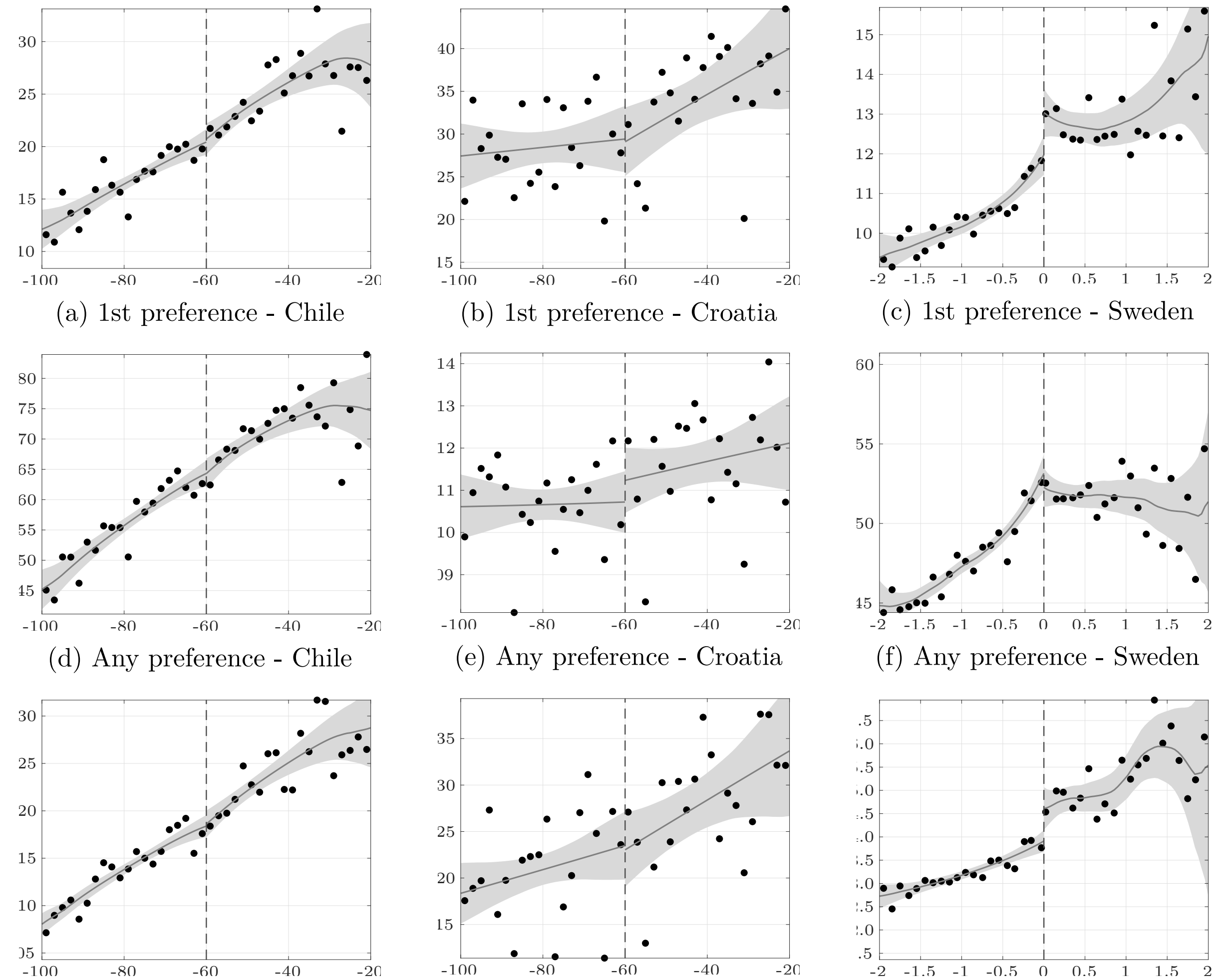
Figure C.XI: Placebo - Probabilities of Applying and Enrolling in Younger Sibling's Target College



(a) Enrolls in Four−Year College (b) Applies to Target College (c) Enrolls in 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.XII: 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.

Table C.I: Robustness of Younger Siblings' College Choices

	2SLS			Reduced Form	
	Baseline specification	Including covariates	Donut	Baseline Specification	Kolésar & Rothe SEs
	(1)	(2)	(3)	(4)	(5)
(A) All students					
Enrolled in target college	0.172***	0.168***	0.272***	0.014	0.014
	(0.054)	(0.054)	(0.070)	(0.004)	(0.004)
Enrolled in 4-year college	0.230*	0.186	0.116	0.019	0.019
	(0.132)	(0.127)	(0.161)	(0.011)	(0.010)
B.A. completion rate	0.180**	0.149**	0.131	0.015	0.015
	(0.080)	(0.076)	(0.098)	(0.006)	(0.006)
Peer quality	0.316**	0.253*	0.279	0.026	0.026
	(0.148)	(0.141)	(0.183)	(0.012)	(0.011)
(B) Uncertain college-goers					
Enrolled in target college	0.257***	0.257***	0.417***	0.019	0.019
	(0.099)	(0.099)	(0.142)	(0.007)	(0.007)
Enrolled in 4-year college	0.531**	0.540**	0.587*	0.036	0.038
	(0.248)	(0.245)	(0.320)	(0.018)	(0.017)
B.A. completion rate	0.473***	0.463***	0.540***	0.034	0.035
	(0.150)	(0.147)	(0.202)	(0.010)	(0.010)
Peer quality	0.699***	0.654***	0.871**	0.051	0.053
	(0.260)	(0.253)	(0.352)	(0.019)	(0.018)

Notes: Heteroskedasticity robust standard errors clustered by family are in parentheses in columns 1 - 4. (* p<.10 ** p<.05 *** p<.01). In column (5), standard errors are computed according to Kolésar & Rothe 2018. Each coefficient in columns 1-3 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. Coefficients in columns 4 and 5 are reduced form estimates of an older sibling's admission to the target colleges on younger siblings' college choices. 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. Columns 1, 4 and 5 use a bandwidth of 93 SAT points and a donut specification that exclude observations exactly at the threshold. Column 2 adds controls, including gender, race, income and parental education, to the regression in column 1. Column 5 contains standard errors as described by Kolésar & Rothe (2018). 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.II: Additional Robustness Checks in the U.S. Sample

	Baseline specification (1)	Quadratic Polynomial (2)	Triangular Kernel (3)	Varying Slope (4)
(A) All students				
Enrolled in target college	0.172***	0.175*	0.171***	0.174***
	(0.054)	(0.098)	(0.060)	(0.054)
Enrolled in 4-year college	0.230*	0.250	0.231	0.235*
	(0.132)	(0.242)	(0.147)	(0.131)
B.A. completion rate	0.180**	0.211	0.186**	0.178**
	(0.080)	(0.147)	(0.089)	(0.079)
Peer quality	0.316**	0.256	0.290*	0.309**
	(0.148)	(0.270)	(0.166)	(0.147)
(B) Uncertain college-goers				
Enrolled in target college	0.257***	0.340*	0.269**	0.258**
	(0.099)	(0.192)	0.106	(0.101)
Enrolled in 4-year college	0.531**	0.321	0.419	0.559**
	(0.248)	(0.443)	(0.262)	(0.252)
B.A. completion rate	0.473***	0.319	0.391**	0.496***
	(0.150)	(0.260)	(0.155)	(0.154)
Peer quality	0.699***	0.334	0.543**	0.741***
	(0.260)	(0.453)	(0.270)	(0.266)

Notes: Heteroskedasticity robust standard errors clustered by family 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 93 SAT points and a donut hole specification that exclude observations on the threshold itself. Column 2 includes a quadratic polynomial for the distance of a student's score from the cutoff. Column 3 uses a triangular kernel instead of a uniform one. Column 4 allows the slope of the running variable to be different for each admissions cutoff. 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 Spillovers on Applications to and Enrollment in Older Sibling's Target College

	Applies in the 1st preference		Applies in any preference		Enrolls	
	P1 (1)	P2 (2)	P1 (3)	P2 (4)	P1 (5)	P2 (6)
Panel A - Chile						
Older sibling enrolls	0.067***	0.060***	0.076***	0.068***	0.038***	0.031**
	(0.012)	(0.015)	(0.014)	(0.017)	(0.011)	(0.013)
Older sibling above cutoff	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)
Older sibling enrolls (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						
Older sibling enrolls	0.075***	0.070**	0.109***	0.102***	0.084***	0.090***
	(0.019)	(0.023)	(0.019)	(0.024)	(0.018)	(0.023)
Older sibling above cutoff	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)
Older sibling enrolls (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						
Older sibling enrolls	0.122***	0.110***	0.132***	0.124***	0.049***	0.040***
	(0.008)	(0.007)	(0.011)	(0.010)	(0.005)	(0.004)
Older sibling above cutoff	0.033***	0.030***	0.035***	0.033***	0.013***	0.011***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.001)	(0.001)
First stage	0.268***	0.270***	0.268***	0.270***	0.268***	0.270***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Older sibling enrolls (Triangular Kernel)	0.143***	0.126***	0.149***	0.138***	0.058***	0.048***
	(0.008)	(0.007)	(0.011)	(0.010)	(0.005)	(0.005)
Observations	378466	903783	378466	903783	378466	903783
Counterfactual mean	0.087	0.082	0.206	0.196	0.032	0.030
Bandwidth	0.360	0.933	0.360	0.933	0.360	0.933
Kleibergen-Paap Wald F-statistic	7215.227	8815.583	7215.227	8815.583	7215.227	8815.583

