CERGE
Center for Economic Research and Graduate Education
Charles University

Essays in Econometrics of Matching Markets: Identification, Estimation and Practice

Marin Drlje

Dissertation

Prague, 22.12.2020.

Dissertation Committee:

Štěpán Jurajda (CERGE-EI, Chair)

Nikolas Mittag (CERGE-EI)

Jan Zapál (CERGE-EI)

Referees:

Seth Zimmerman

Thomas Le Barbanchon

Acknowledgments

The process of producing this dissertation has been a life-changing experience that would not have been as fulfilling or possible without the support of many people.

First of all, I would like to thank my supervisor, Štepan Jurajda, who provided me with never-ending guidance, persistent support, and motivational pushes during the long and curvy PhD road. Štepan facilitated the growth of this dissertation tremendously, but also provided me with practical advice and viewpoints that I will carry for years. Above all, I think we had a lot of fun. Štepan, I cannot thank you enough.

I would like to express my sincere thanks to the members of my committee, Nikolas Mittag and Jan Zapal, for many valuable comments and support throughout my studies. I am thankful to my referees Seth Zimmerman and Thomas Le Barbanchon for their in-depth review and useful comments.

I am thankful to Christopher Neilson and David S. Lee for numerous consultations and advice during my Princeton mobility.

I thank Michal Bauer, Alena Bicakova, Henry Farber, Randall K. Filer, Vasily Korovkin, Dejan Kovac, David Lee, Alexandre Mas, Andreas Menzel, Nikolas Mittag, Daniel Munich, Christopher Neilson, Christian Ochsner, Steven Rivkin, Jan Zápal, Zhuan Pei and Kresimir Zigic for their many useful comments. I am grateful to the Ministry of Education of Croatia and ASHE (AZVO) for access to their administrative data.

I would also like to thank all of the CERGE-EI community – faculty members, library and staff for their assistance.

I would especially like to thank my family. My wife, Ivana, has been extremely supportive of me throughout this entire process and has made countless sacrifices. Last but not the least, a major credit goes to my newborn son, Oliver, whose birth was the biggest motivator to speed up and finish the dissertation.

Marin Drlje

Abstrakt

Rozsáhlá literatura odhaduje různé efekty přijetí a absolvování škol s využitím variace v bodových výsledcích studentů. Při odhadech je předpokládáno quasi-náhodné přijetí studentů kolem spodní hranice přijetí. V této disertaci se zaměřuji na teoretické I praktické aspekty těchto odhadů.

V prvním článku předkládám důkazy naznačující, že vzorky odpovídající obvyklým aplikacím nespojité regrese (dále RDD z anglického regression discontinuity design) v relevantní literatuře nesplňují předpoklady náhodného přiřazení. Rozlišuji ex-post randomizaci (odpovídající loterii uchazečů na hraně přijetí) od ex-ante randomizace, odrážející nejistotu ohledně struktury všech uchazečů v centralizovaném systému, která může být přirozeně kvantifikována opakovaným výběrem z populace uchazečů. S využitím dat z chorvatského centralizovaného systému přijímacích řízení na vysoké školy ukazuji, že ex-ante pravděpodobnosti přijetí se významně liší mezi přijatými a odmítnutými studenty nacházejícími se v obvyklém vzorku používaném pro RDD analýzy. Takový nepoměr v rozdělení pravděpodobnosti přijetí naznačuje, že šířka pásma v okolí hranice přijetí, tj. velikost výběru pro analýzu kvazi-náhodných přiřazení, by měla být oproti současné praxi významně redukována s cílem vyhnout se výběrovému zkreslení. Také ukazuji, že značný podíl kvazi-náhodných přiřazení do přijetí a nepřijetí se nachází mimo typickou šířku RDD pásma, což naznačuje, že odhady nejsou vydatné. Jako alternativu k RDD metodám navrhuji novou odhadovací metodu Propensity Score Discontinuity Design (PSDD), která využívá všechna pozorování s kvazi-náhodným přiřazením a srovnává výsledky uchazečů porovnatelných co do ex-ante pravděpodobností přijetí do daného programu, tj. pravděpodobností přijetí podmíněných bodovými výsledky přijímacích řízení.

V druhém článku zaznamenáváme, že v centralizovaných systémech párujících studenty a univerzity, kde je student přiřazen k nevyhovujícímu studijnímu programu, obvykle následuje zápis do preferovaného programu v roce následujícím po prvotním přiřazení. To vytváří významné náklady plynoucí z neshody studenta a studijního programu. Ukazujeme, že s těmito náklady na neshodu dochází k porušení klíčového předpokladu LATE (local average treatment effect) theoremu a může potenciálně vést ke zkresleným RDD odhadům. Využíváme data z chorvatského systému párujícího studenty a univerzity k ilustraci empirického významu tohoto potenciální zdroje zkreslení a navrhujeme metodu inspirovanou Leem (2009), která umožňuje odhadnout interval treatment efektu za předpokladu, že náklady neshody nesouvisí s konkrétním přiřazením.

Třetí článek analyzuje vliv rodinných vazeb na volbu univerzity. Zatímco se obecně předpokládá, že rodina a sociální sítě mohou ovlivnit důležitá životní rozhodnutí, identifikace jejich kauzálních efektů je pověstně obtížná. Tento článek předkládá důkazy

kauzálního vlivu, kdy studijní směřování starších sourozenců může významně ovlivnit volbu univerzity a studijního programu u mladších sourozenců. Využíváme institucionální charakteristiky systémů přijímacích řízení z Chile, Chorvatska a Švédska, které generují quazi-náhodnou variaci ve studijním směřování starších sourozenců. S použitím RDD ukazujeme, že mladší sourozenci se významně častěji přihlásí a zapíši na stejnou školu a studijní program jako jejich starší sourozenci, kteří byli náhodně přiřazeni. Zjišťujeme, že tento sourozenecký vliv je silnější, pokud se starší sourozenci zapíší a jsou úspěšní ve studijních oborech, které jsou více výběrové, mají nižší podíl neúspěšných studentů a absolventi mají vyšší průměrné příjmy. Výsledky ze Švédska a Chile naznačují, že sourozenecký efekt je větší, pokud starší sourozenec je muž. Zkoumáme řadu možných mechanismů a srovnáváme výsledky napříč zeměmi, které mají výrazné odlišné sociální a ekonomické charakteristiky. Po shromáždění důkazů docházíme k závěru, že výsledky jsou nejvíce konzistentní s mechanismem, kdy starší sourozenec poskytuje jinak nedostupné informace o zkušenostech s univerzitou a potenciálních výnosech ze studia.

Abstract

A large literature estimates various school admission and graduation effects by employing variation in student admission scores around schools' admission cutoffs, assuming (quasi-) random school assignment close to the cutoffs. In this dissertation I focus on this variation, both from the theoretical and practical standpoints.

In the first paper, I present evidence suggesting that the samples corresponding to typical applications of regression discontinuity design (RDD) fail to satisfy these assumptions. I distinguish ex-post randomization (as in admission lotteries applicable to those at the margin of admission) from ex-ante randomization, reflecting uncertainty about the market structure of applicants, which can be naturally quantified by resampling from the applicant population. Using data from the Croatian centralized college-admission system, I show that these ex-ante admission probabilities differ dramatically between treated and non-treated students within typical RDD bandwidths. Such unbalanced admission probability distributions suggest that bandwidths (and sample sizes) should be drastically reduced to avoid selection bias. I also show that a sizeable fraction of quasi-randomized assignments occur outside of the typical RDD bandwidths, suggesting that these are also inefficient. As an alternative, I propose a new estimator, the Propensity Score Discontinuity Design (PSDD), based on all observations with random assignments, which compares the outcomes of applicants matched on ex-ante admission probabilities, conditional on admission scores.

In the second paper, we note that, in centralized student-college matching markets, non-compliance with the matching assignment typically corresponds to enrolling in one's preferred program a year after the initial assignment, introducing significant non-compliance costs. We show that with costly non-compliance, the *exclusion restriction*, the key assumption of the LATE theorem, is violated, potentially leading to biased RDD estimates. We use data from a student-college matching market in Croatia to illustrate the empirical importance of this potential source of bias and propose a method inspired by Lee (2009), which recovers the treatment effect bounds under the assumption that the costs of non-compliance are not related to the treatment assignment.

The third paper analyzes family ties behind the college choice. While it is widely believed that family and social networks can influence important life decisions, identifying their causal effects is notoriously difficult. This paper presents causal evidence from three countries indicating that the educational trajectories of older siblings can significantly influence the college and major choices of younger siblings. In this analysis, we exploit institutional features of the college admissions systems in Chile, Croatia and Sweden that generate quasi-random variation in the educational paths followed by older siblings.

Using regression discontinuity design, we show that younger siblings are significantly more likely to apply and enroll in the same college and major to which their older siblings are randomly assigned. We find that these sibling effects are stronger when older siblings enroll and are successful in majors that are more selective, have lower dropout rates and in which graduates have higher average earnings. We explore several potential mechanisms and compare results across countries that have very different social and economic contexts. Taking the evidence together we conclude the results are most consistent with older siblings transmitting otherwise unavailable information about the college experience and its potential returns.

Contents

List of Tables

List of Figures

Chapter 1

Identification of School Admission Effects Using Propensity Scores Based on a Matching Market Structure

1.1 Introduction

The deferred acceptance (DA) algorithm, a major result in market design, has numerous important practical applications. Several countries operate centralized matching markets that implement the DA algorithm to assign students to colleges. In these markets, college applicants submit their school preferences (rankings) along with their (potentially school-specific) admission scores. A growing empirical literature exploits a feature of these college admission systems whereby students with similar admission scores in a neighborhood of a school's admission threshold are or are not offered admission to the schools based on small differences in admission scores. For students at the margin of admission, treatment (school assignment) is driven by uncertainty in their admission scores (Lee, 2008). The literature relies on data from such centralized markets and on regression discontinuity design (RDD) to estimate the causal effects of attending specific schools on various outcomes.[1]

In a typical matching market setting, students submit their preferences (school rankings) knowing their exact admission scores. The matching market mechanism then compares test scores of the whole population of applicants in the order of their rankings and, given school-specific limits on the number of admitted students, it determines the school-specific admission score cutoffs. At the time of the application, there is therefore no individual-level admission score uncertainty; instead, uncertainty of admission (the source of quasi-random assignments to schools) corresponds to the uncertainty of school-specific score cutoffs, which in turn, are entirely determined by the market-level structure of applications, i.e., by the test scores and school rankings of all applicants in the market. From an analyst's perspective, it is natural to quantify the extent of randomness in this structure (and in the implied score cutoffs) by resampling from the applicant population and recording the simulated matching market outcomes.[2] Such resampling then allows one to form an ex-ante probability of admission for each student-school pair — an admission propensity score.

The RDD approach assumes that students' applications within a limiting neighborhood of the cutoffs, defined by a particular bandwidth, have similar admission probabilities regardless of the admission outcome.[3] Abdulkadiroglu et al. (2019) show that the RDD estimator can identify causal effects in matching-market settings only after controlling

[1] Kirkeboen et al. (2016) and Hastings et al. (2014) estimate labor-market returns of specific fields of study, Lucas and Mbiti (2014) and Abdulkadiroglu et al. (2014) study effects on standardized test scores, and Angrist et al. (2016) evaluate high school attendance effects on college choice. Dustan (2018), Fernandez (2019) and Altmejd et al. (2019a) use this research design to ask about the role of family ties in school choice.

[2] This is feasible in most existing applications of the RDD design to matching markets, in which analysts typically work with the entire applicant population.

[3] In RDD applications outside of the matching market literature, observations away from the cutoffs are often used, after conditioning on local polynomial regressions, to aid in the prediction of conditional means on either side of the cutoff, even if these away-from-the-cutoff observations are not subject to assignment risk themselves. In matching markets, however, these techniques are less applicable, since typically hundreds of schools are pooled to obtain a multi-cutoff RDD estimator, where cutoffs are endogenously determined by the market level structure of applicants. In such a setup, conditioning on school-specific polynomials (i.e. Altmejd et al., 2019b) might be overlooking confounding factors tied to the choice process at a specific school (i.e. higher-ranked schools might have different outcomes than lower-ranked schools) - see Abdulkadiroglu et al. (2019).

for admission propensity scores. However, Abdulkadiroglu et al. (2019) do not estimate propensity scores and instead rely on their limiting distributions. Specifically, they assume that within the bandwidth-defined neighborhoods around admission cutoffs, propensity scores are constant: they assume that the bandwidths are sufficiently narrow so that the random-assignment assumption holds. In this paper, I provide a direct test of the key RDD assumption in these settings, that close to admission cutoffs, admission probabilities are similar regardless of the admission outcome thanks to randomized assignment. The test is based on estimating and comparing the distributions of propensity scores for the treated and non-treated groups. To the best of my knowledge, this is the first study to assess this key identifying assumption of the RDD approach.

Most RDD school-choice studies choose a constant arbitrary bandwidth around all school admission cutoff thresholds.[4] To demonstrate that bandwidth choices do not drive the results, these studies typically employ robustness checks repeating the estimation for alternative values of the bandwidth. As an example, Abdulkadiroglu et al. (2014) use bandwidths ranging from roughly a third of the standard deviation up to the full standard deviation, while Kirkeboen et al. (2016) use all data (impose no bandwidths) in their main specification. Propensity scores, constructed by resampling from the applicant population, allow one to also assess whether these bandwidth choices lead the analyst to study outcomes of students who face no (quasi-) randomness in their school assignment.

For example, it could be the case that applicants to small schools face more uncertainty in their admission offers than applicants to large programs or schools. Therefore, propensity scores can also be used to inspect whether there are quasi-random assignments outside of the typical bandwidths used in RDD studies.

To illustrate the use of propensity scores in these settings, I calculate (and validate) propensity scores for each student-college pair using data from the Croatian college choice matching market from 2014 to 2018.[5] Next, I evaluate the propensity scores of applicants near school-specific admission cutoffs using bandwidth values typically found in the literature and obtain results that are not consistent with the assumptions of the RDD approach. First, I find that the propensity score distributions within typical bandwidths differ considerably across the treated and the non-treated groups. When considering applications of students who have admission scores at most half a standard deviation away from the school admission cutoffs, the average propensity score for the treated group is 85.8%, compared with 7.6% in the non-treated group. Second, I show that a substantial fraction of applications (roughly 40% in the case of half a standard deviation bandwidth) within the typical bandwidths faces no assignment risk at all (i.e., propensity score equal to 1 or 0). Such extensive differences between propensity score distributions for the applications of the treated and the non-treated students contradict the assumed random assignment to treatment near the admission threshold. Furthermore, the fact that almost half of the applications in RDD comparisons face no assignment uncertainty at all directly violates the Lee (2008) non-trivial assignment probability assumption.

[4]To improve the efficiency of the RDD approach, Abdulkadiroglu et al. (2019) use school-specific optimal bandwidths based on Imbens and Kalyanaraman (2012). However, optimality requires the independence of observations assumption, which is not satisfied in the school-choice framework as students apply to multiple schools.

[5]Propensity scores are estimated by adopting the Agarwal and Somain (2018) approach.

I find that only a drastic reduction in bandwidths — considering observations at most 0.01 standard deviations away from the cutoff (i.e., applying a bandwidth size that is 10 to 50 times smaller than those employed in Abdulkadiroglu et al. (2014) — results in comparable distributions of propensity scores, and less than 1% of students with a deterministic assignment. Focusing on narrow neighborhoods around the admission cutoffs, however, comes at the expense of neglecting observations with non-trivial propensity scores that are located outside of the chosen bandwidths. As an example, suppose that we are studying all students who have a probabilistic assignment to school u (i.e., non-trivial propensity score at school u). A student who has a non-trivial propensity score at some school ranked higher than school u, which implies a non-trivial propensity score also at school u, has a probabilistic assignment to school u, despite being far above the u-school cutoff. When identifying the effects of admission to school u, a typical RDD estimator will ignore these observations. To illustrate the extent of this, in the Croatian data I employ here, around 1.7% of the total applicant-school dyadic population has a propensity score between 40% and 60%. However, less than 30% of these highly randomized observations are captured within (RDD) samples defined by a 0.01 standard deviation bandwidth.

In sum, to adhere to the RDD assumptions, one needs to use smaller bandwidths, which, however, lead one to ignore much of the quasi-random applications available in matching market data. Hence, the application of the RDD design to matching market data faces fundamental obstacles. As an alternative approach, I propose a new estimator, the propensity score discontinuity design (PSDD), which applies the Rosenbaum and Rubin (1983) propensity score theorem to the matching market setting. By considering the propensity scores, the PSDD extracts the ex-ante uncertainty contained in the market-level structure of applications, and instead of choosing an arbitrary bandwidth, focuses only on the applications whose assignment is (quasi-) random. Crucially, the PSDD takes advantage of the timing of the matching market, recognizing that any potential selection into treatment must be embedded in the students' submitted preferences and in the admission score. Therefore, the selection-on-observables assumption in the standard propensity score theorem, which might seem unrealistic in the school choice setting, is not needed to employ the PSDD, as I show in Section 3. The PSDD estimator studies the outcomes of only those applications that face ex-ante (quasi-) randomness in their school assignment. Identification is therefore, by construction, based on observations with random assignments, both close to and away from the admission cutoff, and not on assuming randomized assignments as a function of distance from admission cutoffs as in the usual RDDs.

The remainder of this paper is structured as follows. Section 2 develops the matching market framework and proves that admission probabilities do not depend on the admission score in a limiting neighborhood around the cutoff. Section 3 develops the PSDD. Section 4 calculates propensity scores using Croatian matching market data and evaluates the typical bandwidths used in the literature. Section 5 concludes.

1.2 RDD Meets the Matching Market

In this section, I formally define a model of a student-school matching market inspired by Fack et al. (2019), which I then use to apply the quasi-experimental interpretation of RDD (Lee, 2008). The aim is to develop an evaluation tool to assess the appropriateness of the bandwidths used in studies employing the RDD in matching market settings. I adapt the model in Fack et al. (2019) by considering a potentially general form of students' preferences over schools in a market characterized by a finite number of students, and by modelling a student's type as a collection of his admission scores, preferences over schools, and observable and unobservable covariates.[6]

Consider a matching market defined by a set of students I and a set of schools U. Denote the cardinality of the set of students with $|I|$ and suppose that I is constructed by independently drawing $|I|$ students from the distribution H. Next, suppose that a set of schools U is fixed. The objective of the market is to match each student $i \in I$ to exactly one school $u \in U$.

