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 : October 29, 2020

Abstract

This online appendix is organized in two broad sections. The first one 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 presents 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.

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 US colleges using hidden cutoffs in their admissions.

[†]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.

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 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

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

[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.

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

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 receive a later offer.

A.2 College Admission System in Croatia

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

As in Chile, NISpVU uses a deferred acceptance admission system that focuses primarily on students' high-school performance and in a national level university exam.[7] The national exam is taken in late June, approximately one month after the end of the Croatian academic year. However, students are required to submit a free-of-charge online registration form by mid-February.

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

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

[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.

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 because if students do not accept the offer they receive the first time that they apply, they loss the tuition fee waiver offered by the government. This means that in case of re-applying 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]

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

The Swedish admission system has a few important differences compared to the Chilean and Croatian systems. For one thing, the same system is open to applications to full majors and shorter courses alike. To simplify, we will henceforth refer to all these alternatives as *majors*. Moreover, applicants are ranked by different scores separately in a number of *admission groups*. Their best ranking is then used to determine their admission status.[10] Finally, the Swedish admission system

[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.

[10]Admission is essentially determined by a max function of high school GPA and Högskoleprovet score, as compared

has two rounds. After the first round, applicants learn their admission status and they place in the waiting list for all their applications. At this point, they can decide whether to accept the best offer they have or to wait and participate in a second application round. Their scores and lists of preferences do not change between the two rounds, but the cutoffs might. In this project we focus on the variation generated by the cutoff of the second round. Since some applicants decide to accept the offers they received after the first round instead of waiting for the second round, not all the applicants above the second round admission cutoff end receiving an offer. Those who dropout from the waiting list after the first round cannot receive a second round offer, even if their score was above the final admission cutoff. This explains why in the case of Sweden the jump in older siblings' admission and enrollment probabilities is smaller than in the other two countries. 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.

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

After receiving an offer, applicants can either accept or decide to stay on the waiting list for choices

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.

for which they have not yet been admitted to. Should they decide to wait, admissions after the second round will again only include the highest-ranked 45 ECTS, and all lower-ranked alternatives will be discarded, even those that they were previously admitted to.[12]

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

A.4 College Admission System in the United States

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

During the period that we study the SAT was offered seven times a year and could be taken as often as the college application timeline allowed.[13]. As in the case of the admission exams used in the other countries, the SAT has different sections and in terms of application is common for colleges to consider students' "superscores"(Goodman et al., 2020). The "superscores"are the sum of a student's maximum math and maximum critical reading scores, regardless of whether those scores occurred on the same attempt. In order to apply to college students need to submit their SAT scores and any other application material requested by the institutions in which they are interested.

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

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

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

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

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

To search for discontinuities, we estimate local linear regression discontinuity models at each SAT score that might represent a potential threshold for each college in each year.[14] 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:

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

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

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

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

nuities only those instances where discontinuities in enrollment rates exceed five percentage points and where we reject the null hypothesis of no discontinuity with $p > 0.0001$. Finally, we discard any colleges where thresholds are detected in fewer than five years 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 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.

B.1 Older Siblings' Higher Education Trajectories and Spillovers on Total College Enrollment

In this Section we discuss how crossing the admission threshold changes the higher education trajectories of older siblings. As shown in Table B1, in the US being above the hidden admission cutoffs 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 US 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 B2 shows how the characteristics of the college attended by older siblings in the United States changes at the admission cutoff. All the results in the 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 at 50 or more miles from home.

In the case of 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 B3 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 United States.

When focusing on younger siblings, Table B3 shows that having an older sibling enrolling in their 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 United States (discussed in the paper). However, this is not surprising. While in the case of the United States, crossing the threshold moves older and younger siblings from two-year 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 B4 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 that they apply, could try again and gain admission to another college before their younger siblings are ready to apply.

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

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

Table B5 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 siblings pairs in which the older sibling is a male, while Panel D focuses on sibling pairs in which the older sibling is a female. In both cases the interaction is a dummy variable that takes value 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 United States, a setting where after splitting the sample we end with relatively few observations. 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 that we find. 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 never significant and the coefficients are small.

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 extend by younger sisters and brothers.

Table B6 replicates these results but focusing on enrollment instead of on applications. The results follow the same pattern that we just discussed.