Notes: The first and second row of each panel report 2SLS and reduced form estimates. The third row presents the first stage of the 2SLS, and the fourth reports the results of a 2SLS specification that uses a triangular kernel to give more weight to observations close to the cutoff. All specifications use the same set of controls and bandwidths as in Table **??**. In addition, we report models with controls for quadratic polynomials of the running variables in the columns labelled "P2". Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C.IV: Sibling Spillovers on Applications to and Enrollment in Older Sibling's Target College-Major

	Applies in the 1st preference		Applies in any preference		Enrolls	
	P1 (1)	P2 (2)	P1 (3)	P2 (4)	P1 (5)	P2 (6)
			Panel A - Chile			
Older sibling enrolls	0.012***	0.014***	0.023***	0.024***	0.006***	0.007**
	(0.003)	(0.004)	(0.005)	(0.006)	(0.002)	(0.003)
Older sibling above cutoff	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)
Older sibling enrolls (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			
Older sibling enrolls	0.015***	0.014**	0.036***	0.038***	0.013**	0.015**
	(0.004)	(0.005)	(0.009)	(0.011)	(0.004)	(0.005)
Older sibling above cutoff	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)
Older sibling enrolls (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			
Older sibling enrolls	0.020***	0.017***	0.031***	0.025***	0.005***	0.004***
	(0.002)	(0.002)	(0.005)	(0.004)	(0.001)	(0.001)
Older sibling above cutoff	0.006***	0.005***	0.009***	0.007***	0.001***	0.001***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
First stage	0.287***	0.294***	0.287***	0.294***	0.287***	0.294***
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
Older sibling enrolls (Triangular Kernel)	0.025***	0.019***	0.031***	0.028***	0.006***	0.005***
	(0.003)	(0.002)	(0.005)	(0.004)	(0.002)	(0.001)
Observations	482220	1235550	482220	1235550	482220	1235550
Counterfactual mean	0.011	0.009	0.053	0.048	0.003	0.003
Bandwidth	0.386	1.130	0.386	1.130	0.386	1.130
Kleibergen-Paap Wald F-statistic	10406.511	14120.902	10406.511	14120.902	10406.511	14120.902

Notes: The first and second row of each panel report 2SLS and reduced form estimates. The third row presents the first stage of the 2SLS, and the fourth reports the results of a 2SLS specification that uses a triangular kernel to give more weight to observations close to the cutoff. All specifications use the same set of controls and bandwidths as in Table **??**. In addition, we report models with controls for quadratic polynomials of the running variables in the columns labelled "P2". Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C.V: Sibling Spillovers on Applications and Enrollment in Older Sibling's Target College - Different Slope for each Admission Cutoff

	Applies in the 1st preference		**Applies in any preference**		**Enrolls**	
	P1	P2	P1	P2	P1	P2
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
Older sibling enrolls	0.060***	0.056***	0.082***	0.090***	0.054***	0.052***
	(0.015)	(0.020)	(0.018)	(0.023)	(0.013)	(0.017)
Older sibling above cutoff	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			
Older sibling enrolls	0.080**	0.081*	0.107***	0.115**	0.085***	0.096**
	(0.024)	(0.037)	(0.025)	(0.038)	(0.023)	(0.036)
Older sibling above cutoff	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			
Older sibling enrolls	0.147***	0.145***	0.150***	0.149***	0.061***	0.059***
	(0.012)	(0.011)	(0.015)	(0.015)	(0.007)	(0.007)
Older sibling above cutoff	0.039***	0.038***	0.040***	0.040***	0.016***	0.016***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.002)	(0.002)
Observations	367494	891217	367494	891217	367494	891217
Counterfactual mean	0.087	0.082	0.206	0.196	0.032	0.030
Bandwidth	0.360	0.933	0.360	0.933	0.360	0.933
Kleibergen-Paap Wald F-statistic	3557.006	3931.993	3557.006	3931.993	3557.006	3931.993

Notes: The reported specifications use the same set of controls and bandwidths as in Table **??**, but we allow the slope of the running variable to be different for each admission cutoff. In addition, we report models with quadratic polynomials of the running variables in the columns labelled "P2". Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C.VI: Sibling Spillovers on Applications and Enrollment in Older Sibling's Target College-Major - Different Slope for each Admission Cutoff

	Applies in the 1st preference		**Applies in any preference**		**Enrolls**	
	P1	P2	P1	P2	P1	P2
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
Older sibling enrolls	0.013***	0.015***	0.025***	0.025***	0.007***	0.007*
	(0.003)	(0.004)	(0.005)	(0.007)	(0.003)	(0.003)
Older sibling above cutoff	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			
Older sibling enrolls	0.016**	0.016*	0.044***	0.051***	0.014**	0.017**
	(0.005)	(0.007)	(0.010)	(0.013)	(0.005)	(0.006)
Older sibling above cutoff	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			
Older sibling enrolls	0.026***	0.022***	0.040***	0.031***	0.008***	0.007***
	(0.004)	(0.003)	(0.007)	(0.006)	(0.002)	(0.002)
Older sibling above cutoff	0.008***	0.006***	0.011***	0.009***	0.002***	0.002***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)
Observations	470259	1222427	470259	1222427	470259	1222427
Counterfactual mean	0.011	0.009	0.054	0.049	0.003	0.003
Bandwidth	0.386	1.130	0.386	1.130	0.386	1.130
Kleibergen-Paap Wald F-statistic	5767.689	7091.725	5767.689	7091.725	5767.689	7091.725

Notes: The reported specifications use the same set of controls and bandwidths as in Table **??**, but we allow the slope of the running variable to be different for each admission cutoff. In addition, we report models with quadratic polynomials of the running variables in the columns labelled "P2". Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C.VII: Sibling Spillovers on Applications to and Enrollment in Older Sibling's Target College and Target College-Major (Controlling for Covariates)

	Older Sibling's Target College			**Older Sibling's Target College-Major**		
	Applies in the 1st preference (1)	**Applies in any preference** (2)	**Enrolls** (3)	**Applies in the 1st preference** (4)	**Applies in any preference** (5)	**Enrolls** (6)
			Panel A - Chile			
Older sibling enrolls	0.068***	0.076***	0.038***	0.012***	0.023***	0.006**
	(0.012)	(0.015)	(0.011)	(0.003)	(0.005)	(0.002)
Older sibling above cutoff	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	85328	85328	85328	168646	168646	168646
Counterfactual mean	0.22	0.45	0.13	0.02	0.06	0.01
Bandwidth	12.500	12.500	12.500	18.000	18.000	18.000
Kleibergen-Paap Wald F-statistic	5532.71	5532.71	5532.71	14624.31	14624.31	14624.31
			Panel B - Croatia			
Older sibling enrolls	0.074***	0.114***	0.081***	0.016***	0.038***	0.014***
	(0.019)	(0.020)	(0.019)	(0.005)	(0.009)	(0.004)
Older sibling above cutoff	0.062***	0.095***	0.067****	0.013***	0.031***	0.011***
	(0.016)	(0.017)	(0.016)	(0.004)	(0.007)	(0.003)
Observations	12216	12216	12216	34711	34711	34711
Counterfactual mean	0.29	0.52	0.25	0.02	0.11	0.02
Bandwidth	80.000	80.000	80.000	80.000	80.000	80.000
Kleibergen-Paap Wald F-statistic	5884.61	5884.61	5884.61	13631.25	13631.25	13631.25
			Panel C - Sweden			
Older sibling enrolls	0.125***	0.136***	0.050***	0.021***	0.033***	0.005***
	(0.008)	(0.011)	(0.005)	(0.003)	(0.005)	(0.001)
Older sibling above cutoff	0.033***	0.036***	0.013***	0.006***	0.009***	0.001***
	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)	(0.000)
Observations	375488	375488	375488	478421	478421	478421
Counterfactual mean	0.087	0.206	0.033	0.011	0.053	0.003
Bandwidth	0.360	0.360	0.360	0.386	0.386	0.386
Kleibergen-Paap Wald F-statistic	7162.748	7162.748	7162.748	10332.521	10332.521	10332.521

Notes: The reported specifications use the same set of controls and bandwidths as in Table **??**. In addition, we add a vector of individual level controls in each setting. In Chile, these controls include the gender of both siblings, the size of the family group, the number if siblings in higher education, household income level, parental education, health insurance type and administrative dependence of the high school in which the older sibling completed secondary education (i.e. public, voucher, private). In Croatia we control for the gender of both siblings, for the number of siblings and for the size of the city of origin. Finally, in Sweden, we control for gender, household size, immigrant status and origin, disposable income and parental education. Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C.VIII: Sibling Spillovers on Applications to and Enrollment in Older Sibling's Target College and Target College-Major (Donut)

