O Brother, Where Start Thou?

Sibling Spillovers on College and Major Choice in Four Countries†

Adam Altmejd   Andrés Barrios-Fernández   Marin Drlje   Joshua Goodman

Michael Hurwitz   Dejan Kovac   Christine Mulhern   Christopher Neilson   Jonathan Smith

Latest Version : October 26, 2020

Abstract

This paper presents causal evidence from Chile, Croatia, Sweden, and the United States that older siblings' higher education trajectories influence their younger siblings' college and major choices in significant ways. We exploit admission cutoffs that generate quasi-random variation in older siblings' higher education paths and show that they are systematically followed by their younger siblings. Older siblings are followed independently of whether their target and counterfactual options have large, small or even negative differences in expected earnings, peer quality and retention rates. The documented spillover effects disappear if the older sibling drops out of college, suggesting that older siblings' experiences in college matter. Despite the many differences across these four countries, in each case we find evidence that siblings influence important human capital investment decisions. The consistent results that we obtain across these different settings suggest that our findings are not context specific or driven by institutional details.

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

†For granting us access to their administrative data, we thank the the College Board in the United States, the Ministries of Education of Chile and Croatia, the DEMRE, ASHE (AZVO), Riksarkivet, UHR and SCB. All errors are due to our older siblings. Melanie Rucinski and Cecilia Moreira provided excellent research assistance.

1 Introduction

The decisions of whether to go to college, where to enroll and what to specialize in, are among the most consequential an individual will make in their life. Each of these margins can significantly impact a host of important outcomes such as future earnings and broad life outcomes, and in the aggregate can drive economic growth and inequality (Goldin and Katz, 2008).[1] Despite the significance of these choices, we know very little about their determinants. Social context and family background seem to play an important role in shaping higher education trajectories, which suggests that close peers and relatives could significantly influence decisions regarding post-secondary education (Hoxby and Avery, 2013; Chetty et al., 2020). However, causally identifying the influence of family and social network on human capital investment is challenging, and the evidence on how close peers affect crucial post-secondary decisions is still scarce.

This paper provides causal evidence that an older sibling—one of the most relevant members of an individual's social network—influences the educational path of younger siblings. We show in four very different settings—Chile, Croatia, Sweden and the United States—that shocks to older siblings' higher education trajectories impact younger siblings' application and enrollment decisions in meaningful ways. The consistent results that we obtain across these different settings suggest that our findings are not context specific or driven by institutional details.

To overcome the main identification challenges of peer effects (i.e., correlated effects and the reflection problem) we exploit admission cutoffs that generate quasi-random variation in the college or college-major in which older siblings enroll. In each country, we use rich administrative data that allow us to identify siblings and link them to detailed data on college applications and enrollment decisions. In some cases we are also able to identify the older siblings' counterfactual educational paths (i.e., the college-major they would have been admitted to if they had been below their target

[1] Labor economists have accumulated extensive evidence on the causal effects of education on earnings and other life outcomes. The evidence on the returns to education is reviewed in Card (1999) and Card (2001). Altonji et al. (2012) documents the heterogeneity in earnings across college and majors. Altonji et al. (2016) reviews the literature on the returns to college and majors, emphasizing heterogeneity in the effects of education. Hastings et al. (2013) and Kirkebøen et al. (2016) show causal evidence that specific college-major combinations as well as broader fields of study, also significantly impact earnings in both the short and longer term. Heckman et al. (2018) emphasizes heterogeneity in these returns as well as showing impacts on a broader set of outcomes such as smoking and health. It should be noted that the important differences in costs, both in resources and time, make post-secondary human capital investment decisions very important even in the absence of differential earnings outcomes.

admission threshold).

In Chile, Croatia and Sweden, universities coordinate their admissions to jointly process applications through a centralized system that provides students with a single admissions offer.[2] These systems allocate applicants to a unique college-major combination based on their academic performance (i.e., high school GPA and test scores) and on a ranked ordered list of college-major preferences that they submit when applying. The single admissions offer system generates sharp cutoffs at all oversubscribed programs. In addition, the application data used by these systems also allows us to identify the next-best-alternative the applicant would have been assigned to had they not been accepted at their assigned college-major program. This can be used to identify the counterfactual educational trajectory as in Kirkebøen et al. (2016).

In the United States, admissions decisions are decentralized so that students may receive multiple offers. However, using administrative data on applications, test scores and enrollment, we identify a subset of colleges that use SAT scores cutoffs as part of their admission process. These cutoffs generate quasi-random variation in admissions probabilities for a subset of the population, similar to that found in Chile, Croatia and Sweden. In each country, we use the cutoffs to causally identify siblings' influence by comparing, through a regression discontinuity design, the college and major choices of younger siblings whose otherwise identical older siblings were marginally above or below these admission cutoffs.

In all four countries, we find causal evidence that younger siblings systematically follow their older siblings to college. In Chile, Croatia and Sweden, where students are admitted to a specific major within a college, younger siblings also follow their older siblings to the same specific college-major combination. In the United States, we present evidence that older sibling's affect the extensive margin —an older sibling's enrollment in a four year college increases the younger sibling's probability of also enrolling in a four year college.

Sibling spillovers on college application and enrollment decisions can shift younger siblings decisions in important ways. In the United States, older siblings induce younger siblings to enroll in four year college. This is important because not only does attending four year college have a positive

[2] These systems are common in developed and developing countries around the world. Over 40 countries used similar centralized systems in higher education in 2019 (Neilson, 2019).

return on average (Card, 1999), recent evidence shows this is the case even for marginal students, which is likely to be the group we are studying (Zimmerman, 2014; Goodman et al., 2017). Older siblings also affect the college younger siblings enroll in, which can matter a lot in some cases given recent evidence on heterogeneity in returns across colleges in the US (Chetty et al., 2020).

In Chile, Croatia and Sweden, we use data on average earnings of graduates, peer quality and retention rates to compare the college and major programs older siblings are applying to on the margin. We find that older siblings are followed when the difference between the target program and the next best alternative is both large or small. Younger siblings follow their older siblings to the same college and college-major combination even when the target program has lower expected earnings, peer quality and retention rates. The only exception to the general finding that younger siblings follow their older siblings is when the older sibling drops out of college. This eliminates any spillover effect and suggests that older siblings experience in college also matters.

We discuss three broad classes of mechanisms that could explain why older siblings influence the higher education trajectories of their younger siblings. First, an older sibling's educational trajectory could affect the costs of the option. For example, siblings could commute together or siblings could share housing costs. Second, older siblings' choices could affect the utility that younger siblings derive from particular colleges and majors for non-pecuniary reasons. Third, an older sibling could affect younger siblings' availability and awareness of options, either by improving the chances of being admitted or by providing relevant information that would otherwise be difficult to obtain. To explore these potential mechanisms, we leverage institutional differences across countries, our rich data and heterogeneity analyses. Taken together we are able to reject several potential mechanisms, but we are unable to distinguish between older siblings changing their younger siblings preferences or changing the awareness of options in their choice set.

Our results contribute to two major strands of research.

First, we contribute to the literature studying peer effects in human capital investment decisions. We provide some of the first evidence that siblings causally affect very important life decisions such as college enrollment, choice of institution and major. A number of recent papers have studied the influence of siblings on educational choices focusing on primary and secondary education.

For example, Qureshi (2018a) shows that additional schooling for Pakistani eldest sisters induces younger brothers to pursue more years of schooling. Gurantz et al. (2020) find that individuals are more likely to take advanced end-of-year exams if their older siblings do so. Joensen and Nielsen (2018) document that Danish older siblings' pursuit of advanced math and science coursework in high school increases younger siblings' propensity to take such courses. Dahl et al. (2020) show that Swedish older siblings and parents influence the field of study that individuals choose in high school. Finally, Dustan (2018) finds that students are more likely to attend a high-school that their older siblings have attended.[3] Goodman et al. (2015) use administrative data to descriptively document that in the United States one-fifth of younger siblings enroll in the same college as their older siblings, and that younger siblings are more likely to enroll in four-year colleges if their older siblings do.[4]

Second, our work informs the literature studying the determinants of post-secondary education decisions and their implications for inequality. As highlighted by Altonji et al. (2016), the decisions of whether to go to college, where to enroll and what to specialize in, are important determinants of future earnings and the type of jobs that people hold. However, we observe large differences in the higher education trajectories of individuals from different social groups characterized by income, education and race Patnaik et al. (2020).[5] The significant differences across groups has been at least partially attributed to barriers to access such credit constraints or differences in school and teacher quality.[6] More recent work has shown that limited information could also influence human

[3]Some papers have also looked at sibling spillovers on academic performance. These studies have found that individuals experience positive spillovers on academic performance from having older siblings with good teachers (Qureshi, 2018b), older siblings who perform better (Nicoletti and Rabe, 2019), and younger siblings who start school at an older age (Landersø et al., 2017). Karbownik and Özek (2019) find positive spillovers for low socioeconomic status siblings, but negative spillovers for high socioeconomic status siblings.

[4]Two contemporaneous working papers show additional evidence on peer effects and sibling spillovers in post-secondary human capital investment decisions for Chile. Barrios-Fernández (2019) uses a regression discontinuity design to investigate extensive margin spillovers from both close neighbors and siblings. Aguirre and Matta (2020) follows an approach similar to ours and studies siblings' spillovers in college choice in Chile. The results in both papers are consistent with our findings that close social peers influence post-secondary education choices.

[5]In the US, students from the top one percent of the income distribution attend Ivy League colleges at a rate 77 times higher than those in the bottom quintile (Chetty et al., 2020). Even among similarly low income students, enrollment rates vary substantially by geography. For those in the 25th percentile of the local parental income distribution, college enrollment rates range from less than 32 percent in the lowest-attending decile of commuting zones to over 55 percent in the highest decile commuting zones. See Online Appendix Figure VII, panel B from Chetty et al. (2014).

[6]Some examples of research studying the role of credit constraints includes Belley and Lochner (2007); Dynarski (2003); Lochner and Monge-Naranjo (2012); Solis (2017), differences in teacher and school quality (Card and Krueger, 1992; Goldin and Katz, 2008; Chetty et al., 2014), spatial variation in college options (Hillman, 2016).

capital decisions on multiple margins.[7]

We build on this work by providing evidence that there are causal links between the post-secondary paths of close peers. Our findings show that shocks to the education trajectory of an older sibling propagate through their family network. This is important because there is vast evidence that some groups face more obstacles and are exposed to more negative shocks than others. Goldin and Katz (2008) argue that the educational system is failing to provide enough development opportunities particularly for poor, minority and immigrant children. Similarly, Scott-Clayton (2012) discusses institutional, behavioral and information barriers that lower socioeconomic status (SES) students face in their path to college. Recent work by Ang (2020) shows exposure to police violence can lead inner-city students to be less likely to enroll in college.

Our results indicate that the consequences of shocks and barriers to access can be amplified by social influences, exacerbating inequality in specialization in higher education and in longer-term economic outcomes. Our findings also suggest that the effects of policies designed to help individuals to overcome these obstacles can also be amplified. Programs—such as financial aid, information interventions or affirmative action—will likely have larger effects than those typically measured because they indirectly benefit younger siblings and potentially other close peers of the direct beneficiaries.

The rest of the paper is organized as follows. Section 2 describes the higher education systems of Chile, Croatia, Sweden and the United States, along with the data we use, and Section 3 details the empirical strategy. Section 4 presents our main results and Section 5 discusses potential mechanisms. Section 6 concludes.

2 Institutions and Data

This section describes the institutional context and data in Chile, Croatia, Sweden, and the United States.[8] As shown in Table I, the four countries are very different in size, economic development

[7] Some examples include Hastings and Weinstein (2008) on school choice, Jensen (2010) on years of education, Hoxby and Turner (2013) on college applications and Hastings et al. (2016) on college and major choice. Recent research investigating how to help students overcome some of these obstacles has shown that "high-touch"information interventions can make a substantial difference in the education choices of low income students (see for instance Bettinger et al., 2012; Carrell and Sacerdote, 2017; Dynarski et al., 2018).

[8] The Online Appendix presents a more detailed description of the relevant institutions in each country.

and inequality. Their higher education systems are also structured very differently. For example, universities in Chile and the United States charge tuition fees, while in Croatia, students receive a fee waiver if they accept the first offer they receive after applying to college, and higher education is free in Sweden.

Most importantly for our analysis, students in Chile, Croatia and Sweden apply to specific college-major combinations through a centralized platform, and admissions decisions are solely based on academic performance. In the United States, students submit separate applications to each college and each institution has its own admission process (which may take into account many factors beyond academic achievement). Thus, many of our analyses and tables separate the US from the other three countries. We provide details for each country below, followed by a description of how admission score cutoffs generate the discontinuities we exploit for identification, and a summary of how we identify our sibling sample.

2.1 Chile

Chile uses a nation-wide centralized admission system. This system allocates applicants to college-major combinations based only on applicants' preference rankings and academic performance. Students compete for places based on a weighted average of their high school GPA and their scores in different sections of a university admissions exam (PSU).

We use administrative data provided by the Chilean agency in charge of college admissions, DEMRE. They provided individual-level data on all students who registered to take the university admission exam between 2004 and 2018. The data include information on students' performance in high school and on each section of the college admissions exam. The data also contain student-level demographic and socioeconomic characteristics, information on applications, and admissions and enrollment in schools that use the centralized application system.[9]

We complement these data with registers from the Ministry of Education, which record enrollment in all higher education institutions in Chile between 2007 and 2015. This information allows us to build program-year specific measures of retention for the cohorts entering the system in 2006

[9]The centralized admission system is used by 33 out of the 60 Chilean universities. This group includes all selective institutions. Note that this does not affect the internal validity of our analyses. See the Online Appendix for more details.

or later. We also observe some program and institution characteristics, including past students' performance in the labor market (i.e. annual earnings). Finally, we are able to match students to their high schools and observe their academic performance before they start higher education.

2.2 Croatia

Similar to Chile, Croatia has a nation-wide centralized application system through which students rank institutions and compete for places based on their academic performance. In Croatia, students apply to college-major combinations and admissions are based on preference rankings and on a weighted average of their high school GPA and their scores on different sections of the university admission exam.

We use administrative data from the central applications office, NISpVU, and the Agency for Science and Higher Education (ASHE). The data contain information on all individuals completing high school and applying to higher education between 2012 and 2018. We observe students' demographic characteristics, their performance in high school and on the college admissions exam, and their applications and enrollment in any Croatian college.

2.3 Sweden

Sweden also has a centralized application and admissions process. Students rank their college-major preferences and are admitted to programs based on their rankings and academic performance. Most students are admitted based only on their high school GPA. There is also a voluntary exam that provides a secondary path to admission.

Our Swedish data come from the Swedish Council for Higher Education (UHR). They include applications from the current admissions system (2006–2017) and an older system (1993–2005). The centralized platform has been mandatory since 2006. Prior to 2006, universities were not required to select their students through the centralized platform, but the majority of universities used it, especially for their larger programs. Thus, in the early period our sample does not include individuals whose older siblings applied to off-platform options. In the more recent period, our sample includes the universe of applicants.[10] The data also contain information on students' high

[10]Note that given the nature of our empirical strategy, not observing these applications does not affect the internal

school GPAs, their scores on the admission exam, and individual and program unique identifiers that allow us to match students and programs to additional registries from Statistics Sweden.

2.4 United States

In the United States, individuals typically apply to colleges (not to specific college-major combinations), and each college sets its own admission criteria. Most colleges take applicants' SAT scores into account and some require minimum SAT scores.

Our main data come from the College Board, who administer the SAT. We observe all students from the high school classes of 2004–2014 who took the PSAT, SAT, or any Advanced Placement exam (all of which are administered by the College Board). We observe each student's name, home address and high school, as well as self-reported demographic information on gender, race, parental education and family income. We also observe scores from each time a student takes the SAT. We observe all colleges to which students send their SAT scores, and we use these score sends as a proxy for college applications (Pallais, 2015).

We merge the College Board data with data from the National Student Clearinghouse (NSC). NSC tracks student enrollment in almost all institutes of higher education in the US, so we can use NSC data to measure students' initial college enrollment (our focus) and all subsequent enrollments and degrees earned.[11] We combine these data with the federal government's Integrated Postsecondary Education Data System (IPEDS), which contains information on college characteristics such as tuition, median SAT score for enrolled students, and whether the school is public or private and two-year or four-year.

2.5 Admission Cutoffs

Our empirical strategy relies on admissions cutoffs. In each country, crossing a program's admissions threshold boosts the probability of gaining admission to and enrolling in the program.

The centralized admissions systems in Chile, Croatia and Sweden generate sharp admissions cut-

validity of our estimates. In the current system, there are also some programs with special admission rules for which we do not observe applications. For most of these programs high school GPA and test scores are not the most important components for selection (i.e. music, art, and acting degrees).

[11] See Dynarski et al. (2015) for NSC data limitations, many of which are for-profit enrollments that most students in our sample are unlikely to attend.

offs in all oversubscribed college-major combinations.[12] Figure I illustrates how older siblings' admissions and enrollment change at admissions cutoffs. The running variable corresponds to older siblings' application scores centered around their target college-major admission cutoff. In Chile and Croatia, the admissions probability increases from 0 to 1 at the cutoff; in Sweden it increases from 0 to 0.6. The Swedish application system has two rounds: individuals submit their rank of preferences at the beginning of the process, and at the end of the first round they can decide whether to accept the offer that they receive or wait for the results of the next round. Since not all applicants wait, some do not receive an offer to their preferred college-major combination even when their application scores were above the cutoff generated in the second round. This explains why the admission probability above the cutoff is only 0.6.[13] Figure I also shows that receiving an offer for a specific college-major increases the probability of enrolling there. However, in none of these countries does admission translate one-to-one into enrollment.

In the United States, where the higher education system and admissions process are decentralized, we focus on the subset of colleges that clearly apply minimum SAT cutoffs in their admissions process but do not publicly announce this process. Using data on SAT scores, applications and enrollment, we empirically identify 21 colleges that appear to employ SAT cutoffs.[14] These colleges are largely public institutions (16 public, 5 private) with an average enrollment of over 10,000 full-time equivalent students, and they are located in eight states on the East coast. The SAT thresholds for these colleges range from 720 to 1060, with students widely distributed across colleges and thresholds. Figure II illustrates how the probability of enrolling in one of these threshold-using colleges nearly doubles at the identified cutoffs.[15]

[12] Because Croatia and Sweden are members of the European Union, it is easier for their students to enroll in foreign institutions. The samples that we use only include individuals whose older siblings apply to programs offered in their home countries. In the case of Sweden, where we can observe if an individual is enrolled in a foreign institution, only 7.4% of high school graduates who attend higher education study at a foreign institution at some point of their careers. Most of them, however, are exchange students and therefore appear in our data. Note that not observing all older siblings does not affect the internal validity of our results.

[13] Note that since each individual represents only one application in a much larger pool of applicants, he or she cannot predict or manipulate the final cutoffs.

[14] The Online Appendix explains in detail how we identified these colleges. We need to focus on sibling pairs in which the older sibling applies to one of these 21 colleges in order to have quasi-random variation in older siblings education trajectories.

[15] We do not observe admissions outcomes in the United States.