We conclude this section by investigating whether having an older sister admitted into a STEM major increases the probability that a younger sister applies and enrolls in this type of programs. In order 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 doing STEM is relatively low in all the settings that we study, we end 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 B7 presents these results. The coefficients suggest that having an older sister attending a STEM major slightly increases the probability of applying and enrolling in a STEM major as well. However, our estimates are not precise enough to rule out that they are equal to zero.

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

In the paper we show that the difference between older siblings' target and next best options do not make a difference in the size of the response that we find when looking at younger siblings' applications. In addition, we find that younger siblings responses do not change a lot with the characteristics of the target college-major of the older sibling. Here, we replicate these results but focusing on the college and college-major in which younger siblings enroll.

Table B8 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 loss a lot of precision. The average effects in this margin were already small, but now 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 by their siblings to the same college-major. In terms of retention rates, we only find an 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 the case of Sweden the coefficients become small and non-significant in all categories.[15].

It is worth noting that while applications reflect individuals 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 B9 indicate that in terms of enrollment, responses are stronger when the target choice is better. The fact that we observe a positive gradient in the effects here and not in applications might reflect

[15]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

that individuals whose older siblings apply to programs in the top quartile understand better to which options they are more likely to be admitted.

B.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 B10 summarizes these results. We find no increases on younger siblings academic performance no matter of the age difference that they have with their older sibling.

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

In the paper we show that in the United States sibling spillovers are larger for individuals with low probabilities of enrolling in college. 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 less exposed to the target college of their older sibling. We compute exposure as the share of schoolmates completing high school one year before the younger sibling who enroll in the older sibling's target college.

Table B11 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 United States. 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 in the margin of being admitted to a selective program. This positive selections means that we observe few uncertain college-goers in these countries .

Table B12 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.

B.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 closest older instead of on the oldest sibling does not make a big difference in the size of the effects that we find. We conclude by presenting additional specifications looking at sibling spillovers on the choice of major that confirm the results we present in the paper.

B.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 reflect at least in part a change in geographic preferences. This would mean that individuals follow their older siblings to the city and not to the institution or major in which they enroll. To investigate this possibility, we take advantage of the fact that in Chile there are three big cities — Santiago, Valparaíso and Concepción— that not only contain an important share of the population, but also multiple universities.[16]

Table B13 presents the results of an exercise in which we estimate the baseline specification on a sample of Chilean students from Santiago, Valparaíso and Concepción whose older siblings apply to institutions in their hometowns. If the effects documented in the paper were driven only by geographic preferences, we should not find sibling spillovers on the choice of college for this subsample. However, the coefficients that we obtain in this case are very similar to the main results discussed in the main body of the paper.

B.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 B14 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

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

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 B15 we once more find estimates that are very similar to the ones that we present in the paper.

B.6.3 Sibling Spillovers on Major Choice - Additional Specifications

Finally, in this Section we present additional specifications looking at sibling spillovers on the choice of major. Apart from the results presented in the main body of the paper, Table B16 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 B1: 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 B2: 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 B3: 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 major-college 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 major-college 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/his 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 B4: Probability of Enrolling in any College by Older Siblings' Admission to their Target Major-College

	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 major-college 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 onwards. 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 B5: 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 ≥ 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 × Age difference ≥ 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 × 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 × 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 × 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 B6: 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 B7: 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 B8: 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 B9: 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 B10: Sibling Effects on Academic Performance by Age Difference

	Major-college 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 B11: Sibling Effects on College and Major Choice by Socioeconomic Status (Bottom $\leq$ 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 $\times$ 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 B12: 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 $\times$ 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 B13: 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 B14: 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 familiy 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 B15: 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 siblinigs 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 B16: Sibling Effects on Applications to and Enrollment in Older Sibling's Target Major

	Ranks 1st (1)	(2)	**Applies** (3)	(4)	**Enrolls** (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 runniing 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.

References

Andrews, R. J., S. A. Imberman, and M. F. Lovenheim (2017, July). Risky Business? The Effect of Majoring in Business on Earnings and Educational Attainment. NBER Working Papers 23575, National Bureau of Economic Research, Inc.

Goodman, J., O. Gurantz, and J. Smith (2020). Take Two! SAT Retaking and College Enrollment Gaps. *American Economic Journal: Economic Policy 12*(2), 115–58.

Goodman, J., M. Hurwitz, and J. Smith (2017). Access to 4-Year Public Colleges and Degree Completion. *Journal of Labor Economics 35*(3), 829–867.