	Older Sibling's Target College			**Older Sibling's Target College-Major**		
	Applies in the first preference	**Applies in any preference**	**Enrolls**	**Applies in the first preference**	**Applies in any preference**	**Enrolls**
	(1)	(2)	(3)	(4)	(5)	(6)
				Panel A - Chile		
Older sibling enrolls	0.067***	0.078***	0.043***	0.012***	0.023***	0.006***
	(0.013)	(0.015)	(0.011)	(0.003)	(0.005)	(0.002)
Older sibling above cutoff	0.032***	0.038***	0.021***	0.006***	0.012***	0.003***
	(0.006)	(0.007)	(0.005)	(0.001)	(0.003)	(0.001)
Observations	84708	84708	84708	168286	168286	168286
Counterfactual mean	0.22	0.45	0.13	0.02	0.06	0.01
Bandwidth	12.500	12.500	12.500	18.000	18.000	18.000
Kleibergen-Paap Wald F-statistic	5179.10	5179.10	5179.10	13978.84	13978.84	13978.84
				Panel B - Croatia		
Older sibling enrolls	0.072***	0.113***	0.078***	0.016***	0.038***	0.014***
	. (0.019)	(0.020)	(0.019)	(0.005)	(0.009)	(0.004)
Older sibling above cutoff	0.060***	0.094***	0.065***	0.013***	0.031***	0.011***
	(0.016)	(0.017)	(0.016)	(0.004)	(0.007)	(0.003)
Observations	12216	12216	12216	34710	34710	34710
Counterfactual mean	0.29	0.52	0.25	0.02	0.11	0.02
Bandwidth	80.000	80.000	80.000	80.000	80.000	80.000
Kleibergen-Paap Wald F-statistic	5900.74	5900.74	5900.74	13634.55	13634.55	13634.55
				Panel C - Sweden		
Older sibling enrolls	0.036**	0.034*	0.013	0.013***	0.029***	0.002
	(0.014)	(0.020)	(0.009)	(0.004)	(0.009)	(0.003)
Older sibling above cutoff	0.008**	0.007*	0.003	0.003***	0.007***	0.001
	(0.003)	(0.004)	(0.002)	(0.001)	(0.002)	(0.001)
Observations	305669	305669	305669	394716	394716	394716
Counterfactual mean	0.089	0.207	0.033	0.011	0.054	0.003
Bandwidth	0.360	0.360	0.360	0.386	0.386	0.386
Kleibergen-Paap Wald F-statistic	2046.843	2046.843	2046.843	3162.516	3162.516	3162.516

Notes: The reported specifications use the same set of controls and bandwidths as in Table **??**. Observations exactly at the cutoff are excluded from the estimation sample. Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C.IX: Sibling Spillovers on Applications to and Enrollment in Older Sibling's Target College and Target College-Major (Target $\times$ Next Best College-Major Fixed Effects)

	Older Sibling's Target College			**Older Sibling's Target College-Major**		
	Applies in the first preference	**Applies in any preference**	**Enrolls**	**Applies in the first preference**	**Applies in any preference**	**Enrolls**
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
Older sibling enrolls	0.041**	0.056***	0.034**	0.018***	0.029***	0.006*
	(0.018)	(0.021)	(0.016)	(0.004)	(0.007)	(0.003)
Older sibling above cutoff	0.019**	0.026***	0.016**	0.009***	0.015***	0.003*
	(0.009)	(0.010)	(0.007)	(0.002)	(0.004)	(0.002)
Observations	64886	64886	64886	128112	128112	128112
Counterfactual mean	0.230	0.460	0.140	0.020	0.070	0.010
Bandwidth	12.500	12.500	12.500	18.000	18.000	18.000
Kleibergen-Paap Wald F-statistic	2639.50	2639.50	2639.50	5003.480	5003.480	5003.480
			Panel B - Croatia			
Older sibling enrolls	0.053	0.106***	0.078**	0.012	0.038***	0.011
	(0.033)	(0.032)	(0.033)	(0.008)	(0.014)	(0.007)
Older sibling above cutoff	0.047	0.094***	0.069***	0.010	0.033***	0.010
	(0.030)	(0.028)	(0.029)	(0.006)	(0.012)	(0.006)
Observations	6743	6743	6743	23076	23076	23076
Counterfactual mean	0.355	0.588	0.319	0.033	0.144	0.027
Bandwidth	80.000	80.000	80.000	80.000	80.000	80.000
Kleibergen-Paap Wald F-statistic	2517.738	2517.738	2517.738	10630.120	10630.120	10630.120
			Panel C - Sweden			
Older sibling enrolls	0.135***	0.126***	0.056***	0.026***	0.033***	0.009***
	(0.013)	(0.017)	(0.008)	(0.004)	(0.009)	(0.002)
Older sibling above cutoff	0.034***	0.032***	0.014***	0.007***	0.009***	0.002***
	(0.003)	(0.004)	(0.002)	(0.001)	(0.002)	(0.001)
Observations	303452	303452	303452	372778	372778	372778
Counterfactual mean	0.088	0.204	0.033	0.011	0.052	0.003
Bandwidth	0.360	0.360	0.360	0.386	0.386	0.386
Kleibergen-Paap Wald F-statistic	2982.010	2982.010	2982.010	3770.740	3770.740	3770.740

Notes: The reported specification s use the same set of controls and bandwidths as in Table **??**, but we include fixed effects for each target and counterfactual admission cutoff combination. Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C.X: Sibling Spillovers on Applications to and Enrollment in Older Sibling's Target College (Fixing Target and Next Best Option Major)

	Applies in the 1st preference		**Applies in any preference**		**Enrolls**	
	P1	P2	P1	P2	P1	P2
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
Older sibling enrolls	0.087***	0.075**	0.073***	0.058*	0.052***	0.054**
	(0.020)	(0.029)	(0.024)	(0.035)	(0.018)	(0.027)
Older sibling above cutoff	0.059***	0.045**	0.050***	0.035*	0.036***	0.033**
	(0.014)	(0.018)	(0.016)	(0.021)	(0.013)	(0.016)
Observations	15803	19203	15803	19203	15803	19203
Counterfactual mean	0.20	0.20	0.44	0.43	0.13	0.13
Bandwidth	12.500	20.500	12.500	20.500	12.500	20.500
Kleibergen-Paap Wald F-statistic	3197.65	1377.94	3197.65	1377.94	3197.65	1377.94
			Panel B - Croatia			
Older sibling enrolls	0.080**	0.067	0.111***	0.105***	0.065	0.064
	(0.040)	(0.047)	(0.042)	(0.050)	(0.040)	(0.048)
Older sibling above cutoff	0.071**	0.060	0.099***	0.093**	0.058	0.056
	(0.036)	(0.019)	(0.037)	(0.020)	(0.036)	(0.043)
Observations	3100	3980	3100	3980	3100	3980
Counterfactual mean	0.31	0.31	0.54	0.55	0.27	0.27
Bandwidth	80.000	120.000	80.000	120.000	80.000	120.000
Kleibergen-Paap Wald F-statistic	2779.47	2080.48	2779.47	2080.48	2779.47	2080.48
			Panel C - Sweden			
Older sibling enrolls	0.098***	0.104***	0.100***	0.106***	0.030***	0.034***
	(0.013)	(0.012)	(0.019)	(0.017)	(0.008)	(0.008)
Older sibling above cutoff	0.031***	0.033***	0.032***	0.034***	0.010***	0.011***
	(0.004)	(0.004)	(0.006)	(0.005)	(0.003)	(0.002)
Observations	101522	192791	101522	192791	101522	192791
Counterfactual mean	0.085	0.081	0.215	0.207	0.031	0.029
Bandwidth	0.360	0.933	0.360	0.933	0.360	0.933
Kleibergen-Paap Wald F-statistic	2635.556	3190.000	2635.556	3190.000	2635.556	3190.000

Notes: The table shows estimates based on the sample of older siblings who are on an admission margin such that their counterfactual alternative is the same major but at a different college. The reported specifications use the same set of controls and bandwidths as in Table III in the paper. In addition, we report models with quadratic polynomials of the running variables in the columns labelled "P2". Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C.XI: Sibling Effects on Applications to and Enrollment in Older Sibling's Target College-Major (Fixing Target and Next Best Option College)