Student $i \in I$ is described by a random vector $i = \{R_{i,v\ v\in U}, >_i, W_i, X_i\}$, where $R_{i,v}$ is the admission score of the student i at school v, $>_i$ describes preferences of student i over (some) schools in U,[7] and W_i and X_i are students i's unobservables and observables, respectively. School $u \in U$ is described by a fixed scalar $u = \{q_u\}$, where q_u is a fixed quota for school u. Given two applications by students i and j, school u gives admission priority to student i if and only if $R_{i,u} > R_{j,u}$.

The timeline of the matching market is as follows:

1. A set of students I is constructed by $|I|$ independent draws from the distribution H.
2. Each student i learns his admission scores $R_{i,v,\ v\in U}$.
3. Each student i, based on his preferences $>_i$, admission score $R_{i,v,\ v\in U}$, observable (X_i) and unobservable (W_i) covariates, submits an ordered priority list $S_i := S_{i,l},\ l \in \{1, ..., L_i\}$, where l denotes the priority of the school, and L_i denotes cardinality of the set of schools submitted by student i. For example, if students i's admission score exceeds the school's $S_{i,1}$ cutoff, he is offered admission to school $S_{i,1}$. If it is below the school's $S_{i,1}$ cutoff, he is considered for eligibility in school $S_{i,2}$ and so on.
4. Given observed students' priority lists $\{S_{i,l}\}, \forall i, l$, and schools' quotas $\{q_u\}, \forall u \in U$, the matching market defines a mapping $T_i : I \mapsto U$, assigning exactly one school $u \in U$ to each student $i \in I$. Denote with c_u the admission cutoff for school u,

[6] In Fack et al. (2019), the student's type does not include observable and unobservable covariates and preferences are assumed to be guided by the von Neumann-Morgenstern utilities. The results do not depend on these extensions of the original model; they are implemented solely due to expositional purposes.

[7] The outside option for a student is not going to any school. If a student's preference over a particular school is not defined, I assume that the outside option is preferred to this school.

which takes a non-zero value only for programs that filled its quota[8]:

$$c_u(I) = \begin{cases} \min_{i:T_i=u} R_{i,u} & \text{if } \sum_i 1_{T_i=u} = q_u \\ 0 & \text{if } \sum_i 1_{T_i=u} < q_u \end{cases}$$

Assume the following properties of the mapping T_i:

- **Non-wastefulness**: $\nexists i \in I$ such that $\exists u \in U$ so that $S_{i,k} = u$ and $S_{i,l} = T_i$ while $k < l$ and $R_{i,u} > c_u$. In words, schools are required to admit students until there are no unfilled vacancies.
- **Transparent assignment**: $\nexists$ a pair of students $i \in I$ and $j \in I$ such that $R_{i,u} > R_{j,u}$ and $T_i = p$, for some $p \in U$, and $T_j = u$ while $S_{i,k} = u$ and $S_{i,l} = p$ while $k < l$. In words, schools are allowed to rank the students only with respect to the (school-specific) admission score.

Denote with I_{-i} the set of students excluding student i. Then, the admission offer of student i at school u is defined as:

$$a_u(S_i, R_i, I_{-i}) = \begin{cases} \mathbb{1}(\text{if } S_{i,l} = u \text{ for some } l \text{ and } R_{i,S_{i,r}} < c_{S_{i,r}}(I)\ \forall r < l \text{ and } R_{i,u} \geq c_u(I)) \\ 0 \text{ otherwise} \end{cases}$$

In words, student i is only offered admission at school u if he listed school u on his priority list, was rejected by all the schools listed above school u, and finally met admission criteria for school u. Further, given that the admission decision for student i, given his submitted priority list and admission scores, depends exclusively on the set of students I, the (ex-ante) probability of student i receiving an admission offer at school u is:

$$P(T_i = u) = \int a_u(S_i, R_i, J_{-i})\ dH(J_{-i}). \tag{1.1}$$

The matching market implementing the deferred acceptance algorithm corresponds to this framework.

Next, I use the framework described above to adapt the quasi-experimental interpretation of RDD (Lee, 2008) to the case of the matching market. Denote with $g(\cdot)$ the density of the unobservable W_i and introduce the following definition:

Definition 1: Selection on unobservables. *The treatment is subject to selection on unobservables if there is a non-zero correlation between unobservable W_i and the treatment assignment T_i.*

Lee (2008) assumes that there are no discontinuities in the density of unobservables when the admission score equals the cutoff:

Assumption 1: Continuity of unobservables. *The conditional density $g(\cdot|R_{i,u} = c_u)$ is continuous at c_u.*

[8]The cutoffs depend on the market level structure of applicants, as the student with the lowest admission score admitted to a particular school defines that school's cutoff.

In Lee (2008), the cutoff is *perfectly known* before the treatment assignment, and thus the assumption essentially assumes away selection on unobservables. In contrast, in our framework, cutoffs are not known at the time of the application — there is a *cutoff uncertainty*. For example, suppose that the admission score $R_{i,u}$ at school u is drawn from the continuous distribution for each student i. Then the distribution of the cutoff c_u, the q_u-order statistics of school-u applicants, is also continuous. In this case, the Continuity of unobservables assumption is satisfied *almost surely* if the conditional density $g(\cdot|R)$ is discontinuous for, at most, a finite number of admission score values R.

Suppose that we are interested in measuring the effect of attending school u on some outcome Y. RDD strategies focus on those applications, for which $a_u(S_i, R_i, I_{-i}) \equiv 1_{R_{i,u}>c_u}$, where the assignment to school u is, from the ex-post perspective, a function of the school u specific admission score $R_{i,u}$ (from now on I refer to these observations as the *RDD estimation sample*).[9] More formally, the RDD estimation sample is defined as follows:

Definition 2: RDD Estimation Sample. *Suppose that student i is assigned to school $S_{i,k}$, $T_i = S_{i,k}$ for some k. Application $S_{i,l}$ is included in the RDD estimation sample only if $l \leq k$.*

As recognized in Kirkeboen et al. (2016), the outcome Y for student i who enrolled at school u, can only be estimated relative to the outcome at the counterfactual school, which the student would be admitted to if he was rejected admission to school u. To formalize this idea, denote with D_i a dummy variable indicating student i's treatment status (enrolling at school u), i.e. $D_i := 1_{T_i=u}$. Suppose that the outcome Y_k at any school k is a function of the unobservable W_i, i.e. $Y_k = Y(W_i|T_i = k)$. Specifically, for students with applications in the RDD estimation sample, denote with Y_1 and Y_0 the outcomes when attending and not-attending school u, respectively: $Y_1 := Y(W_i|D_i = 1)$ and $Y_0 := Y(W_i|D_i = 0)$. Using the law of iterated expectations, we obtain $E[Y_1 - Y_0|R_{i,u} = c_u] = \sum_{k \in U}(\widehat{Y_{i,u}} - \widehat{Y_{i,k}})P(T_i = k|T_i \neq k)$ where $P(T_i = k|T_i \neq u)$ is the probability of student i being admitted to school k, given that he was marginally rejected admission to school u.

The following proposition provides the experimental interpretation of the RDD:

Proposition 1. *Suppose that the Continuity of unobservables assumption holds. Then, the following holds for each school-student pair in the RDD estimation sample:*

$$E[Y|R_{i,u} = c_u] - \lim_{x \to c_u^-} E[Y|R_{i,u} = x] = E[Y_1 - Y_0|R_{i,u} = c_u]$$

Proof. See Lee (2008). □

Under the continuity assumption, the proposition establishes causality by comparing the outcomes of students, with different treatment assignments, in the RDD estimation sample with the school u-specific admission scores within a limiting neighbourhood around the cutoff. A critical aspect of the RDD implementation is choosing a bandwidth to define this limiting neighbourhood — reducing the bandwidths excessively, while improving the credibility of Proposition 1, results in diminishing sample sizes. To compromise this

[9] The usual justification for this is that assignment to these schools is not possible.

empirical tradeoff, a series of papers have proposed procedures to calculate *optimal* bandwidths (i.e., Imbens and Kalyanaraman, 2012). This literature concentrates on cases in which there is only one running variable and does not cover our case of multiple (potentially correlated) cutoffs and admission scores. Therefore, studies applying RDD to the matching markets setting generally use arbitrary bandwidths that are not supported with theory. The following simple adaptation of Proposition 1 can be used to assess the appropriateness of the bandwidth. Intuitively, in the limiting neighbourhood of Proposition 1, as long as the student's admission scores are continuous, the admission score does not change the admission probability significantly. Therefore, to support the bandwidth choice, an analyst can verify that the admission probability distributions of those just above the cutoff (treated students) are similar to admission probability distributions of those just below the cutoff (non-treated students).

Lemma 1. *Suppose that the school u-specific admission score $R_{i,u}$ is drawn from the continuous distribution for each student i. Then, for any application in the RDD estimation sample the following holds:*

$$P(T_i = u | R_{i,u} = c_u) = \lim_{x \to c_u^-} P(T_i = u | R_{i,u} = x).$$

Proof. In the RDD estimation sample, the ex-ante probability of admission to school u for a student with admission score $R_{i,u} = c_u$,

$$P(T_i = u | R_{i,u} = c_u) = \int a_u(S_i, c_u, J_{-i}) \, dH(J_{-i}),$$

which simplifies to

$$P(T_i = u | R_{i,u} = c_u) = \int 1_{c_u > \widehat{c_u}} \, dH(J_{-i}),$$

where the integral goes over the support of possible cutoffs $\widehat{c_u}$ defined by the distribution of $H(J_{-i})$ through:[10]

$$c_u(I) = \min_{j:T_j=u} R_{j,u} \qquad \text{where } \sum_j 1_{T_j=u} = q_u$$

Note that, as the cutoff c_u is simply a q_u-order statistic, the continuity of $R_{j,u}$ for each j, implies the continuity of c_u, so that:

$$\int 1_{c_u > \widehat{c_u}} \, dH(J_{-i}) = \int 1_{c_u > \widehat{c_u}} \, dQ(\widehat{c_u})),$$

for some continuous distribution Q. The lemma now follows from the continuity of Q. □

[10] Note that the schools where the cutoff is not imposed do not belong to the RDD estimation sample.

1.3 Propensity Score Discontinuity Design

A key empirical challenge for the identification of school admission effects is accounting for the potential selection on unobservables. For example, unobservable levels of motivation of students might confound the estimate of school-specific labour market returns. In practice, there are two distinct RDD approaches used to account for the selection on unobservables. The more common one is nonparametric, which estimates a local linear regression around the cutoff using pre-defined bandwidth values. This approach assumes that the unobservables W_i are balanced in a local neighbourhood of a fixed admission score:

Assumption 2: Nonparametric identification. *For each school u-specific admission score $\hat{c}_u$, there exists δ such that $Y_0, Y_1 \perp W_i$, $\forall i \in I$ for $R_{i,u} \in \langle \hat{c}_u - \delta, \hat{c}_u + \delta \rangle$.*[11]

Alternatively, there is a parametric regression approach typically using polynomials to model the running variable (in our case the admission score) over its entire support. This approach assumes that the outcome is orthogonal to the unobservables W_i conditional on polynomials in the admission score:

Assumption 3: Parametric identification. *Outcome is orthogonal to the unobservables conditional on polynomials in admission score: $Y_0, Y_1 \perp W_i | R_{i,u}, R_{i,u}^2, \dots R_{i,u}^p$, $\forall i \in I$.*

As discussed in the previous section, the RDD uses the *RDD estimation sample*, excluding all the alternatives ranked below a student's (ex-post) admission outcome (as the student did not compete at these schools, the assignment was *impossible* from the ex-post perspective.) This is because the RDD assumes, through Assumption 2 or Assumption 3, that students around the cutoff are "the same" in every aspect except the admission outcome, which is deterministically linked to the school-specific admission score (i.e. a student is offered admission if and only if his admission score is above the school-specific cutoff). For schools ranked below the ex-post admission outcome, this deterministic link between admission score and the assignment is broken — the student is never considered for admission (even if he is above the cutoff for these schools). For this reason, these observations are not included in the RDD estimation sample. However, from the ex-ante perspective, excluding these observations is unnecessary as students who are randomly accepted to a higher-ranked school are also randomly not assigned to a lower-ranked school.

For example, suppose a student ranked school A as his first choice, followed by school B. Suppose that ex-ante he had a 50% probability of being admitted to school A, and a 50% chance of being admitted at school B, and assume that, ex-post, he was just above school A's cutoff. The RDD concludes that the assignment to school B is ex-post impossible, and therefore excludes this choice from the estimation sample to ensure that the assignment is a deterministic function of only the admission score and the school-specific cutoff. The estimator I am proposing below understands that the assignment to school A was ex-ante just as probable as the assignment to school B, and thus it includes both choices in the

[11] All results provided in this section hold under the weaker Continuity of unobservables assumption from the previous section, which would require showing that the local PSDD is well defined (analogously to Proposition 1). I decided to focus on the stronger assumption, due to the intuitive appeal, and elegant exposition.

estimation sample. In the following paragraph, I provide the motivation for the new estimator, by naturally generalizing simple estimators used in settings where admissions are resolved using a lottery at the margin of admission, to incorporate also the uncertainty in cutoffs at the time of the application.

In a typical high-school admission system, admission scores are coarse, and students are divided into a small number of groups with different admission priorities. Admission decisions for the students in the group at the margin of admission are then implemented by breaking the ties (within the group) by an admission lottery. The literature on high-school admission effects typically focuses only on the marginal group where the assignment is explicitly random (e.g., Abdulkadiroglu et al., 2017), which effectively conditions on the market structure of applications that assigned a specific group of applicants to the marginal group. While this ex-post assignment randomization is clearly ideal for the purpose of identification within the marginal group, there are additional sources of ex-ante assignment uncertainty corresponding to uncertainty at the time of the application. In settings where admission scores are highly coarse, the additional ex-ante uncertainty is negligible. To illustrate this, suppose that there are only two admission-score groups: A and B, and only one school prioritizing group A over group B. Suppose also that the school capacity is such that everybody in group A is highly likely to be admitted, while a lottery determines admission from group B. In this case, for a specific student in group B, the dominant component of total ex-ante admission uncertainty is the ex-post lottery draw, since group B is almost certainly the marginal group. In contrast, when admission scores are less coarse, such that there are numerous admission groups, the ex-ante probability depends not only on the ex-post lottery draw applicable to the (smaller) group of marginal applicants, but significantly also to the ex-ante uncertainty of ending up in the marginal group, which depends on the market-level structure of applications. The method I propose below incorporates the ex-ante uncertainty, employing larger sample sizes than the conventional lottery-based estimators since it considers students facing ex-ante probabilistic assignments outside of the marginal admission group. In other words, the method also includes the applications of students outside of the marginal group, as long as their admission is ex-ante sufficiently probabilistic. The remainder of this section formalizes this intuition and generalizes it to the case of a continuous admission score (a typical case in college-admission systems).

Assume the matching market from the previous section and suppose we are interested in the effect of attending a college u on student i's outcome $Y_i,\ i \in I$. Denote with D_i a dummy variable indicating treatment assignment, $D_i = 1_{T_i=u}$ and let $Y_1 = Y(W_i|D_i = 1)$ and $Y_0 = Y(W_i|D_i = 0)$. Assume the following:

Assumption 4: Ignorability of cutoffs. *The outcome Y is independent of the cutoff c_u, conditional on the submitted priority list S_i and admission scores $R_{i,v\ v\in U}$:*

$$Y_1, Y_0 \perp C_u|(S_i, R_{i,v\ v\in U})).$$

In words, the Ignorability of cutoffs assumes that the outcome $Y(W_i)$ does not depend on the realization of the cutoff, conditional on the student's observable characteristics. For example, this will hold if student i's outcome $Y(W_i)$ does not depend on the other

students' school assignments, i.e. $Y(W_i) \perp T_j,\ j \in I_{-i}$. It is worth noting that this assumption is also crucial for the identification of the conventional RDD.[12]

Notice that the original Rosenbaum and Rubin (1983) propensity theorem requires the strong ignorability of treatment assumption — it assumes that the selection into treatment is not affected by unobservables.[13] In Proposition 2 below, I adopt the propensity score theorem (Rosenbaum and Rubin, 1983) to the school choice setting, acknowledging that unobservables are reflected in the submitted ordered priority lists and the admission scores. More precisely, given the student's submitted ordered priority list and his admission scores, the admission decision depends only on the realization of the cutoff. Therefore, the student's observable characteristics X_i and the student's unobservable characteristics W_i do not affect his admission outcome D_i, other than through his admission score $R_{i,v\ v\in U}$ and submitted priority list S_i. Consequently, under the assumption that the cutoff realization does not influence the outcome, matching on the propensity scores accounts for the possible selection on unobservables. In other words, unlike in the original propensity score theorem where it is assumed, strong ignorability of treatment follows from the Ignorability of cutoffs.

Proposition 2. *Suppose Ignorability of cutoffs holds. Then treatment assignment is strongly ignorable, in the sense of Rosenbaum and Rubin (1983), given the submitted ordered priority list and the student's admission score:*

$$Y_1, Y_0 \perp D_i | (S_i, R_{i,v\ v\in U}).$$

Therefore, the expected difference in observed outcomes conditional on $P(D_i|(S_i, R_{i,v\ v\in U}))$ is equal to the average treatment effect at $P(D_i|(S_i, R_{i,v\ v\in U}))$:

$$E[Y_1|T=1, P(D_i|(S_i, R_{i,v\ v\in U}))] - E[Y_0|T=0, P(D_i|(S_i, R_{i,v\ v\in U}))] =$$

$$E[Y_1 - Y_0 | P(D_i|(S_i, R_{i,v\ v\in U}))]$$

.