2.6 Identifying Siblings

Our research question relies on identifying siblings. In Chile, students provide their parents' national ID numbers when registering for the university admission exam. We can use this unique identifier to match all siblings that correctly reported these numbers for at least one parent.[16] Nearly all students graduating high school in Chile register for the college entrance exam. Although registering for the admission exam costs around USD 40, students graduating from subsidized high schools—93% of total high school enrollment—are eligible for a fee waiver that is automatically activated when they register for the exam. Thus, even students who do not plan to apply to college typically register for the exam.We complement this data with registers from the Ministry of Health that contain records for people born since 1992 and their mothers. We use the the national IDs from these data to link siblings in cohorts completing their secondary education in 2010 or later.

In Croatia and the United States, we identify siblings through home addresses and surnames. In Croatia, we rely on individual reports generated by high schools at the end of each academic year. In the United States, we use the information provided by students when they register for a College Board exam. We identify siblings as pairs of students from different high school classes whose last name and home address match perfectly. We refer to anyone for whom we fail to identify a sibling as an "only child". This approach should yield few false positives, such as cousins living together. This approach, however, likely generates many false negatives in which we mistakenly label individuals with siblings as only children. False negatives come from two sources. First, and unlikely to generate many false negatives, siblings may record their last names or home address differently.[17] Second, in the United States where we observe students' addresses only when they register for an admission exam, we fail to identify siblings in families that change residential addresses. Failing to identify siblings will have no impact on the internal validity of our estimates, but it does affect both sample size and the characteristics of the population we study.

Statistics Sweden provided family linkages for our full sample in Sweden. Thus, we observe the full set of sibling pairs regardless of whether they registered for an admission exam.

[16] 79.4% of students report a valid national ID number for at least one of their parents. 77.2% report their mother's national ID number.

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

Because some families have more than two siblings, we use each family's oldest applying sibling to determine the treatment status of all younger siblings. The vast majority of siblings in our data appear in pairs, but some come from families where we identify three or more siblings.[18] We define families' demographic characteristics based on the oldest sibling for consistency across siblings and because treatment status is determined when the oldest sibling applies to college. We structure the data so that each observation is a younger sibling, whose characteristics and treatment status are assigned based on their oldest sibling. If older siblings applied to college multiple times, we only use the first set of applications he or she submitted.

Our sample consists of approximately $140,000$ sibling pairs in Chile, $17,000$ in Croatia, $220,000$ in Sweden, and $40,000$ in the United States. In Chile, Croatia, and Sweden, these are the number of younger siblings who had an older sibling with least one active application to an oversubscribed program and an application score within the relevant bandwidths for our regression discontinuity design. In the United States, these are the younger siblings with an older sibling who applied to at least one of the 21 cutoff using colleges in our sample, and had an SAT score near the admissions cutoff.

Table II presents summary statistics for these sibling pairs and for the full set of potential applicants.Individuals with older siblings who already applied to higher education are slightly younger when they apply to college than the rest of applicants and, not surprisingly, they come from bigger households. Since our sample is based on families with at least one college-applying child, it is not surprising that some differences also arise when we look at socioeconomic and academic variables. In Chile and the United States, individuals in the discontinuity sample come from wealthier and more educated households than the rest of the potential applicants. They are also more likely to take the admission exam, and with the exception of the United States, perform better on it.

3 Empirical Strategy

We use admission score cutoffs to identify the impacts of older siblings' college trajectories on younger siblings' college and major choice. In Chile, Croatia and Sweden we exploit thousands

[18]In the Online Appendix we present alternative specifications in which we focus instead on the closest older sibling. We also present specifications in which we focus only on the first- and second-born children in the family. The results are remarkably similar to the ones we report in the body of the paper.

of cutoffs generated by the deferred acceptance admission systems universities use to select their students. In the United States we exploit the variation generated by cutoffs that 21 colleges use in their admission processes (and do not disclose to students).

We use these admission cutoffs in a regression discontinuity (RD) design, which helps us overcome typical challenges in identifying sibling effects. The RD compares younger siblings whose older siblings are similar to one another across most dimensions except that some older siblings score just above an admission cutoff and others score just below it. These small differences in test scores change the educational trajectories of the older siblings and have the potential to influence the younger siblings. Since individuals whose older siblings are near an admission threshold are very similar, the RD allows us to rule out that the estimated effects are driven by differences in individual or family characteristics, which eliminates concerns about correlated effects. We can also rule out concerns related to the reflection problem Manski (1993) because the variation in older siblings' education paths comes only from being above or below the cutoff, and thus cannot be affected by the choices of their younger siblings.

3.1 Method

This section describes the specification we use to estimate how older siblings' higher education trajectories influence the colleges and majors to which their younger siblings apply and enroll. We separately estimate sibling spillovers in each country. For each sample, we pool observations from all applicants to the relevant colleges and college-majors (which includes all oversubscribed college-majors in Chile, Croatia and Sweden and "cutoff-using" colleges in the United States). We center older siblings' application scores around the admission cutoff of their "target" college or "target" college-major depending on the setting, and estimate the effect of an older sibling being above the relevant cutoff. The following equation describes our baseline specification:[19]

$$y_{icmt\tau} = \beta \times \text{above-cutoff}_{icm\tau} + f(a_{icm\tau}; \theta) + \mu_{cm\tau} + \varepsilon_{icmt\tau}. \tag{1}$$

[19]In the United States the variation is at the college level, so we can eliminate the major subscript. In addition, the cutoffs are constant over time in the United States. Thus, the term $\mu_{cm\tau}$ is replaced by μ_c and μ_τ. See the Online Appendix for a detailed description of the procedure we use to identify these cutoffs in the United States.

$y_{icmt\tau}$ indicates whether the younger sibling from sibling-pair i and birthyear t whose older sibling was near the admission cutoff of major m in college c in period τ applies to or enrolls in the target college-major, college or major of the older sibling. above-cutoff$_{icm\tau}$ is a dummy variable indicating whether the older sibling from sibling-pair i had an admission score $a_{icm\tau}$ above the cutoff ($c_{cm\tau}$) of major m offered by college c in year τ ($a_{icm\tau} \geq \bar{c}_{cm\tau}$). $f(a_{icm\tau})$ is a function of the application score of the older sibling of the sibling-pair i for major m offered by college c in year τ. $\mu_{cmt\tau}$ is a fixed effect for the older sibling's cohort and target college-major, and ε_{icmt} is an error term.

By including fixed effects $\mu_{cmt\tau}$ for each cutoff, our identification variation only comes from individuals whose older siblings applied to the same target college in the United States and the same target college-major in Chile, Croatia and Sweden.

Our main results are based on local linear regressions in which we use a uniform kernel and control for the running variable with the following linear function:

$$f(a_{imc\tau};\theta) = \theta_0 a_{imc\tau} + \theta_1 a_{imc\tau} \times 1[a_{imc\tau} \geq c_{mc\tau}].$$

This specification allows the slope to change at the admission cutoff. In Appendix B we show that our results are robust to using a quadratic polynomial of $a_{imc\tau}$, a triangular kernel, and to allowing the slope of the running variable to be different for each admission cutoff. To study the effect of enrollment—instead of the effect of admission—we instrument older siblings' enrollment (enrolls$_{imc\tau}$) with an indicator for admission (above-cutoff$_{imc\tau}$).

We compute optimal bandwidths according to Calonico et al. (2014). In the United States analyses we use a bandwidth of 93 SAT points, which is the median (and mean) optimal bandwidth for the main outcomes that we study. In Chile, Croatia and Sweden, we compute the optimal bandwidth for our three main outcomes: ranking the older sibling's target option in the first preference, ranking it in any preference, and enrolling in it. For each country, we use the smallest of these bandwidths, so that our bandwidths are consistent across outcomes and specifications.[20]

In the centralized admission systems used in Chile, Croatia and Sweden individuals can be admitted

[20]In principle, optimal bandwidths should be estimated for each admission cutoff independently. However, given the number of cutoffs in our sample, doing this would be impractical. Therefore, we compute optimal bandwidths pooling all the cutoffs. Appendix B shows that our estimates are robust to different bandwidth choices.

at most in one college-major. However, they can narrowly miss several options ranked higher in their application list. This means that in principle they may belong to more than one college-major marginal group. We cluster standard errors at the family level to account for the fact that each older sibling may appear several times in our estimation sample if she/he is near two or more cutoffs, or if she/he has more than one younger sibling.

In Appendix B we present a variety of additional robustness checks. As expected, changes in the admission status of younger siblings do not have an effect on older siblings, our estimates are robust to different bandwidth choices, and placebo cutoffs do not generate a significant effect on any of the outcomes that we study.

3.2 Estimation Samples

In Chile, Croatia and Sweden, we use information on older siblings' next best option to define three estimation samples that we use to study sibling spillovers on college choice, college-major choice, and major choice (across all colleges).[21]

- College-Major Sample—Since college-major combinations are unique, being above or below a cutoff always changes the college-major combination to which an older sibling is admitted.[22] This sample includes all individuals whose older siblings are within a given bandwidth for a target cutoff.

- College Sample—Our estimates of sibling spillovers on college choices are based on individuals whose older siblings' target and next best college-major preferences are taught at different colleges. For these older siblings, being below or above the admission threshold changes the college to which they are assigned.[23]

[21] Appendix A presents a more detailed description of these samples.

[22] In some cases, universities use slightly different names for similar majors or change them over time. Thus, in order to make majors comparable across institutions, time, and settings, we classify them into three digit-level ISCED codes. An individual whose older sibling enrolls in economics at the University of Chile is said to choose the same major as her older sibling if she/he applies to Economics (0311) in any college. She/he is said to choose the same college-major combination as her/his older sibling only if she/he applies to the exact same degree—Economics—in the exact same college—University of Chile.

[23] In Appendix B we present additional results that investigate sibling spillovers on college choice in a modified 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.

- Major Sample—To investigate sibling spillovers in major choices, we exclude all individuals whose older siblings' target and next best college-major option correspond to the same major.[24]

3.3 Identifying Assumptions and Alternative Specifications

As in any RD setting, our estimates rely on two key assumptions. First, individuals should not be able to manipulate their application scores around the admission cutoff. Since the exact cutoffs are not known when students apply to college, such manipulation is very unlikely. We find no indication of manipulation when we study the distributions of the running variable in each setting (see Appendix B for more details).

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

To investigate the effect of an older sibling's enrollment on younger siblings choices we rely on a fuzzy regression discontinuity design. This approach can be thought of as an instrumental variable strategy, meaning that in order to interpret our estimates as a local average treatment effect (LATE) we need to satisfy the assumptions discussed by Imbens and Angrist (1994).[25] In addition to the usual IV assumptions we also need to assume that receiving an offer for a specific college or college-major does not make enrollment in a different option more likely. [26] Given the structure of the admission systems that we study, this additional assumption is not very demanding.[27]

[24] In Appendix B we 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 older sibling's major, but not college.

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

[26] Appendix A presents a detailed discussion of these identification assumptions.

[27] In Chile—where not all colleges use centralized admissions—or in the United States—where each school runs its own admission system—this assumption could be violated if, for instance, other colleges were able to offer scholarships or other types of incentives to attract students marginally above the admission cutoffs for other institutions. Although it does not seem very likely that colleges would define students' incentives based on admission cutoffs that they only observe ex-post or do not observe at all, we cannot completely rule out this possibility. In Croatia—where students lose their funding if they reject an offer—and Sweden—where there are no tuition fees and where all the universities

We also show, in Appendix B, that older siblings' marginal admission to their target college-major does not generate a relevant difference in their younger siblings' total enrollment in Chile, Croatia and Sweden. This result relieves concerns about increases in applications and enrollment in an older sibling's target choice being driven by a general increase in college enrollment. This issue is more relevant in the United States, where we document that older siblings crossing an admission threshold induce an increase in 4-year college enrollment among younger siblings. Decomposing this extensive margin response among those following their older siblings to the same college and those going somewhere else helps us understanding how siblings influence higher education decisions. In section 4 we discuss this decomposition in more detail and show that the increase that we find in younger siblings' enrollment in the target college of their older siblings in the United States is much larger than the increase we would observe in the absence of sibling spillovers in the choice of college.

Kirkebøen et al. (2016) argue that when estimating returns to fields of study, controlling for the next best option is important both for identification and for interpreting the results. Since we observe older siblings' next best options in Chile, Croatia and Sweden, in Appendix B we present results that include controls for two-way interacted fixed effects for both target and next-best major-college. These estimates are very similar to the ones presented in Section 4, even though including two-way fixed-effects puts a considerable strain on statistical power. It is important to note, however, that our research question is very different from the one addressed in Kirkebøen et al. (2016). Thus, while in their context it is important to identify the baseline against which returns are computed, it is less important here because we are interested in whether individuals are more likely to apply to and enroll in a college program if an older sibling enrolls there independently of the counterfactual option of the older sibling.[28]

Our baseline specification compares the higher education choices of individuals whose older siblings are marginally above or below specific admission cutoffs. Since we pool many admission cutoffs, our estimates represent a weighted average of the effect of having an older sibling crossing an admission

allocate places through the centralized platform—violations of this assumption seem unlikely.

[28] Appendix A discusses in detail the identifying assumptions that we require in this setting. Considering that in our case there are thousands of college-major combinations available, it is not feasible to follow the approach of Kirkebøen et al. (2016) and independently estimate responses with respect to each next best option. We discuss next an extension of our baseline specification that deals with some of the identification and interpretation concerns raised in their work.

thresholds and gaining admission to their target program as a consequence. At each admission cutoff the counterfactual is a mix of the next best options for each older sibling. By using the samples that we defined earlier in this section, we guarantee that the next best option for the older sibling is a different major-college, a different college, or a different major depending on the outcome we are investigating.[29]

In order to gain a better understanding of what is driving the average effects we document, we exploit the information we have on the target and next best options of older siblings in Chile, Croatia and Sweden. We estimate the following specification:

$$y_{icmt} = \alpha_0 + \sum_{j=1}^{4} \beta_j \text{above-cutoff}_{icm\tau} \times Q_j + f(a_{icm\tau}; \theta) + \mu_{cm\tau} + \mu_{c'm'\tau} + \varepsilon_{imct\tau} \tag{2}$$

As before, y_{icmt} is a dummy variable that indicates whether younger siblings apply to or enroll in their older sibling's target program. However, this time we estimate the effect of crossing the admissions threshold for four groups. To define these groups we first compute the difference between older siblings' target and next best option along a relevant dimension (expected earnings, peer quality or first year retention rate). Each group Q_j corresponds to a quartile in the distribution of this difference. While the differences in the bottom quartile are negative, in the top quartile they are positive (Figure VII illustrates the distributions of these differences). This specification also controls for target ($\mu_{cm\tau}$) and next-best ($\mu_{c_{nb}m_{nb}\tau}$) option fixed effects.

For older siblings, crossing the admission threshold of their target program changes the characteristics of the college-major to which they are allocated. This specification allows older siblings' effects on their younger siblings to vary with the size of the change they experience when crossing the threshold. We further investigate heterogeneous responses by estimating a similar specification in which we construct quartiles from the levels of characteristics in older siblings' target programs instead of the differences with respect to their next best options.

[29] In the United States we do not observe next best options. However, since applications are made at the college level, crossing the threshold changes the college to which individuals are admitted. In the Online Appendix we show that in this setting crossing the threshold increases older siblings' probability of attending a 4-year college by 36 percentage points. The probability of enrolling in some college—either a 2-year or 4-year college—is not affected. This means that for an important share of US compliers, the next best option is a 2-year college.

4 Results

This section presents results on sibling spillovers. First, we show that younger siblings are likely to follow the same higher education trajectory as their older siblings. Second, we show that following an older sibling can be of great consequence, sometimes dramatically shifting the type of college in which a student enrolls. In some instances, this shift impacts the quality of the younger sibling's college choice, as measured by peer achievement, expected earnings and degree completion rates.

4.1 Following an Older Sibling

Across all four countries, an older siblings' admission to a college increases their younger sibling's probability of applying to and enrolling in that same college. We illustrate this causal relationship in Figure IV for Chile, Croatia and Sweden and in Figure III for the US. These figures show the reduced-form relationships, separately for each country, between an older siblings' admissions score and the younger siblings' application to and enrollment in the same college. Each figure indicates a sharp discontinuity in the younger sibling's outcome as a function of the older sibling's admissions score. In Chile, Croatia, and Sweden, younger siblings are more likely to rank a college first in their application portfolio if their sibling is admitted. The rows labeled "older sibling above cutoff"in Table III show the reduced form estimates for Chile, Croatia, and Sweden. In the US, younger siblings are 2.3 pp more likely to apply to and 1.4 pp more likely to enroll in the older sibling's target college if the older sibling scores above the admission cutoff.

Figure V shows that individuals are more likely to apply to and enroll in a college-major combination if an older sibling was admitted to it. Figure VI, however, shows that older siblings' admission into their target major does not significantly impact the probability that their younger siblings apply to or enroll in that major (at any institution). Thus, the influence on major choice seems very local; individuals only follow majors in the same college of the older sibling.

Next, we combine these reduced form estimates with our first stage results (i.e. Figure I) to obtain the fuzzy-RD estimates in Tables III and IV. These estimates represent the effect of an older sibling's enrollment in a target college, college-major, or major on the younger sibling's probability of applying to or enrolling in the same program.[30]

Younger siblings are more likely to rank a college as their first preference, to apply to the college, and to enroll in it when the older sibling enrolls (as a result of barely gaining admission). Columns (4)–(6) in Table III summarize these results in Chile, Croatia and Sweden. In these countries, individuals are 6.7 pp to 12 pp more likely to rank their older siblings' target college as their first preference and between 7.6 pp and 13.2 pp more likely to apply (with it in any preference rank) when the older sibling enrolls there. The increase in applications to the older sibling's target college also translates into an increase in enrollment between 3.8 pp and 8.4 pp.

Older siblings have larger effects on applications and enrollment in the U.S. Table IV shows that younger siblings are 27.9 pp more likely to apply to and 17.2 pp more likely to enroll in their older siblings' target college if the older sibling was admitted and enrolled there. Thus, in all four countries, an older sibling's enrollment in a particular college increases the likelihood of applying to and enrolling in that college.[31]

We also leverage the rich data on college-major and major preferences in Chile, Croatia, and Sweden to examine whether an older sibling's college-major or major choice leads the younger sibling to follow them in these margins as well. In these countries, an older sibling's enrollment in her/his target college-major combination makes younger siblings between 1.2 pp to 2.0 pp more likely to rank the exact same option in their first preference, between 2.3 pp and 3.6 pp more likely to rank it in any preference, and between 0.5 pp to 1.3 pp more likely to enroll in it. These estimates are smaller than those for enrollment in the same college, indicating that many students who follow an older sibling to a college do not choose the same major. These, however, are still meaningful effects, especially when taking into account the low baseline levels in the control group.

[30] If an older sibling's admission to a target option affect younger sibling choices even when the older sibling does not to enroll there, the IV estimates we present would overstate the effects of an older sibling's enrollment on younger sibling choices. Note, however, that the reduced form results will still be valid.

[31] In the next section, we show that older siblings' enrollment in their target college increases enrollment in any 4-years college. This means that the effect that we document here could be in part a mechanical consequence of the increase in the share of individuals going to any 4-years college. However, the size of the effects makes it unlikely that our results are only a mechanical consequence. On the left of the admission cutoffs the share of individuals enrolling in the target college of their older sibling is 1.58% (0.006/0.38). On the right hand side it is 29.2% (0.178/0.609). If preferences were stable around the cutoff and older siblings did not affect preferences for specific colleges, we should find 1 pp (1.58% $\times$ 60.5%) of the younger siblings on the right side enrolling in the target college of their older sibling. However, the increase that we find in enrollment is 17.2 pp, well above the 0.4 pp increase that we should find in the absence of spillovers.