	Applies in the 1st preference		**Applies in any preference**		**Enrolls**	
	P1	P2	P1	P2	P1	P2
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
Older sibling enrolls	0.007	0.013*	0.002	0.005	0.005	0.009
	(0.006)	(0.007)	(0.010)	(0.012)	(0.005)	(0.006)
Older sibling above cutoff	0.004	0.007*	0.001	0.003	0.003	0.005
	(0.004)	(0.004)	(0.006)	(0.007)	(0.003)	(0.003)
Observations	41432	64079	41432	64079	41432	64079
Counterfactual mean	0.03	0.03	0.09	0.09	0.02	0.02
Bandwidth	12.500	20.500	12.500	20.500	12.500	20.500
Kleibergen-Paap Wald F-statistic	4619.84	3137.99	4619.84	3137.99	4619.84	3137.99
			Panel B - Croatia			
Older sibling enrolls	0.005	0.007	0.024**	0.025*	0.008	0.011
	(0.006)	(0.007)	(0.012)	(0.014)	(0.006)	(0.007)
Older sibling above cutoff	0.004	0.006	0.020**	0.021*	0.007	0.009
	(0.005)	(0.006)	(0.010)	(0.012)	(0.005)	(0.006)
Observations	22197	29354	22197	29354	22197	29354
Counterfactual mean	0.02	0.02	0.12	0.12	0.02	0.02
Bandwidth	80.000	120.000	80.000	120.000	80.000	120.000
Kleibergen-Paap Wald F-statistic	8148.29	5890.01	8148.29	5890.01	8148.29	5890.01
			Panel C - Sweden			
Older sibling enrolls	0.009	0.011**	-0.005	-0.003	-0.001	-0.001
	(0.007)	(0.005)	(0.013)	(0.010)	(0.004)	(0.003)
Older sibling above cutoff	0.003	0.004**	-0.002	-0.001	-0.001	-0.001
	(0.003)	(0.002)	(0.005)	(0.004)	(0.002)	(0.001)
Observations	73585	169560	73585	169560	73585	169560
Counterfactual mean	0.018	0.017	0.077	0.073	0.007	0.006
Bandwidth	0.386	1.130	0.386	1.130	0.386	1.130
Kleibergen-Paap Wald F-statistic	2170.078	3296.574	2170.078	3296.574	2170.078	3296.574

Notes: The table shows estimates based on the sample of older siblings who are on an admission margin such that their counterfactual alternative is in the same college. The reported specifications use the same set of controls and bandwidths as in Table III in the paper. In addition, we report models with quadratic polynomials of the running variables in the columns labelled "P2". Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C.XII: Sibling Spillovers on Applications to and Enrollment in Older Sibling's Target College-Major - Eligible Younger Siblings whose Older Siblings' Target and Next Best Option are taught in the Same College

	Applies in the 1st preference (1)	**Applies in any preference** (2)	**Enrolls** (3)
		Panel A - Chile	
Older sibling enrolls	0.031*	0.093***	0.038*
	(0.017)	(0.032)	(0.020)
Older sibling above cutoff	0.017*	0.051***	0.021*
	(0.009)	(0.017)	(0.011)
Observations	9042	9042	9042
Counterfactual mean	0.04	0.14	0.04
Bandwidth	18.000	18.000	18.000
Kleibergen-Paap Wald F-statistic	714.78	714.78	714.78
		Panel B - Croatia	
Older sibling enrolls	0.020*	0.033	0.023**
	(0.011)	(0.023)	(0.010)
Older sibling above cutoff	0.016*	0.026	0.018**
	(0.009)	(0.018)	(0.008)
Observations	6513	6513	6513
Counterfactual mean	0.02	0.15	0.02
Bandwidth	80.000	80.000	80.000
Kleibergen-Paap Wald F-statistic	2400.02	2400.02	2400.02
		Panel C - Sweden	
Older sibling enrolls	0.040	0.074	0.023
	(0.030)	(0.054)	(0.025)
Older sibling above cutoff	0.015	0.027	0.008
	(0.011)	(0.020)	(0.009)
Observations	10106	10106	10106
Counterfactual mean	0.045	0.184	0.030
Bandwidth	0.386	0.386	0.386
Kleibergen-Paap Wald F-statistic	343.503	343.503	343.503

Notes: The table presents estimates based on the sample of older siblings who are on an admission margin such that their counterfactual alternative is in the same college. We include only those younger siblings who are eligible for the college-major chosen by their older sibling. The reported specifications use the same set of controls and bandwidths as in Table **??**. Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

D Additional Results

This Section presents additional results that either complement some of the analyses presented in the main body of the paper or extend them in new directions. First, we show how older siblings' enrollment in their preferred choices change at the cutoff and how this affects the characteristics of the college or college-major that they attend. Second, we extend the discussion on heterogeneous effects by age and gender. Third, we show how the characteristics of the programs in which younger siblings enroll are affected by older siblings. Fourth, we study heterogeneous effects on academic performance by age difference. Fifth, we investigate how the effects change depending on socioeconomic status and on exposure to the older siblings' college. We conclude by presenting additional robustness checks.

D.1 Older Siblings' Higher Education Trajectories and Spillovers on Enrollment in Any College

In this Section we discuss how crossing the admission threshold changes the higher education trajectories of older siblings. As shown in Table D1, being above the hidden admission cutoffs in the U.S. increases the likelihood of enrolling in any four-year college by 3 pp and the probability of enrolling in one's target college by 8.3 pp. This last figure is the first stage coefficient that we use along the paper for the U.S. outcomes. In Chile, Croatia and Sweden, the applications are made at the college-major level. Crossing the admission threshold in these countries increases the probability of enrolling in one's target college-major by 53.6 pp, 82.6 pp and 29.1 pp respectively.

Table D2 shows how the characteristics of the college attended by older siblings in the U.S. changes at the admission cutoff. The results in this table correspond to 2SLS estimates in which actual enrollment is instrumented with an indicator of being above the target college's admission cutoff. According to these results, scoring above one of these admission cutoffs increases the likelihood of enrolling in any four-year college, and improves the quality of the college attended. In addition, it makes it more likely to enroll in a college 50 or more miles from home.

In Chile, Croatia and Sweden, being admitted into their target college-major also seems to improve the quality of the institution in which older siblings enrolls. According to the results presented in Table D3 older siblings who enroll in their target college-major have better peers and higher expected earnings. However, these average differences are considerably smaller than in the case of the U.S..

When focusing on younger siblings, Table D3 shows that having an older sibling enrolling in her target college-major does not generate a significant change on the average quality of the higher education trajectories followed by younger siblings. This result is different from the one that we find for the U.S. (discussed in the paper). This is not surprising because while in the U.S. crossing the threshold moves older and younger siblings from two- to four-year colleges, in the other countries older siblings usually experience a much smaller change at the cutoff. In addition, in the paper we show that individuals follow their older siblings when the next best option is worse, but also when it is better. This average estimate combines both cases.

Finally, in Table D4 we show that although being admitted into their target college-major increases the probability of attending any college among older siblings, this does not translate into any relevant change on the probability of attending college among younger siblings. This in part reflects the fact that some of the older siblings not admitted to one of their preferred options the first time, could try again and gain admission to another college before their younger siblings are ready to apply.

D.2 Sibling Spillovers on College and College-Major Choice by Age and Gender

In the paper we present a heterogeneity analysis that investigates whether the strength of sibling spillovers on applications vary depending on age difference and gender. This section extends these results by presenting a more detailed discussion on how the gender of siblings affect applications, and also by looking at heterogeneity in enrollment along this dimension.

Table D5 summarizes our findings when focusing on applications. Panels A and B are the same that we present in the paper. Now we add two new panels where we explore heterogeneity by gender in greater detail. Panel C focuses on sibling-pairs with a male older sibling, while Panel D focuses on sibling pairs with a female older sibling. In both cases, the interaction is a dummy variable that takes the value of 1 if both siblings are of the same gender.

In terms of applications to any four-year college, our results suggest that both males and females are more likely to be followed by their younger brothers. Indeed, younger sisters do not seem to follow their older sisters in this dimension. Note however, that this is something that we can only investigate in the U.S., a setting where splitting the sample reduces our sample significantly. Thus, we interpret this finding with caution.

When focusing on the choice of college we, find that siblings' gender does not seem to make a huge difference in the strength of the spillovers found. Older siblings are followed to a similar extent by both younger brothers and sisters. As before, when focusing on the older sisters sample we find a negative coefficient for the interaction in all the settings, but it is small and never significant.

We find a more significant difference when focusing on applications to the exact same college-major combination. In this case we find that males are more likely to be followed by their younger brothers than by their younger sisters. Females on the other hand seem to be followed to a similar extent by younger sisters and brothers.

Table D6 replicates these results but focusing on enrollment instead of applications. The results follow the same pattern discussed above.

We conclude this section by investigating whether having an older sister admitted into a STEM major increases the probability that a younger sister applies to and enrolls in this type of programs. To study this, we create a new sample in which we only keep females whose older sisters have a STEM major as their target major, but a major in a different field as their next best option. This guarantees that crossing the threshold makes a difference in the type of program followed by the older sister. Since the number of females pursuing STEM degrees is relatively low in all the settings that we study, we are left with very few observations in our sample. Since in Croatia we already had fewer observations to start with, we leave this country out of the analysis because we end with too few observations.

Table D7 presents these results. The coefficients suggest that having an older sister attend a STEM major slightly increases the probability of applying to and enrolling in a STEM major as well. However, our estimates are not precise enough to rule out that they are equal to zero.

D.3 Sibling Spillovers by Differences between Older Sibling's Target and Next Best Options

As shown in Figure D1, older siblings' counterfactual options are often very similar. However, we find that younger siblings not only change their application decisions when the older siblings is on the margin of very similar alternatives, but also when the differences between these options are large. Similarly, we show that younger siblings' responses do not change much with the characteristics of the target college-major of the older sibling. Here, we replicate these results but focus on the college and college-major in which younger siblings enroll.