Proof. The treatment assignment D_i is determined by a mapping Ψ, where $D_i \equiv \Psi(S_i, R_{i,v\ v\in U}, c_{v,v\in U})$; that is, knowing the student's submitted priority list S_i, his admission scores $R_{i,v\ v\in U}$ and the school-specific cutoffs $c_{v,v\in U}$, one can determine treatment assignment with certainty. Therefore, we obtain:

$$Y_1, Y_0 \perp D_i | (S_i, R_{i,v\ v\in U}) \iff$$
$$Y_1, Y_0 \perp \Psi(S_i, R_{i,v\ v\in U}, c_{v,v\in U}) | (S_i, R_{i,v\ v\in U})$$

By assuming Ignorability of cutoffs, i.e. $Y_1, Y_0 \perp c_{v,\ v\in U} | (S_i, R_{i,v\ v\in U})$, the second line above follows directly. Therefore, treatment assignment D_i is strongly ignorable given

[12] If the Ignorability of cutoffs does not hold, the RDD treatment effect could be driven by the cutoff proximity of the treated students, as their counterfactuals (schools they are assigned if they are just below the treatment) are not necessarily around the cutoff. For example, a student who was marginally declined admission to the cutoff school may be well above the cutoff at his next highest ranked school.

[13] In the school choice setting this might seem unrealistic. For example, suppose that a student is incentivized into a law school, being brought up by a lawyer mother and a lawyer father.

$(S_i, R_{i,v\ v\in U})$, in the sense of Rosenbaum and Rubin (1983). Using theorem 3 from the same paper, and the Strong ignorability of treatment just obtained, we get:

$$E[Y_1|T=1, P(D_i|(S_i, R_{i,v\ v\in U}))] - E[Y_0|T=0, P(D_i|(S_i, R_{i,v\ v\in U}))] =$$

$$E[Y_1 - Y_0|P(D_i|(S_i, R_{i,v\ v\in U}))].$$

□

While Proposition 2 uncovers the treatment effect for students with a particular value of the propensity score, it does not guarantee that there is no heterogeneity in admission scores among these students. More precisely, Theorem 1 in Rosenbaum and Rubin (1983) says that the propensity score is a balancing score, i.e., $(S_i, R_{i,v\ v\in U}) \perp D_1, D_0|P(D_i|(S_i, R_{i,v\ v\in U}))$. Intuitively, given a particular value of the propensity score, the distributions of the submitted priority lists and the admission scores do not differ depending on the treatment assignment. Proposition 2 uses this fact, after proving that submitted priority lists and the admission scores are strongly ignorable, i.e, $Y_1, Y_0 \perp D_i|(S_i, R_{i,v\ v\in U})$, to identify the treatment effect at a particular value of the propensity score (by applying Theorem 3 in Rosenbaum and Rubin, 1983). Therefore, the identification of the average treatment effect at the specified propensity score value is guaranteed due to the balancing property of the propensity score, even though the students who have the same propensity score could potentially have different submitted priority lists and different values of the admission scores. The resulting treatment effect in Proposition 2 is therefore the average over all students with the same propensity score. The following example demonstrates that admission scores can be significantly different for students with the same propensity score.

Example 1. *Suppose that there are two schools A and B, which rank students according to the same admission score. Suppose that student 1, with admission score 100, lists school A as his first priority. Suppose that school A's cutoff is uniformly distributed, with the mean value of 100, so that student 1 has a propensity score for school A of 50%. Suppose that student 2, with admission score 200, lists school A as his second priority, only after school B, which has a uniformly distributed cutoff with the mean value of 200. Even though student 2 has an admission score two times larger than student 1, their propensity scores for school A are the same.*

Example 1 shows that students with the same propensity score can have uncomparable values of the admission score. Therefore, to account for the potential admission score heterogeneity conditional on propensity score (henceforth referred to as heterogeneity), similarly to the conventional RDD, I propose controlling for the admission score, therefore adding the admission score to the conditioning set of Proposition 2:

$$E[Y_1 - Y_0|P(D_i|(S_i, R_{i,v\ v\in U})), R_{i,v\ v\in U}] \tag{1.2}$$

Note that, while the RDD utilizes admission scores to deal with both the selection on the unobservables and the heterogeneity, Equation 1.2 eliminates selection by matching on propensity scores, and uses the admission score only to account for the heterogeneity of students.

Depending on the treatment of the admission score, I define two versions of the PSDD: *Local PSDD* which, similarly to the Nonparametric identification assumption, identifies

effects for students with similar admission scores, and the *Global PSDD* which, similarly to the Parametric identification assumption, utilizes the whole sample while controlling for the admission score.

Definition 3: PSDD. *Fix the school u-specific admission score $\hat{c}$, the propensity score value k and an admission score bandwidth δ. There are two versions of the PSDD, depending on the admission score treatment:*

- ***Local PSDD:***

$$PSDD(u, k, \hat{c}, \delta) = E[Y_1 - Y_0 | P(D_i) = k, R_{i,u} \in \langle \hat{c} - \delta, \hat{c} + \delta \rangle],$$

  and

- ***Global PSDD:***

$$PSDD(u, k) = E[Y_1 - Y_0 | P(D_i) = k, R_{i,u}].$$

In applications, since there is a limited number of observations at the exact specified propensity score value k, I evaluate the average PSDD over an interval of propensity score values around k. Fix a propensity score bandwidth ϵ and define the *average global PSDD*, $PSDD(u, k, \epsilon)$:

$$PSDD(u, k, \epsilon) = \int_{k-\epsilon}^{k+\epsilon} PSDD(u, \hat{k}) \, d\hat{k}. \tag{1.3}$$

Average local PSDD, $PSDD(u, k, \hat{c}, \delta, \epsilon)$ is defined analogously:

$$PSDD(u, k, \hat{c}, \delta, \epsilon) = \int_{k-\epsilon}^{k+\epsilon} PSDD(u, \hat{k}, \hat{c}, \delta) \, d\hat{k}. \tag{1.4}$$

To estimate the average local $PSDD(u, k, \hat{c}, \delta, \epsilon)$, I run the following regression:

$$y_i = \alpha + \rho \cdot D_i, \text{where } PS_i \in \langle k - \delta, k + \delta \rangle \text{ and } R_{i,u} \in \langle \hat{c} - \epsilon, \hat{c} + \epsilon \rangle \,, \tag{1.5}$$

where ϵ is the admission score bandwidth for some fixed value of the admission score $\hat{c}$, δ is the propensity score bandwidth for some fixed value of the propensity score k, PS_i is the propensity score of individual i and propensity score, respectively, and ρ is the global average $PSSD(u, k, \delta)$ defined in Equation (1.4). From a practical perspective, to implement the local PSDD, an analyst can run the usual RDD specification, while restricting the sample to the applications with similar propensity scores. However, unlike in the RDD, where proximity to the cutoff is assumed to eliminate the potential selection on unobservables into schools, the average local PSDD (Equation (1.5)) eliminates selection on unobservables by restricting the sample to the applications who also have a similar propensity score. Therefore, some applications that are *close to the cutoff* in the RDD sense, might not be included in the PSDD sample. The purpose of restricting the sample to applications with a similar admission score is thus only eliminating heterogeneities in admission scores of students who have similar propensity scores.

To estimate the average global $PSDD(u, k, \delta)$, denote with $Pl_w(\cdot)$ an operator transforming a variable to a polynomial of a fixed degree w, i.e. $Pl_3(PS_i) = \alpha_1 \cdot PS_i + \alpha_2 \cdot PS_i^2 + \alpha_3 \cdot PS_i^3$, and run the following regression:

$$y_i = \alpha + Pl_w(R_{i,u}) + Pl_w(PS_i) + \rho \cdot D_i, \text{where } PS_i \in \langle k - \delta, k + \delta \rangle \tag{1.6}$$

In some applications, analysts use the Parametric identification assumption 1.3 employing RDD on the whole data, without using a bandwidth to restrict the RDD sample to those close to the cutoff. In these cases, RDD uses polynomials in the admission score to account for selection into schools. In contrast, the global PSDD (equation (1.6)) considers only applications with similar propensity scores, i.e. $PS_i \in \langle k - \delta, k + \delta \rangle$, thus explicitly modelling selection on unobservables. Similarly to the above, the purpose of the polynomial in the admission score is thus only to pick up heterogeneity in students' admission scores. In other words, the global PSDD can move away from the cutoff and identify treatment effects by comparing only the applications of students whose assignment, from the ex-ante perspective, is (quasi-) random.

I conclude the section with outlining the procedure for calculating propensity scores, which is a direct adoption of the Agarwal and Somain (2018). As demonstrated in the previous section, the (ex-ante) probability of student i receiving admission to school u is:

$$P(T_i = u) = \int a_u(S_i, R_{i,u}, J_{-i})\, dH(J_{-i}).$$

Consider the following estimator of the admission probability:

$$\widehat{P(T_i = u)} = \frac{\sum_{r=1}^{N_b} a_u(S_i, R_i, J_{-i,r})}{N_b}, \tag{1.7}$$

where $J_{-i,r}$ are independent draws from $H^{|I-1|}$, H is the student's distribution (i.e., H determines the student's admission scores and unobservable and observable characteristics), and $|I|$ is the number of students in the application year. The central limit theorem then guarantees the consistency of the estimator (1.7) as $N_b \to \infty$. Since the set of students I corresponds to the whole population, I assume that the distribution H is completely determined by I, i.e. $H = I$. Then, the independent $H^{|I-1|}$ draws can be constructed by bootstrapping students with replacement from the set of students I. Below, I provide steps for calculating propensity scores.

Procedure: Calculating assignment probability. Denote with N the cardinality of the set of students I, $N := |I|$ and create a vector of zeros $A_{i,u} = 0$ for each $i \in I, u \in U$. Repeat the following steps N_b times :

1. Draw N students with replacement from the set I. Denote the generated student sample with $\hat{I}$.
2. Given the school's quotas $\{q_u\}, \forall u \in U$ assign exactly one school to each student $i \in \hat{I}$ using the matching algorithm $T_i : \hat{I} \mapsto U$.
3. For each student i, matched with some school u, update value $A_{i,u}$ according to $A_{i,u} = A_{i,u} + 1$.

The bootstrapped probability estimate of student i's assignment to a school u is then $P(T_i = u) = A_{i,u}/N_b$.

To implement the proposed bootstrapping procedure, one has to observe a population of schools with their admission quotas (maximum number of admitted students) and a population of students with their submitted priority lists and school-specific admission scores. Since the school-students matching markets under consideration are centralized, this information is usually available to analysts.

1.4 Empirical Application - DA in Croatia

A large literature estimates various school-graduation effects by employing the RDD around school admission score cutoffs, assuming (quasi-) random school assignment within the implemented RDD bandwidths. In this section, I present evidence suggesting that this key assumption is violated for the bandwidths typically used, using data from the Croatian centralized college admission system from 2014 to 2018. After describing the institutional setup and data, I calculate the propensity scores and evaluate various bandwidths using Lemma 1.

1.4.1 Institutional Setup and the Data

In Croatia, admissions to all college programs are implemented through a national online platform. Since its introduction in 2010, this platform operates a deferred acceptance (DA) algorithm that ranks students based on their high-school grades and subject-specific elective national-level exams that take place in June, a month after high-school graduation. Students register on the platform in the early spring of their high-school graduation year when universities also list on the platform their program admission quotas along with program-specific weights of subject-specific grades and exams. Students are free to submit their ranked priority lists of up to 10 programs as of registration and update these preference rankings until the system closes for clearing at a predetermined date in mid-July (in 2019, the final deadline was 2 pm on 15th July).

Students first receive information on their position in various admission queues one week before the final deadline, immediately after receiving their state exam scores and hence, admission scores. The DA algorithm is then regularly updated to show students their current admission position.

I analyze the first preference submission after receiving national exam scores when students are fully aware of their admission scores but do not yet receive the signal about market demand from observing their position in admission queues. This choice is meant to focus on a decision referencing the one-off preference ranking decision in a conventional static DA mechanism with no updating. In addition, by focusing on the first applications students submit after learning their exam performance, I avoid endogeneity issues in admission results that may arise from some students learning about their current admission rankings and being more active in modifying their applications before the deadline. [14] In a recent multi-national study, Altmejd et al. (2019b) argue that the Croatian first preference submissions are structurally similar to the static DA submissions in Sweden and Chile, and find similar sibling spillover effects on college applications and enrollment in each of these countries.

Appendix Table 1.1 shows basic summary statistics for the Croatian DA matching market throughout 2014-2018. The year 2015 is excluded as only the RDD estimation sample is available, which excludes the observations ranked below the admission school — this is not sufficient to calculate propensity scores. Annually, approximately 35,000 students enter the matching market, choosing between approximately 620 programs belonging to 49 distinct colleges. An average student applies to approximately six programs, and

[14] I obtained virtually the same results when focusing on the last preference submission.

the average admission rate, calculated as the number of admissions over the number of applications, is just under 0.2.

1.4.2 Propensity Scores and the RDD

In this section, I generate a function[15] that takes program-specific quotas, and student-specific preferences and admission scores as inputs, performs the DA algorithm, and returns, as output, the matched program T_i for each student i. I validate the function by correctly and completely replicating the actual DA assignments in Croatia. I calculate propensity scores as described in procedure 1.3, iteratively redrawing the student population, running the DA algorithm, and recording simulated admission outcomes. Student-program-specific propensity scores are then calculated as simple averages of the simulated admission outcomes.[16] The goal of the propensity scores is to extract the ex-ante admission probability for each student-program pair from the complex probability space generated by different programs (and their sizes), students' admission scores, and students' submitted preferences.

The propensity scores predict the actual DA assignments almost perfectly (propensity scores explain 97% of the variation in admission offers), in large part since the majority of the sample (almost 85%) has a deterministic assignment (i.e., propensity score either 0 or 1). In these cases, the propensity score is, by construction, a perfect indicator of the admission.[17]

The RDD approach assumes that students within a limiting neighbourhood of the cutoff have similar admission probabilities regardless of the admission outcome. Using calculated propensity scores and Lemma (1) I evaluate the random assignment assumption using samples of Croatian data defined by the bandwidths typically used in the literature.[18] Most of the RDD studies in the literature choose a constant arbitrary bandwidth, applied to each program-specific cutoff to define a limiting neighborhood. To demonstrate that the bandwidth choice does not drive the results, these studies typically employ robustness checks repeating the estimation for alternative values of the bandwidth. As an example, Abdulkadiroglu et al. (2014) use bandwidths ranging from roughly a third of the standard deviation up to the full standard deviation, while Kirkeböen et al. (2016) use all data (impose no bandwidths) in their main specification.

Figure 1.1 plots the distribution of propensity scores for the treated and the non-treated group separately[19], and strongly suggests that the bandwidths exceeding half a standard deviation are excessively large, since using them results in samples in which a sizeable fraction of students is deterministically assigned to a program. More precisely, when considering students with applications that have admission scores at most half a stan-

[15] Codes in Python and R are available upon request

[16] I use 10,000 iterations, ensuring that at the end of the algorithm each additional iteration changes a particular propensity score by at most 0.0001.

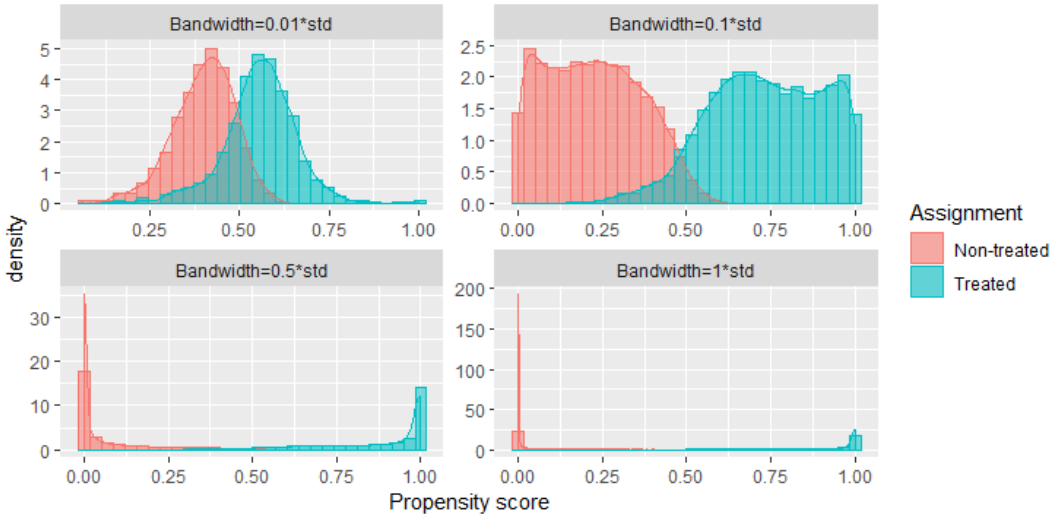
[17] After restricting the sample to applications with propensity scores between 10% and 90%, propensity scores are still a very strong predictor of the admission offer, explaining 81% of the variation

[18] In the appendix I perform the typical Cattaneo et al. (2019) manipulation test around the cutoff, which finds no evidence of discontinuity.

[19] If an application of student i to school s resulted in an admission offer (student i was offered a place in school s), I consider the application treated. Otherwise, I consider the application non-treated.

dard deviation away from the school admission cutoffs, 36% of the applications face no assignment risk at all (i.e., their propensity score equals either 1 or 0). When using a full standard deviation bandwidth, 56% of the applications have trivial (0 or 1) propensity scores. Even reducing the bandwidth to 0.1 of the standard deviation results in a sample in which a sizeable fraction of applications face almost no admission risk - 20% of applications have a propensity score higher than 90%.

Figure 1.1: Distributions of propensity scores for the treated and the non-treated group by RDD bandwidths choice



The figure plots the kernel density of propensity scores in RDD estimation samples defined by different bandwidth values around the cutoffs. The blue (red) histogram plots the distribution for students who were (not) offered admission to the applied school — (non-)treated students. According to Lemma (1), these two distributions should be *similar*. As we increase the bandwidths (to the values typically employed in the literature), the differences between these two distributions become striking.