Finally, in columns (7)–(9) of Table III, we study whether preferences for majors—independent of the college that offers them—are influenced by older siblings' choices. We focus on the major sample defined in Section 3, which only includes individuals whose older sibling's target and next-best option correspond to different majors. In contrast to the strong college-choice spillover effects, we find almost no influence on major choices. None of the estimates are statistically significant at conventional levels and, in general, the coefficients are small.

These results show that younger siblings' major choices are only locally affected. Younger siblings are not more likely to apply or enroll in the older sibling's major in any college, but they do follow the older sibling to the same college-major. In order to further investigate these effects on major choices, we build a new sample that only includes 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 timing means that if, after receiving their scores, younger siblings believe they are unlikely to gain admission to their older sibling's college-major they might not apply there. Thus, for this exercise we further restrict the sample to individuals who are likely to be admitted to their older siblings' target college-major if they apply (we present these results in Appendix B). Although our estimates are not always precise, the sibling spillovers that we find on college-major choices in this sample are larger than the ones we present in this section.[32]

We find evidence across all four countries that an older sibling's educational trajectory has a causal effect on the younger sibling application and enrollment decisions. Next, we examine the consequences of this behavior.

4.2 Does Following an Older Sibling Matter?

In this section, we examine when students are most likely to follow their older siblings and whether this changes the types of colleges and majors that younger siblings attend.

[32]In Appendix B we present results from a similar exercise in which we investigate spillovers on college choice focusing only on individuals whose older siblings' target and next best options correspond to the same major, but are offered at different colleges. The results that we find are very similar to the ones we document for college choice in the current section.

First, we show that younger siblings follow older siblings independent of the characteristics of the program attended by the older sibling. We use the full rank of preferences observed for applicants in Chile, Croatia and Sweden to estimate how younger siblings' choices vary with the characteristics of their older sibling's counterfactual options. We estimate specification 2, which allows younger siblings' responses to change depending on the difference between the older sibling's target and next best options along three dimensions: expected earnings, peer quality, and first year retention rates. We classify the differences in quartiles and allow the effect to be different for sibling pairs in each quartile.[33]

Older siblings' counterfactual options are often very similar (see Figure VII). However, we find that younger siblings not only follow their older siblings when the older sibling is on the margin of very similar alternatives, but also when the differences between these options are large. Table V summarizes these results. It indicates that, independent of the difference between older siblings' target and next best options, having an older sibling admitted to a given college or college-major increases the probability that their younger siblings apply there. This means that some individuals follow their older siblings to institutions with worse peers, lower retention rates and lower expected earnings than the older sibling's next best option.

We find similar patterns when we estimate specification 2, but define the quartiles based on the levels of the characteristics in older siblings' target options and not on differences. Table VI shows that an older sibling's admission to her target college-major increases the probability that the younger sibling applies to the same college, independent of the quality of the older sibling's target. The effects are remarkably stable across groups in Croatia and Sweden. The results in Chile are, for the most part, positive and significant. The only individuals for whom we find no significant effects are those whose older siblings enroll in a college-major with very low retention rates. We also find positive and significant effects when looking at applications to the older sibling's target college-major.[34]

Overall, our results show that individuals follow their older siblings' both when crossing the admis-

[33] We estimate this specification in samples that are slightly different from the samples in the previous section. Here we only include observations in which we observe both the older sibling's target and next best options. This restriction, for instance, excludes individuals whose older siblings were marginally rejected from their last preference.

[34] The Online Appendix presents similar results focusing on enrollment instead of applications.

sion threshold implies a gain and when it implies a loss in expected earnings, peer quality or first year retention rates. These results suggest that individuals do not learn from their older siblings about all available alternatives and their relative quality; instead, they seem to learn about the institution the older sibling enrolls in. These findings also suggest that social spillovers are likely to amplify the effects of frictions and barriers that prevent individuals from making optimal education choices. By affecting the choices of close peers, these obstacles add to the inequality that we observe in educational trajectories.

In the U.S. we do not observe applicants' counterfactual college options. However, we find that crossing an admissions threshold increases older siblings' likelihood of enrolling in a four-year college. When measuring the outcomes of older and younger siblings we focus on their initial enrollment decisions; we study what they do the year after completing high school. This increase is largely due to these students being more likely to attend their target (four-year) college than a two-year college.[35] IV estimates indicate that nearly half of the marginal older siblings induced to attend target colleges by admissions thresholds would not have attended four-year colleges if they had not passed the admissions threshold.[36] This behavior contrasts with what we observe in Chile, Croatia and Sweden, where most of the older siblings in our sample enroll in a 4-year college.[37]

Older siblings' increased access to four-year colleges has important consequences for younger siblings in the U.S. Plot (c) of Figure III indicates that older siblings' marginal admission to their target college substantially increases younger siblings' enrollment in four-year colleges. The IV estimate in column (1) of Table IV shows that applicants whose older siblings enroll in their target four-year college are 23 pp more likely to enroll in any four-year college than students whose older siblings just miss the cutoff. Column (2) shows a small and insignificant decrease in two-year college enrollment. This decrease indicates that the older sibling's admission to her target college leads to some younger sibling movement from two-year to four-year colleges, as well as increased enrollment among younger siblings who would not have attended college otherwise.

[35]Figure II shows that older siblings with SAT scores above the target college's admission cutoff are 8.5 pp more likely to attend that college than students with scores just below the threshold.

[36]The Online Appendix shows that older siblings' scoring above the cutoff in their target college are 36 pp points more likely to attend a 4-year college, and 28 pp less likely to attend a 2-year college. Thus, only a small fraction of the marginal older siblings would not have attended college if they had not crossed the threshold.

[37]The Online Appendix shows that in these three countries older siblings' admission to their target option does not affect younger siblings' enrollment in 4-year colleges.

This increase in enrollment is also evident in columns (3) and (4) of Table IV, which show that older siblings' admission to target colleges improves the quality of the educational path followed by younger siblings. Here we define quality as the bachelor's degree completion rate and the standardized PSAT scores for students attending the institution.[38] We assign students who do not enroll in college a bachelor's degree completion rate of zero, and the mean PSAT score for all students who do not enroll in college. Younger siblings whose older sibling attended the target college enroll in colleges with graduation rates 18 pp higher and peer quality 0.31 standard deviations higher than the colleges they would have chosen otherwise.

Our results also indicate that the most responsive younger siblings are the "uncertain college-goers."These are students whose predicted probability of attending college—based on observable characteristics—is in the bottom third of our sample.[39] Older siblings appear to have little impact on the type of institution attended by younger siblings who are probable college goers. Overall, these results are consistent with older siblings providing general college information, which makes younger siblings—especially those less likely to know about their college options—more likely to enroll in a four-year college.[40]

The results discussed in this section show that shocks affecting an older sibling's education trajectory can be of great consequence for their younger siblings. Across all four countries, younger siblings follow their older siblings even when there are large differences in their counterfactual options. In the U.S., where many of the younger siblings in our sample are on the margin of attending college, an older sibling's enrollment in a 4-year college induces them to follow the same path.

[38] We build a peer quality measure following Smith and Stange (2016) and compute the average standardized PSAT score of initial enrollees for each college. This peer quality measure allows for comparisons between two- and four-year institutions; two-year colleges do not require SAT scores and thus lack a peer quality measure in IPEDS.

We build a second quality index using the NSC data to compute the fraction of initial enrollees at each college who earn a B.A. from any college within six years. Unlike the IPEDS graduation rate measures, this accounts for transfers between institutions and allows for direct comparisons of two- and four-year colleges.

[39] To predict the likelihood of enrolling in a four-year college, we use the sample of "only children" and the socioeconomic and demographic characteristics that we observe in the College Board data.

[40] Additional results in the Online Appendix show that the strength of sibling spillovers does not vary by socioeconomic status for siblings in Chile, Croatia and Sweden. However, in these countries most older siblings in our samples are likely to enroll in 4-year colleges, suggesting that the individuals we study are not marginal college goers.

5 Mechanisms

Our results in Section 4 show that older siblings' higher education trajectories influence the trajectories of their younger siblings. Older siblings' pathways play an important role in the younger sibling's decisions both to attend college and which college to attend. We also find spillover effects on the choice of major, though they seem to be relevant only for individuals that can follow their older siblings to the exact same college-major combination.

To properly identify causal effects, our analyses focus on changes in older siblings' educational paths that arise from admissions cutoffs. This is likely to capture only a small part of siblings' influence on education trajectories. Considering the source of variation that we exploit, and the fact that an oldest sibling is only one member of an individual's social network, our estimated effects are large.

Our results are also large compared to the effects of previously studied college-going interventions. Most nudge-style informational interventions at the state or national scale fail to meaningfully affect college enrollment choices. Higher touch interventions that complement information with some type of personalized support have been more effective. Bettinger et al. (2012), for instance, finds that helping families apply for funding increases college enrollment by 8pp, while Carrell and Sacerdote (2017) finds that assigning females to a mentoring program increases college enrollment by 15pp; among those who actually took part in the program the effect is twice as large. These estimates are similar to the increase we document in 4-year college enrollment in the United States. In terms of college choice, Hoxby and Turner (2015) shows that providing students with customized information about different dimensions of the college experience and reducing application costs increases enrollment in institutions with similar peers by 5.3 pp. This effect is smaller than our estimate of sibling spillovers on college choices in the United States, and is of similar magnitude to our estimates from the other three countries.

In the rest of this section, we estimate heterogeneity in sibling effects across settings and outcomes to investigate the mechanisms behind sibling spillovers. We focus on three broad classes of mechanisms through which older siblings are likely to affect the choices of their younger siblings. First, the older sibling's educational trajectory could affect the household budget constraint, and possibly the value proposition of a specific institution or major. Second, the older sibling's outcomes could affect the utility the younger siblings derive from different higher education trajectories. Finally, older siblings' experiences could affect younger siblings' choice sets by making some options more salient, or by providing information that would otherwise be difficult to obtain.

5.1 Heterogeneity in Sibling Spillovers

This section presents several heterogeneity analyses to help us investigate potential mechanisms driving our results.

First, we explore differences in younger siblings' responses to their older siblings' college choices based on siblings' age differences and genders.[41] Table VII summarizes these results. Column (1) investigates differences by sibling age gap and siblings' gender on enrollment in any four-year college; columns (2)–(5) focus on the probability that younger siblings apply to their older sibling's target college; and columns (6)–(8) focus on the probability that they apply to their older sibling's target college-major.

Results from the United States suggest that the effects on the decision to enroll in a 4-year college and on the specific college chosen are stronger for siblings born five or more years apart. These results contrast with our findings for Chile, Croatia and Sweden, where we find that the probability of following an older sibling to her target college decreases with the age gap. Despite this decrease, there is still a significant and meaningful effect even for siblings born more than five years apart. We find a similar pattern when looking at the choice of college-major. In this case, the magnitude of the effect also decreases with the age gap, but there is still a significant effect for siblings with large age differences.

The fact that siblings who are more than five years older than their younger sibling still influence their college choices means that sibling spillovers are not just about a younger sibling wanting to be on campus with their older sibling. In addition, the shrinking size of spillover effects as age gaps grow in Chile, Croatia and Sweden might indicate that individuals pay more attention to what happens with a sibling who is more similar to them.[42]

[41] The analyses presented in this section focus on **applications**. We present similar results for **enrollment** in the Online Appendix. The Online Appendix also includes a more detailed discussion on gender differences.

[42] Even if age difference does not explain how close two siblings are, the experience of an older sibling closer in age might be a better proxy for what younger siblings could expect from a college.

To further explore how siblings' similarity affects the strength of the sibling spillovers that we document, we next investigate whether responses vary by siblings' gender. In the US, effects on 4-year college enrollment are stronger among siblings of opposite genders, but we find no gender differences in the probability of applying to the older sibling's target college. In Chile, Croatia and Sweden, we do not find either heterogeneous effects by gender in the probability of following an older sibling to college. However, when looking at the probability of applying to the older sibling's target college-major, we find that individuals are more likely to follow an older sibling of the same gender.[43]

Next, we explore whether sibling spillover effects persist if the older sibling has a negative experience in college. We estimate the effect of older siblings' college enrollment for older siblings who drop out of their target program. Since the decision to leave college could be affected by having a younger sibling at the same school, we focus on first year dropouts and siblings who are at least two years apart in age.

Table VIII shows that siblings' effects disappear if the older sibling drops out. This result is consistent with the hypothesis that individuals learn from their older siblings' college experiences whether a specific college-major or college would be a good match for them. The results of this exercise should be interpreted with caution because dropping out of college is not random. Although controlling for the baseline effect of dropout helps us capture some of the differences between individuals who remain at or leave a particular college, there could still be differences we are unable to control for. In addition, we can only build the dropout variable for older siblings who actually enroll somewhere.[44]

These results suggest that younger siblings are more likely to follow their older sibling if the older sibling has a positive experience in college. However, in light of the results from section 4.1, this effect primarily operates through dimensions that are not related to a program's average expected

[43]The Online Appendix presents a more detailed discussion of heterogeneous effects by gender. The heterogeneous effects we find in the probability of following an older sibling to the same college-major is driven by males being more likely to follow older brothers. Indeed, we do not find evidence of females' college-major choices affecting or being affected by a sibling.

[44]The Online Appendix shows that in Chile and Sweden, marginal admission does not translate into increases in older siblings' college enrollment. Thus, in these countries, we focus on older siblings who enroll in college. In the United States, on the other hand, marginal admission increases older siblings' enrollment and we include everyone in the estimation sample. Since we can only define dropouts for older siblings who enroll, this specification does not control for its main effect.

earnings, peer quality and retention rates. Thus, the specific experience of the sibling seems much more important than the average experience of students in the program.

5.2 Sibling Spillovers on Academic Performance

Next we study older siblings' effects on younger siblings' college preparation and academic performance. We estimate our baseline specification using various measures of younger siblings' academic performance as the outcomes. When looking at changes in younger siblings' scores we focus on the subset of individuals who actually take the test. Since not all younger siblings take an admissions exam, these results need to be interpreted with caution. We use the same bandwidths as in the previous sections.

Table IX shows that an older sibling's enrollment in her/his target program does not significantly change younger siblings' high school grade point average. We also find no significant increases in the probability of taking the college admission exam.[45] In Chile and Croatia, we do not find spillovers on younger siblings' performance on the college admission exam. In Sweden and the U.S., younger siblings perform better when their older siblings enroll in their target program. The results in Sweden should be interpreted with caution because we find a decrease in test-taking rates, so this result could be driven by selection. The increased exam performance in the US is imprecisely estimated, but large enough that it may be economically meaningful.

Finally, we do not find significant increases in college applications. In Chile, Croatia and Sweden, where we study the effect on applications using a dummy variable for whether younger siblings submit at least one application, we find a small and insignificant decrease in applications. In the United States we look instead at the total number of applications submitted. In this setting, we find that an older sibling's enrollment in her/his target college increases the number of applications the younger sibling submits by 0.159. This is also a small and insignificant effect.

On balance, these results suggest that sibling effects on college and college-major choices are not driven by an improvement in the academic performance or college preparation of younger siblings.

[45] In Sweden, where students do not need to take the admission exam to apply, we find a small (significant) decrease in the share of younger siblings taking it. In the United States we find that individuals whose older siblings enroll in their target college are 7.3 pp more likely to take the SAT, but this coefficient is not statistically significant.

5.3 Discussion

We discuss and explore the three broad classes of mechanisms we introduced at the beginning of this section that could drive siblings effects.

First, older siblings' college enrollment can affect the family budget constraint. On the extensive margin, an older sibling's attendance at her target college could reduce the resources available for financing the younger sibling's education. However, our results from the United States indicate that older siblings' enrollment increases younger siblings' four-year college enrollment. This indicates that the additional costs faced by families when one child enrolls in college do not outweigh the positive effects on the younger sibling's college enrollment.[46]

An older sibling's enrollment in a particular college campus may affect the costs faced by younger siblings in other ways. For instance, siblings attending the same college may save on commuting and living costs. An older sibling's enrollment may also increase the amount of financial aid available for the younger sibling, or colleges may offer siblings a tuition discount. In the four countries that we study, sibling spillovers persist even among siblings who, due to age differences, are unlikely to attend college at the same time. In addition, universities do not charge tuition in two of the four settings we study. Thus price effects seem unlikely to explain much of the observed spillovers.[47]

Sibling spillovers could arise if colleges offer family members an advantage in the admissions process. In the United States, legacy effects are common because some colleges give admissions preferences to students whose family members have previously enrolled. Hurwitz (2011) noted that this practice is more frequent among colleges seeking to increase donations. Legacy effects are, however, unlikely to explain the spillovers we find because the target colleges we identify in the United States are largely public, non-flagship institutions, and legacy admissions are concentrated in more prestigious colleges. In addition, colleges in Chile, Croatia and Sweden select their students based only on their previous academic performance, so legacy effects play no role in these countries.

Second, an older sibling's enrollment in a specific college or major could affect individual preferences. Preferences may change if younger siblings experience utility gains from being close to their older

[46]The Online Appendix shows that in Chile, Croatia and Sweden having an older sibling enrolling in her/his target college-major does not reduce total enrollment among younger siblings.

[47]In the Online Appendix we show that the effects do not seem to be driven by location preferences either.

sibling, perhaps because they enjoy the company of their older sibling or because they think their older sibling can support them and make their college experience easier. Preferences may also be affected if older siblings are seen as role models and younger siblings are inspired by them, if siblings are competitive, or if parental pressure changes as a consequence of older sibling enrollment.

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

The lack of effects on younger siblings' academic performance and college preparation also suggests that individual aspirations and parental pressure to apply to and enroll in college are not important drivers of our findings. If this were an important channel, we would expect to see younger siblings exerting additional effort in preparation for college. Joensen and Nielsen (2018) argue that the fact that their results (on spillovers in high school) are driven by brothers who are close in age and in academic performance is evidence that competition is driving their results. This does not appear to be the case in our setting because our results persist even among siblings with large age differences and among opposite gender siblings.

Finally, an older sibling enrolling in a specific college or college-major could affect the choice set of their younger siblings by making some options more salient or by providing information about relevant attributes of the available options.[48] Since applicants face a huge number of college and major options, both hypotheses could play an important role. An older sibling's enrollment at a particular college may generate information for parents or a younger sibling that would otherwise be costly or impossible to obtain.

Evidence on when individuals are most likely to follow their older sibling suggests that their older siblings' experiences are more relevant than the average experiences of other students on campus. Our results for Chile, Croatia and Sweden show that individuals follow their older siblings when there are both positive and negative differences between the older sibling's target and next best

[48] Hastings et al. (2015) and Conlon (2019) show evidence from a randomized control trial that information about earnings of graduates could potentially affect college and major choice.

options in terms of expected earnings, peer quality and first year retention rates. While we do not observe older siblings' counterfactual options in the U.S., our estimates indicate that crossing an admissions threshold moves many older siblings from two-year to four-year colleges. This large change in older siblings' educational trajectories also impacts their younger siblings' choices, especially among uncertain college-goers.