Table D8 shows that as in the paper the effects that we find in terms of the college in which younger siblings enroll do not change much by the difference between older siblings' target and next best options. Older siblings are followed to their college when their next best option is worse, but also when it is better. When looking at the exact college-major in which younger siblings enroll we lose precision. The average effects in this margin were already small, but after splitting the sample the coefficients become very noisy.

In Chile, the effects seem to be stronger when the differences in expected earnings and peer quality are positive. However, in the case of expected earnings, we also find that even when the difference is negative individuals are followed to the same college-major by their siblings. In terms of retention rates, we only find a significant effect when the difference is negative (i.e. when dropout rates are higher in the target than in the next best option). In Croatia, where the first stage is stronger, we do find that individuals are followed to the same college-major independently of the size of the difference in peer quality. Finally, in Sweden the coefficients become small and non-significant in all categories.[20].

It is worth noting that while applications reflect individual preferences, enrollment is a mix between preferences and availability of places.

When looking at heterogeneity by the quality of the target choice of older siblings, the results in D9 indicate that, in terms of enrollment, responses are stronger when the target choice is better.

[20]The sample used in this analyses is a subset of our baseline sample. Here we only keep older siblings for whom we observe the best next option and for whom we are able to compute the characteristics along which we implement the heterogeneity analyses. Thus, sample sizes are considerably smaller

The fact that we observe a positive gradient in the effects and not in applications might reflect that individuals whose older siblings apply to programs in the top quartile understand better to which options they are more likely to be admitted.

D.4 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 D10 summarizes these results. We find no increases on younger siblings academic performance no matter the age difference with their older sibling.

D.5 Sibling Spillovers by SES and Exposure to Older Sibling's College

In the paper we show that sibling spillovers are larger for individuals with low probabilities of enrolling in college in the U.S.. Here we investigate heterogeneity by socioeconomic characteristics of individuals in Chile and Sweden. We find no evidence of heterogeneous effects along this dimension. We do find, however, that in Chile and Sweden the effects seem to be stronger for individuals who have the least exposure 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.

Table D11 shows that the socioeconomic status of individuals does not generate a significant difference in the effects that we find on the probabilities of applying to and enrolling in older siblings' target college and college-major. It is worth highlighting that these results are not directly comparable with the ones we find for the U.S.. As discussed in the main body of the paper, the individuals we observe in Chile and Sweden are positively selected. They come from families where at least one child is on the margin of being admitted to a selective program. This positive selection means that we observe few uncertain college-goers in these countries .

Table D12 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.

D.6 Additional Robustness Checks

This Section presents additional robustness checks. We first investigate whether the sibling spillovers that we document in the paper are driven by a change in geographic preferences. Then, we show that focusing on the older sibling closest in age instead of the oldest sibling of the family does not make a big difference in the size of the effects. We conclude by presenting additional specifications looking at sibling spillovers on the choice of major that confirm the results we present in the paper.

D.6.1 Sibling Spillover on College Choices and Location Preferences

One hypothesis that may explain the large effects that we find on the choice of college is that they could at least partly reflect 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.[21]

Table D13 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 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 body of the paper.

D.6.2 Sibling Spillovers on College and College-Major Choice - Closest Siblings

In this Section we replicate the results presented in the main body of the paper, but this time we define the treatment at the closer older sibling level instead of at the oldest sibling level. Table

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

D14 summarizes the results in all the countries that we investigate. The coefficients that we find are very similar to the ones we discuss in the paper. We further complement these analyses by focusing only on first and second children. We implement this last exercise only in Sweden where we are able to identify all the family members of each individual. As shown in Table D15 we once more find estimates that are very similar to the ones presented in the paper.

D.6.3 Sibling Spillovers on Major Choice - Additional Specifications

Finally, in this Section we present additional specifications that look at sibling spillovers on the choice of major. Apart from the results presented in the main body of the paper, Table D16 presents results from specifications that control for a second order polynomial of the running variable, and also from specifications in which we use a triangular instead of a uniform kernel. As in the paper, we do not find evidence of sibling spillovers in this dimension.

Table D1: Older Siblings' Enrollment Probabilities in Any College, in their Target College and in their Target College-Major

	Any 4-year College	**Target College**	**Target College-Major**		
	US	US	CHI	CRO	SWE
	(1)	(2)	(3)	(4)	(5)
Older Sibling Above Cutoff = 1	0.030***	0.083***	0.536***	0.826***	0.291***
	(0.011)	(0.007)	(0.004)	(0.007)	(0.003)
Observations	44190	44190	170886	36757	331178
Bandwidth	93.000	93.000	18.000	80.000	0.386
Counterfactual mean	0.693	0.140	0.118	0.002	0.056

Notes: All specifications in the table control for a linear polynomial of older siblings' application score centered around the admission cutoff of the target choice. Fixed effects for older siblings' application year, each admission cutoff and younger siblings' birth year are included. Bandwidths are the same used Tables III and IV in the paper. Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D2: 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**	0.239***	0.440***	2.688	0.298**
	(0.128)	(0.116)	(0.070)	(0.113)	(2.193)	(0.136)
Counterfactual mean	0.64	0.28	0.43	-0.06	10.93	0.26
Uncertain college-goers	0.386	-0.477**	0.328***	0.634***	2.263	0.294
	(0.237)	(0.225)	(0.125)	(0.212)	(3.647)	(0.237)
Counterfactual mean	0.61	0.48	0.30	-0.38	11.26	0.25
Probable college-goers	0.268*	-0.117	0.148*	0.296**	2.014	0.248
	(0.155)	(0.139)	(0.087)	(0.138)	(2.820)	(0.170)
Counterfactual mean	0.73	0.12	0.54	0.15	11.70	0.32

Notes: 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 93 SAT points, a donut 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. Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D3: Change in College-Major Characteristics of Older and Younger Siblings

	Chile			**Croatia**	**Sweden**		
	Expected Earnings (USD 000)	Peer Quality (z-scores)	First Year Retention	Peer Quality (z-scores)	Expected Earnings (USD 000)	Peer Quality (z-scores)	First Year Retention
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A - Changes in the characteristics of the college-major in which older siblings enroll							
Older Sibling Enrolls	2.004***	0.121***	0.016***	0.313***	0.706***	0.129***	0.022***
	(0.019)	(0.007)	(0.001)	(0.021)	(0.132)	(0.005)	(0.003)
Observations	113697	113697	113697	30191	115511	130764	122401
Kleibergen-Paap Wald F-statistic	16436.429	16436.429	16436.429	12520.6599	4798.875	5306.640	5162.964
Counterfactual mean	38.317	0.874029	0.867621	-0.267	40.815	0.562	0.777
Panel B - Changes in the characteristics of the college-major in which younger siblings enroll							
Older Sibling Enrolls	-0.504*	-0.010	0.003	-0.045	-0.250	-0.011	-0.008
	(0.295)	(0.008)	(0.002)	(0.028)	(0.508)	(0.023)	(0.011)
Observations	102,484	102,484	102,484	31179	46149	56968	40891
Kleibergen-Paap Wald F-statistic	14057.194	14057.194	14057.194	12851.004	1148.264	1416.464	1020.207
Counterfactual mean	37.137	0.734	0.851	-0.071	41.678	0.484	0.724

Notes: These specifications investigate how the characteristics of the programs in which older and younger siblings enroll change when the older sibling is admitted in her target college-major. We use the same set of controls and bandwidths used in the 2SLS specifications described in the main body of the paper. Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D4: Probability of Enrolling in any College by Older Siblings' Admission to their Target College-Major

	Younger sibling enrolls in any college		**Older sibling enrolls in any college**	
	P1	P2	P1	P2
	(1)	(2)	(3)	(4)
		Panel A - Chile		
Older sibling above cutoff	-0.010*	-0.009	0.044***	0.044***
	(0.005)	(0.007)	(0.005)	(0.007)
Observations	114424	136355	78655	93826
Counterfactual mean	0.760	0.757	0.804	0.799
Bandwidth	12.000	14.500	12.000	14.500
		Panel B - Croatia		
Older sibling above cutoff	-0.003	0.000	0.123***	0.131***
	(0.007)	(0.008)	(0.007)	(0.008)
Observations	36757	48611	36757	48611
Counterfactual mean	0.90	0.90	0.88	0.85
Bandwidth	80	120	80	120
		Panel C - Sweden		
Older sibling above cutoff	-0.002	-0.005*	0.110***	0.109***
	(0.003)	(0.003)	(0.004)	(0.003)
Observations	482220	1235550	482220	1235550
Counterfactual mean	0.590	0.567	0.468	0.404
Bandwidth	0.386	1.130	0.386	1.130

Notes: The table presents estimates for the effect of older siblings' marginal admission in a target college-major on their own and on their younger siblings' probability of enrolling anywhere. We use the same set of controls and bandwidths used in the 2SLS specifications described in the main body of the paper. Columns 3 and 4 also include quadratic polynomials of the running variable. We do not include cohort fixed effects for the younger sibling in columns 3 and 4. In Chile, we observe enrollment for all the colleges of the system from 2007 onward. Thus, the sample is adjusted accordingly. Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D5: Sibling Spillovers on Applications to College and College-Major by Age Difference and Gender