Further, motivated by Lemma (1), which says that the propensity score, in the limiting neighbourhood around the cutoff, does not depend on the admission score, Figure 1.1 compares the distributions of the propensity score for the treated and the non-treated groups.[20] It is impossible to compare the applications exactly at the cutoff (all of these applications belong to the treated group). However, one would expect that in a reasonable RDD estimation sample, defined by an appropriate bandwidth, the distributions of the propensity score for the treated and the non-treated groups, as per Lemma (1), are similar. Again, Figure 1.1 suggests that using a bandwidth of 0.5 standard deviations or more is inappropriate as it results in entirely unbalanced propensity score distributions. For example, when using the bandwidth of 0.5 standard deviations, the average propensity score for the treated group is 85.8%, compared with the 7.6% in the non-treated group. Only a drastic reduction of the bandwidth to a value of around 0.01 standard deviations results in comparable propensity score distributions. Additionally, the figure indicates that the implementation of PSDD is feasible for propensity score values close to 50%, since, for these values, both treated and non-treated applications are common (i.e., it is a common support).

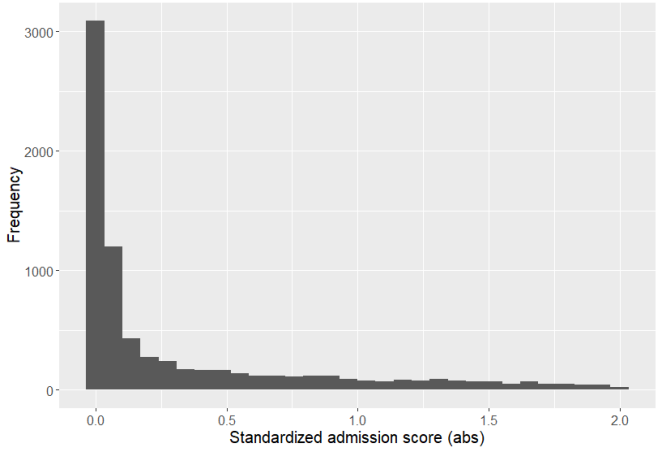
[20]By construction (definition of the cutoff), there is a mass of applications exactly at the cutoff (at least 1 application per program). I exclude these applications before plotting the figure.

The extensive differences between propensity score distributions for the treated and the non-treated students contradict the assumption of random assignment to treatment near the admission threshold. Furthermore, that almost half of applications in "RDD" comparisons face no assignment uncertainty at all directly violates the Lee (2008) non-trivial assignment probability assumption. Since bandwidths used for robustness checks are typically in the range from 0.1 standard deviations to 1 standard deviation, these conclusions hold for them too. Therefore, even if the RDD estimates look stable across robustness checks typically employed, they could be different when using the sample of (quasi-) randomized applications, as the estimation samples differ significantly.

Applying a drastically reduced bandwidth is not a solution either, as any constant bandwidth cannot reflect potentially program-specific cutoff uncertainty. For example, considering only the applications that are at most 0.01 standard deviations away from the cutoff (i.e., 10-50 times smaller bandwidth than in Abdulkadiroglu et al., 2014), which results in 99% of applications within the bandwidths having non-trivial propensity scores, still leaves only 35% of propensity scores between 40% and 60%, as Figure 1.1 suggests.

Additionally, focusing on narrow neighborhoods around the admission cutoffs comes at the expense of neglecting observations with non-trivial propensity scores that are located outside of the chosen bandwidths. Figure 1.2 plots the histogram of the distance of the absolute admission score (divided by the standard deviation) from the cutoff for the observations with the propensity scores between 40% and 60% (around 1.7% of the whole sample). Less than 30% of these "highly randomized" observations are captured within the RDD estimation sample defined by a bandwidth of 0.01 standard deviation.

Figure 1.2: Histogram of admission scores for the applications with random assignment



The figure plots the distribution of the absolute value of the standardized (divided by its standard deviation) admission score for the observations with propensity scores between 40% and 60%. The figure shows that there is a sizeable portion of applications with randomized assignment whose admission score is far from the cutoff. Therefore, the figure suggests that reducing the bandwidths excessively comes at the expense of excluding randomized observations.

In the Croatian case, to adhere to RDD assumptions, one needs to use smaller bandwidths, which, however, leads one to ignore much of the quasi-random assignments available in matching market data. Hence, the application of the RDD design to matching market

data faces fundamental obstacles. As described in the previous section, by considering both the admission score and the propensity score, the PSDD, in particular the global PSDD, extracts the ex-ante uncertainty contained in the market-level structure of applications, and instead of choosing an arbitrary bandwidth, focuses on all the applications of students whose assignment is quasi-random, similarly to the lottery-based estimators.

1.5 Conclusion

This paper provides a new empirical perspective on the nature of assignment uncertainty in centralized matching markets by distinguishing between ex-post randomization reflecting uncertainty after submitting an application, and ex-ante randomization capturing uncertainty at the time of the application. In a typical student-school matching market setting, students submit their applications after learning their admission scores. Therefore, at the time of submitting the application, students are aware of their admission score, and thus the ex-ante uncertainty of admissions is contained in the school-specific score cutoff uncertainty, which in turn, is determined by the admission scores and submitted applications of all the market participants. Using this insight, I propose a resampling procedure, which generates the uncertainty of cutoffs by redrawing with replacement from the applicant population and recording the simulated matching market outcomes, to calculate the propensity score for each student-school pair.

I use data from the Croatian DA matching market to compare the distributions of admission propensity scores for treated and non-treated applicants within RDD bandwidths typically used in the literature. I find striking differences that are not in line with the randomized assignment assumption employed in RDD studies, which is particularly important in multi-cutoff settings where cutoffs are endogenously determined. This suggests a drastic reduction of RDD bandwidths and sample sizes. However, the data also implies that the sizeable fraction of quasi-randomized assignments occurs outside of the typical RDD bandwidths. This introduces a trade-off into the RDD implementation. To comply with the RDD assumptions, smaller bandwidths need to be employed. However, small bandwidths ignore a considerable portion of quasi-random variation available in matching market data.

As an alternative approach, I propose a new estimator, the propensity score discontinuity design (PSDD), which applies the Rosenbaum and Rubin (1983) propensity score theorem to the matching market setting. Instead of running the regression using an (arbitrary) bandwidth, identification in PSDD is based on propensity score matching. Therefore, the PSDD focuses exclusively on applications with (quasi-) random treatment assignment, regardless of the proximity to cutoff, extracting the whole ex-ante uncertainty contained in the matching market. Furthermore, while the original propensity score theorem utilizes a strong selection-on-observables assumption, the PSDD elegantly avoids this by acknowledging that any potential selection into treatment must be embedded in the student's submitted preferences and the admission score.

A natural direction for future research is to replicate the results of this paper on a dataset with an outcome variable of interest, such as the Norway college choice dataset where Kirkeboen et al. (2016) estimate labor market returns on different fields of study, and assess the sensitivity of treatment effects with respect to propensity scores. Typical

studies run a series of robustness checks, attempting to replicate the main estimates often using bandwidths that are not sufficiently narrow to exclude all the applications with non-randomized assignments. Therefore, the PSDD estimates might look different from the RDD estimates, even if the RDD estimates look stable across robustness checks typically employed.

Other than being based on ex-ante randomness in assignments, the PSDD offers two advantages over the standard RDD. First, as argued in Kirkeboen et al. (2016), the RDD identifies attendance effects relative to a counterfactual school at the cutoff. To approximate the choice margin, they control for the next most preferred school. In the possible case the next most preferred school is not the correct counterfactual, because this can bias the estimates. As an example, consider a student who was marginally declined in school A, with school B as his next preference. If the student has no chance of being admitted to it, school B bears no information about his counterfactual. Since the propensity scores generate the distribution of all the possible counterfactuals, future research can also concentrate on inspecting and resolving this potential bias. Second, unlike the RDD, which identifies treatment effects only around the cutoffs, the global PSDD specification incorporates all applications with a probabilistic assignment, including those that are potentially well above the cutoff. Future research can use this feature of the PSDD for quasi-experimental identification of away-from-the-cutoff treatment effects.[21] Furthermore, since the PSDD is not tied to the (existence of) cutoffs, it could be used for identification of school-specific treatment effects even in undersubscribed schools that do not have a cutoff.

[21] Angrist and Rokkane (2016) present a method for estimating treatment effects away from the cutoffs in the college choice setting. However, their technique is based on the strong assumption that, conditional on the available covariates, the running variable is ignorable.

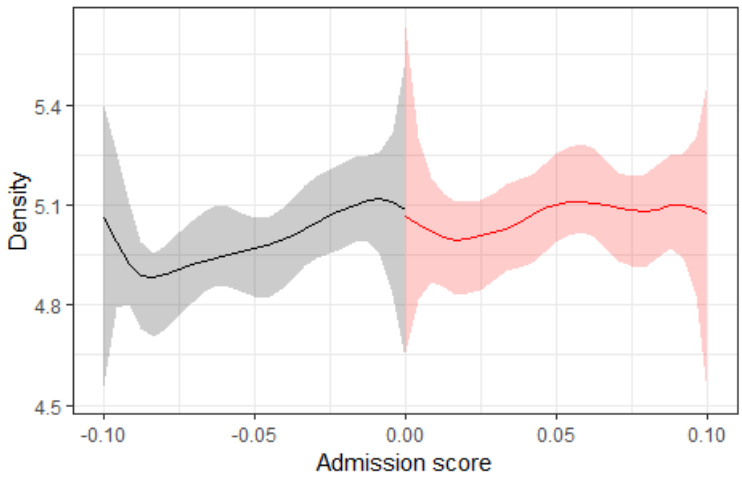
1.6 Appendix

Table 1.1: Summary statistics

	Year 2014 (1)	Year 2016 (2)	Year 2017 (3)	Year 2018 (4)
Number of programs	616	620	620	614
Number of colleges	49	49	49	49
Number of applicants	34,305	34,518	36,466	33,503
Avg. admission score	632.22	648.95	624.47	636.02
	(122.44)	(118.66)	(117.83)	(120.60)
Avg. length of choice list	6.40	6.23	5.58	5.13
	(3.53)	(3.45)	(3.23)	(3.01)
Avg. admission rate	0.15	0.15	0.16	0.18

Notes: The first panel shows the number of programs, colleges and students for each year. The second panel shows the average admission score calculated over all applications in a particular year. The third panel shows the average length of the submitted choice list in a particular year. The final panel shows the average admission rate calculated as the ratio between the number of all applications over the number of all admissions in a particular year. The values in the brackets are standard deviations.

Figure 1.3: Density of the standardized admission score



The figure implements the manipulation test around the cutoffs employing the local polynomial density estimation method as in Cattaneo et al. (2019). The figure plots a kernel density of standardized admission score centered around school-specific cutoff values using all applications in the data.

Chapter 2

LATE Estimators under Costly Non-compliance in Student-College Matching Markets

Co-authored with Štepan Jurajda (CERGE-EI)

2.1 Introduction

The instrumental variable (IV) estimator is widely used to account for unmeasured confounding factors and to identify causal effects (Angrist and Krueger, 1991). It is predominantly implemented in the form of the 2SLS estimator, which, under certain assumptions, identifies the local average treatment effect (LATE) for individuals whose treatment is manipulated by (quasi-) random instrumental variation—the so-called *compliers*. In this paper, we consider the properties of the 2SLS estimator in a setup where non-compliance with quasi-random treatment assignment is costly, which violates the *exclusion restriction*, one of the crucial assumptions necessary for the causal interpretation of the 2SLS estimator. We build on the LATE theorem (Imbens and Angrist, 1994) to show that in the case of costly non-compliance, the IV estimator can be interpreted as LATE only after assuming that both the costs and the probabilities of non-compliance do not depend on the instrument's value. Intuitively, if the costs depend on the instrument's value, the instrument affects non-compliers through the costs of non-compliance, and becomes correlated with the outcome not only for compliers, but also for non-compliers, which biases the 2SLS estimator.

We apply this insight to the growing literature exploiting a feature of centralized college admission systems where students with similar admission scores in a neighbourhood of a school's admission threshold are or are not offered admission based on small differences in admission scores. Assuming that the students at the margin of admission differ only in their treatment assignment, this literature relies on an indicator of whether a student is above the school-specific admission threshold (admission score cutoff) to instrument for graduation or admission. The LATE theorem is then invoked to interpret these IV estimates (e.g., Kirkeboen et al. (2016)).

A basic feature of centralized college admission systems (as operated, e.g., in Chile, Croatia, Norway, and Sweden) is that a student who intends to not comply with his school assignment can choose to drop out of the system or can accept the initial admission offer, but apply to and enrol in his preferred school in the following year(s). The former happens rarely as it typically means not enrolling in any college in a given year; the latter happens frequently and it delays graduation and labor market entry by at least a year. Hence, in centralized matching markets of this type, non-compliance costs arise naturally, at least for always takers, i.e., those ultimately enrolling in a given school regardless of the initial application outcome.

Our analysis implies that when the admission offer is used to instrument for graduation (as in, e.g., Kirkeboen et al., 2016), these non-compliance costs originate before treatment status (graduation) is resolved, and therefore bias the LATE estimator. A plausible strategy to solve the problem of non-compliance is to change the treatment of interest. In the context of school-program evaluation, instead of estimating the effect of graduation, this would correspond to estimating the effect of admission into the first year of a given program. This strategy may trade off gains in terms of identification credibility with economic relevance of the treatment effect. When the admission offer is used to instrument for admission (as in, e.g., Altmejd et al., 2019a), non-compliance costs originate only after initial-application admission treatment status is determined. In this case, there is no bias since the instrument-treatment mapping occurs before costs are realized, and is

thus unaffected by the non-compliance costs. Nevertheless, when studies in the literature interpret admission as extended attendance, the interpretation of the treatment effect is similarly impaired as in the case of graduation effects.

As a prime example of this literature, consider Kirkeboen et al. (2016), who estimate the returns to graduating in different fields of education in Norway by instrumenting for graduation with the initial quasi-random admission offer, and by measuring labor market returns eight years after the initial application. Enrolling in a program other than the one initially assigned a year or more after the initial application (we refer to such situation as 're-enrolling') results in deferred graduation and thus reduces labor market experience as labor market returns are measured eight years after the initial application regardless of the actual graduation date. This in turn implies costly non-compliance.[1] Therefore, according to the results provided here, the estimates in Kirkeboen et al. (2016) can be interpreted as returns to fields of study only if the costs of foregoing labor market experience are not field-specific and if the probability of non-compliance with the initial assignment does not depend on the initial assignment. Using data from the centralized college-student matching market in Croatia spanning the period from 2012 to 2018, we show this is not the case by documenting that the probabilities of non-compliance do depend on the initial assignment.

Using the Croatian data, we consider the same instrument as Kirkeboen et al. (2016) and document a sizeable re-application rate.[2] Importantly, the rate of applying to programs other than the one initially assigned within two years of the initial application (referred to as re-applying) for those just below the treatment program's admission score cutoff is 18.3% compared with 12% for those just above the treatment program's cutoff.[3] This discontinuity in the re-application rate at the cutoff translates into discontinuity in the non-compliance (re-enrollment) rate at the cutoff: there are 14.6% of non-compliers just below the cutoff, compared with 10.1% just above the cutoff. The higher share of non-compliers just below the cutoff compared with non-compliers just above the cutoff breaks the exclusion restriction, as the instrument now affects the outcome through channels other than the treatment assignment since it also affects the non-compliers due to non-compliance costs.

To deal with this issue, we propose a method inspired by Lee (2009), which recovers the treatment effect bounds under the *homogenous non-compliance costs* assumption, i.e., when all non-compliers pay the same cost. The method consists of two steps. The first involves trimming the data (excluding observations) until the non-compliance rates for those assigned and those not assigned to the treatment program are the same. For example, suppose that the fraction of always takers (those ultimately receiving the treatment

[1] Non-compliance implies net costs if the negative effect of lower labor market experience outweighs the potential benefits of temporary enrolment in a non-preferred program. Providing evidence on this issue is beyond the scope of our analysis; for the purpose of our analysis, we assume the benefits are small.

[2] In order to re-enroll in the year(s) after the initial college application, a student needs to re-apply. Therefore, we analyze the re-application rate (intent to non-comply) and the re-enrollment rate (non-compliance) separately.

[3] Following Dustan (2018), Fernandez (2019) and Kirkeboen et al. (2016), we define the *treatment program* for a particular applicant as the program for which he was close to the admission score cutoff, i.e., either just above the cutoff or just below the cutoff.

regardless of the instrument's value from the initial application year) is larger than the fraction of never takers (those never getting the treatment) so that the fraction of non-compliers who were assigned not to get the treatment is disproportionally large. The first step of the proposed method balances the fraction of always takers assigned away from the treatment and the fraction of never takers assigned to the treatment by dropping a fraction of always takers assigned away from the treatment. Due to the homogenous non-compliance costs assumption, the effect of the non-compliance costs of the remaining always takers is then offset by (the same amount of) non-compliance costs of never takers.

However, excluding the non-compliers based on their treatment assignment and treatment indicator induces selection bias (as selection into non-complying can generally be non-random). By selectively excluding only the always takers who are not assigned to the treatment, the instrument now gains predictive power over outcomes of non-compliers—the probability of an individual being an always taker becomes higher for those assigned to the treatment in the trimmed sample, compared to the original sample, and always takers may have different outcomes than never takers. Therefore, the second step of the proposed procedure accounts for the sample selection by adapting the Lee (2009) treatment effect bounds, additionally trimming individuals in order to ensure that the instrument does not predict the outcome for non-compliers; in our case, this involves trimming individuals who were assigned to the treatment. The final ingredient of the method is to select individuals for trimming in each stage from the upper/lower tails of the outcome distribution in order to ensure the most conservative treatment effect bounds.

The homogenous costs assumption is plausible in the school choice setting since non-compliance costs here originate in large part in reduced labor market exerience due to re-enrolling in another program a year or more after the initial quasi-random assignment, and therefore postponing graduation. [4] Whether these costs are homogenous can be tested empirically by asking whether the slopes of experience wage profiles of always takers and never takers who did not comply with the treatment assignment are similar.

This paper contributes to several strands of the literature. First, it is relevant to the literature employing 2SLS-type estimators in centralized school-student matching markets, in which non-compliance costs arise naturally. Using 2SLS near admission cutoffs or the closely related regression discontinuity design (RDD) estimators, Kirkeboen et al. (2016) analyze school-specific labor-market returns, Lucas and Mbiti (2014) and Abdulkadiroglu et al. (2014) study school-specific attendance achievement effects (measured through standardized test scores), Kaufmann et al. (2013) study marriage market returns, while Dustan (2018), Fernandez (2019) and Altmejd et al. (2019a) analyze the role of family ties in school choice. These applications are potentially affected by the non-compliance cost issue.