Our results are consistent with individuals placing particularly high weight on their family members' college experiences because the educational success of a close relative is more salient and predictive of one's own success than more general sources of information. The fact that sibling spillovers vanish if the older sibling drops out suggests that older siblings' experiences matter, and that the younger sibling updates her choices accordingly. These results also suggest that some of the information transmitted between siblings is related to quality aspects that we do not measure. In line with this reasoning, recent research suggest that non-pecuniary aspects of college life matter more than labor market prospects for applicants' preferences (Wiswall and Zafar, 2014; Patnaik et al., 2020). It might very well be that younger siblings learn about the social life and general satisfaction of students at their older sibling's institution, and this information could be more important than information readily available about other programs.

Although these results are consistent with information transmission, we cannot rule out that part of the effects are driven by changes in younger siblings' preferences. Finding that older siblings are followed when the shocks affecting their higher education trajectories move them to better, but also to worse options may indicate that for younger siblings there is an intrinsic value in following the path of an older sibling. This could also explain why some of them follow their older siblings to what seem worse educational paths.

Even though the evidence discussed in this section does not allow to perfectly identify the mechanisms behind our findings, it suggests that information about the college experience of someone close to the applicant plays a relevant role in their college related choices. Further research is required to learn what individuals learn from the higher education experience of siblings and other close peers.

6 Conclusion

The education and earning trajectories of individuals from the same social group are highly correlated. However, it is challenging to identify whether the influence of family and social networks in important life decisions could explain part of these correlations. This paper presents causal evidence that shocks to the educational trajectories of older siblings impact relevant human capital investment decisions of their younger family members. We use rich administrative data from four countries to identify siblings and link them to detailed data on college applications and enrollment decisions. Our empirical strategy exploits admission cutoffs that generate quasi-random variation in the education trajectory of older siblings.

We show that in four very different settings—Chile, Croatia, Sweden and the United States—shocks to older siblings' higher education trajectories impact younger siblings' application and enrollment decisions in meaningful ways. Having an older sibling crossing the admission threshold of a four-year college makes younger siblings more likely to attend a four-year college as well. Older siblings also influence the institution and program that their younger siblings attend. An older sibling's admission to a college increases the younger sibling's enrollment in the same college. Similarly, an older sibling's admission to a specific college-major combination makes their younger sibling's more likely to enroll in that same program. Using information on the older sibling's counterfactual option, we find that this phenomenon occurs even when the older sibling's target and counterfactual options differ significantly in expected earnings, peer quality and retention rates. However, younger siblings do not always follow their older siblings; the effects that we document disappear when the older sibling has a negative experience in college and drops out. This suggests that individuals learn from their older siblings about the institutions they enroll in and about the experience they could have there.

The four countries that we study vary in size, economic development, and education institutions. The GDP per capita of Sweden and the United States is twice as large as that of Chile and Croatia. The share of adults with post secondary degrees varies significantly across these countries, and while colleges in Chile and the United States charge high tuition fees, in Croatia and Sweden they are free. Despite these differences, we consistently find that older siblings' higher education trajectories influence the application and enrollment decisions of their younger siblings. Finding consistent results across these four different settings strongly suggests that the effects that we document are not context-specific or driven by institutional details.

These results are important because they show that relatives and potentially other close peers causally influence the consequential decisions of whether to go to college, where to study and what to specialize in. The available evidence suggests that all of these margins are relevant for future earnings and life outcomes, therefore, gaining a better understanding of what drives these decisions is critical.

These findings also shed new light on how policymakers should assess both the drivers of inequality and policies to mitigate it. Our results confirm that there is a *causal* component to the correlations that we observe on the educational choices of individuals from the same social group. Especially in contexts where some groups are more likely to face barriers and negative shocks in their path to higher education, these social spillovers could amplify inequality in educational trajectories. On the other hand, our findings suggest that the effects of policies designed to mitigate this inequality could have multiplier effects through social networks. Programs that improve individuals' educational trajectories—such as financial aid, information interventions or affirmative action—will likely have larger effects than those typically estimated because they indirectly benefit younger siblings and potentially other close peers of the direct beneficiaries.

References

Aguirre, J. and J. J. Matta (2020). Walking in Your Footsteps: Sibling Spillovers in Higher Education Choices. unpublished manuscript.

Altonji, J. G., P. Arcidiacono, and A. Maurel (2016). The Analysis of Field Choice in College and Graduate school: Determinants and Wage Effects. In *Handbook of the Economics of Education*, Volume 5, pp. 305–396. Elsevier.

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

Ang, D. (2020, 09). The Effects of Police Violence on Inner-City Students*. *The Quarterly Journal of Economics*. qjaa027.

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

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

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

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

Card, D. (1999). The causal effect of education on earnings. In D. Card and O. Ashenfelter (Eds.), *Handbook of labor economics*, Volume 3, pp. 1801–1863. Elsevier.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica 69*(5), 1127–1160.

Card, D. and A. B. Krueger (1992). Does school quality matter? returns to education and the characteristics of public schools in the united states. *Journal of political Economy 100*(1), 1–40.

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

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

Chetty, R., J. N. Friedman, E. Saez, N. Turner, and D. Yagan (2020, 02). Income Segregation and Intergenerational Mobility Across Colleges in the United States*. *The Quarterly Journal of Economics 135*(3), 1567–1633.

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

Conlon, J. J. (2019). Major malfunction: A field experiment correcting undergraduates' beliefs about salaries. *Journal of Human Resources*, 0317–8599R2.

Dahl, G., D.-O. Rooth, and A. Stenberg (2020). Family spillovers in field of study. NBER Working Papers 27618, National Bureau of Economic Research, Inc.

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

Dynarski, S., C. Libassi, K. Michelmore, and S. Owen (2018, December). Closing the Gap: The Effect of a Targeted, Tuition-Free Promise on College Choices of High-Achieving, Low-Income Students. NBER Working Papers 25349, National Bureau of Economic Research, Inc.

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

Dynarski, S. M., S. W. Hemelt, and J. M. Hyman (2015). The missing manual: Using national student clearinghouse data to track postsecondary outcomes. *Educational Evaluation and Policy Analysis 37*(1_suppl), 53S–79S.

Goldin, C. and L. Katz (2008). The race between technology and education. *Cambridge, MA: Harvard*.

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.

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

Gurantz, O., M. Hurwitz, and J. Smith (2020). Sibling Effects on High School Exam Taking and Performance. Technical Report 20-265, Annenberg Institute at Brown University.

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

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

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

Hastings, J. S. and J. M. Weinstein (2008). Information, school choice, and academic achievement: Evidence from two experiments. *The Quarterly journal of economics 123*(4), 1373–1414.

Heckman, J. J., J. E. Humphries, and G. Veramendi (2018). Returns to education: The causal effects of education on earnings, health, and smoking. *Journal of Political Economy 126*(S1), S197–S246.

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

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

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

Hoxby, C. M. and S. Turner (2015). What High-Achieving Low-Income Students Know About College. *American Economic Review 105*(5), 514–517.

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

Imbens, G. W. and J. D. Angrist (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica 62*(2), 467–475.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics 125*(2), 515–548.

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

Karbownik, K. and U. Özek (2019). Setting a Good Example? Examining Sibling Spillovers in Educational Achievement Using a Regression Discontinuity Design. Working Paper 26411, National Bureau of Economic Research.

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

Kolesár, M. and C. Rothe (2018, August). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review 108*(8), 2277–2304.

Landersø, R., H. S. Nielsen, and M. Simonsen (2017). How Going to School Affects the Family. Technical Report 2017-01, Department of Economics and Business Economics, Aarhus University.

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

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

Neilson, C. A. (2019). The rise of centralized mechanisms in education markets around the world. Technical report, Technical Report August.

Nicoletti, C. and B. Rabe (2019). Sibling spillover effects in school achievement. *Journal of Applied Econometrics 34*(4), 482–501.

Pallais, A. (2015). Small Differences that Matter: Mistakes in Applying to College. *Journal of Labor Economics 33*(2), 493–520.

Patnaik, A., M. J. Wiswall, and B. Zafar (2020). College Majors. Working Paper 27645, National Bureau of Economic Research.

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

Qureshi, J. A. (2018b). Siblings, Teachers, and Spillovers on Academic Achievement. *Journal of Human Resources 53*(1), 272–297.

Scott-Clayton, J. (2012). Do high-stakes placement exams predict college success? ccrc working paper no. 41. *Community College Research Center, Columbia University*.

Smith, J. and K. Stange (2016, Fall). A New Measure of College Quality to Study the Effects of College Sector and Peers on Degree Attainment. *Education Finance and Policy 11*(4), 369–403.

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

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

Zimmerman, S. D. (2014). The returns to college admission for academically marginal students. *Journal of Labor Economics 32*(4), 711–754.

Table I: Institutional Characteristics

	Chile (1)	Croatia (2)	Sweden (3)	US (4)
	A. Countries Characteristics			
Population	17,969,353	4,203,604	9,799,186	320,742,673
GDP per Capita	$22,688	$23,008	$48,436	$56,803
GINI Index	47.7	31.1	29.2	41.5
Human Development Index	0.84	0.827	0.929	0.917
Adults with Postsecondary Ed.	15.17%	18.30%	34.56%	39.95%
	B. University System Characteristics			
Colleges	33/60	49/49	36/36	21/3004
College-Major Combinations	1,423	564	2,421	
Tuition Fees	Yes	Yes	No	Yes
Funding for Tuition Fees	Student loans and scholarships.	Fee waiver[1].	NA	Student loans and scholarships.
Application Level	College-Major	College-Major	College-Major	College

Notes: The statistics presented in Panel A come from the World Bank (`https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD`) and from the United Nations (`http://hdr.undp.org/en/data`) websites. All statistics reported correspond to 2015 data, with the exception of the share of adults who completed a postsecondary education, which we observe in 2011. The share of adults who completed a postsecondary education is computed using the educational attainment level of individuals who were at least 25 years old in 2011.. The row "Colleges" shows the ratio of colleges that use a centralized admissions system (or which we identified to use admission cutoff rules) to the total number of colleges. In the United States the total number of colleges includes 2-year colleges. "College-Major combinations" refers to the total number of alternatives available for students through centralized admission systems in 2015.

[1] While Croatian universities charge tuition fees, first-time applicants who accept their offer receive a fee waiver. The applicant loses the fee waiver if they reject the offer.

Table II: Summary Statistics

	Chile		Croatia		Sweden		US	
	Younger Siblings (1)	All Potential Applicants (2)	Younger Siblings (3)	All Potential Applicants (4)	Younger Siblings (5)	All Potential Applicants (6)	Younger Siblings (7)	All Potential Applicants (8)
A. Demographic characteristics								
Female	0.522	0.525	0.563	0.567	0.586	0.595	0.530	0.533
Age when applying	19.028	20.059	18.880	19.158	20.486	20.823		
Household size[1]	4.632	4.322	2.790	1.925	3.104	2.950	2.250	1.288
Race: White							0.570	0.543
B. Socioeconomic characteristics								
High income[2]	0.373	0.113			0.350	0.339	0.19	0.15
Mid income[2]	0.387	0.286			0.259	0.289	0.27	0.21
Low income[2]	0.240	0.478			0.391	0.371	0.16	0.23
Parental ed: 4-year college[3]	0.434	0.207			0.571	0.519	0.650	0.595
C. Academic characteristics								
High school track: academic[4]	0.905	0.582	0.439	0.416				
Takes admission test	0.995	0.864	0.865	0.835	0.667	0.624	0.850	0.963
High school GPA score	-0.147	-0.757	268.373	265.298	0.713	0.437		
Admission test avg. score	-0.322	-0.534	312.800	286.247	0.288	-0.049	987.19	1026.095
Applicants	140,043	3,889,550	16,721	199,475	237,663	877,610	44,191	14,432,122

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

[1] In Croatia and in the United States *Household Size* refers only to the number of children in the household.

[2] In Chile, the High Income category includes households with monthly incomes greater or equal than CLP 850K (USD 2,171 of 2015 PPP); the Mid Income category includes households with monthly incomes between CLP 270K - 850K; and the Low Income category includes households with monthly incomes below CLP 270K (USD 689.90 of 2015 PPP). In Sweden, the High Income category includes households in the top quintile of the income distribution; the Mid Income category includes households in quintiles 3 and 4; and the Low Income category households in quintiles 1 and 2. The average monthly disposable income in the Swedish households is USD 5,664 (2015 PPP) in the siblings sample and USD 5,265 (2015 PPP) among all applicants. In the US, low income refers to students from families earning less than $50,000 USD per year. Middle income refers to families with $50,000-$100,000 and high income refers to families with incomes above $100,000. In the US, incomes are self-reported by the students and are missing for many students.

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

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

Table III: Sibling Spillovers on Applications to and Enrollment in Older Sibling's Target Choice

	Older Sibling's Target College			**Older Sibling's Target College-Major**			**Older Sibling's Target 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)	**Applies in the 1st preference** (7)	**Applies in any Preference** (8)	**Enrolls** (9)
Panel A - Chile									
Older sibling enrolls	0.067***	0.076***	0.038***	0.012***	0.023***	0.006***	0.012	0.017*	-0.001
	(0.012)	(0.014)	(0.011)	(0.003)	(0.005)	(0.002)	(0.007)	(0.010)	(0.006)
Older sibling above cutoff	0.033***	0.037***	0.018***	0.006**	0.012***	0.003***	0.005	0.010*	-0.000
	(0.006)	(0.007)	(0.005)	(0.001)	(0.003)	(0.001)	(0.003)	(0.005)	(0.003)
Observations	86521	86521	86521	170886	170886	170886	106085	106085	106085
Counterfactual mean	0.222	0.447	0.132	0.019	0.064	0.012	0.079	0.179	0.054
Bandwidth	12.500	12.500	12.500	18.000	18.000	18.000	16.000	16.000	16.000
Kleibergen-Paap Wald F-statistic	5576.25	5576.25	5576.25	14765.19	14765.19	14765.19	4833.50	4833.50	4833.50
Panel B - Croatia									
Older sibling enrolls	0.075***	0.109***	0.084***	0.015***	0.036***	0.013**	0.008	0.010	0.004
	(0.019)	(0.019)	(0.018)	(0.004)	(0.009)	(0.004)	(0.007)	(0.012)	(0.006)
Older sibling above cutoff	0.063***	0.091***	0.070***	0.012***	0.030***	0.011**	0.007	0.008	0.003
	(0.016)	(0.016)	(0.015)	(0.004)	(0.007)	(0.003)	(0.005)	(0.009)	(0.005)
Observations	12950	12950	12950	36757	36757	36757	31698	31698	31698
Counterfactual mean	0.293	0.523	0.253	0.022	0.111	0.017	0.059	0.218	0.054
Bandwidth	80.000	80.000	80.000	80.000	80.000	80.000	80.000	80.000	80.000
Kleibergen-Paap Wald F-statistic	6459.56	6459.56	6459.56	14512.30	14512.30	14512.30	10158.25	10158.25	10158.25
Panel C - Sweden									
Older sibling enrolls	0.122***	0.132***	0.049***	0.020***	0.031***	0.005***	0.000	-0.002	-0.001
	(0.008)	(0.011)	(0.005)	(0.002)	(0.005)	(0.001)	(0.006)	(0.009)	(0.004)
Older sibling above cutoff	0.033***	0.035***	0.013***	0.006***	0.009***	0.001***	0.000	-0.001	0.000
	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)	(0.000)	(0.002)	(0.002)	(0.001)
Observations	378466	378466	378466	482220	482220	482220	355885	355885	355885
Counterfactual mean	0.087	0.206	0.032	0.011	0.053	0.003	0.049	0.101	0.016
Bandwidth	0.360	0.360	0.360	0.386	0.386	0.386	0.389	0.389	0.389
Kleibergen-Paap Wald F-statistic	7215.227	7215.227	7215.227	10406.511	10406.511	10406.511	6643.373	6643.373	6643.373

Notes: "Applies in the 1st preference" looks at the probability that the younger sibling ranks the target choice of the older sibling in her/his first preference; "Applies in any preference" looks at the probability of ranking the older sibling's target choice in any preference; and "Enrolls" looks at the probability of enrolling in the target choice of the older sibling. The first row of each panel presents 2SLS estimates in which older siblings' enrollment is instrumented with them being above an admission cutoff. The second row presents reduced form estimates. All the 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, admission cutoffs and younger siblings' birth year are included. Among the three outcomes in each sample, we use the smallest Calonico et al. (2014) optimal bandwidth. Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table IV: Sibling Spillovers on College Choice and College Quality in the US

	College type		College quality		Price, location		Older Sibling's Target College	
	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)	Younger sibling applies (7)	Younger sibling enrolls (8)
Panel A - All Students								
Older sibling enrolls	0.230*	-0.002	0.180**	0.316**	2.263	0.135	0.279**	0.172***
	(0.132)	(0.114)	(0.080)	(0.148)	(2.321)	(0.127)	(0.103)	(0.054)
Counterfactual mean	0.38	0.20	0.30	-0.21	8.74	0.21	0.10	0.01
Panel B - Uncertain college-goers								
Older sibling enrolls	0.531**	0.055	0.473***	0.699***	11.245***	0.429**	0.271	0.257***
	(0.248)	(0.214)	(0.150)	(0.260)	(4.168)	(0.221)	(0.179)	(0.099)
Counterfactual mean	0.09	0.09	0.01	-0.67	-0.28	-0.04	0.13	-0.08
Panel C - Probable college-goers								
Older sibling enrolls	0.049	-0.033	0.011	0.091	-2.893	-0.066	0.291**	0.131*
	(0.159)	(0.139)	(0.098)	(0.186)	(2.968)	(0.162)	(0.131)	(0.068)
Counterfactual mean	0.57	0.25	0.49	0.08	14.07	0.41	0.07	0.06

Notes: Each coefficient is a 2SLS estimate of the impact of an older sibling's enrollment in her/his target college on younger siblings' college choices, using admissibility as an instrument. Each estimate comes from a local linear regression with a bandwidth of 93 SAT points, a donut 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 panel includes all students, while the Panels B and C 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 V: Sibling Spillovers on College and College-Major Choice by Differences between Older Siblings' Target and Next Best Options