	Any 4-year College	Younger sibling follows to same college				Younger sibling follows to same college-major		
	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 (E)	0.217*	0.092**	0.109**	0.141**	0.268***	0.025**	0.039**	0.038**
	(0.130)	(0.015)	(0.020)	(0.011)	(0.102)	(0.005)	(0.009)	(0.006)
E $\times$ Age difference $\geq$ 5	0.136	-0.035**	0.000	-0.019*	0.104	-0.004	-0.018	-0.016**
	(0.142)	(0.011)	(0.026)	(0.010)	(0.107)	(0.004)	(0.013)	(0.004)
Observations	44190	86364	12950	378466	44190	170570	36756	482220
F-statistic	64.892	2767.580	3230.667	3562.527	64.892	7330.470	7225.706	5147.083
Panel B: Same gender								
Older sibling enrolls (E)	0.310**	0.070**	0.114***	0.129***	0.304***	0.017**	0.026*	0.028**
	(0.137)	(0.016)	(0.022)	(0.012)	(0.106)	(0.005)	(0.009)	(0.005)
E $\times$ Same gender	-0.152**	0.011	-0.007	0.007	-0.052	0.011**	0.023	0.006
	(0.071)	(0.012)	(0.020)	(0.010)	(0.056)	(0.004)	(0.009)	(0.005)
Observations	44190	86521	12950	378466	44190	170886	36757	482220
F-statistic	65.114	2788.470	3229.534	3607.870	65.114	7383.02	7220.184	5204.123
Panel C: Same gender, older brothers only								
Older sibling enrolls (E)	0.363	0.078**	0.124***	0.120**	0.624***	0.016**	0.025	0.017
	(0.257)	(0.022)	(0.033)	(0.020)	(0.218)	(0.007)	(0.015)	(0.009)
E $\times$ Same gender	0.065	0.001	0.001	0.036**	-0.039	0.020**	0.044*	0.041**
	(0.122)	(0.017)	(0.032)	(0.016)	(0.103)	(0.006)	(0.016)	(0.007)
Observations	17881	39919	5008	134815	17881	81072	14203	180192
F-statistic	19.214	1435.860	1405.970	1571.713	19.214	4150.72	4025.070	2633.160
Panel D: Same gender, older sisters only								
Older sibling enrolls (E)	0.265*	0.079**	0.098**	0.140**	0.142**	0.018*	0.031	0.040**
	(0.160)	(0.024)	(0.031)	(0.016)	(0.121)	(0.007)	(0.013)	(0.008)
E $\times$ Same gender	-0.217**	-0.004	-0.027	-0.008	-0.021	-0.000	0.007	-0.021**
	(0.088)	(0.018)	(0.027)	(0.013)	(0.068)	(0.006)	(0.012)	(0.006)
Observations	26296	44222	7545	233839	26296	87895	22239	291078
F-statistic	45.708	1223.530	1651.529	1711.833	45.708	7383.02	3662.675	2251.639

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in the main body of the paper. In addition they include interactions with dummies for age difference and if the siblings are of the same gender. These variables are also included separately as controls. Standard errors clustered at the family level are reported in parenthesis. The table reports the Kleibergen-Paap Wald F-statistic. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D6: Sibling Spillovers on Enrollment in College and College-Major by Age Difference and Gender

	Any 4-year College	**Younger sibling follows to same college**				**Younger sibling follows to same college-major**		
	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 (E)	0.217*	0.052**	0.089**	0.054**	0.163**	0.006**	0.013*	0.007**
	(0.013)	(0.011)	(0.019)	(0.005)	(0.053)	(0.002)	(0.004)	(0.002)
E $\times$ Age difference $\geq$ 5	0.136	-0.029**	-0.029	-0.012**	0.098	-0.001	0.001	-0.004**
	(0.142)	(0.009)	(0.026)	(0.005)	(0.060)	(0.002)	(0.006)	(0.001)
Observations	44190	86364	12950	378466	44190	170570	36756	482220
F-statistic	64.892	2767.580	3230.667	3562.527	64.892	7330.470	7225.706	5147.083
Panel B: Same gender								
Older sibling enrolls (E)	0.310*	0.033**	0.065**	0.044**	0.184**	0.002	0.007	0.002
	(0.137)	(0.011)	(0.021)	(0.005)	(0.055)	(0.002)	(0.009)	(0.001)
E $\times$ Same gender	-0.152**	0.010	0.037	0.010*	-0.022	0.007**	0.013*	0.006**
	(0.071)	(0.009)	(0.019)	(0.005)	(0.029)	(0.002)	(0.004)	(0.001)
Observations	44190	86521	12950	378466	44190	170886	36757	482220
F-statistic	65.114	2788.47	3229.534	3607.870	65.114	7383.020	7220.184	5204.123
Panel C: Same gender, older brothers only								
Older sibling enrolls (E)	0.363	0.041*	0.066	0.048**	0.284***	-0.001	0.008	0.004
	(0.257)	(0.016)	(0.034)	(0.009)	(0.109)	(0.003)	(0.007)	(0.003)
E $\times$ Same gender	0.065	0.014	0.014	0.010	-0.034	0.015**	0.031**	0.008**
	(0.122)	(0.012)	(0.031)	(0.008)	(0.052)	(0.003)	(0.008)	(0.002)
Observations	17881	39919	5008	134815	17881	81072	14203	180192
F-statistic	19.214	1435.860	1405.970	1571.713	19.214	4150.072	4025.070	2633.160
Panel D: Same gender, older sisters only								
Older sibling enrolls (E)	0.265*	0.034	0.044	0.046**	0.134**	0.006	0.006	0.002
	(0.160)	(0.018)	(0.029)	(0.007)	(0.064)	(0.003)	(0.006)	(0.002)
E $\times$ Same gender	-0.217**	0.008	0.046	0.007	-0.006	-0.002	0.004	0.003*
	(0.088)	(0.013)	(0.026)	(0.006)	(0.036)	(0.003)	(0.005)	(0.002)
Observations	26296	44222	7545	233839	26296	87895	22239	291078
F-statistic	45.708	1223.530	1651.529	1711.833	45.708	3096.64	3662.675	2251.639

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in the main body of the paper. In addition they include interactions with dummies for age difference and if the siblings are of the same gender. These variables are also included separately as controls. Standard errors clustered at the family level are reported in parenthesis. The table reports the Kleibergen-Paap Wald F-statistic. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D7: Sisters Spillovers on Applications to and Enrollment in STEM Majors

	Chile		**Sweden**	
	Ranks STEM 1st (1)	**Enrolls in STEM** (2)	**Ranks STEM 1st** (3)	**Enrolls in STEM** (4)
Older sibling enrolls (2SLS)	0.024	0.004	0.028	0.031
	(0.034)	(0.031)	(0.043)	(0.034)
Above cutoff (RF)	0.013	0.002	0.011	0.012
	(0.019)	(0.017)	(0.016)	(0.013)
Observations	9419	9419	12284	12284
Counterfactual mean	0.24	0.19	0.18	0.10
Bandwidth	22.00	22.00	0.386	0.386
F-statistic	906.85	906.85	461.260	461.260

Notes: In this table we report how an older sibling's marginal enrollment in a STEM subject impacts the likelihood that the younger sibling will apply to or enroll in a STEM program at any college. The specifications use the same set of controls and bandwidths used in the 2SLS specifications described in the main body of the paper. Standard errors clustered at the family level are reported in parenthesis. The table reports the Kleibergen-Paap Wald F-statistic. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D8: Sibling Spillovers by Differences between Older Siblings' Target and Next Best Option

	Chile			**Croatia**	**Sweden**		
	Expected Earnings (USD 000)	Peer Quality (z-score)	First Year Retention Rate	Peer Quality (z-score)	Expected Earnings (USD 000)	Peer Quality (z-score)	First Year Retention Rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A - Younger sibling enrolls in older sibling's target college							
Older sibling enrolls (ΔX in Q1)	0.047**	0.066***	0.040*	0.083***	0.038***	0.038***	0.029***
	(0.021)	(0.023)	(0.022)	(0.030)	(0.012)	(0.010)	(0.011)
Older sibling enrolls (ΔX in Q2)	0.038**	0.037*	0.046***	0.119***	0.032***	0.026***	0.021**
	(0.021)	(0.020)	(0.021)	(0.029)	(0.010)	(0.010)	(0.010)
Older sibling enrolls (ΔX in Q3)	0.056***	0.043**	0.035*	0.085***	0.031***	0.027**	0.031***
	(0.021)	(0.021)	(0.020)	(0.031)	(0.010)	(0.010)	(0.010)
Older sibling enrolls (ΔX in Q4)	0.033	0.034*	0.053**	0.104***	0.036***	0.026**	0.049***
	(0.022)	(0.020)	(0.021)	(0.032)	(0.012)	(0.012)	(0.011)
Observations	32987	32987	32987	9610	147190	167290	159146
F-statistic	722.509	744.566	740.276	1089.054	613.193	676.879	673.860
Counterfactual mean	0.134	0.134	0.134	0.232	0.031	0.031	0.032
Panel B - Younger sibling enrolls in older sibling's target college-major							
Older sibling enrolls (ΔX in Q1)	0.007*	0.003	0.011***	0.017***	-0.001	0.000	0.000
	(0.004)	(0.004)	(0.004)	(0.006)	(0.003)	(0.003)	(0.003)
Older sibling enrolls (ΔX in Q2)	0.005	0.006	0.005	0.010*	0.005	0.002	0.001
	(0.004)	(0.004)	(0.004)	(0.005)	(0.003)	(0.003)	(0.003)
Older sibling enrolls (ΔX in Q3)	0.005	0.008*	0.006	0.012*	0.004	-0.001	0.001
	(0.004)	(0.004)	(0.004)	(0.006)	(0.003)	(0.003)	(0.003)
Older sibling enrolls (ΔX in Q4)	0.010**	0.010**	0.005	0.011*	0.003	0.005*	0.002
	(0.004)	(0.004)	(0.004)	(0.006)	(0.003)	(0.003)	(0.003)
Observations	81849	81849	81849	32288	214143	248297	230709
F-statistic	2384.614	2437.617	2439.986	3137.876	1151.517	1280.638	1262.027
Counterfactual mean	0.013	0.013	0.013	0.016	0.004	0.004	0.004