More generally, this approach can be applied in other empirical settings. For example, when programs are offered through a randomized list and applicants can apply to several lotteries (de Chaisemartin and Behaghel (2020)), or in college applications without

[4] Postponing graduation could also produce certain gains (i.e maturation effect), or different types of costs (i.e. the (cognitive) costs of preparing and re-taking the state exam). In this paper, we interpret the *net costs* after "aggregating" all gains and costs, thus abstracting away from potential cost breakdowns.

matching markets (see e.g. Zimmerman, 2014, Goodman et al. (2017), Goodman et al. (2020) and Kozakowski (2020)).

Second, it adds to the literature on exclusion restriction violation. Heckman (1997) establishes that any selection into treatment based on individual-specific unobserved characteristics breaks the exclusion restriction and results in economically un-interesting parameters. Similarly, Jones (2015) identifies economically plausible potential violations of exclusion restriction for infra-marginal individuals (always takers and never takers) in cases where treatment may change their outcomes, which loosely fits our framework. However, Jones (2015) only constructs isolated theoretical examples, in which the exclusion restriction is likely violated, without presenting empirical content or developing a solution, while we develop a general non-compliance setup and tie it directly to a large literature. We also provide an alternative estimator, which addresses the underlying issue. Moreover, in our setup the cost is generated endogenously to the IV model - by the decision of agents to not comply - and not by external spillovers of treatment assignment as in Jones (2015).

The remainder of this paper is structured as follows. In the next section, we develop the procedure for bounding the treatment effect in the case of costly non-compliance. In the third section, we demonstrate that the Croatian college-student matching market is subject to differing probabilities of complying depending on the college assignment. The fourth section concludes.

2.2 Treatment Effect Bounds

In this section, we develop a general framework that supports the assumption of costly non-compliance and analyzes the behavior of the LATE estimator. We use the typical LATE notation and introduce an additional parameter γ, which denotes non-compliance costs. As a result, we produce a practical framework that can be straightforwardly used in typical LATE applications.

Our illustrative empirical school-choice analysis presented in the next section is based on a dynamic setup where the costs of non-compliance are embodied in the time needed to alter the treatment assignment by re-enrolling at another school. A more complicated, structural model could attempt to elicit the gains (i.e., maturation effects) and losses (i.e., foregone labor-market experiences) from this non-compliance process. Our model collapses the net non-compliance costs into the parameter γ, and applies the newly developed LATE framework. Such an approach allows one to divide the analysis into two steps. First, to analyze the components of the non-compliance costs embodied in the parameter γ, and second, to analyze the LATE conditional on a specific value of the non-compliance costs γ.

We show that in presence of non-compliance costs, the exclusion restriction is likely violated, thus biasing the LATE estimator. We address this issue by developing a treatment-bounds method inspired by Lee (2009), and discuss the assumptions needed to recover treatment effect bounds.

Suppose we are interested in the causal effect of treatment D_i on the outcome y_i. Denote with Y_{1i} (Y_{0i}) potential outcomes of individual i when $D_i = 1$ ($D_i = 0$). An instrument $Z_i = \{0, 1\}$ (treatment assignment) is assumed to shift the treatment indicator D_i. In particular, denote with D_{1i} (D_{0i}) the treatment indicator of individual i when $Z_i = 1$ ($Z_i = 0$). The outcome of interest y_i is now indexed against two variables, the value of the treatment indicator D_i and the value of the instrument Z_i as $y_i = Y_i(D_i, Z_i)$.

Define an indicator t_i describing an individual i's type as:

$$t_i = \begin{cases} N \text{ if } D_{1i} = 0 \text{ and } D_{0i}{=}0 \text{ (\textit{Never taker})}, \\ A \text{ if } D_{1i} = 1 \text{ and } D_{0i}{=}1 \text{ (\textit{Always taker})}, \\ C \text{ if } D_{1i} = 1 \text{ and } D_{0i}{=}0 \text{ (\textit{Complier})}, \\ D \text{ if } D_{1i} = 0 \text{ and } D_{0i}{=}1 \text{ (\textit{Defier})}, \end{cases}$$

and denote with $P(t_i = x)$ the probability that individual i's type is x. The LATE theorem of Imbens and Angrist (1994) is widely used to identify local average treatment effects in (quasi-) experimental studies:

Theorem 1. *Assume the following LATE assumptions:*

- *Independence - The instrument is independent:*

$$\{Y_i(D_{1i}, 1), Y_i(D_{0i}, 0), D_i(1), D_i(0)\} \perp Z_i$$

- *Exclusion restriction - The instrument affects the outcome only through the treatment indicator:*

$$Y_i(d,0) = Y_i(d,1) \equiv Y_{di} \text{ for } d = 0,1$$

- *First stage - The instrument has predictive power over assignment:*

$$E[D_{1i} - D_{0i}] \neq 0$$

- *Monotonicity - There are no defiers:*

$$D_{1i} - D_0 \geq 0 \text{ or vice versa}, \forall i$$

Then, the Wald estimator equals the average treatment effect on the treated:

$$\frac{E[y_i|Z=1] - E[y_i|Z=0]}{E[D_i|Z=1] - E[D_i|Z=0]} = E[Y_{1i} - Y_{0i}|D_{1i} - D_{0i} > 0]$$

Proof. See Imbens and Angrist (1994). □

Under the LATE assumptions, the Wald estimator equals the average treatment effect for compliers (individuals with $t_i = C$). Intuitively, non-compliers, i.e., always takers (those with $t_i = A$) and never takers (those with $t_i = N$), do not contribute to the IV estimator for two reasons. First, this is due to the exclusion restriction as the instrument does not change their treatment assignment. Second, this is due to the *independence* assumption, as the instrument is independent from their treatment decisions D_i. Therefore, the instrument has no predictive power over the outcomes of non-compliers.

In contrast, if non-compliance with the quasi-random treatment assignment is costly, non-compliers generally do contribute to the IV estimator of LATE. For example, if always takers with $Z = 0$ have to pay a cost to get treatment, they are no longer the same as the always takers with $Z = 1$ (the exclusion restriction does not hold). Generally, this implies predictive power of the instrument over the outcome for the non-compliers, which violates the assumptions of the LATE theorem.

Proposition 1. *Assume that Independence, First stage and Monotonicity assumptions from Theorem 1 hold and assume heterogenous non-compliance costs accross t:*

$$E[Y_i(1,1) - Y_i(1,0)] = \gamma_A \text{ and } E[Y_i(0,1) - Y_i(0,0)] = \gamma_N, \gamma_A \neq \gamma_N.$$

Let $\bar{\gamma} = \frac{\gamma_A + \gamma_N}{2}$. *The Wald estimator now equals:*

$$\begin{aligned}\frac{E[y_i|Z=1] - E[y_i|Z=0]}{E[D_i|Z=1] - E[D_i|Z=0]} =& E[Y_{1i} - Y_{0i}|D_{1i} - D_{0i} > 0] \\ &+ \frac{\bar{\gamma} \cdot (P(t_i = A) - P(t_i = N))}{P(t_i = C)} \\ &+ \frac{\frac{P(t_i=A)+P(t_i=N)}{2} \cdot (\gamma_A - \gamma_N)}{P(t_i = C)}.\end{aligned}$$

Proof. Applying the Independence and Monotonicity assumption to the first term of the Wald estimator we obtain:

$$\begin{aligned} E[y_i|Z=1] = E[Y_{1i}|D_{1i}=1, D_{0i}=1] \cdot \overbrace{P[D_{1i}=1, D_{0i}=1]}^{\text{Always takers, } t_i=\text{A}} \\ + E[Y_{1i}|D_{1i}=1, D_{0i}=0] \cdot \overbrace{P[D_{1i}=1, D_{0i}=0]}^{\text{Compliers, } t_i=\text{C}} \\ + E[Y_{0i}|D_{1i}=0, D_{0i}=0] \cdot \overbrace{P[D_{1i}=0, D_{0i}=0]}^{\text{Never takers, } t_i=\text{N}}. \end{aligned}$$

After performing an analogous decomposition of $E[y_i|Z=0]$, and using the *Heterogenous non-compliance costs* assumption, the numerator of the Wald estimator, after some algebra, becomes:

$$\begin{aligned} E[Y_{1i} - Y_{0i}|D_{1i} - D_{0i} > 0] + \bar{\gamma} \cdot (P(t_i = A) - P(t_i = N)) \\ + \frac{P(t_i = A) - P(t_i = N)}{2} \cdot (\gamma_A - \gamma_N). \end{aligned}$$

A similar argument shows that

$$E[D_i|Z=1] - E[D_i|Z=0] = E[D_{1i} - D_{0i}] = P[D_{1i}=1, D_{0i}=1]) = P[t_i = C].$$

□

Proposition 1 says that under costly non-compliance with the (quasi-) random treatment assignment, the Wald estimator equals the average treatment effect for compliers if the costs as well as the probabilities of non-compliance are the same for always takers and never takers (i.e., if $\gamma_A = \gamma_N$ and $P(t_i = A) = P(t_i = N)$).

In the remainder of this section we propose sharp bounds of the LATE[5] for the simple homogenous non-compliance costs case (i.e, $\gamma_A = \gamma_N$).[6] In the next section, we apply the bounding procedure to the Croatian centralized student-school matching system, arguing that in these types of settings assuming homogenous non-compliance costs may be reasonable.

At an intuitive level, the proposed bounding method mechanically equates the probabilities $P(t_i = A)$ and $P(t_i = N)$ by excluding individuals leading to the highest upper (lowest lower) bound. Suppose, WLOG, that $P(t_i = A) > P(t_i = N)$ - there are more always takers than never takers. Therefore, to calculate the upper LATE bound, we trim a proportion of always takers (individuals with $D = 1$ and $Z = 0$) until $P(t_i = A) = P(t_i = C)$, starting with those with the highest Y values (to obtain the highest possible value of the Wald estimator). This solves the problem of differing probabilities of non-compliance for

[5]The bounds are sharp in the Lee (2009) sense that they are the largest (smallest) lower (upper) LATE bounds consistent with the data.

[6]The homogenous costs assumption in the school choice setting is testable with data on labor market outcomes since the costs originate in large part in the reduced labor market experience due to re-enrolling in another program. One can test the equality of slopes of the experience profiles of always takers and never takers who did not comply with the initial treatment assignment by comparing their realized experience curves.

always takers and never takers, but it also introduces a selection problem by selectively excluding always takers with $Z = 0$ values. Intuitively, in the new sample, individuals with $Z = 0$ are less likely to be always takers than individuals with $Z = 1$. Therefore, in addition to predicting treatment, the instrument now predicts the non-compliance types, and potentially also the outcome (if the selection into non-compliance is non-random), which breaks the exclusion restriction.

To account for this, we aim to drop the same number of always takers who were assigned to treatment (i.e., $Z = 1$). The problem is that among the individuals with $Z = 1$, we cannot distinguish compliers from always takers — both of them accept the treatment assignment. However, by trimming individuals with the lowest Y values (of those with $Z = 1$), we generate the upper LATE bound. This result is formalized in the following adaptation of Proposition 1 from Lee (2009).

Proposition 2. *Let Y be a continuous random variable. Assume that Independence, First stage and Monotonicity assumptions from Theorem 1 hold and assume Homogenous non-compliance costs:*

$$E[Y_i(1,1) - Y_i(1,0)] = \gamma = E[Y_i(0,1) - Y_i(0,0)]$$

Assume, WLOG, that $P(t_i = A) > P(t_i = N)$ and introduce $R = \frac{P(t_i=A)-P(t_i=N)}{P(t_i=A)}$. Next, set $y_{q|E} = G^{-1}(q)$, where G is the cdf of Y conditional on an event E, which defines the value of treatment D_i and instrument Z_i. Under these assumptions, Δ_{LB} and Δ_{UB} are sharp lower and upper bounds for the average LATE effect $E[Y_{1i} - Y_{0i}|D_{1i} - D_{0i} > 0]$:

$$\Delta_{LB} = \frac{E\left[Y|Z=1, Y \leq y_{1-R\cdot p(t_i=A)|Z=1}\right] - E\left[Y|(Z=0,D=0)\cup(Z=0,D=1,Y\geq y_{R|(Z=1,D=0)})\right]}{P_L(T_i = C)},$$

$$\Delta_{UB} = \frac{E\left[Y|Z=1, Y \geq y_{R\cdot p(t_i=A)|Z=1}\right] - E\left[Y|(Z=0,D=0)\cup(Z=1,D=0,Y\leq y_{1-R|(Z=1,D=0)})\right]}{P_U(T_i = C)}$$

and P_L (P_U) is a probability measure evaluated on the trimmed sample used when calculating Δ_{LB} (Δ_{UB}). The bounds are sharp in the sense that Δ_{LB} (Δ_{UB}) is the largest (smallest) lower (upper) bound that is consistent with the observed data.

Proof. First, draw a random proportion R of individuals with $Z = 0$ and $D = 1$ and assign them values $S_{0ir} = 0$, where r indexes the random seed generating this variable. Assign the remaining individuals with values $S_{0ir} = 1$. To simplify notation, assume that the variable $S_{1ir} = 1$ for each individual i and introduce:

$$\begin{aligned} S_{ir} &= S_{1ir}Z + S_{01r}(1 - Z) \\ Y_i^* &= S_{ir} \cdot \{Y_{1i}Z + Y_{0i}(1 - Z)\} \end{aligned} \tag{2.1}$$

Next, assume that the variable Y_i^* is only observed when $S_{ir} = 1$ and is, in that case, equal to Y_i. In other words, model 2.1 treats S_{ir} as a sample selection indicator. Denote with Y_{1i}^* (Y_{0i}^*) the outcome of the individual i when $Z_i = 1$ ($Z_i = 0$). According to the Lee (2009) theorem, the sharp lower ($\Delta_{LB,r}$) and upper ($\Delta_{UB,r}$) bounds for the intention to treat estimator ($E[Y_i|Z = 1, S_{1i} = 1, S_{0i} = 1] - E[Y_i|Z = 0, S_{1i} = 1, S_{0i} = 1]$) are:

$$\Delta_{LB,r} = E\left[Y|Z=1, S=1, Y^* \leq y^*_{1-R\cdot p(t_i=A)}\right] - E\left[Y|Z=0, S=1\right],$$

$$\Delta_{UB,r} = E\left[Y|Z=1, S=1, Y^* \geq y^*_{R\cdot p(t_i=A)}\right] - E\left[Y|Z=0, S=1\right].$$

We index the bounds with r to emphasize the dependence on the seed corresponding to the random draw of R individuals.

Note that on the sample of individuals with $S_{ir} = 1$, $P(t_i = A) = P(t_i = N)$. Therefore, according to Proposition 1:

$$\frac{\Delta_{LB,r}}{P_L(T_i = C)} \leq E[Y_{1i} - Y_{0i} | D_{1i} - D_{0i} > 0, S_{1ir} = 1, S_{0ir} = 1] \leq \frac{\Delta_{UB,r}}{P_U(T_i = C)}$$

$$\Longrightarrow$$

$$\min_r \left(\frac{\Delta_{LB,r}}{P_L(T_i = C)} \right) \leq E[Y_{1i} - Y_{0i} | D_{1i} - D_{0i} > 0] \leq \max_r \left(\frac{\Delta_{UB,r}}{P_U(T_i = C)} \right),$$

where P_L (P_U) is the probability evaluated on the trimmed sample used when calculating Δ_{LB} (Δ_{UB}). Finally, note that the treatment bounds depend on the random draw R only through the outcome values of randomly sampled individuals with $Z = 1$ and $D = 0$ (i.e., they do not the depend on the randomly sampled subset of those with $D = 1$). The proposition now follows from observing that, for example, the lowest $\Delta_{LB,r}$ is achieved when trimming those individuals with $Z = 1$ and $D = 0$ who have the highest y values. □

To demonstrate the value of Proposition 2, we compare the 2SLS estimator to the proposed LATE bounds on a simulated dataset. We generate N individuals according to the following steps:[7]

- The type of individual i is drawn from the following distribution:

$$t_i = \begin{cases} A \text{ with probability } p_a, \\ N \text{ with probability } p_n, \\ C \text{ with probability } 1 - p_a - p_n. \end{cases} \tag{2.2}$$

- The treatment assignment Z_i is a Bernoulli random variable with parameter 0.5.
- The outcome of interest y_i is defined as:

$$y_i = \epsilon - \gamma \cdot 1_{Z_i \neq D_i},$$

  where ϵ is $N(0, 1)$.

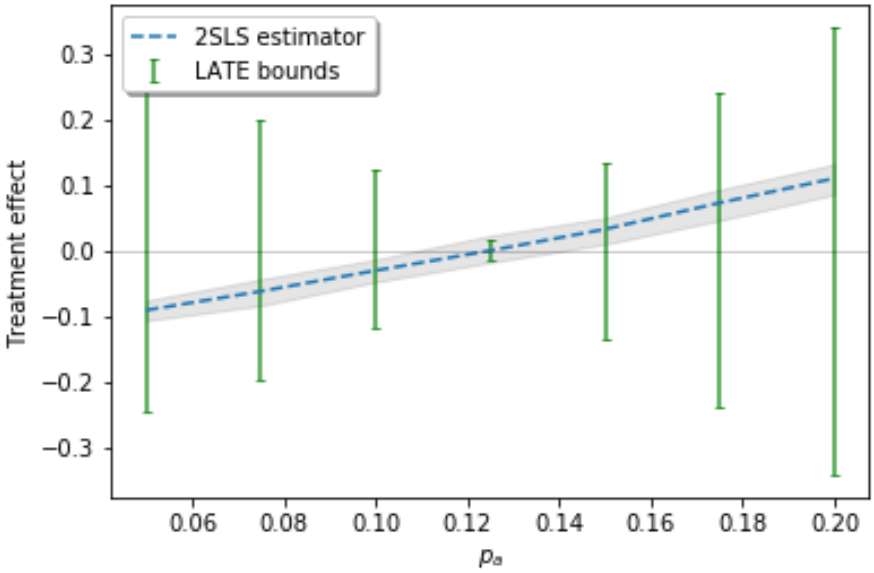
The procedure generates a population with no treatment effect (i.e., the treatment effect is zero) where assignment to treatment is equiprobable for each individual. Individuals differ only with respect to their type, which defines their attitude towards treatment assignment (i.e., $t_i = A$ individuals always get treated regardless of the assignment status, $t_i = N$ never get treated and $t_i = C$ comply with the assignment).