	Chile			Croatia	Sweden		
	Effect of older siblings' enrollment on younger siblings' applications by differences in:						
	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 Applies to Older Sibling's Target College							
Older sibling enrolls (ΔX in 1st quartile)	0.096***	0.146***	0.083***	0.064*	0.098***	0.109***	0.115***
	(0.028)	(0.030)	(0.028)	(0.033)	(0.025)	(0.021)	(0.023)
Older sibling enrolls (ΔX in 2nd quartile)	0.117***	0.102***	0.102***	0.146***	0.129***	0.097***	0.103***
	(0.027)	(0.026)	(0.027)	(0.031)	(0.021)	(0.020)	(0.020)
Older sibling enrolls (ΔX in 3rd quartile)	0.091***	0.096***	0.105***	0.122***	0.117***	0.091***	0.089***
	(0.027)	(0.026)	(0.025)	(0.031)	(0.022)	(0.022)	(0.021)
Older sibling enrolls (ΔX in 4th quartile)	0.090***	0.082***	0.112***	0.105***	0.168***	0.123***	0.106***
	(0.029)	(0.026)	(0.027)	(0.032)	(0.028)	(0.025)	(0.024)
Observations	32987	32987	32987	9610	147190	167290	159146
Kleibergen-Paap F-statistic	722.509	744.566	740.276	1089.054	613.193	676.879	673.860
Counterfactual mean	0.443	0.443	0.443	0.502	0.221	0.220	0.222
Panel B - Younger Sibling Applies to Older Sibling's Target College-Major							
Older sibling enrolls (ΔX in 1st quartile)	0.020**	0.018**	0.030***	0.037***	0.024**	0.027***	0.023**
	(0.009)	(0.009)	(0.009)	(0.011)	(0.011)	(0.010)	(0.010)
Older sibling enrolls (ΔX in 2nd quartile)	0.022***	0.017**	0.011	0.030***	0.040***	0.023**	0.016*
	(0.008)	(0.008)	(0.008)	(0.012)	(0.010)	(0.009)	(0.009)
Older sibling enrolls (ΔX in 3rd quartile)	0.012	0.018**	0.018**	0.040***	0.035***	0.017*	0.027***
	(0.008)	(0.008)	(0.008)	(0.012)	(0.010)	(0.009)	(0.009)
Older sibling enrolls (ΔX in 4th quartile)	0.018**	0.024***	0.020**	0.039***	0.026**	0.028***	0.026**
	(0.009)	(0.009)	(0.009)	(0.013)	(0.012)	(0.010)	(0.010)
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	0.072	0.072	0.072	0.112	0.067	0.063	0.062

Notes: We investigate how the probability of applying to an older sibling's target alternative changes with enrollment and with quality differences between the older sibling's target and next-best options. Quality is measured in terms of expected earnings, peer quality and first year retention rates. Differences in these variables between older siblings' target and next best options are classified in four quartiles. The effect of an older sibling's enrollment is allowed to be different in each quartile. The reported specifications use the same set of controls and bandwidths as the 2SLS specifications described in Table III. In addition, we include next-best options fixed effects. Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table VI: Sibling Spillovers on Younger Siblings' Application by Older Siblings' Target Option Characteristics

	Chile			Croatia	Sweden		
	Effect of older siblings' enrollment on younger siblings' applications by target option's:						
	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 applies to older sibling's target college							
Older sibling enrolls (X in 1st quartile)	0.095***	0.087**	0.052*	0.058*	0.156***	0.118***	0.103***
	(0.030)	(0.042)	(0.028)	(0.033)	(0.032)	(0.031)	(0.026)
Older sibling enrolls (X in 2nd quartile)	0.058***	0.082***	0.038	0.156***	0.109***	0.044*	0.115***
	(0.029)	(0.029)	(0.027)	(0.033)	(0.025)	(0.025)	(0.022)
Older sibling enrolls (X in 3rd quartile)	0.111***	0.061***	0.102***	0.104***	0.101***	0.123***	0.095***
	(0.028)	(0.025)	(0.026)	(0.035)	(0.022)	(0.022)	(0.021)
Older sibling enrolls (X in 4th quartile)	0.056***	0.086***	0.097***	0.119***	0.116***	0.115***	0.110***
	(0.021)	(0.021)	(0.022)	(0.038)	(0.021)	(0.019)	(0.021)
Observations	39960	39960	39960	9610	169619	178814	175951
Kleibergen-Paap F-statistic	824.637	626.324	926.147	1098.798	588.205	651.385	723.051
Counterfactual mean	0.444	0.444	0.444	0.502	0.219	0.220	0.220
Panel B - Younger sibling applies to older sibling's target college-major							
Older sibling enrolls (X in 1st quartile)	0.015*	0.023**	0.023***	0.046***	0.022*	0.020	0.028**
	(0.009)	(0.011)	(0.008)	(0.013)	(0.013)	(0.013)	(0.011)
Older sibling enrolls (X in 2nd quartile)	0.011	0.022**	0.017**	0.016	0.024**	0.012	0.026***
	(0.008)	(0.009)	(0.008)	(0.012)	(0.011)	(0.011)	(0.009)
Older sibling enrolls (X in 3rd quartile)	0.028***	0.014*	0.024***	0.036**	0.033***	0.026***	0.021**
	(0.008)	(0.008)	(0.008)	(0.014)	(0.010)	(0.010)	(0.009)
Older sibling enrolls (X in 4th quartile)	0.020***	0.023***	0.017**	0.047***	0.030***	0.029***	0.021**
	(0.007)	(0.007)	(0.008)	(0.014)	(0.010)	(0.009)	(0.010)
Observations	97321	97321	97321	32228	247960	264527	256565
Kleibergen-Paap F-statistic	2501.594	1819.772	2883.727	3046.997	1002.833	1090.406	1340.660
Counterfactual mean	0.073	0.073	0.073	0.112	0.065	0.063	0.063

Notes: We investigate how the probability of applying to older sibling's target choice changes depending on older siblings' enrollment and on the quality of her/his target option. Quality is measured in terms of expected earnings, peer quality and first year retention rates. Older siblings' target options are classified in four quartiles in each of these dimensions. The effect of an older sibling's enrollment is allowed to differ along these quartiles. The reported specifications use the same set of controls and bandwidths as the 2SLS specifications described in Table III. In addition, we include next-best option fixed effects. Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

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

	Younger sibling applies to any 4-year college	**Younger sibling applies to older sibling's target college**				**Younger sibling applies to older sibling's target college-major**		
	US	CHI	CRO	SWE	US	CHI	CRO	SWE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Interaction with $\mathbb{1}$(Age Difference between Siblings ≥ 5)								
Older sibling enrolls	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)
Older sibling enrolls $\times$ Age diff. ≥ 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	378446	44190	170570	36756	482220
Kleibergen-Paap Wald F-statistic	64.892	2767.580	3230.667	3562.527	64.892	7330.470	7225.706	5147.083
Panel B: Interaction with $\mathbb{1}$(Siblings are of the Same Gender)								
Older sibling enrolls = 1	0.310**	0.070***	0.114***	0.129***	0.304***	0.017***	0.026**	0.032***
	(0.137)	(0.016)	(0.022)	(0.012)	(0.106)	(0.005)	(0.009)	(0.006)
Older sibling enrolls $\times$ Same gender = 1	-0.152**	0.011	-0.007	0.007	-0.052	0.011***	0.023*	0.008
	(0.071)	(0.012)	(0.020)	(0.010)	(0.056)	(0.004)	(0.009)	(0.005)
Observations	44190	86521	12950	378446	44190	170886	36757	482220
Kleibergen-Paap Wald F-statistic	65.114	2788.470	3229.534	3607.870	65.114	7383.02	7220.184	5204.123

Notes: The reported specifications use the same set of controls and bandwidths as the 2SLS specifications described in Tables IV and III. In addition, they include an interaction between the treatment and a dummy variable that indicates if siblings are 5 or more years apart (Panel A) or between the treatment and a dummy variable that indicates if siblings are of the same gender (Panel B). In both cases the variable defining the interaction is also included as control. Younger siblings are counted as applying to the same alternative as their older siblings if they include that alternative at any rank in their application. Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table VIII: Sibling Spillovers on College and College-Major Choice by Older Sibling's Dropout

	Any 4-year College	**Older Sibling's Target College**			**Older Sibling's Target College-Major**	
	US	CHI	SWE	US	CHI	SWE
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A - Younger Sibling Applies to Target Alternative						
Older sibling enrolls	0.454***	0.089***	0.180***	0.448***	0.020***	0.050***
	(0.123)	(0.018)	(0.015)	(0.101)	(0.006)	(0.007)
Older sibling enrolls $\times\mathbb{1}$(Drops out)	-0.505***	-0.129***	-0.108***	-0.049	-0.036***	-0.042***
	(0.087)	(0.020)	(0.014)	(0.071)	(0.006)	(0.006)
Observations	37330	49183	359300	37330	99104	457505
Kleibergen-Paap Wald F-statistic	48.875	1950.41	2450.027	48.875	5444.24	3521.871
Panel B - Younger Sibling Enrolls in Target Alternative						
Older sibling enrolls	0.508***	0.058***	0.064***	0.183***	0.008***	0.008***
	(0.130)	(0.014)	(0.007)	(0.052)	(0.003)	(0.002)
Older sibling enrolls $\times\mathbb{1}$(Drops out)	-0.618***	-0.085***	-0.032***	0.007	-0.014***	-0.006***
	(0.091)	(0.014)	(0.006)	(0.037)	(0.003)	(0.002)
Observations	37330	49183	359300	37330	99104	457505
Kleibergen-Paap Wald F-statistic	48.875	1950.41	2450.027	48.875	5444.24	3521.871

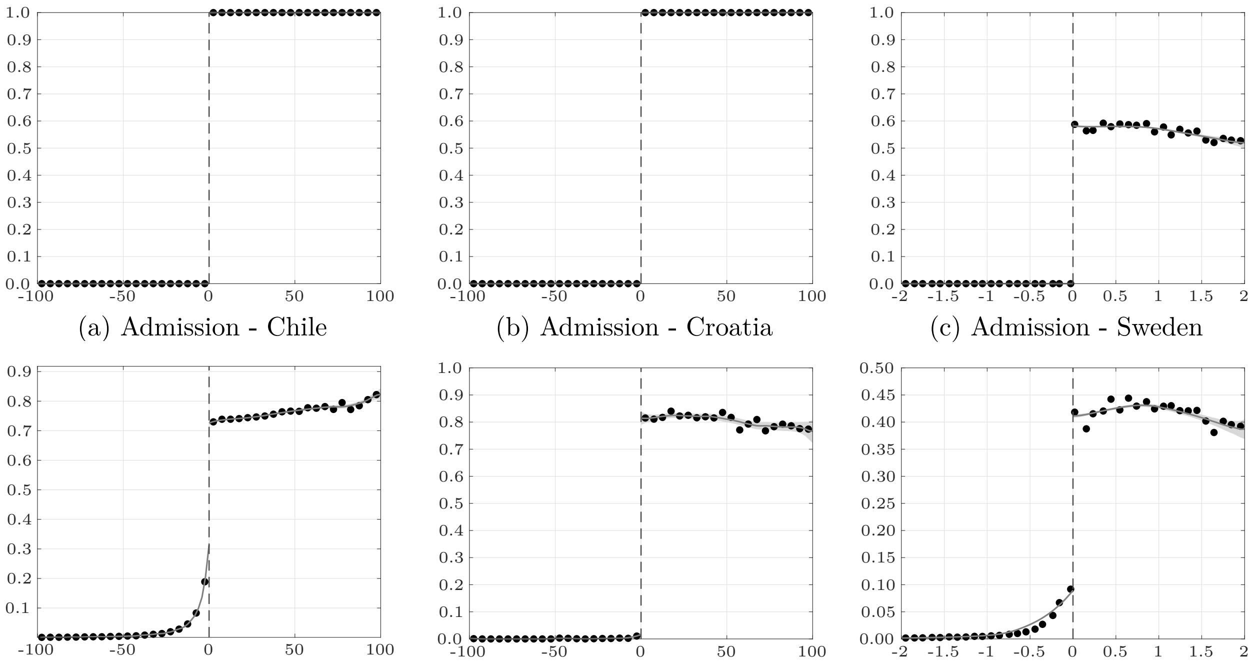
Notes: The reported specifications use the same set of controls and bandwidths as the 2SLS specifications described in Tables IV and III. In addition, they include an interaction between the treatment and a dummy variable that takes value 1 if the older sibling drops out after the first year. This dummy variable is also included as a control. We exclude siblings that are less than 2 years apart in age. Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table IX: Sibling Spillovers on Academic Performance

	High school GPA (1)	Takes an admission exam (2)	Average score on admissions exam (3)	Applies to college (4)
		Panel A - Chile		
Older sibling enrolls	-0.009	0.001	-0.011	-0.002
	(0.019)	(0.001)	(0.011)	(0.005)
Observations	170,886	170,886	170,886	170,886
Counterfactual mean	-0.170	0.995	-0.240	0.930
Kleibergen-Paap F-statistic	14765.190	14765.190	14765.190	14765.190
		Panel B - Croatia		
Older sibling enrolls	-0.043	-0.013	-0.054	-0.008
	(0.045)	(0.017)	(0.043)	(0.009)
Observations	12,443	12,443	10,233	36,757
Counterfactual mean	-0.030	0.810	-0.035	0.866
Kleibergen-Paap F-statistic	4498.481	4498.481	3728.910	14512.30
		Panel C - Sweden		
Older sibling enrolls	0.011	-0.031***	0.068**	-0.009
	(0.022)	(0.011)	(0.030)	(0.010)
Observations	421268	482220	227976	482220
Counterfactual mean	0.218	0.494	0.040	0.654
F-statistic	9714.124	10406.511	6660.104	10406.511
		Panel D - United States		
Older sibling enrolls		0.073	46.9	0.159
		(0.096)	(43.0)	(0.125)
Observations		44,190	37,554	44,190
Counterfactual mean		0.830	951.000	0.545
Kleibergen-Paap F-statistic		129.730	120.758	129.730

Notes: The table presents 2SLS estimates for the effect of older siblings' enrollment in their preferred college-major (Chile, Croatia and Sweden) or college (United States) on younger siblings' high school GPA (column 1), probability of taking the admission exam (column 2), average performance on the admission exam (column 3) and applying to college (column 4). For the US, (2) looks at the number of applications submitted. The reported specifications use the same set of controls and bandwidths as the 2SLS specifications described in Table IV and Table III. Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

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

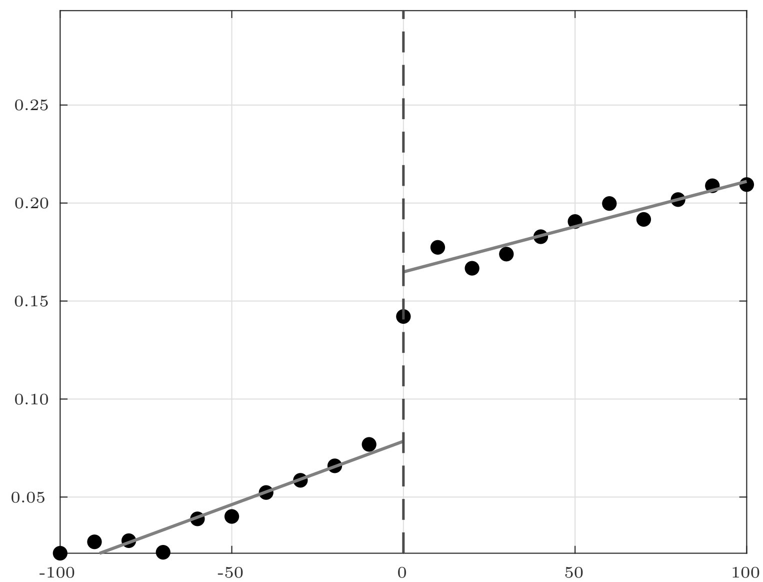


(a) Admission - Chile (b) Admission - Croatia (c) Admission - Sweden

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

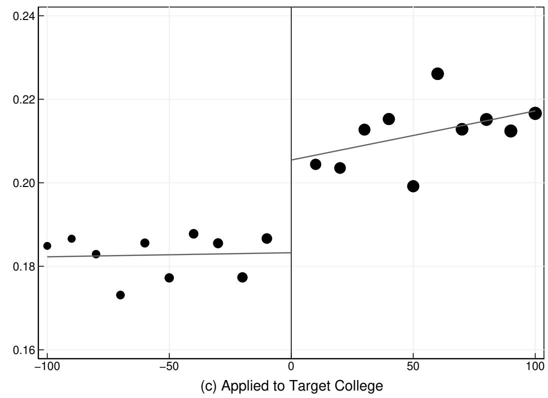
This figure illustrates older siblings' admission and enrollment probabilities around the admission cutoffs of their target majors in Chile, Croatia and Sweden. Figures (a) and (d) illustrate these probabilities for Chile, figures (b) and (e) for Croatia and figures (c) and (f) for Sweden. Gray lines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

Figure II: Older Siblings' Enrollment Probability in in the Target College at the Admission Cutoff (First Stage)

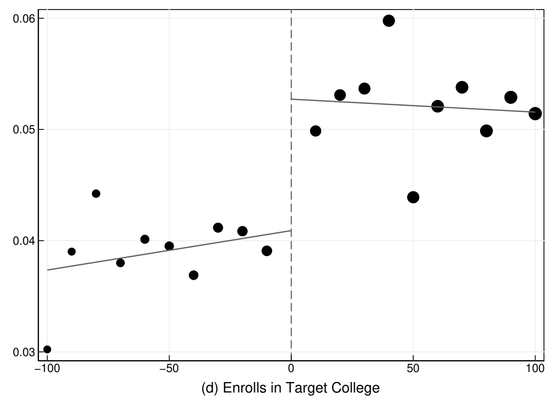


This figure illustrates older siblings' enrollment probability in their target college around the admission cutoffs in the United States. Gray lines represent local linear polynomials. Black dots represent sample means of the dependent variable at different values of older siblings' SAT score.