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in the main body of the paper. In addition, each column includes dummy variables for the quality difference between the older sibling's preferred and counterfactual alternatives divided into quartiles. For example, the top-left cell shows that a younger sibling is 4.7 p.p. more likely to enroll in their older sibling's target college when looking only at older siblings who were in the bottom quartile in terms of the difference in expected earnings between their target and next-best. Likely in this case, the difference is negative and the earnings would have been higher for the older sibling should he or she not be admitted. Standard errors clustered at the family level are reported in parenthesis. The table reports the Kleibergen-Paap Wald F-statistic. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D9: Sibling Spillovers on Younger Siblings' Enrollment by Older Siblings' Target Option Characteristics

	Chile			Croatia	Sweden		
	Expected Earnings (USD 000)	Peer Quality (z-score)	First Year Retention Rate	Peer Quality (z-score)	Expected Earnings (USD 000)	Peer Quality (z-score)	First Year Retention Rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A - Younger sibling enrolls in older sibling's target college							
Older Sibling Enrolls (X in Q1)	0.032	-0.008	0.011	0.064**	0.026*	0.005	0.017
	(0.022)	(0.031)	(0.021)	(0.027)	(0.014)	(0.015)	(0.012)
Older Sibling Enrolls (X in Q2)	0.021	0.049**	0.045**	0.098***	0.038***	0.011	0.033***
	(0.022)	(0.036)	(0.021)	(0.028)	(0.012)	(0.011)	(0.010)
Older Sibling Enrolls (X in Q3)	0.057***	0.036*	0.028	0.081***	0.024**	0.040***	0.030***
	(0.022)	(0.025)	(0.020)	(0.030)	(0.011)	(0.010)	(0.010)
Older Sibling Enrolls (X in Q4)	0.031*	0.035**	0.044**	0.061*	0.041***	0.041***	0.046***
	(0.017)	(0.016)	(0.018)	(0.033)	(0.010)	(0.009)	(0.010)
Observations	39960	39960	39960	11776	169619	178814	175951
F-statistic	824.637	626.324	926.147	1423.162	588.205	651.385	723.051
Counterfactual mean	0.134	0.134	0.134	0.254	0.031	0.031	0.031
Panel B - Younger sibling enrolls in older sibling's target college-major							
Older Sibling Enrolls (X in Q1)	0.002	0.002	0.007*	0.019***	-0.003	-0.003	-0.003
	(0.004)	(0.005)	(0.004)	(0.006)	(0.004)	(0.004)	(0.003)
Older Sibling Enrolls (X in Q2)	0.004	0.006*	0.005	0.010*	0.002	-0.005	0.003
	(0.004)	(0.004)	(0.004)	(0.006)	(0.003)	(0.003)	(0.003)
Older Sibling Enrolls (X in Q3)	0.011***	0.005	0.005	0.012*	-0.001	0.002	0.003
	(0.004)	(0.004)	(0.004)	(0.006)	(0.003)	(0.003)	(0.003)
Older Sibling Enrolls (X in Q4)	0.008**	0.010***	0.009**	0.013**	0.007**	0.005*	0.003
	(0.004)	(0.004)	(0.004)	(0.006)	(0.003)	(0.003)	(0.003)
Observations	97321	97321	97321	34424	247960	264527	256565
F-statistic	2501.594	1819.772	2883.727	3259.789	1002.833	1090.406	1340.660
Counterfactual mean	0.013	0.013	0.013	0.017	0.004	0.004	0.004

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in the main body of the paper. In addition, each column includes dummy variables for the quality of the older sibling's preferred alternatives divided into quartiles. For example, the top-left cell shows that a younger sibling is 3.2 p.p. more likely to enroll in their older sibling's target college when looking only at older siblings who had applied to a target alternative that was in the bottom quartile in terms of expected earnings when compared to all other alternatives during that admission round. Standard errors clustered at the family level are reported in parenthesis. The table reports the Kleibergen-Paap Wald F-statistic. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D10: Sibling Effects on Academic Performance by Age Difference

	College-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
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
F-statistic	1461.978	1461.978	659.829	659.829
		Panel C - Sweden		
Older sibling enrolls	0.013	0.038	0.030	0.071
	(0.026)	(0.036)	(0.032)	(0.047)
Older sibling enrolls × 2 < Δ Age ≤ 4	0.027	0.083**	-0.005	0.084*
	(0.024)	(0.034)	(0.032)	(0.046)
Older sibling enrolls × 4 < Δ Age	-0.016	0.008	-0.020	0.000
	(0.024)	(0.034)	(0.031)	(0.044)
Observations	421268	227976	329598	176765
Counterfactual mean	0.218	0.040	0.217	0.043
Bandwidth	0.386	0.386	0.360	0.360
F-statistic	3202.331	2196.951	2193.582	1453.763
		Panel D - United States		
Older sibling enrolls × Δ Age ≤ 2				36.914
				(28.823)
Older sibling enrolls × Δ Age ≤ 4				51.224
				(39.571)
Older sibling enrolls × Δ Age ≤ 10				-12.860
				(56.439)
Observations				37554
Counterfactual mean				953.162
Bandwidth				93
F-statistic				38.960

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in the main body of the paper. Each column has a different outcome variable, measuring the academic performance of the younger sibling. In addition the effect is allowed to vary with the age difference between the siblings. Standard errors clustered at the family level are reported in parenthesis. The table reports the Kleibergen-Paap Wald F-statistic. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D11: Sibling Effects on College and Major Choice by Socioeconomic Status (Bottom ≤ 40% of Income Distribution)

	Sibling follows to same college				Sibling follows to same 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 (E)	0.081***	0.123***	0.036***	0.045***	0.024***	0.028***	0.004	0.004**
	(0.201)	(0.012)	(0.015)	(0.005)	(0.006)	(0.005)	(0.003)	(0.002)
E × Bottom 40%	-0.008	0.022**	0.002	0.009*	-0.002	0.008	0.003	0.003*
	(0.016)	(0.011)	(0.011)	(0.005)	(0.005)	(0.005)	(0.002)	(0.001)
Observations	86521	377004	86521	377004	170886	480338	170886	480338
F-statistic	4921.048	3598.968	4921.048	3598.968	7225.228	5186.527	7225.228	5186.527

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in the main body of the paper. 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. Standard errors clustered at the family level are reported in parenthesis. The table reports the Kleibergen-Paap Wald F-statistic. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D12: 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 (E)	0.099***	0.108***	0.054***	0.028***
	(0.018)	(0.013)	(0.012)	(0.006)
E × Peer exposure	-0.317***	-0.323***	-0.192	0.056
	(0.138)	(0.055)	(0.137)	(0.039)
Observations	84911	316799	84911	316799
Avg. exposure in the sample	0.075	0.070	0.075	0.070
F-statistic	2775.363	3101.136	2775.363	3101.136

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in the main body of the paper. 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 (exposure). The main effect is also included in the specifications as a control. Standard errors clustered at the family level are reported in parenthesis. The table reports the Kleibergen-Paap Wald F-statistic. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D13: Sibling Effects on Applications to and Enrollment in Older Sibling's Target College: Cities with Multiple Colleges (Chile)

	Ranks 1st (1)	Applies (2)	Enrolls (3)
Older sibling enrolls (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
F-statistic	3353.800	3353.800	3353.800

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in the main body of the paper. 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. Standard errors clustered at the family level are reported in parenthesis. The table reports the Kleibergen-Paap Wald F-statistic. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D14: Sibling Effects on Applications to and Enrollment in Older Sibling's Target College and Target College-Major (Closest Sibling)