We conduct two exercises: First asking about the performance of the 2SLS estimator and of the treatment-effect bounds under fixed costs of non-compliance and varying gaps between p_a and p_n, second allowing the cost of non-compliance to vary but keeping non-compliance probabilities fixed. Specifically, in the first exercise we set the non-compliance

[7] N is set at 50,000 to resemble Kirkeboen et al. (2016), where $N = 50,083$.

cost to equal the outcome standard deviation ($\gamma = 1$). Next, we set p_n at 0.12[8] and for each value $p_a \in \{0.05, 0.075, \ldots, 0.175, 0.2\}$, we generate 1,000 independent populations and plot the average LATE bounds and the average of the 2SLS estimates in Figure 1. Even though the treatment effect is 0 by construction, the 2SLS estimator reflects the asymptotic bias $\frac{\gamma \cdot (p_a - p_n)}{1 - p_a - p_n}$ and would lead one to reject the zero treatment effect even for small differences between p_a and p_n, while the LATE bounds correctly include 0 and remain smaller than one half of the treatment standard deviation even for large differences between p_a and p_n.

Figure 2.1: LATE bounds vs. 2SLS estimates - varying non-compliance probabilities



Note: The figure plots 2SLS estimates and LATE bounds (y-axis) against the probability (of being an always taker) p_a, holding the probability (of being a never taker) p_n fixed at 0.125 on a series of simulated datasets. For each parameter value $p_a \in \{0.05, 0.075, \ldots, 0.175, 0.2\}$ we generate 1,000 independent populations using parameters $\gamma = 1, N = 50000$, $p_n = 0.125$ under no treatment effect (LATE= 0), and plot the average LATE bounds from Proposition 2 and the average 2SLS estimates as well as the corresponding average 95% confidence intervals.

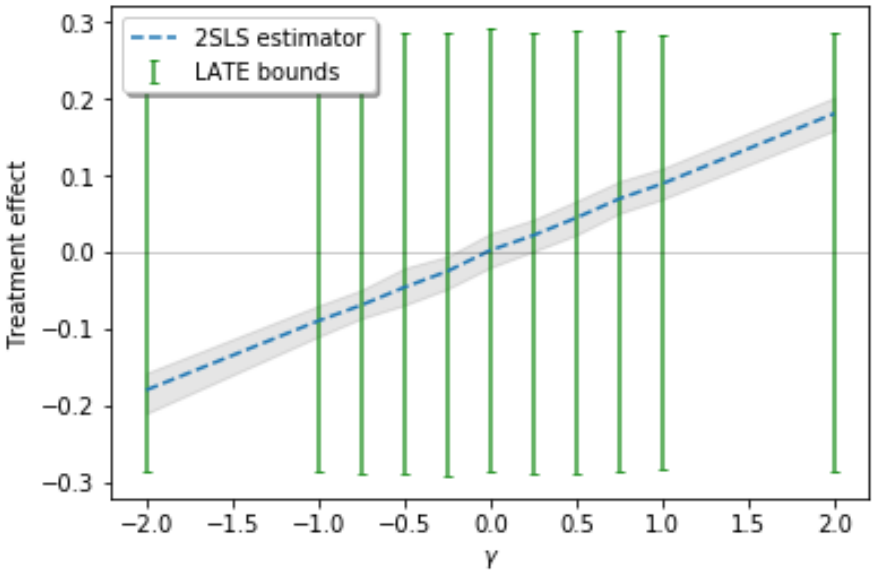
In the second exercise presented in Figure 2.2, we vary non-compliance costs γ while holding p_a fixed at 18.3% and p_n at 12%.[9] Again, the 2SLS estimator coincides with its asymptotic bias and reports significant estimates even for reasonably small values of γ, while the LATE bounds correctly include 0.[10]

[8] This probability correspond to an empirical estimate obtained in section 2.3.

[9] Again, these probabilities correspond to empirical estimates obtained in section 2.3.

[10] The LATE bounds do not depend on the γ value, since they neutralize the effect of the non-compliance cost by trimming enough individuals so that the costs of always takers and never takers cancel.

Figure 2.2: LATE bounds vs. 2SLS estimates - varying γ



Note: The figure plots 2SLS estimates and LATE bounds (y-axis) against the costs of non-compliance γ, while holding fixed the probability (of being a never taker) $p_n = 0.12$ as well as the probability (of being an always taker) $p_a = 0.183$ on a series of simulated datasets. For each parameter value $\gamma \in \{-2, -1, 5, \ldots, 2\}$ we generate 1,000 independent populations using parameters $p_a = 0.183, p_n = 0.12, N = 50000$ under no treatment effect (LATE$= 0$), and plot the average LATE bounds from Proposition 2 and the average 2SLS estimates as well as the corresponding average 95% confidence intervals.

2.3 Empirical Application to Croatian College Matching Market

In a recent study, Kirkeboen et al. (2016) use RDD to instrument for graduation and estimate returns to graduating in different fields of education in Norway by instrumenting for the graduation with the initial admission offer and measuring labor market returns eight years after the initial college application. In such a setup, re-enrolling in another field, potentially years after the initial application, results in deferred graduation and reduces labor market experience. In the likely case that the length of labor market experience affects labor market returns, Proposition 1 implies that Kirkeboen et al. (2016) identify unbiased returns to fields only in the homogenous non-compliance costs case and if the probabilities of non-compliance do not depend on the initial treatment-program assignment. In this section, we show that the latter is not the case in Croatia, where the probabilities of non-compliance do depend on the initial assignment. Therefore, the LATE bounds from Proposition 2 should be applied when estimating LATE effects in the Croatian matching market. A similar issue arises naturally in studies that rely on quasi-random admission offers to instrument for graduation or other outcomes occuring years after the initial offer of admission (e.g. Hastings et al., 2014; Lucas and Mbiti, 2014; Abdulkadiroglu et al., 2014; Kaufmann et al., 2013; Dustan, 2018; Fernandez, 2019; Altmejd et al., 2019a).

We begin the section with a brief summary of the estimation strategy employed in Kirkeboen et al. (2016) and similar student-school assignment studies. We proceed with a note on the institutional setup in Croatia, followed by a subsection rejecting equal probabilities of non-compliance for students who were or were not (quasi-randomly) offered admission to their treatment program (i.e., the program where they were just below or just above the program-specific admission cutoff). We conclude the section with a discussion of the homogeneous non-compliance costs assumption.

2.3.1 Empirical Strategy

The literature studying school-student centralized matching markets frequently exploits a feature of these systems in which students with similar admission scores in a neighborhood of a school's admission threshold are or are not offered admission to the schools based on small differences in admission scores. Taking advantage of these discontinuities, the literature typically uses regression discontinuity design (RDD) to instrument for admission/graduation , assuming that students around the cutoff are 'the same' in every aspect except the assigned school (program). The assigned school is assumed to be deterministically linked to the school-specific admission score (i.e., a student is offered admission if and only if his admission score is above the school-specific admission score cutoff). For schools ranked below the assigned school, this deterministic link between admission score and the assignment is broken — the student is never considered for admission even if he is above the cutoff for these schools. For this reason, applications to these schools are not included in the *RDD estimation sample*, which consists of applications at the margin of admission, i.e., within a bandwidth neighbourhood of school-specific admission score cutoffs.

More formaly, let c_{jt} be the admission score cutoff of program j in year t. If program

j is ranked above program j' in student i's preference list, we write $(j) \succ_i (j')$. Denote the school-j-specific application score of individual i as a_{ijt}. Student i's application to program j belongs to the RDD estimation sample if student i:

1. listed program j as his choice, such that all programs preferred to j had a higher cutoff score than c_{jt} (otherwise assignment to j is impossible):
   $c_{jt} < c_{j't} \; \forall \; (j') \succ_i (j)$,
2. had a score a_{ijt} sufficiently close to j's cutoff score to be within a given bandwidth bw around the cutoff:
   $|a_{ijt} - c_{jt}| \leq bw$.

The following regression, applied to the RDD estimation sample, is a typical specification used in the school-choice literature to estimate various graduation effects:

$$y_{i\tau} = \beta \cdot graduated_{ijt} + f(a_{ijt}; \gamma) + \mu_\tau + \mu_{jt} + \varepsilon_{ijt} \tag{2.3}$$

where $y_{i\tau}$ is the outcome of interest measured at time $\tau > t$ of student i who was near the program j admission cutoff in year t, $graduated_{ijt}$ is an indicator variable that takes value 1 if student i graduated from program j where he initally applied in year t, $f(a_{ijt}; \gamma)$ is a function of the application score of student i for program j in year t, μ_{jt} and $\mu_\tau$are fixed effects corresponding to application year-program combinations and outcome years, respectively, and where ε_{ijt} is an error term. Since $graduated_{ijt}$ is likely influenced by various unobserved, potentially endogenous factors, researchers typically use admission offer ($1_{a_{ijt} \geq c_{jt}}$) to instrument for graduation. In the language of the previous section, being just above the cutoff corresponds to the instrument value $Z = 1$, and being just below the cutoff corresponds to the instrument value $Z = 0$.

2.3.2 Institutional Setup

In Croatia, admissions to all college programs are implemented through a national online platform. Since its introduction in 2010, this platform operates a deferred acceptance (DA) algorithm that ranks students based on their high-school grades and subject-specific elective national-level exams that take place in June, a month after high-school graduation. Students register on the platform in early spring of their high-school graduation year when universities also list on the platform their program-specific admission quotas along with program-specific weights of subject-specific grades and exams. Students are free to submit their ranked priority lists of up to 10 programs as of registration and update these preference rankings until the system closes for clearing at a predetermined date in mid-July (i.e., in 2019, the final deadline was 2 pm on July 15$^{\text{th}}$).

Students first receive information on their position in various admission queues one week before the final deadline, immediately after receiving their admission scores. Admission scores, which are a function of student's high school grades and national exam scores, are the only factor determining admission rankings. The DA algorithm is then regularly updated to show students their current admission rankings. Students update their preference rankings continuously until the system closes for clearing in mid-July.

During the application period applicants often drop their previously highly ranked alter-

natives they are unlikely to get admitted to.[11] Therefore, in order to study a case similar to the typical centralized college admission system, where students are not able to get signals on the current school-specific demand, we analyze admission outcomes implied by the first preference submissions after receiving the national exam scores (5 days before the system closes), when students are fully aware of their admission scores but are not yet able to learn about the market demand structure. We thus consider that a student applied to a particular program if this program was on the student's preference list five days before the admission deadline.

In centralized college admission systems, it is not feasible for always takers to not comply with their initial assignment out of their treatment program within the year of initial application. They can, however, apply to their preferred program in the following years. Further, in Croatia, there is only limited scope for never takers to not comply with their initial-application assignment to their treatment program.[12] Therefore, since we do not observe enrollment, we assume that the final admission offer translates to enrollment one-to-one. Hence, we abstract from non-compliance within the year of initial application and focus on non-compliance through re-applications in years following the initial college application.

In sum, we analyze applications (based on the ranking lists submitted 5 days before the system closes) which resemble the applications at typical centralized college admission systems, and enrollments (based on the final ranking lists) separately. We consider that a student re-applied (attempted non-compliance) if he applied to a program different from the one initially assigned in the two years following the initial application year. While we observe re-applications, we do not observe re-enrollment, so again, we assume that a re-applying student always re-enrolled in a particular program if this program was his final DA admission assignment.

2.3.3 Data and Results

We use complete administrative data corresponding to the Croatian centralized college admission system from years 2012-2018. In these data, we consider a student to be a non-complier if, following a re-application, he was assigned by the DA algorithm to a program different from the one initially assigned at most two years after his initial college enrollment. As the re-application window is two years, we exclude the boundary years of the data[13] and generate the RDD estimation sample using applications from 2014-2016 that are at most 0.4 standard deviations (60 admission score points) away from

[11] Due to the dynamic nature of the admission system, students can get hourly updates about their admission rankings, and therefore resolve a significant part of the admission uncertainty. They can do this only after they receive their admission scores, approximately 1 week before the admission deadline.

[12] According to the Ministry of Science and Education, 95% of Croatian college applicants comply with their DA admission assignment, enrolling at their assigned program. If they decide not to comply, they lose their tuition waiver, otherwise covered by the Ministry. This introduces an additional (constant) cost of non-compliance.

[13] We exclude the first two years to ensure that we work with only initial college applicants who have not applied in previous years. We exclude the last two years to observe re-applications following initial applications.

program-specific admission cutoffs.[14]

Table 2.1 shows basic summary statistics for the Croatian DA matching market and the RDD estimation sample defined by a 60 admission score points bandwidth, throughout 2014-2016. Annually, approximately 35,000 students enter the system, choosing between about 600 programs belonging to 49 distinct universities. The RDD estimation sample appears to have similar average characteristics to the unrestricted sample.

Using the RDD estimation sample, we estimate the following regression:

$$y_i = \alpha_0 \cdot a_{ij} + \alpha_1 \cdot a_{ij} 1_{a_{ij} \geq c_j} + \delta \cdot 1_{a_{ij} \geq c_j} + f(a_{ij}) + \mu_j + \varepsilon_i, \tag{2.4}$$

where y_i is a non-compliance indicator for applicant i (i.e., a dummy variable taking the value 1 if the applicant i re-enrolled into a program different from the initially quasi-randomly assigned program within two years of his initial enrollment), a_{ij} is the initial-application admission score of applicant i at program j, c_j is the cutoff of program j, $f(a_{ij})$ is a polynomial in admission scores, and μ_j are program fixed effects. The time index, which should denote the year of the applicant's first (initial) college application, is surpressed. We study not only re-enrollment, but also re-application (non-compliance intent) by estimating a version of regression (2.4) with the dependent variable y_i indicating if applicant i re-applied after his initial enrollment. These regressions are also estimated on subsamples where program j is (or is not) the applicant's first priority, and where the applicant re-applies to program j (or not). A significant estimate of δ is interpreted as evidence that the probabilities of non-complying (re-applying) depend on the initial assignment.

The first column of Table 2.2 shows statistically as well as economically significant estimates of δ both when considering re-application (-6.2 p.p. compared to the baseline of 18.3%) and re-enrollment (-4.5 p.p. compared to the baseline of 14.6%).[15] Hence, there are 14.6% of non-compliers just below the admission cutoff, compared with 10.1% just above the cutoff. Approximately half of these non-compliance gaps is attributable to students who re-apply to the same program after they were marginally declined at their initial application (i.e., always takers). The effects are most pronounced when students are around the cutoff at their initial-application-ranking top-priority program (-7.9 p.p. when considering re-applications and -6.4 p.p. when considering re-enrollment). These results can also be seen in Figure 2.3 (Figure 2.4), which plots the re-application (re-enrollment) probability against the application score distance from the initial-application cutoff.[16]

In sum, being just below the admission score cutoff of a program during one's initial college application disproportionately incentivizes students to re-apply, and subsequently re-enroll, relative to students just above an initial-application program cutoff. If Croatian

[14]Each cutoff is defined as the admission score of the applicant with the lowest admission score who was offered admission. The optimal bandwidth according to Imbens and Kalyanaraman (2012) is 60 admission points. We replicated the analysis for numerous bandwidth values (10, 20, 30, 40, 50, 60, 70, 80, 90 and 100) and obtained similar results.

[15]On average, around 70% of the re-applying students succeed in changing their initial school assignment, such that the re-application effects largely translate into re-enrollment effects.

[16]The distance from cutoff is defined as admission score centered around the cutoff.

students are subject to non-complying (re-application and re-enrollment) costs, Proposition 1 implies that RDD induced estimates cannot be interpreted simply as graduation treatment effects.

2.3.4 Discussion

In the Croatian case, the probabilities of non-compliance for applicants just above the cutoff ($Z = 1$) are significantly lower (4.5 p.p.) than for those just below the cutoff ($Z = 0$). This, according to Proposition 1 violates the LATE theorem, invoked in, e.g., Kirkeboen et al. (2016). In order to apply the LATE bounds from Proposition 2, one needs to assume the homogenous non-compliance costs assumption. In our case, the costs of non-compliance originate in the reduced labor market experience due to re-enrolling in another program.[17] For example, an always taker with $Z = 1$ is expected to graduate five years after admission, while an always taker with $Z = 0$ is going to use at least one additional year due to re-enrollment. Therefore, the homogenous costs assumption translates into assuming equal slopes of the experience wage profiles of always takers and never takers who did not comply with the treatment assignment—this can be tested empirically by directly comparing experience profiles of always takers and never takers who did not comply. If the gradients of these experience profiles are not the same, one can assume that the experience profile is multiplicative, and perform the same test using the logarithm of returns (or some other transformation of the outcome variable)

2.4 Conclusion

In this paper, we consider a quasi-experimental intention-to-treat setup where non-compliance with treatment assignment is costly (affects the outcome), which violates the exclusion restriction — one of the crucial LATE assumptions. We generalize the LATE theorem to include the case of costly non-compliance and show that the IV estimator can be interpreted as LATE only under the strong assumption that both the costs and the probability of non-compliance do not depend on treatment assignment. We recover treatment effect bounds with an alternative method, inspired by Lee (2009), under the homogenous non-compliance costs assumption, i.e., if the costs do not depend on the initial assignment.

To illustrate the relevance of this design, we consider the recent study by Kirkeboen et al. (2016), who estimate returns to graduating in different fields of education in Norway by instrumenting for graduation with the initial (random) admission offer and measuring labor market returns eight years after the initial application. In such a setup, re-enrolling in another field year(s) after the initial application results in deferred graduation, which reduces labor market experience (as labor market returns are measured eight years after the initial application regardless of the actual graduation date). In the likely case that the length of the labor market experience affects labor market returns, the estimates in Kirkeboen et al. (2016) can be interpreted as returns to fields of study only if the cost of foregoing labor market experience is not field-specific and if the probabilities of non-compliance do not depend on the initial assignment. We show that the latter is not the case in Croatia, using data on the Croatian student-college matching market from 2012 to

[17] If Croatian students re-enroll, they also lose their national-level tuition waiver (otherwise covered by the Ministry of Science and Education), which is constant (homogenous) across programs.