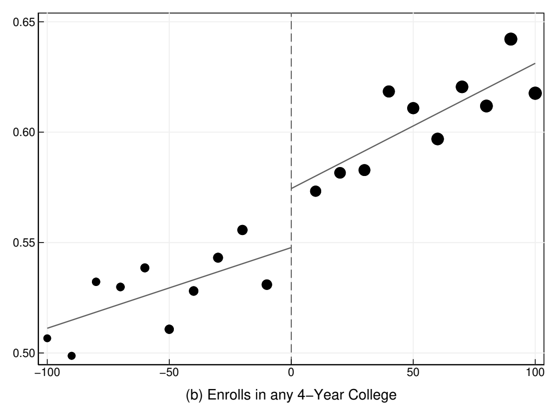
Figure III: Probabilities of Enrolling in any 4-year College and in the Older Sibling's Target College



(a) Applies to Target College



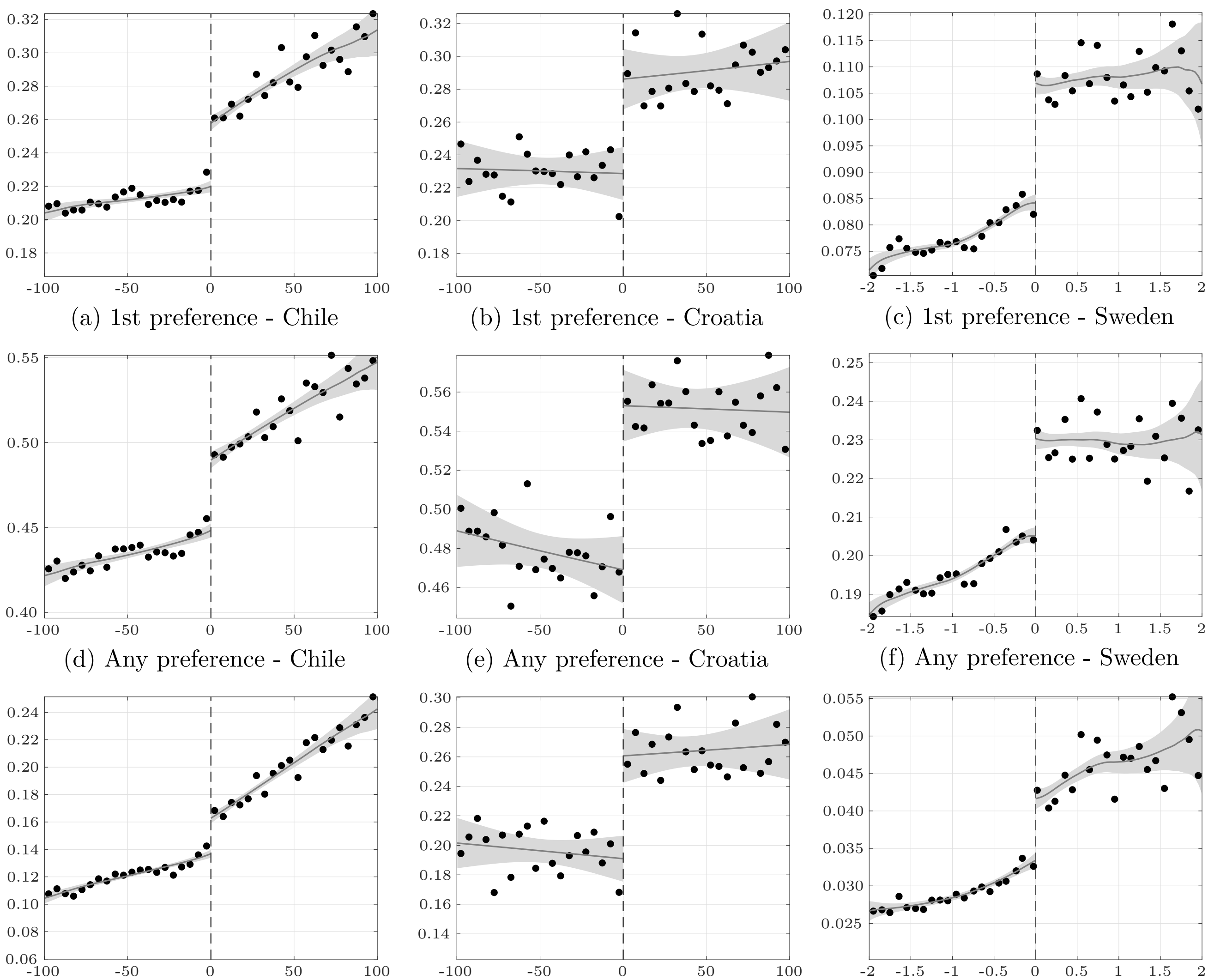
(b) Enrolls in Target College



(c) Enrolls in any 4-year College

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

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



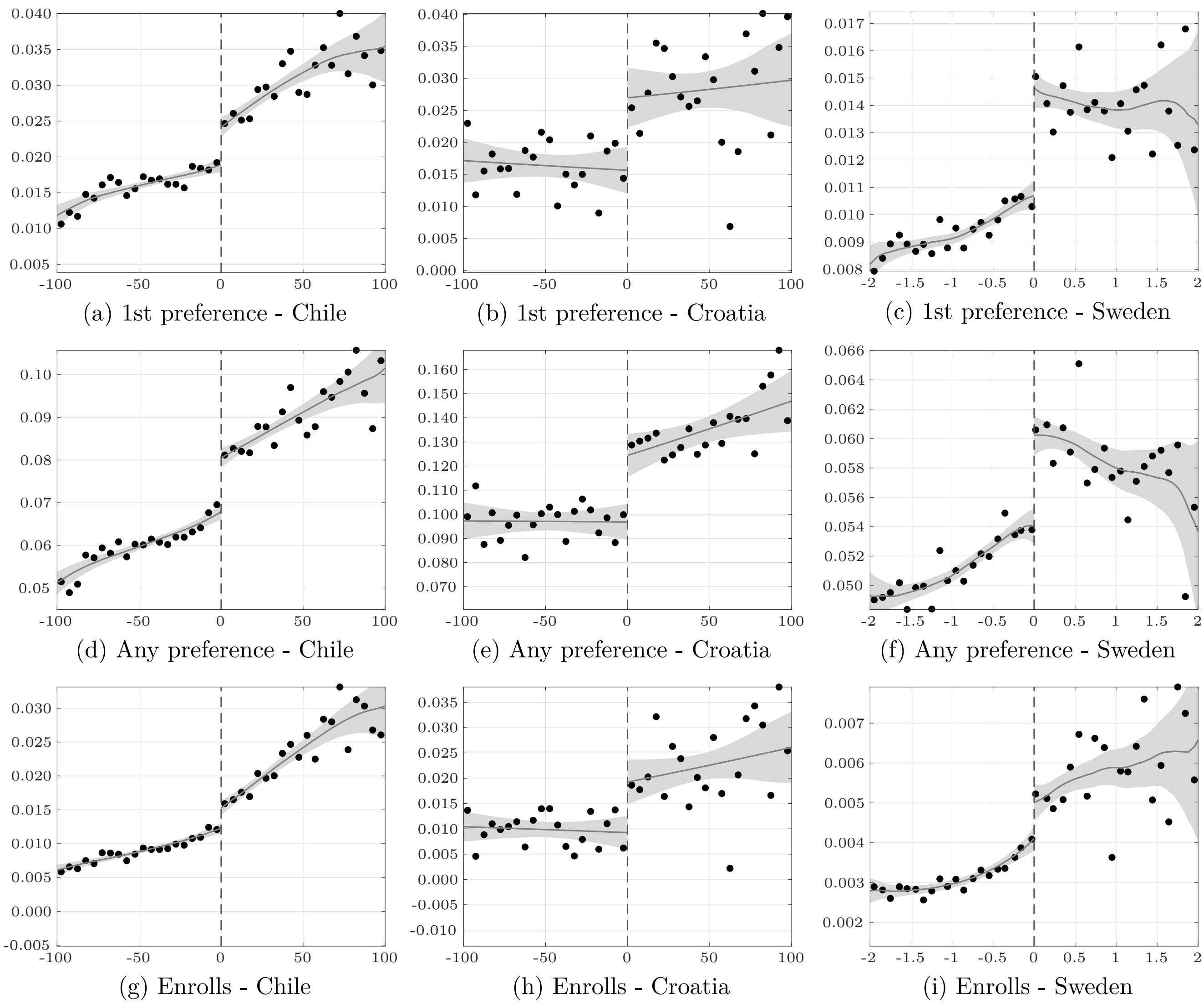
(a) 1st preference - Chile (b) 1st preference - Croatia (c) 1st preference - Sweden

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

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

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

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



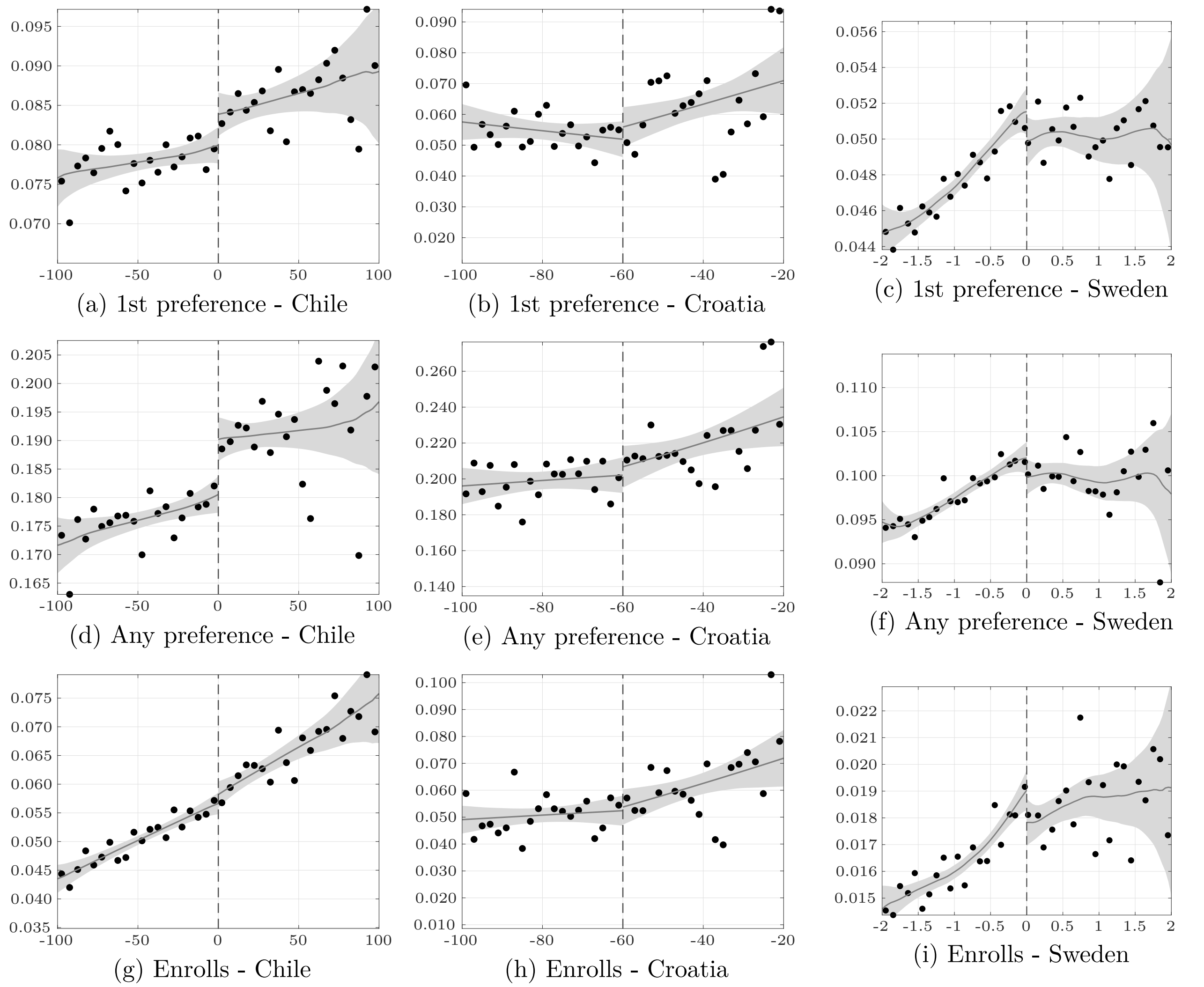
(a) 1st preference - Chile (b) 1st preference - Croatia (c) 1st preference - Sweden

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

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

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

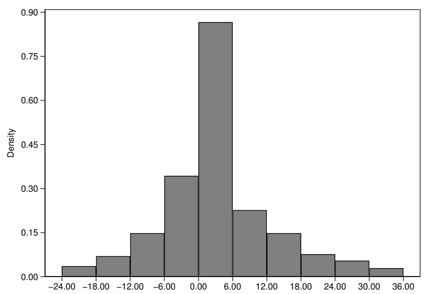
Figure VI: Probabilities of Applying and Enrolling in Older Sibling's Target Major



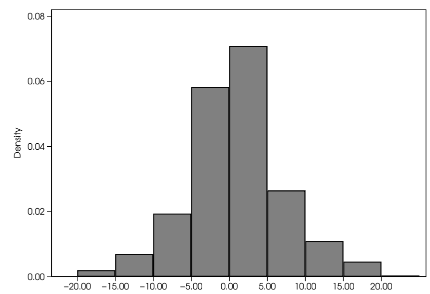
(a) 1st preference - Chile
(b) 1st preference - Croatia
(c) 1st preference - Sweden
(d) Any preference - Chile
(e) Any preference - Croatia
(f) Any preference - Sweden
(g) Enrolls - Chile
(h) Enrolls - Croatia
(i) Enrolls - Sweden

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

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

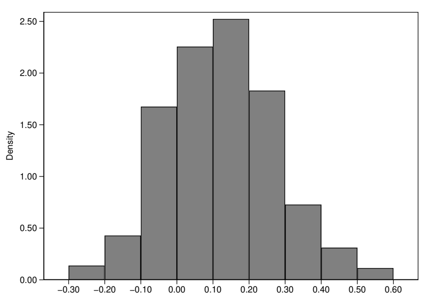


(a) Chile

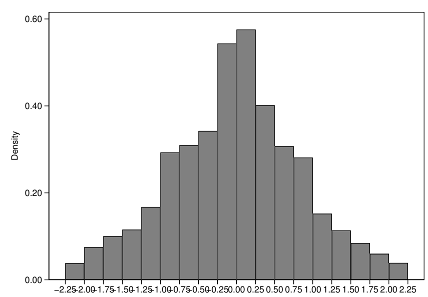


(b) Sweden

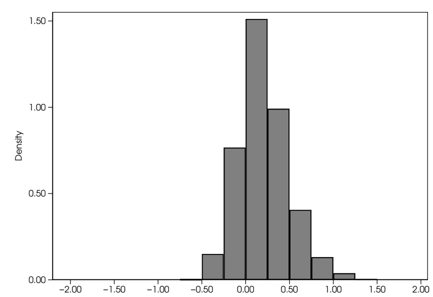
Panel A - Δ in Annual Earnings (USD 000)



(c) Chile

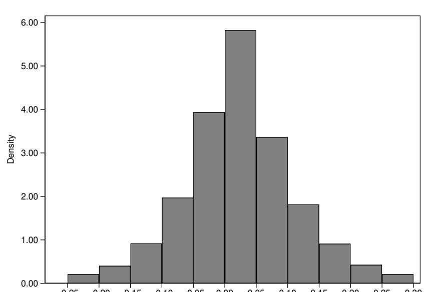


(d) Croatia

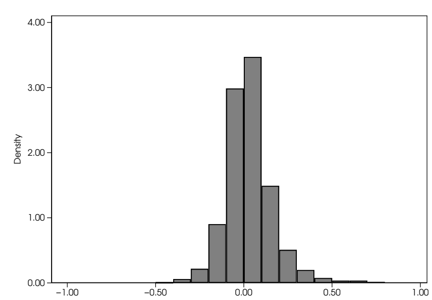


(e) Sweden

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).

A Identification Strategy: Further Discussion

A.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.

A.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 $(m, c) \succ (m', c')$.[49] 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/his applications. This implies that the same individual may belong to more than one college-major marginal group.

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

A.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.[50]

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

[50] In Appendix B 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.

attended).

This means that we remove from this sample all older siblings whose target and next best option correspond to the same major.[51]

A.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 3.3, a fuzzy RD can be thought as an IV. In what follows, and for ease of notation, we drop time and individual indices t, i, τ and focus our analysis on a specific major-college u. Following this notation, the treatment in which we are interested is:

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

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

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

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

[51] In Section 4 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.

j when his older sibling enrolls in k, and let's introduce the following notational simplification:

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

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

1. Independence of the instrument:

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

2. Exclusion restriction:

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

3. First stage:

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

4. Monotonicity:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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 3 becomes:

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

Taken together, 4 and 5 imply:

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

□

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

B Robustness Checks

This section investigates if the identification assumptions of our empirical strategy are satisfied. We start by checking if there is any evidence of manipulation of the running variables. Next, we check if other variables that could affect individuals' application and enrollment decisions present jumps at the cutoff and if the results are robust to different bandwidths. We continue by performing two types of placebo exercises. In the first, we study if similar effects arise when looking at placebo cutoffs (i.e. cutoffs that do not affect older siblings' admission). In the second, we analyze if similar effects arise when looking at the effect of the younger sibling enrollment on older siblings decisions. We then investigate if our conclusions change when using a second order polynomial of the running variable, when using a triangular kernel and when allowing the slope of the running variable to vary by major-college and year. Finally, we end this section by showing that there are no extensive margin responses of younger siblings (i.e. increases in total enrollment) in Chile, Croatia and Sweden that could explain our findings.

B.1 Manipulation of the Running Variable

A first condition for the validity of our RD estimates is that individuals should not be able to manipulate their older siblings' application scores around the admission cutoff. The structures of the admission systems in Chile, Croatia and Sweden make the violation of this assumption unlikely. Something similar occurs in the United States, where the cutoffs that we exploit are hidden. To confirm this we study whether the distribution of the running variable (i.e. older sibling's application score centered around the relevant cutoff) is continuous at the cutoff. As discussed in Section 2, in Sweden the admission exam is voluntary and institutions select their students using either their high school GPA or their scores in the admission exam. Both of these measures are not fully continuous and in addition, the admission exam suffered some transformations in 2013. Therefore, to investigate manipulation of these scores we present independent histograms for each one of these variables. Figure B.I illustrates the density of the relevant running variables for all the countries that we study. These histograms do not show any evidence of manipulation.

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

B.2 Discontinuities in Potential Confounders

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

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

Figure B.II summarizes the result of this analysis for Chile, Croatia and Sweden. It plots the estimated discontinuities at the cutoff and their 95% confidence intervals. To estimate these discontinuities at the cutoff we use the same specification described in the main body of the paper. This means that we control for a linear polynomial of the running variable and allow the slope to change at the cutoff. Using the same bandwidths reported for linear specifications in Section 4, 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 B.III presents similar results to the United States. Here instead of presenting the estimated jump at the cutoff we illustrate how the variable on the y-axis evolves with the running variable. None of the potential confounders studied in this figure seem to jump at the cutoff.

B.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 B.IV and B.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 B.VI does the same exercise for the outcomes that we investigate in the United States. In this case, the coefficients remain also very stable when using smaller bandwidth; when we increase it, the coefficients begin to drop what suggests a non-linear relationship between the running variable and the outcomes outside the 100 SAT points window used in our analyses.

B.4 Placebo Exercises

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

First, in Figures B.VIII and B.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 case of the United States we do not perfectly observe the actual cutoffs; instead, we estimate them from the data. Figure B.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 B.X and B.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 B.XI does something similar for the United States. Since younger siblings apply to college after their older siblings, being marginally admitted or rejected from a major or college should not affect what happens with their older siblings. These figures show that this is indeed the case. Even though when looking at the placebo on college choice in Sweden we find small discontinuities at the cutoff, their size is considerably smaller than the ones we document in the main body of the paper.

B.5 Alternative Specifications and Total Enrollment

We conclude this section by presenting the results to alternative specifications.

Tables B.I and B.II summarize the results for the US. The first table presents results of additional 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 Kolesár and Rothe (2018) to compute standard errors (columns 4 and 5). The second table, present 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 loss 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 4-year college and on the choice of college.

We present similar analyses for Chile, Croatia and Sweden distributed along multiple tables. First, Tables B.IV and B.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 B.VI and B.V we show that our results are robust to allowing the running variable to have cutoff-major specific slopes, and in Table B.VII we show that our main results are robust to control by covariates. Table B.VIII presents results from specifications that 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 B.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 thing that changes at the margin is the college to which older siblings are allocated to. 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 Section 4.

B.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 to. The results that we find—summarized in Table B.X— are very similar to the ones we document for college choice in the current section.In the case of Croatia, the country for which we have less observations, these estimates become less precise, but still they are similar in magnitude to the ones we present in Section 4.