	Sibling follows to same college			Sibling follows to same college-major		
	Ranks 1st	Applies	Enrolls	Ranks 1st	Applies	Enrolls
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A - Chile			
Older sibling enrolls	0.065***	0.073***	0.036***	0.012***	0.023***	0.006***
	(0.012)	(0.015)	(0.011)	(0.003)	(0.005)	(0.002)
Above cutoff (RF)	0.032***	0.035***	0.018***	0.006***	0.012***	0.003***
	(0.006)	(0.007)	(0.005)	(0.001)	(0.003)	(0.001)
Observations	86009	86009	86009	169885	169885	169885
Counterfactual mean	0.22	0.45	0.13	0.02	0.06	0.01
Bandwidth	12.500	12.500	12.500	18.000	18.000	18.000
F-statistic	5506.23	5506.23	5506.23	14639.55	14639.55	14639.55
			Panel B - Croatia			
Older sibling enrolls	0.077***	0.117***	0.087***	0.016***	0.038***	0.013***
	(0.019)	(0.020)	(0.019)	(0.004)	(0.009)	(0.004)
Above cutoff (RF)	0.064***	0.097***	0.072***	0.013***	0.032***	0.011***
	(0.016)	(0.016)	(0.016)	(0.004)	(0.007)	(0.003)
Observations	12197	12197	12197	34693	34693	34693
Counterfactual mean	0.29	0.52	0.25	0.02	0.11	0.02
Bandwidth	80.000	80.000	80.000	80.000	80.000	80.000
F-statistic	6250.51	6250.51	6250.51	14677.22	14677.22	14677.22
			Panel C - Sweden			
Older sibling enrolls	0.116***	0.110***	0.044***	0.019***	0.024***	0.005***
	(0.008)	(0.011)	(0.005)	(0.003)	(0.005)	(0.001)
Above cutoff (RF)	0.031***	0.029***	0.012***	0.006***	0.007***	0.001***
	(0.002)	(0.003)	(0.001)	(0.001)	(0.002)	(0.000)
Observations	353079	353079	353079	452834	452834	452834
Counterfactual mean	0.088	0.206	0.033	0.011	0.055	0.004
Bandwidth	0.360	0.360	0.360	0.386	0.386	0.386
F-statistic	8367.474	8367.474	8367.474	12035.142	12035.142	12035.142
			Panel D - United States			
2SLS		0.232**	0.136**			
		(0.108)	(0.057)			
Above cutoff (RF)		0.019**	0.011**			
		(0.009)	(0.005)			
Observations		39,214	39,214			
Counterfactual mean		0.137	0.041			
Bandwidth		93				
F-statistic		136.475	136.475			

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in the main body of the paper, but looking only at the sibling pair in each family that is closest in age. Standard errors clustered at the family level are reported in parenthesis. The table reports the Kleibergen-Paap Wald F-statistic. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D15: Sibling Effects on Applications to and Enrollment in Older Sibling's Target College and Target College-Major (First and Second Children Only, Sweden)

	Sibling follows to same college			**Sibling follows to same college-major**		
	Ranks 1st	**Applies**	**Enrolls**	**Ranks 1st**	**Applies**	**Enrolls**
	(1)	(2)	(3)	(4)	(5)	(6)
Older sibling enrolls	0.130***	0.134***	0.051***	0.026***	0.040***	0.008***
	(0.011)	(0.015)	(0.007)	(0.004)	(0.008)	(0.002)
Above cutoff (RF)	0.036***	0.037***	0.014***	0.008***	0.012***	0.002***
	(0.003)	(0.004)	(0.002)	(0.001)	(0.002)	(0.001)
Observations	175696	175696	175696	230233	230233	230233
Counterfactual mean	0.096	0.234	0.035	0.013	0.066	0.004
Bandwidth	0.360	0.360	0.360	0.386	0.386	0.386
F-statistic	5740.818	5740.818	5740.818	8642.139	8642.139	8642.139

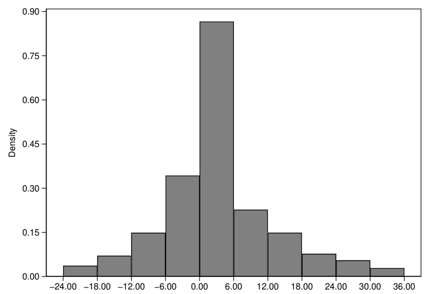
Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in the main body of the paper, but looking only at first-born older siblings and second-born younger siblings. Standard errors clustered at the family level are reported in parenthesis. The table reports the Kleibergen-Paap Wald F-statistic. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table D16: Sibling Effects on Applications to and Enrollment in Older Sibling's Target Major

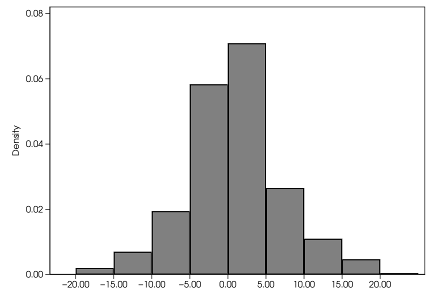
	Ranks 1st		**Applies**		**Enrolls**	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A - Chile						
2SLS	0.012	0.009	0.017*	0.014	-0.001	-0.004
	(0.007)	(0.009)	(0.010)	(0.012)	(0.006)	(0.007)
Reduced form	0.005	0.005	0.010*	0.009*	0.000	-0.001
	(0.003)	(0.003)	(0.005)	(0.005)	(0.003)	(0.003)
First stage	0.478***	0.449***	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-statistic	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-statistic	10158.245	7440.903	10158.245	7440.903	10158.245	7440.903
Panel C - Sweden						
2SLS	0.000	-0.007	-0.002	-0.011*	-0.001	-0.006**
	(0.006)	(0.005)	(0.009)	(0.007)	(0.004)	(0.003)
Reduced Form	0.000	-0.002	-0.001	-0.003*	0.000	-0.002**
	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)
First stage	0.273***	0.284***	0.273***	0.284***	0.273***	0.284***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
2SLS (Triangular Kernel)	-0.002	-0.006	-0.005	-0.008	-0.001	-0.003
	(0.006)	(0.005)	(0.009)	(0.007)	(0.004)	(0.003)
Observations	355885	1033836	355885	1033836	355885	1033836
Counterfactual mean	0.049	0.044	0.101	0.096	0.016	0.014
Bandwidth	0.389	1.213	0.389	1.213	0.389	1.213
F-statistic	6643.373	9843.804	6643.373	9843.804	6643.373	9843.804

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in the main body of the paper, but columns 2,4,6 also include 2nd degree polynomials of the running variable. 2SLS (Triangual Kernel) specifications use a triangular kernel to give more weight to observations close to the cutoff. Standard errors clustered at the family level are reported in parenthesis. The table reports the Kleibergen-Paap Wald F-statistic. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Figure D1: Differences between Older Siblings' Target and Next Best Choices

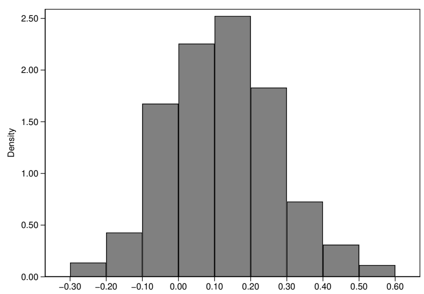


(a) Chile

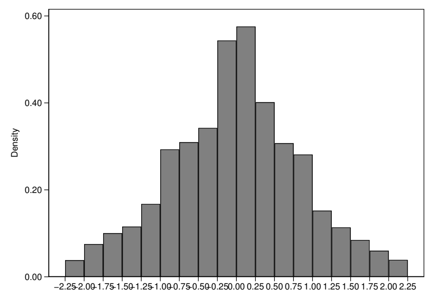


(b) Sweden

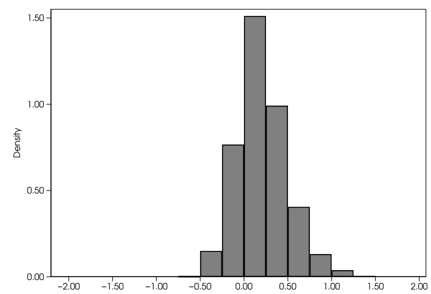
Panel A - Δ in Annual Earnings (USD 000)



(c) Chile

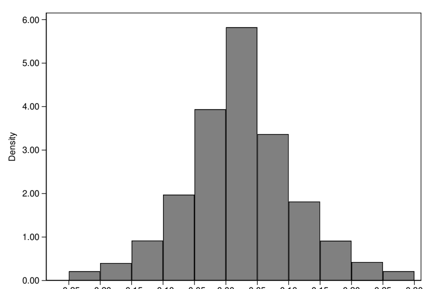


(d) Croatia

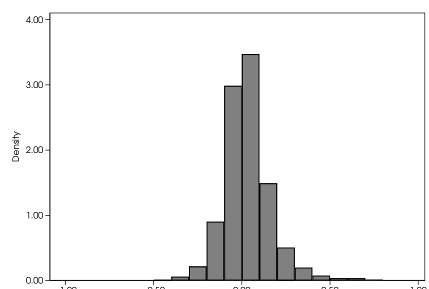


(e) Sweden

Panel B - Δ in Peer Quality (z-scores)



(f) Chile



(g) Sweden

Panel C - Δ in First Year Retention Rates

These figures illustrate the differences between older siblings' target and next best options in terms of expected earnings (Panel A), peer quality (Panel B) and first year retention rates (Panel C).