2018, where both the probability of non-compliance (re-enrollment) and the probability of re-taking the national exam (re-application) do depend on the initial assignment.

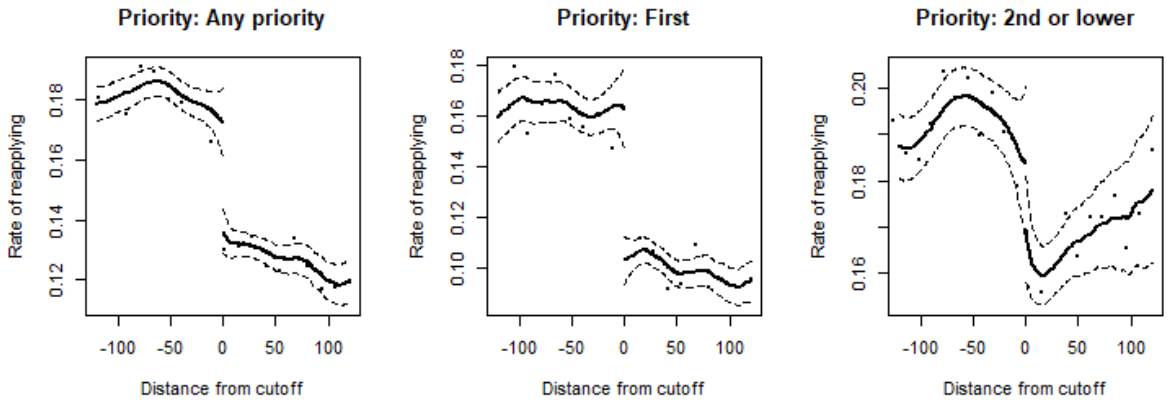
It is reasonable to assume that in the school-college matching market framework, non-compliance with the initial assignment comes at a cost. Not only does it likely imply deferred graduation, but, as demonstrated in the case of Croatia, it also often results in retaking the national exam which is, potentially, also costly (in terms of the cognitive costs of preparation).

The bounding method developed in this paper can be applied in other empirical settings where non-compliance costs arise. For example, when programs are offered through randomized list and applicants can apply to several lotteries (de Chaisemartin and Behaghel (2020)), or in college applications without matching markets (see e.g. Zimmerman, 2014, Goodman et al. (2017), Goodman et al. (2020) and Kozakowski (2020)).

Therefore, our analysis suggests that RDD based IV studies relying on centralized student-school matching markets should first test whether the probabilities of non-compliance with treatment assignment depend on the assignment. If treatment assignment does affect the probability of non-compliance, and if the homogenous costs assumption is not rejected, we suggest employing sharp LATE bounds.

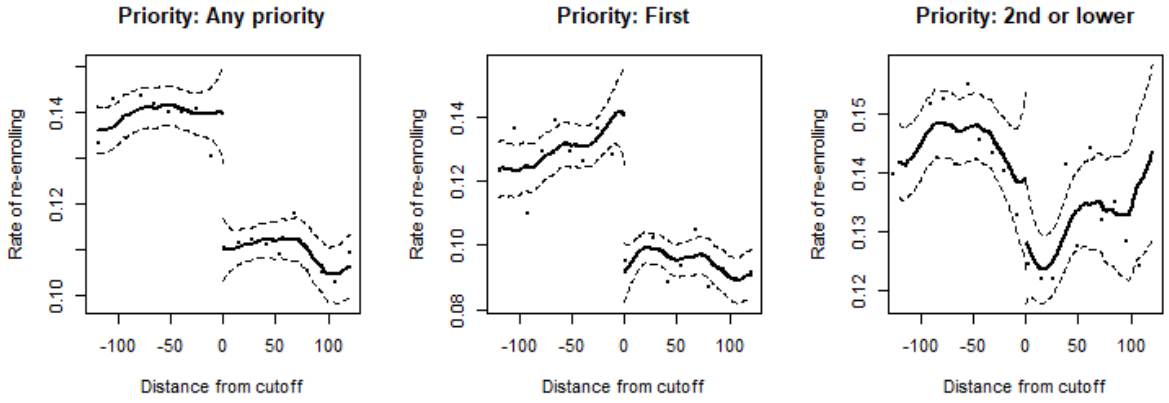
2.5 Appendix - Tables and Figures

Figure 2.3: Re-application probability at the initial-application admission cutoff



Notes: The graphs show re-application probabilities, defined using a two-year window following on the initial-application year, around the admission cutoff in the initial application year. The bandwiths used for the local polynomials correspond to optimal bandwidths according to Imbens and Kalyanaraman (2012). The three graphs show cases when the cutoff school (the school where an applicant was near the school admission cutoff at the initial application) was anywhere on the student's ranked choice list, when it was the student's first priority, and when it was his lower-ranked priority, respectively.

Figure 2.4: Re-enrollment probability at the initial-application admission cutoff



Notes: The graphs show re-enrollment probabilities, defined using a two-year window following on the initial-application year, around the admission cutoffs in the initial application year. The bandwiths used for the local polynomials correspond to optimal bandwidths according to Imbens and Kalyanaraman (2012). The three graphs show cases when the cutoff school (the school where an applicant was near the school admission cutoff at the initial application) was anywhere on the student's ranked choice list, when it was the student's first priority, and when it was his lower-ranked priority, respectively.

Table 2.1: Summary statistics

	All data	RDD estimation sample
	(1)	(2)
Number of programs	620	620
Number of universities	49	49
Number of applicants	101,484	22,383
Number of applications	571,354	80,702
Average admission score	634.53	619.19
	(122.76)	(143.98)
Average GPA	4.01	3.96
	(0.62)	(0.58)
Fraction male	0.47	0.45
Average re-applying rate	0.13	0.16
	(0.33)	(0.36)
Average re-enrolling rate	0.10	0.13
	(0.31)	(0.34)

Notes: The table presents summary statistics calculated for the entire administrative dataset and for the RDD estimation sample (based on the bandwidth of 60 admission score points corresponding to 0.5 of standard deviations, calculated on the ranking lists reported 5 days before the final admission deadline). Standard errors are in the parentheses. The first panel shows the number of programs, universities, applicants, and applications. The second panel shows the average admission score calculated over all applicant-program pairs and the average GPA and fraction male calculated over all applicants. The third panel shows the rates of re-applying and re-enrolling (within a two-year window after the initial-application year) calculated over applicant-program pairs.

Table 2.2: Probability of re-application and re-enrollment

	Cutoff program : Any priority			**Cutoff program : 1^{st} priority**			**Cutoff program : 2^{nd} or lower priority**		
	Any program	Same program	Different program	Any program	Same program	Different program	Any program	Same program	Different program
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A - Probability of re-applying									
Admission offer	-0.062^{***}	-0.036^{***}	-0.025^{***}	-0.079^{***}	-0.050^{***}	-0.028^{***}	-0.043^{***}	-0.018^{***}	-0.025^{**}
	(0.008)	(0.003)	(0.007)	(0.009)	(0.004)	(0.009)	(0.013)	(0.003)	(0.013)
Observations	59,495	59,495	59,495	28,966	28,966	28,966	30,529	30,529	30,529
Baseline	0.183	0.036	0.148	0.178	0.054	0.123	0.192	0.016	0.175
	(0.005)	(0.002)	(0.005)	(0.007)	(0.003)	(0.006)	(0.007)	(0.002)	(0.007)
Panel B - Probability of re-enrolling									
Admission offer	-0.045^{***}	-0.029^{***}	-0.016^{**}	-0.064^{***}	-0.041^{***}	-0.023^{***}	-0.024^{**}	-0.013^{***}	-0.011
	(0.007)	(0.002)	(0.007)	(0.009)	(0.004)	(0.008)	(0.012)	(0.003)	(0.011)
Observations	59,495	59,495	58,216	28,966	28,966	28,345	30,529	30,529	29,871
Baseline	0.146	0.030	0.116	0.153	0.045	0.108	0.140	0.014	0.126
	(0.004)	(0.002)	(0.004)	(0.006)	(0.003)	(0.006)	(0.006)	(0.002)	(0.006)
Program FE	Y	Y	Y	Y	Y	Y	Y	Y	Y

Notes: The table presents RDD estimates corresponding to equation (4) of the effect of students' marginal admission to a program (in the initial application year) on the probability of re-application (Panel A) and re-enrollment (Panel B) in the following two years. The first three columns of the table consider marginal admissions to all programs, the middle three columns focus on marginal admissions to programs ranked as top priority on students' school choice lists, and the last three columns focus on lower-ranked programs from students' ranked choice lists. For each of these specifications, we consider separately re-applying/re-enrollment to any program, to the 'cutoff program', i.e. the program where in their initial-application year they were near the program's admission score cutoff, and to a program other than the cutoff program. All specifications use bandwidths calculated according to the Imbens and Kalyanaraman (2012) optimal bandwidth procedure. All specifications control for a local quadratic polynomial in students' admission score centered around program admission cutoffs. Application year fixed effects and program fixed effects are also used in all specifications. A triangular kernel is used to give more weight to observations close to the cutoffs.
*p-value<0.1 **p-value<0.05 ***p-value<0.01.

Chapter 3

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

Co-authored with Adam Altmejd (Stockholm School of Economics and SOFI), Andrés Barrios-Fernández (Centre for Economics Performance (LSE)), Dejan Kovac (Princeton University, Woodrow Wilson) and Christopher Neilson (Princeton University, Department of Economics).

3.1 Introduction

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

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

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

To causally identify spillover effects, we exploit the fact that all three countries have centralized admission systems that employ a deferred acceptance (DA) mechanisms to allocate applicants to majors depending on their stated preferences and previous academic performance. These selection systems give rise to sharp admission cutoffs in all oversubscribed majors. Taking advantage of the quasi-random variation generated by

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

these cutoffs, we implement a fuzzy Regression Discontinuity Design to investigate how having an older sibling enrolling in a specific major, college or field of study affects individuals' probabilities of applying and enrolling in them.

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

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

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

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

Despite all the research on family and peer effects in education, little is known about how siblings affect human capital investment decisions.[3] Recent evidence shows that older

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

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

siblings can affect high school related choices. Dustan (2018) uses an approach similar to ours and finds that older siblings' influence the choice of high school in Mexico. Joensen and Nielsen (2018), on the other hand, exploit quasi-random variation induced by a policy change in Denmark and find that siblings affect participation in advanced mathematics and science courses during high school.

Much less is known about the role of siblings in higher education specialization choices. Goodman et al. (2015) investigate the relationship between siblings' college choices in the United States and find that the correlation between siblings' applications is much stronger than among similar classmates.[4] Barrios-Fernandez (2018) studies spillovers from both neighbors and siblings in the access to university in Chile, and finds that having a close neighbor or sibling going to university increases the probability of reaching this level of education, especially in areas where university attendance has traditionally been low. Our paper complements this work by exploiting a different source of variation and by focusing on the choice of college and major, rather than in the decision to attend college. Aguirre and Matta (2019) and Goodman et al. (2019), two contemporaneous working papers, also investigate siblings' spillovers in college choices in Chile and the US and provide similar results.

More generally, this paper also contributes to the literature that studies how individuals choose colleges and majors. This has been an active area of research in recent decades that has investigated the role of costs, information, and more recently, of some behavioral barriers.[5] This paper adds a new element by analyzing the role of family networks on

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

[5]The role of funding and liquidity constraints has been investigated by Dynarski (2000), Seftor and Turner (2002), Dynarski (2003), Long (2004), van der Klaauw (2002), and Solis (2017). Misinformation and biased beliefs can also be important determinants of college and major choices Wiswall and Zafar, 2015. Hoxby and Avery (2013) show that low-income, high-achieving students do not apply to selective colleges in the US, even if they are likely to be admitted and would receive more generous funding than they receive from the non-selective colleges to which they currently apply. Mismatches in higher education have also been studied by Griffith and Rothstein (2009), Smith et al. (2013), Black et al. (2015) and Dillon and Smith (2017). Hoxby and Turner (2013) find that providing low-income students with targeted information on their college options, the application process and funding opportunities significantly increased their applications and actual enrollment in selective institutions. In the context of Chile, Hastings et al. (2016) and Hastings et al. (2015), respectively, show that students are uninformed about the costs and benefits of majors and colleges, and that individuals from lower socioeconomic backgrounds are more likely to choose majors with lower earnings. The latter also shows that providing disadvantaged applicants with information about the labor market outcomes of graduates in different programs changed their applications towards majors with higher net of costs earnings. Similarly, Busso et al. (2017) find that information on funding and labor market opportunities improves the quality of the majors to which Chilean students apply in comparison to their baseline preferences. However, there is also research indicating that only providing information is not enough to change applicants' decisions. Bettinger et al. (2012) find that a pure information intervention in the US does not increase college applications or enrollment, and Pekkala Kerr et al. (2015) find that information on labor market

these choices.

The rest of the paper is organized in seven sections. Section 3.2 describes the higher education systems of Chile, Croatia and Sweden, Section 3.3 the data, and Section 3.4 the empirical strategy and the samples that we use. Section 3.5 presents the main results and Section 3.6 places them in the context of previous findings and discusses potential mechanisms. Finally, Section 3.7 concludes.

3.2 Institutions

This section describes the college admission systems of Chile, Croatia and Sweden, emphasizing the rules that generate the discontinuities that we later exploit to identify spillovers among siblings. Despite the differences that exist among these three countries in terms of size, economic development and inequality (Table 3.1), a common feature is that a significant share of each countries' universities select students using centralized admission systems that allocate applicants to majors only considering their preferences rank and previous academic performance. These systems generate sharp admission cutoffs in all oversubscribed programs that we later exploit to identify siblings' spillovers.

prospects of postsecondary education programs does not significantly affect Finnish students' applications or enrollment decisions. Lavecchia et al. (2016); French and Oreopoulos (2017) discuss a host of frictions and behavioral barriers that could explain why some individuals do not take full advantage of educational opportunities. Along this line, Carrell and Sacerdote (2017) argue that college-going interventions work not because of their information component, but because they compensate for the lack of support that disadvantaged students receive from their families and schools.

Table 3.1: Differences across Countries

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

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

3.2.1 College Admission System in Chile

In Chile, all of the public universities and 9 of the 43 private universities are part of the Council of Chilean Universities (CRUCH).[6] All CRUCH institutions, and since 2012 an additional eight private colleges, select their students using a centralized deferred acceptance admission system that only takes into account students' academic performance in high school and in a college admission exam similar to the SAT (Prueba de Selección Universitaria, PSU).[7] Students take the PSU in December, at the end of the Chilean academic year, but they typically need to register before mid-August.[8] As of 2006, all public and voucher school graduates are eligible for a fee waiver that makes the PSU free

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

[7]The PSU has four sections: language, mathematics, social sciences and natural sciences. The scores in each section are adjusted to obtain a normal distribution of scores with a mean of 500 and a standard deviation of 110. The extremes of the distribution are truncated to obtain a minimum score of 150 and a maximum score of 850. In order to apply to university, individuals need to take the language, and the mathematics sections and at least one of the other sections. Universities set the weights allocated to these instruments for selecting students in each program.

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

for them.[9]

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

With this information available and after receiving their PSU scores, students apply to their majors of interest using an online platform. They are asked to rank up to 10 majors according to their preferences. Places are then allocated using an algorithm of the Gale-Shapley family that matches students to majors using their preferences and scores as inputs. Once a student is admitted to one of her preferences, the rest of her applications are dropped. As shown in panel (a) of Figure 3.1, this system generates a sharp discontinuity in admission probabilities in each major with more applicants than vacancies.

Colleges that do not use the centralized system have their own admission processes.[10] Although they could use their own entrance exams, the PSU still plays an important role in the selection of their students, mostly due to the existence of strong financial incentives for both students and institutions.[11] For instance, the largest financial aid programs available for university studies require students to score above a certain threshold in the PSU.

The coexistence of these two selection systems means that being admitted to a college that uses the centralized platform does not necessarily translate into enrollment. Once students receive an offer from a college, they are free to accept or reject it without any major consequence. This also makes it possible for some students originally rejected from a program to receive a later offer. Panel (d) of Figure 3.1 illustrates how the admission to a major translates into enrollment.

3.2.2 College Admission System in Croatia

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

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

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

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

As in Chile, NISpVU uses a deferred acceptance admission system that focuses primarily on students' high-school performance and in a national-level university exam.[12] The national exam is taken in late June, approximately one month after the end of the Croatian academic year. However, students are required to submit a free-of-charge online registration form by mid-February.

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

Once registered, students are able to submit a preference ranking of up to 10 majors. The system allows them to update these preferences until mid-July. At this point, students are allocated to programs based on their current ranking. As in Chile, vacancies are allocated using a Gale-Shapley algorithm, giving rise to similar discontinuities in admission probabilities (Figure 3.1).

Before the final deadline, the system allows students to learn their position in the queue for each of the majors to which they applied. This information is regularly updated to take into account the changes that applicants make in their list of preferences. In this paper, we focus on the first applications submitted by students after receiving their scores on the national admission test. Since some of them change their applications before the deadline, admission based on these applications does not translate one-to-one into enrollment (Figure 3.1).[13]

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

3.2.3 Higher Education Admission System in Sweden

Almost all higher academic institutions in Sweden are public. Neither public nor private institutions are allowed to charge tuition or application fees. Our data include 40 academic institutions, ranging from large universities to small specialized schools.[14]

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

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

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

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

The Swedish admission system has a few important differences compared to the Chilean and Croatian systems. For one thing, the same system is open to applications to full majors and shorter courses alike. To simplify, we will henceforth refer to all these alternatives as *majors*. Moreover, applicants are ranked by different scores separately in a number of *admission groups*. Their best ranking is then used to determine their admission status.[15] Finally, the Swedish admission system has two rounds. After the first round, applicants learn their admission status and they place in the waiting list for all their applications. At this point, they can decide wether to accept the best offer they have or to wait and participate in a second application round. Their scores and lists of preferences do not change between the two rounds, but the cutoffs might. In this project we focus on the variation generated by the cutoff of the second round. Since some applicants decide to accept the offers they received after the first round instead of waiting for the second round, not all applicants above the second round admission cutoff end up receiving an offer. Those who dropout from the waiting list after the first round cannot receive a second round offer, even if their score was above the final admission cutoff. This explains why, in the case of Sweden, the jump in older siblings' admission and enrollment probabilities is smaller than in the other two countries (see Figure 3.1).