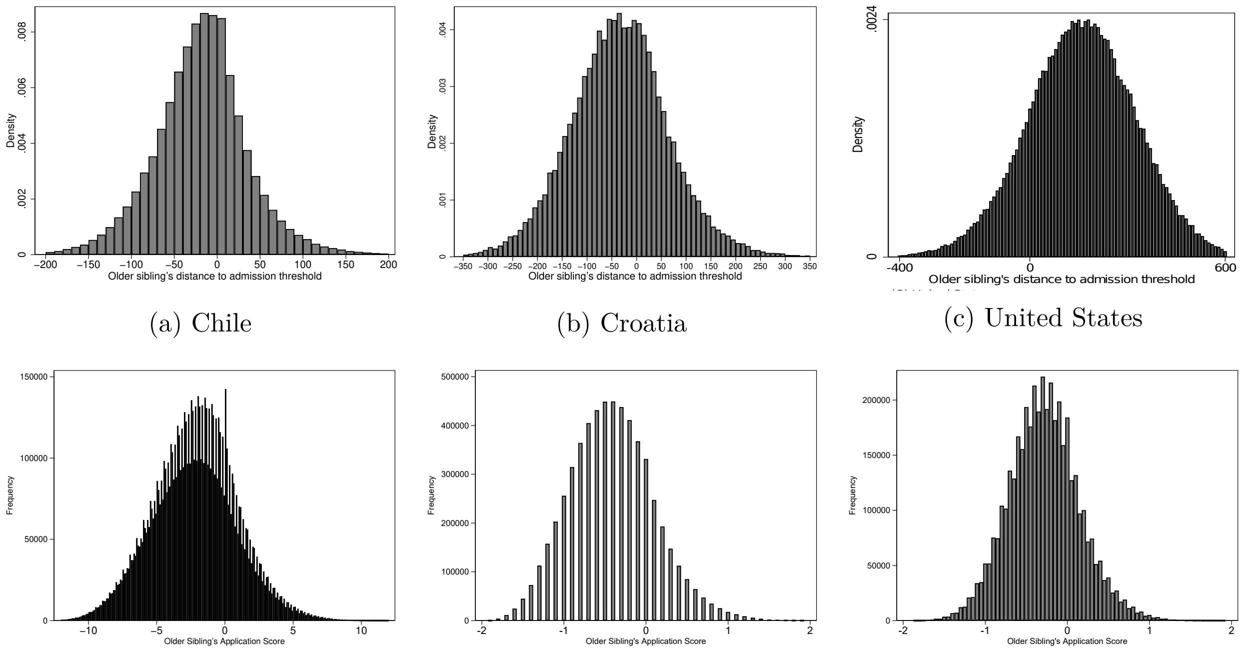
In order to investigate if sibling spillovers in the choice of major are only local—i.e. only affect preferences for the major in the same college of the older sibling— we build a new sample in which we only include individuals whose older sibling's target and next best option are offered by the same college (i.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.[52]

Table B.XII summarizes the results of this exercise. We find that when eligible for the older sibling's

[52] 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 5.2, 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.

college-major choice, individuals' 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.

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

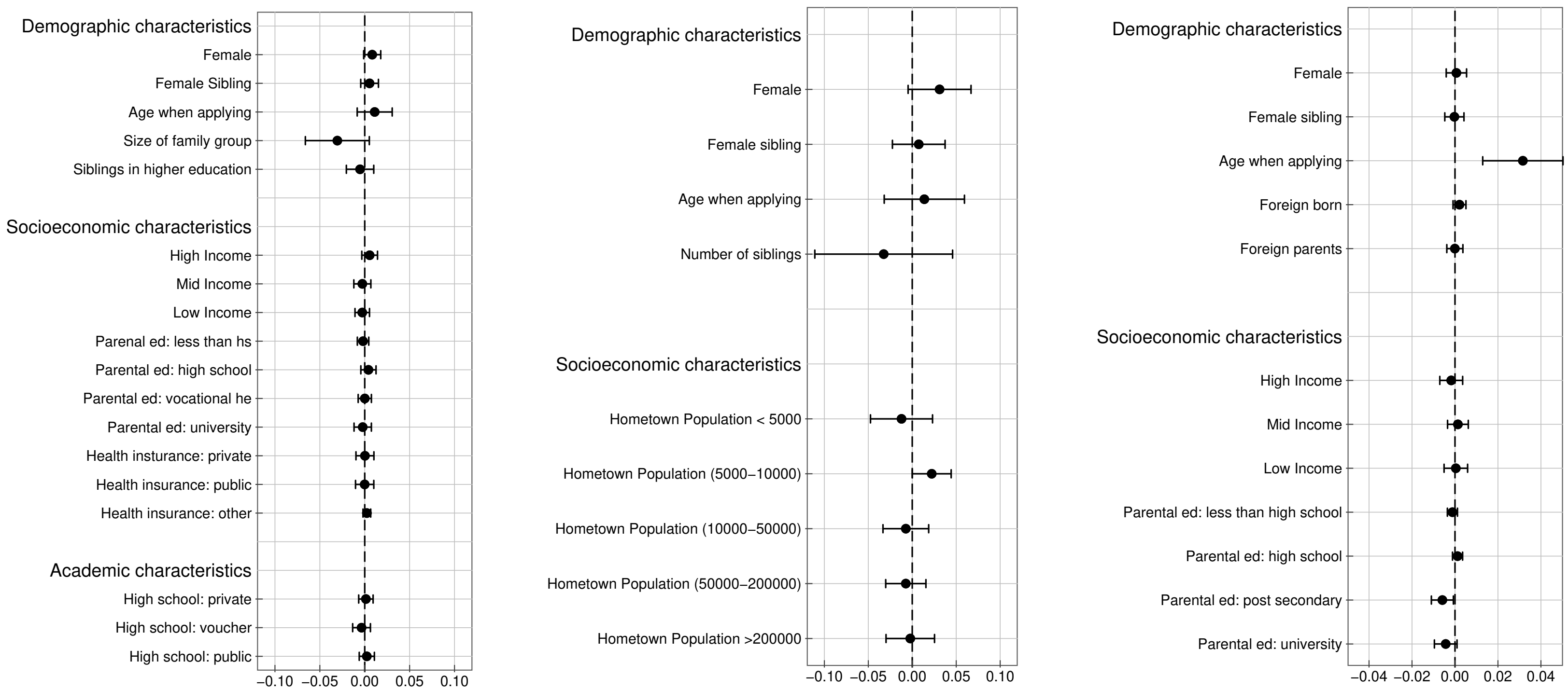


(a) Chile (b) Croatia (c) United States

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

These histograms illustrate distributions of older siblings' admission exam and high school GPA around admission cutoffs for Chile, Croatia, Sweden and the United States. Panels (a), (b) and (c) illustrate the distribution of admission exam scores in Chile, Croatia and the United States respectively. Panel (d) illustrates the distribution of high school GPA in Sweden and panel (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 B.II: Disconitnuities in other Covariates at the Cutoff



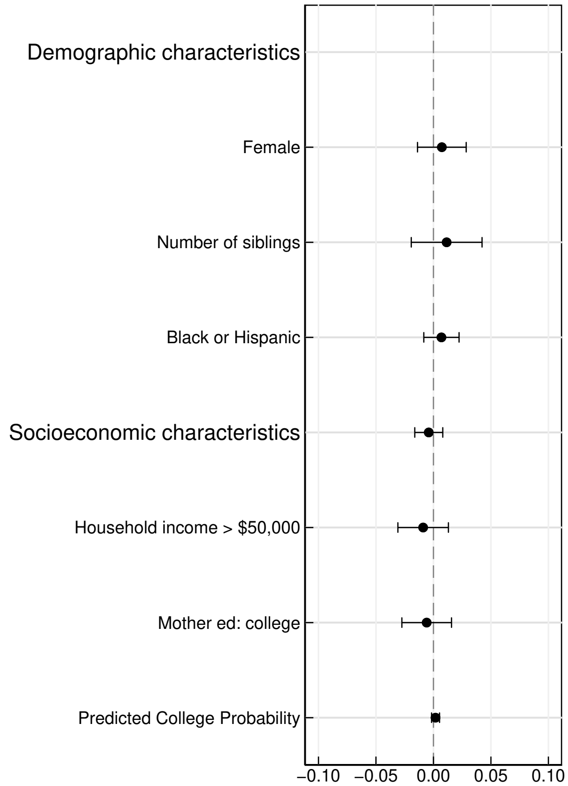
(a) Chile

(b) Croatia

(c) Sweden

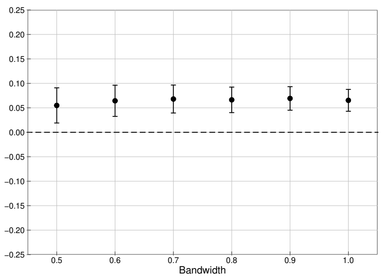
This figure illustrates the estimated jumps at the cutoff for a vector of socioeconomic and demographic characteristics. These estimates come from parametric specifications that control for a linear polynomial of the running variable. As the main specifications, these also include major-college-year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

Figure B.III: Disconitnuities in other Covariates at the Cutoff (United States)

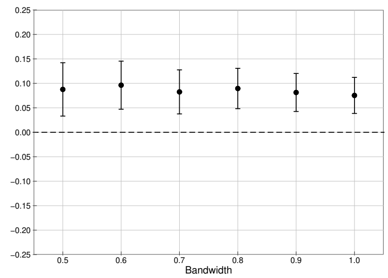


This figure illustrates how demographic and socioeconomic characteristics vary at the admissions cutoff in the United States. The range of the running variable corresponds to the bandwidth used in our main specifications. The points represent the estimated coefficient , while the lines represent 95%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

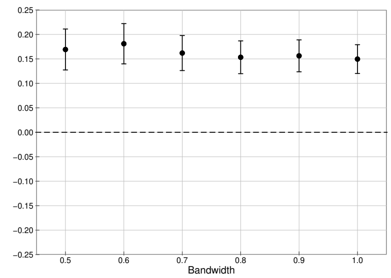
Figure B.IV: Probabilities of Applying and Enrolling in Older Sibling's Target College - Different Bandwidths



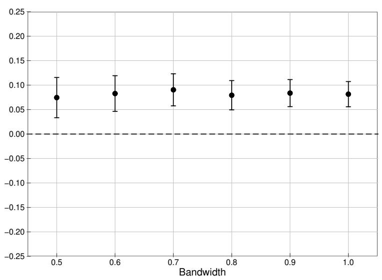
(a) 1st preference - Chile



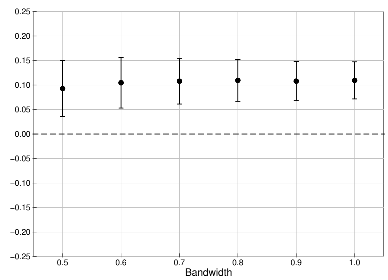
(b) 1st preference - Croatia



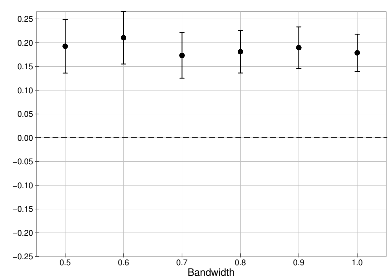
(c) 1st preference - Sweden



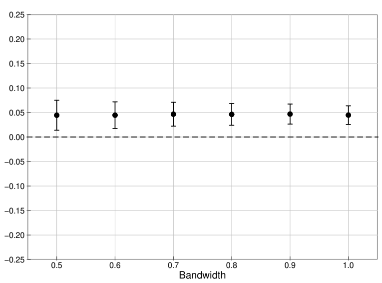
(d) Any preference - Chile



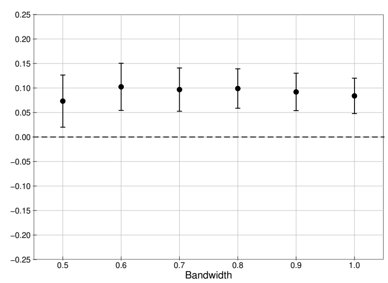
(e) Any preference - Croatia



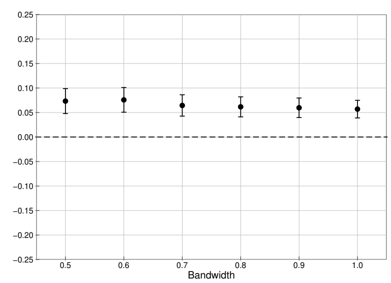
(f) Any preference - Sweden



(g) Enrolls - Chile



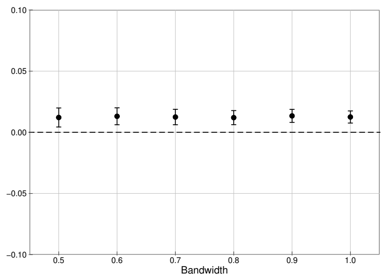
(h) Enrolls - Croatia



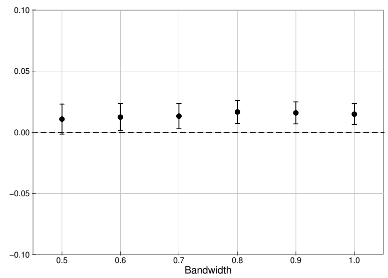
(i) Enrolls - Sweden

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

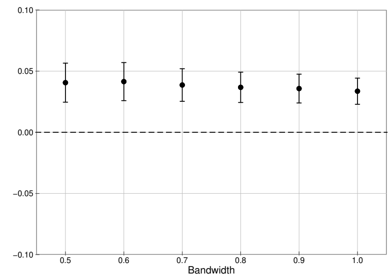
Figure B.V: Probabilities of Applying and Enrolling in Older Sibling's Target Major-College - Different Bandwidths



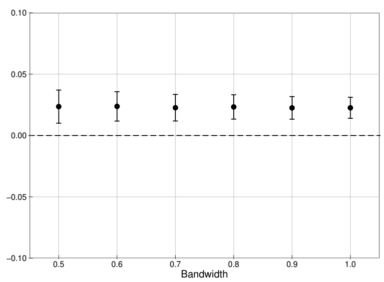
(a) 1st preference - Chile



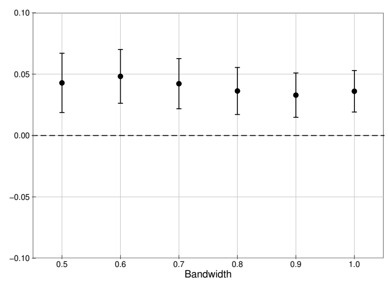
(b) 1st preference - Croatia



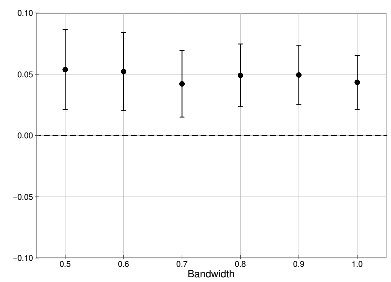
(c) 1st preference - Sweden



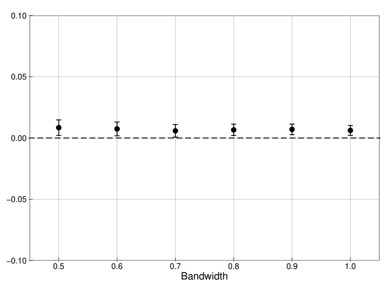
(d) Any preference - Chile



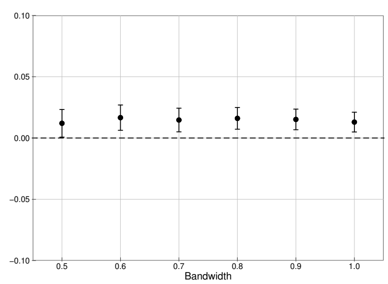
(e) Any preference - Croatia



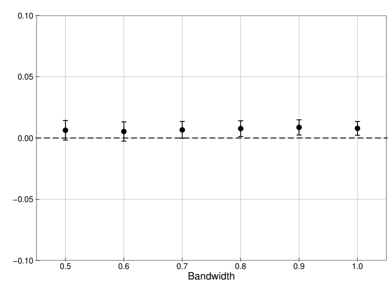
(f) Any preference - Sweden



(g) Enrolls - Chile



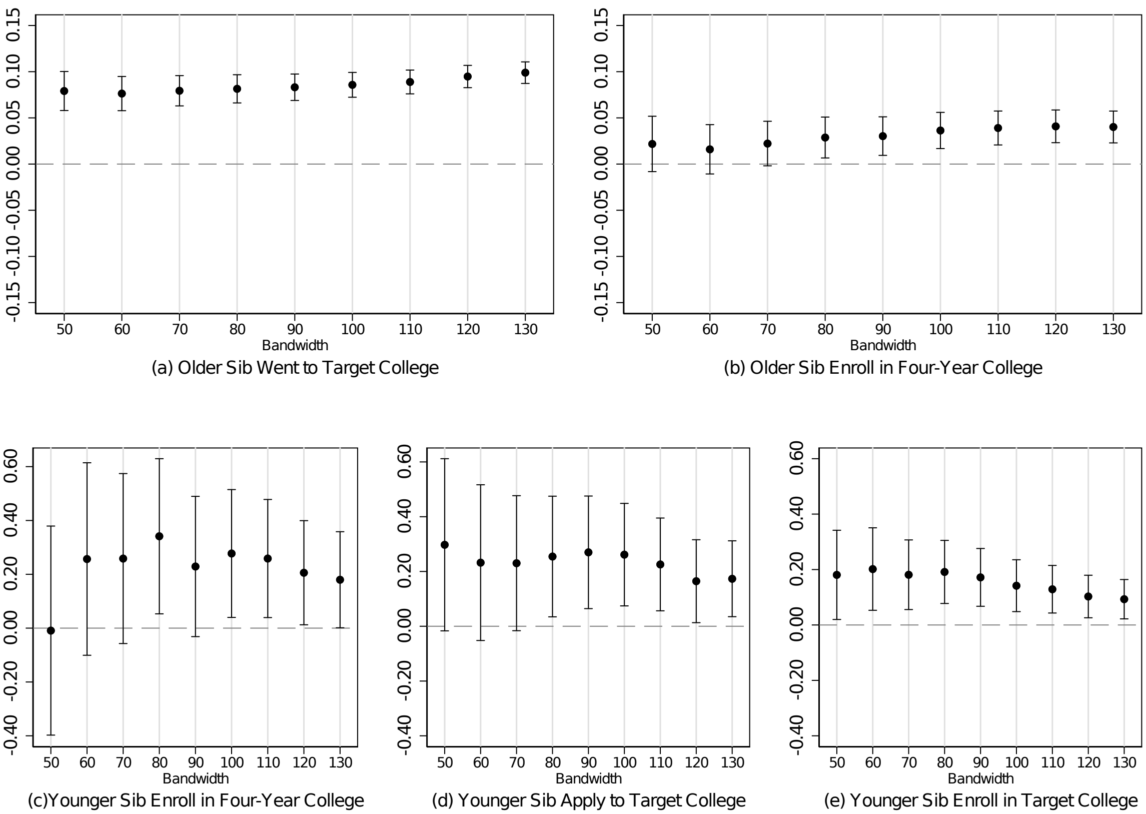
(h) Enrolls - Croatia



(i) Enrolls - Sweden

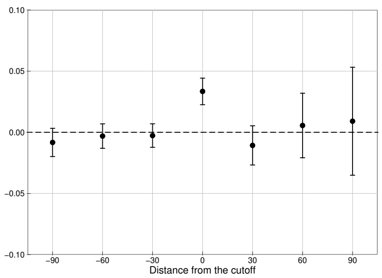
This figure illustrates how being admitted to a specific program changes younger siblings' probabilities of applying and enrolling in the same major. The x-axis corresponds to different bandwidths used to build these figures, chosen as multiples of the optimal bandwidths computed following Calonico et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

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

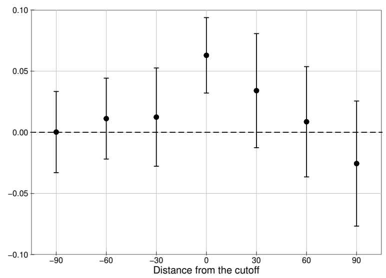


This figure illustrates how an older sibling's marginal enrollment in her target college changes a younger sibling's probability of enrolling in any 4-year college and in the older sibling's target college. The x-axis corresponds to different bandwidths used to build these figures. The dots represent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

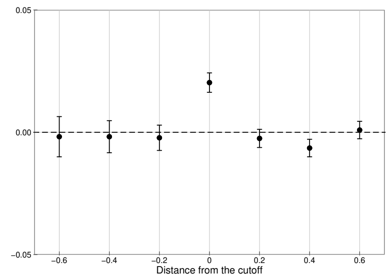
Figure B.VII: Placebo Cutoffs - Probabilities of Applying and Enrolling in Older Sibling's Target College



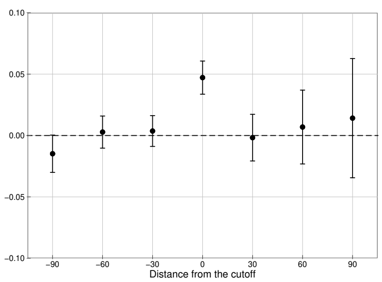
(a) 1st preference - Chile



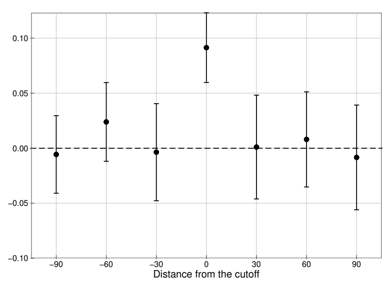
(b) 1st preference - Croatia



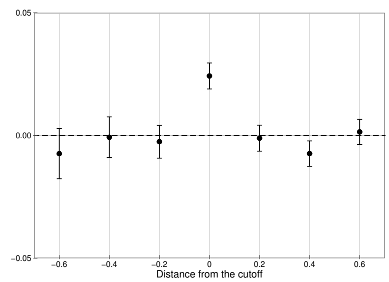
(c) 1st preference - Sweden



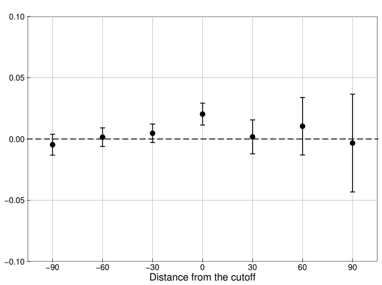
(d) Any preference - Chile



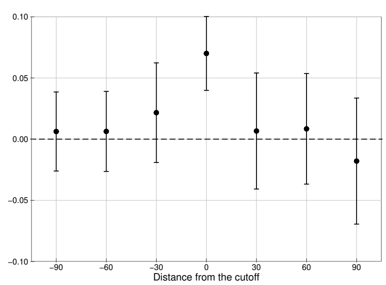
(e) Any preference - Croatia



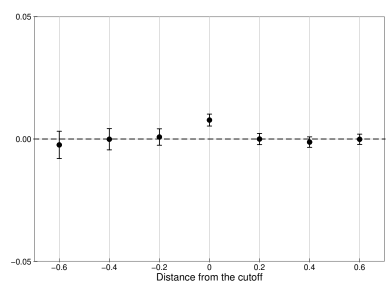
(f) Any preference - Sweden



(g) Enrolls - Chile



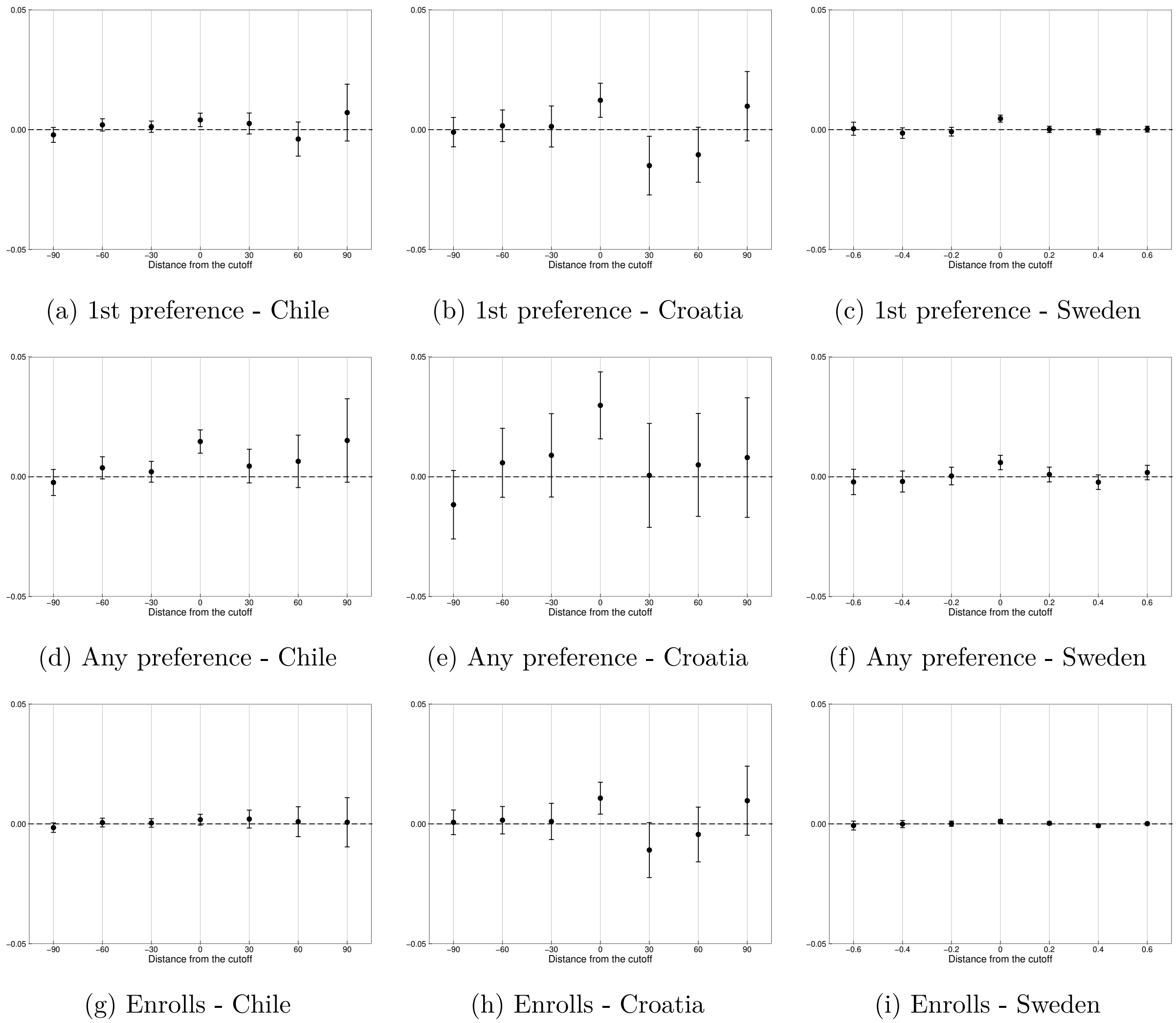
(h) Enrolls - Croatia



(i) Enrolls - Sweden

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

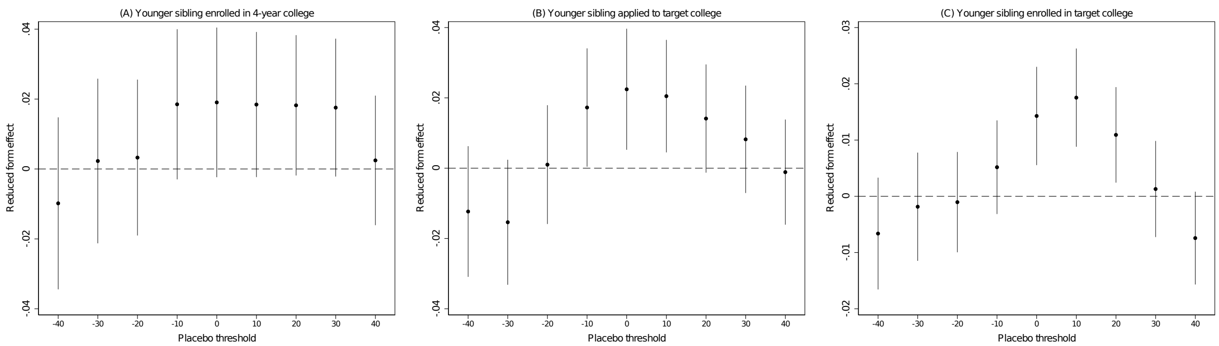
Figure B.VIII: Placebo Cutoffs - Probabilities of Applying and Enrolling in Older Sibling's Target Major-College



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

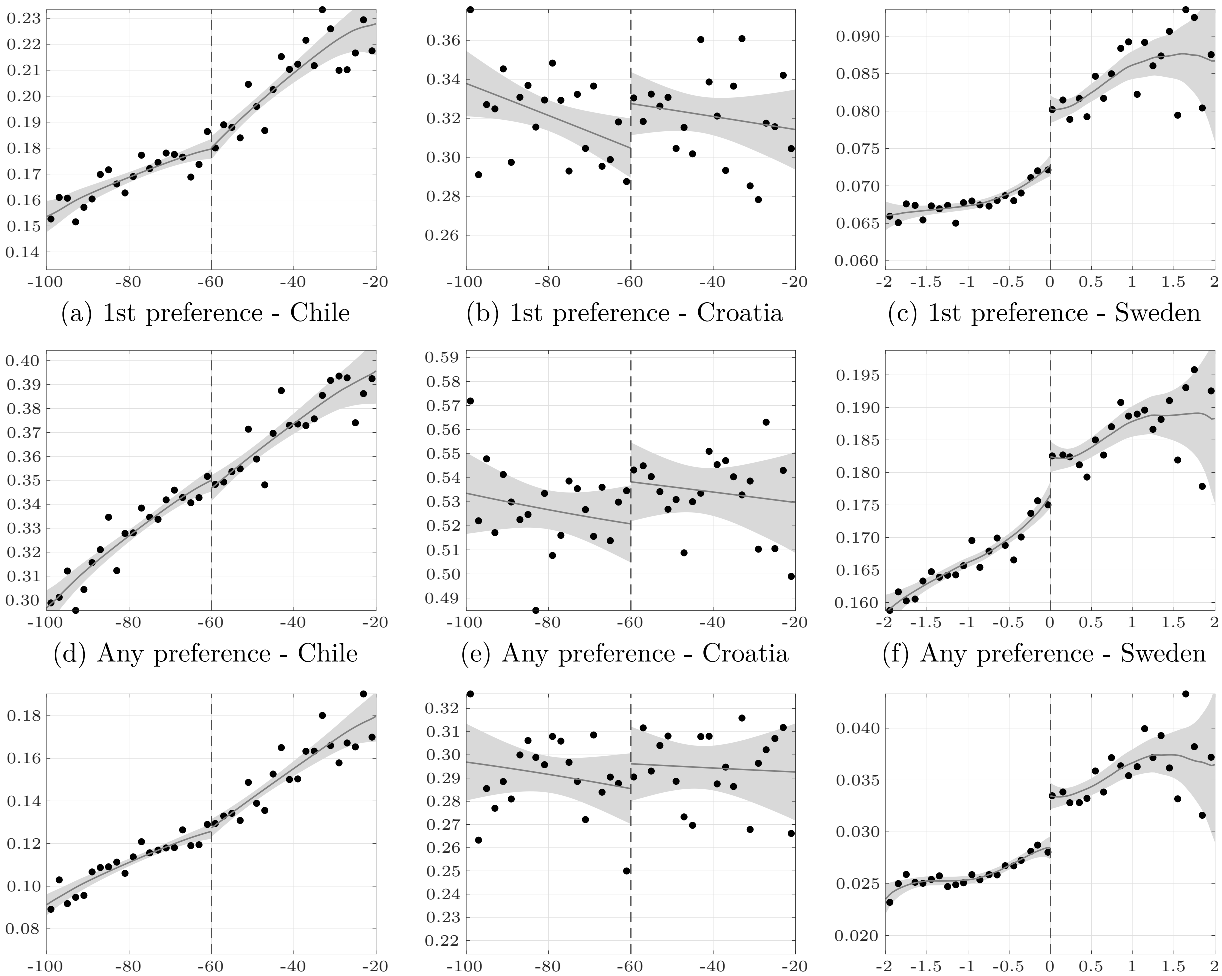
This figure illustrates the results of a placebo exercise that investigates if effects similar to the ones documented in figure V arise at different values of the running variable. Therefore, the x-axis corresponds to different (hypothetical) values of cutoffs - 0 corresponds to the actual cutoff used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

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



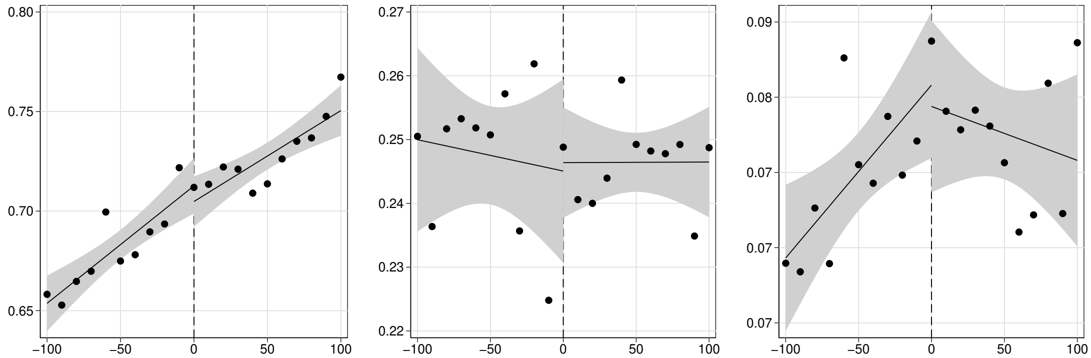
(a) 1st preference - Chile (b) 1st preference - Croatia (c) 1st preference - Sweden

(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.

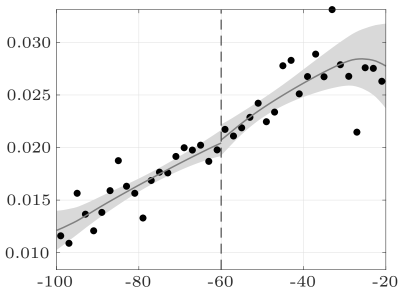
Figure B.XI: Placebo - Probabilities of Applying and Enrolling in Younger Sibling's Target College



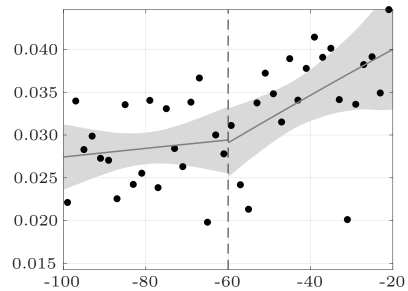
(a) Enrolls in Four−Year College (b) Applies to Target College (c) Enrolls in Target College

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.

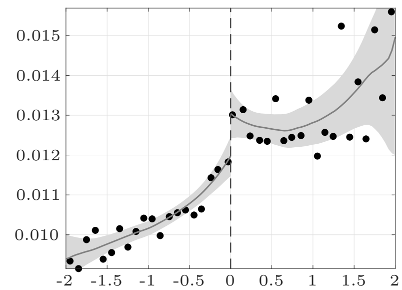
Figure B.XII: Placebo - Probabilities of Applying and Enrolling in Younger Sibling's Target Major-College



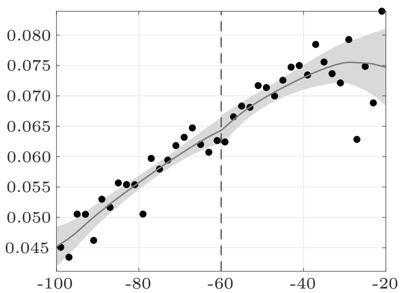
(a) 1st preference - Chile



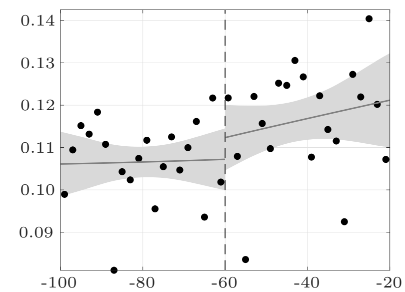
(b) 1st preference - Croatia



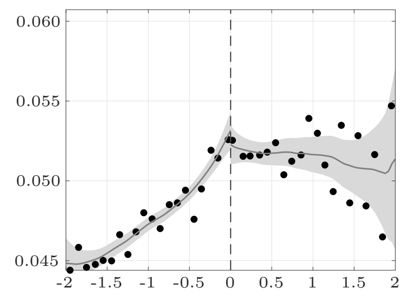
(c) 1st preference - Sweden



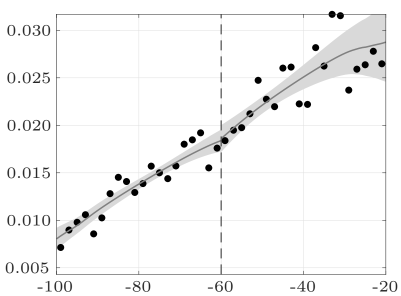
(d) Any preference - Chile



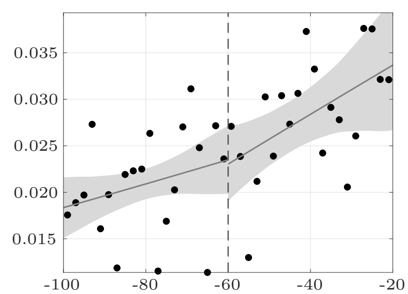
(e) Any preference - Croatia



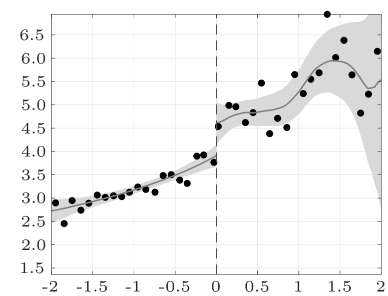
(f) Any preference - Sweden



(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 B.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 B.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 B.III: 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	P2	P1	P2	P1	P2
	(1)	(2)	(3)	(4)	(5)	(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 III. 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 B.IV: 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 III. 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 B.V: Sibling Spillovers on Applications and Enrollment in Older Sibling's Target Major-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	0.013***	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 III, 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 B.VI: 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 (1)	P2 (2)	P1 (3)	P2 (4)	P1 (5)	P2 (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 III, 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 B.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 III. 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 B.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 III. 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 B.IX: Sibling Spillovers on Applications to and Enrollment in Older Sibling's Target College and Target College-Major (Target × Next Best College-Major Fixed Effects)

	Older Sibling's Target College			**Older Sibling's Target College-Major**		
	Applies in the first preference (1)	**Applies in any preference** (2)	**Enrolls** (3)	**Applies in the first preference** (4)	**Applies in any preference** (5)	**Enrolls** (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 III, 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 B.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 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 B.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 (1)	P2 (2)	P1 (3)	P2 (4)	P1 (5)	P2 (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 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 B.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 III. Standard errors clustered at the family level are reported in parenthesis. *p-value<0.1 **p-value<0.05 ***p-value<0.01.