For each program, at least a third of the vacancies are reserved for the high school GPA admission group. No less than another third is allocated based on results from the Högskoleprovet exam. The remaining third of vacancies are mostly also assigned by high school GPA, but can sometimes be used for custom admission.[16]

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

Students can apply to majors starting in the fall or spring semesters, and the application occurs in the previous semester. In each application, they rank up to 20 alternatives (students were able to rank 12 alternatives until 2005). Full-time studies correspond to

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

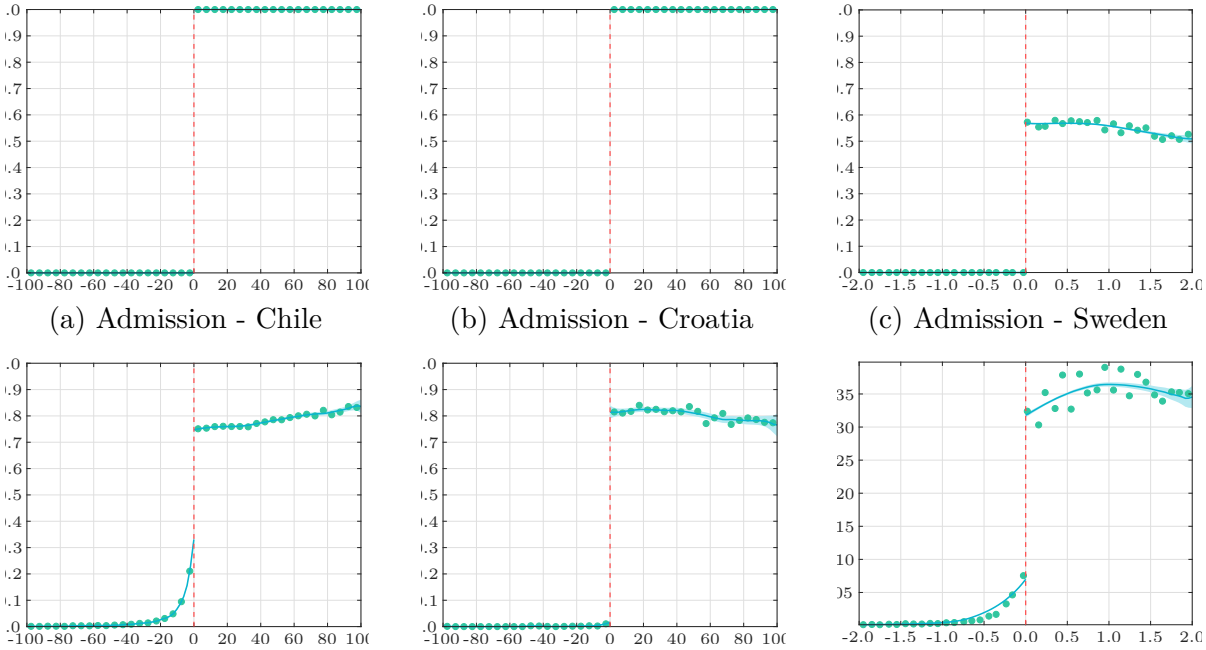
[16] This is the case in some highly selective majors, where an additional test or an interview is sometimes used to allocate this last third of vacancies. We do not include admissions through such groups in our analysis.

30 credits per semester, but students who apply to both full-time majors and courses in the same application receive offers for the highest-ranked 45 credits in which they are above the threshold.

After receiving an offer, applicants can either accept or decide to stay on the waiting list for choices which they have not yet been admitted. Should they decide to wait, admissions after the second round will again only include the highest-ranked 45 ECTS, and all lower-ranked alternatives will be discarded, even those that they were previously admitted to.[17]

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

Figure 3.1: Older Siblings' Admission and Enrollment Probabilities in Target Major-College at the Admission Cutoff (First Stage)



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

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

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

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

3.3 Data

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

From DEMRE we obtani individual-level data on all the students registered to take the PSU between 2004 and 2015. These datasets contain information on students' performance in high school and in the different sections of the college admission exam. They also contain student-level demographic and socioeconomic characteristics, information on students' application, college acceptances through the centralized application system, and college enrollment. To identify siblings, we exploit the fact that when registering for the exam, students provide the national identification number of their parents. Using this unique identifier we can match all siblings that correctly reported this number for at least one of their parents.[18]

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

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

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

The Swedish application data consists of two parts. We obtain data on applications from

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

the modern system, for the years 2008 to 2016, directly from the Swedish Council for Higher Education (UHR). Applications for the years 1992–2005 are from an older system and are obtained from the Swedish National Archives (Riksarkivet). While the modern system contains the universe of applications to higher education in Sweden, institutions were not required to participate in centralized admissions before 2006.[19] Family connections and all demographic and socioeconomic variables that we use are provided by Statistics Sweden.

Using these data, we identify around 83,000, 17,000, and 300,000 pairs of siblings in Chile, Croatia, and Sweden, respectively, where the older sibling had at least one active application to an oversubscribed major with an application score within the minimum bandwidth used in each country. Table 3.2 presents summary statistics for these subsets of siblings and also for the full set of potential applicants.[20]

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

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

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

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

Table 3.2: Summary Statistics

	Chile		Croatia		Sweden	
	Siblings Sample (1)	Whole Sample (2)	Siblings Sample (3)	Whole Sample (4)	Siblings Sample (5)	Whole Sample (6)
			A. Demographic characteristics			
Female	0.521	0.520	0.563	0.567	0.579	0.595
	(0.500)	(0.499)	(0.496)	(0.495)	(0.493)	(0.490)
Age when applying	18.786	19.829	18.880	19.158	20.589	20.872
	(0.606)	(2.484)	(0.654)	(0.963)	(2.374)	(2.562)
Household size[1]	4.756	4.625	2.790	1.925	3.086	2.946
	(1.498)	(1.607)	(1.243)	(1.198)	(1.142)	(1.186)
			B. Socioeconomic characteristics			
High income[2]	0.279	0.128			0.349	0.339
	(0.449)	(0.334)			(0.477)	(0.473)
Mid income[2]	0.403	0.325			0.262	0.290
	(0.490)	(0.469)			(0.440)	(0.454)
Low income[2]	0.318	0.546			0.389	0.371
	(0.466)	(0.498)			(0.488)	(0.483)
Parental ed: < high school	0.095	0.254			0.038	0.056
	(0.294)	(0.435)			(0.191)	(0.229)
Parental ed: high school	0.333	0.386			0.339	0.481
	(0.471)	(0.487)			(0.471)	(0.481)
Parental ed: vocational HE	0.149	0.115			0.067	0.063
	(0.356)	(0.319)			(0.250)	(0.244)
Parental ed: university	0.413	0.234			0.562	0.517
	(0.492)	(0.423)			(0.496)	(0.500)
			C. Academic characteristics			
High school track: academic[3]	0.846	0.673	0.439	0.416		
	(0.361)	(0.469)	(0.496)	(0.496)		
High school: vocational[3]	0.154	0.327	0.561	0.584		
	(0.361)	(0.469)	(0.496)	(0.496)		
Takes admission test	0.956	0.868	0.835	0.835	0.679	0.628
	(0.204)	(0.338)	(0.371)	(0.372)	(0.467)	(0.483)
High school GPA score	-0.107	-0.465	-1.191	-1.238	0.673	0.432
	(1.235)	(1.357)	(2.728)	(2.763)	(0.766)	(0.773)
Admission test avg. score	0.241	-0.512	-0.779	-1.027	0.281	-0.061
	(1.619)	(1.708)	(1.835)	(2.034)	(0.991)	(1.000)
Applicants	83,379	2,823,897	16,721	199,475	301,967	3,822,188

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

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

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

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

3.4 Empirical Strategy

The identification of siblings' effects is challenging. In the first place, since siblings share genetic characteristics and grow up under very similar circumstances, it is not surprising to find that their outcomes —including the major and college that they attend— are highly correlated. Thus, the first identification challenge consists in distinguishing these correlated effects from the effects generated by interactions among siblings. In addition, if

siblings' outcomes simultaneously affect each other, this gives rise to what Manski (1993) described as the reflection problem. In our setting, given that older siblings decide to apply and enroll in college before their younger siblings, this is less of a concern (i.e. decisions that have not yet taken place should not affect current decisions). However, there could still be cases in which siblings decide together the college and major that they want to attend and therefore we need an empirical strategy to address this potential threat.

To overcome these identification challenges, we exploit thousands of cutoffs generated by the deferred acceptance admission (DA) systems that Chilean, Croatian and Swedish universities use to select their students. Taking advantage of the discontinuities created by these cutoffs on admission, we use a Regression Discontinuity (RD) design to investigate how older siblings' admission to their target major affects the probability that their younger siblings will apply and enroll in the same major, college or field of study.[21]

Since individuals whose older siblings are marginally admitted or rejected from a specific major are very similar, the RD allows us to rule out the estimated effects being driven by differences in individual or family characteristics, eliminating concerns about correlated effects. Moreover, considering that the variation we exploit in the major-college in which older siblings enroll comes only from their admission status and cannot be affected by the choices that their younger siblings will make in the future, we can abstract from the reflection problem.[22]

As discussed in Section 3.2, rejecting an offer does not have any major consequence for Chilean students. As a result, there is a non-negligible share of applicants who, despite being admitted to a particular college or major, decide not to enroll. Thus, when studying how older siblings' actual enrollment affects their younger siblings, we use a fuzzy RD in which older siblings' enrollment in a specific major is instrumented with an indicator of admission.

We follow a similar approach for Croatia. Although in this setting rejecting an offer is costly, we use a fuzzy and not a sharp RD because, as explained in Section 3.2, we focus our attention on the first application students submit after receiving their results in the college admission exam. Since some individuals modify their applications in the weeks following the exam results, admission to the first set of preferences does not translate

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

[22] We show that this is indeed the case in a series of placebo exercises that we present in Appendix .2.

one-to-one into enrollment.[23]

For Sweden, we focus our attention on the applications submitted during the first round of the admission process. Since students can reject these offers there is no perfect compliance either.[24] Thus, as in the previous two cases, we also use a fuzzy-RD to identify the siblings' spillovers.

This paper investigates how individuals' probabilities of applying and enrolling in specific majors, colleges and fields of study change when their older siblings are marginally admitted and enroll in them. The basic idea behind our empirical design consists in defining for each major, college and field of study the sample of older siblings marginally admitted and marginally rejected from them, and then comparing how this affects their younger siblings' choices. Therefore, each observation in our estimation sample corresponds to a pair of siblings in which the older one is close enough to the admission cutoff of a specific major. Given that in the three countries individuals are allowed to apply to multiple programs, this means that the same pair of siblings could eventually appear several times in the sample. In cases where multiple older siblings are identified, we focus on the one close in age to the potential applicant of interest. In addition, if older siblings have applied multiple times to college, we only take the first set of applications he or she submitted.

We define major as a specific combination of major and college, and field of study as the three digit-level ISCED code of these majors.[25] This means that in each country we consider around 80 different fields of study.

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

3.4.1 Major Sample

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

As discussed earlier, the assignment mechanism used in Chile, Croatia and Sweden results in cutoff scores for each major with more applicants than available places; these cutoffs

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

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

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

correspond to the lowest score among the admitted students. Let c_{jfut} be the cutoff for major j belonging to field of study f in college u in year t. If the major j of field f offered in college u is ranked before the major j' of field f' offered by college u' in student i's preference list, we write $(j, f, u) \succ (j', f', u')$.[26] Denoting the application score of individual i as a_{ijfut}, we can define marginal students in the major sample as those whose older siblings:

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

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

Note that this sample includes individuals whose older siblings were rejected from (j, u) ($a_{ijfut} < c_{jfut}$) and those whose older siblings scored above the admission cutoff ($a_{ijfut} \geq c_{jfut}$). Since the application list in general contains more than one preference, this means that the same individual may belong to more than one major-college marginal group. Figure 3.1 illustrates the probability of admission and enrollment in a given major around the admission cutoff in Chile, Croatia and Sweden.

3.4.2 College Sample

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

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

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

3.4.3 Field of Study Sample

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

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

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

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

3.4.4 Identifying Assumptions

As in any other RD setting, the validity of our estimates relies on two key assumptions. First, individuals should not be able to manipulate their application scores around the admission cutoff. The structures of the admission systems in Chile, Croatia and Sweden make the violation of this assumption unlikely. However, to confirm this, we show that the distribution of the running variable (i.e. older sibling's application score) is continuous at the cutoff (see Appendix .2 for more details).

Second, in order to interpret changes in individuals' outcomes as a result of the admission status of their older siblings, there cannot be discontinuities in other potential confounders at the cutoff (i.e. the only relevant difference at the cutoff must be older siblings' admission). Appendix .2 shows that this is indeed the case for a rich set of socioeconomic and demographic characteristics.

As previously mentioned, we use a fuzzy RD to study the effect of older siblings' enrollment (instead of admission) on younger siblings' outcomes. This approach can be thought of as an IV strategy, meaning that in order to interpret our estimates as a local average treatment effect (LATE) we need to satisfy the assumptions discussed by Imbens and Angrist (1994).[27] In this setting, in addition to the usual IV assumptions, we also need to assume that receiving an offer for a specific major does not make the probability

[27]Independence, relevance, exclusion and monotonicity. In this setting, independence is satisfied around the cutoff. The existence of the first stage is shown in Figure 3.1. The exclusion restriction implies that the only way through which older siblings' admission to a major affects younger siblings' outcomes is by the increase it generates in older siblings' enrollment in that major. Finally, the monotonicity assumption means that admission to a major weakly increases the probability of enrollment in that major (i.e. being admitted into a major does not reduce the enrollment probability in that major).

of enrolling in a different major bigger than in the absence of the offer. [28] Given the structure of the admission systems that we study, this additional assumption does not seem very demanding.[29]

An additional issue related to the interpretation of our estimates is that as noted by Cattaneo et al. (2016), by pooling together different cutoffs, our estimates correspond to a weighted average of LATEs across programs. This weighted average gives more importance to programs with more applicants in the vicinity of the admission cutoff. Since there could be heterogeneity in the characteristics of individuals around each admission cutoff, and also on the effect of admission and enrollment at each admission cutoff, we need to be careful with the interpretation of this weighted average. [30]

A final consideration for the interpretation of our results relates to the findings of Barrios-Fernandez (2018). According to these, the probability of attending university increases with close peers' enrollment. If marginal admission to the programs that we study translates into an increase in total university enrollment, then our estimated results could simply reflect that individuals whose older siblings attend college are more likely to enroll. We address this concern in Appendix .2 where we show that older siblings' marginal admission to their target majors does not generate a difference in younger siblings' total enrollment. [31]

Appendix .2 presents multiple additional robustness checks. We show that, as expected, changes in the admission status of younger siblings do not have an effect on older siblings; that our estimates are robust to different bandwidth choices and that placebo cutoffs do not significantly effect any of the outcomes that we study.

[28] Appendix .1 presents a detailed discussion of the the identification assumptions.

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

[30] In order to understand what is driving our results we perform a detailed heterogeneity analysis along multiple dimensions including both individual and program characteristics. In Appendix .2 we study how our results vary when we re-weight observations around each cutoff by the inverse of the total number of applicants around it. Although the estimates are slightly smaller, the main conclusions still hold.

[31] In Chile, we find only a small increase in the total enrollment of older siblings. This result is not surprising. As discussed in Section 3.2, the colleges that use the centralized admission system in Chile are, on average, more selective than the rest. This means that individuals rejected from these institutions still have many other alternatives available. In Croatia, we find that marginal admission translates into a more significant increase in older siblings total enrollment. However, we do not find an extensive margin response among younger siblings. Finally, in Sweden we once again find a small increase in older siblings' total enrollment, but as in the previous cases it does not translate into any significant difference in the total enrollment of their younger siblings.

3.5 Results

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

3.5.1 Method

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

$$y_{ijut\tau} = \beta admitted_{ijut\tau} + f(a_{ijut\tau};\gamma) + \mu_t + \mu_{jut\tau} + \varepsilon_{ijut\tau} \tag{3.1}$$

where,

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

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

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

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

We estimate two versions of this specification. In both cases, $f(a_{ijut\tau};\gamma)$ corresponds to a linear or a quadratic polynomial of $a_{ijut\tau}$ whose slope is allowed to change at the admission cutoff. However, while in one specification we use a uniform kernel, in the second one we use instead a triangular kernel to give more weight to observations close to the cutoff.[32] Our analysis of younger siblings' responses to older siblings' marginal enrollment focuses on three levels: first preference in the application list, all the preferences in the application

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

list, and enrollment. Depending on the margin of interest (i.e. major, college or field) we use one of the samples described in Section 3.4. We compute optimal bandwidths according to Calonico et al. (2014) for each sample and level being investigated , but then we use a single bandwidth per sample: the smallest one among the three computed. [33]

Since all the specifications that we use focus on individuals whose older siblings are near an admission cutoff, our estimates represent the average effect of older siblings' marginal admission compared to the counterfactual of marginal rejection from a target major.[34]

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

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

To study heterogeneous effects, we add to the baseline specification an interaction between older siblings' admission and the characteristic along which heterogeneous effects are being investigated (i.e. $admitted_{ij u \tau} \times x_{ij u t \tau}$). This interaction is also used as an instrument for the interaction between the older sibling's enrollment and $x_{ij u t \tau}$. In both cases, $x_{ij u t \tau}$ is also included as a control.

3.5.2 Effects of Older Siblings on Major Choice

This section discusses how older siblings' admission and enrollment in a specific major-college combination affect their younger siblings' probabilities of applying to and enrolling in it. To investigate changes in this margin, we use the Major Sample defined in Section 3.4.2.

The RD estimates illustrated in Figure 3.2 provide consistent causal evidence that students are more likely to apply to and enroll in a major if an older sibling was admitted to it before.[35]

As discussed in Section 3.4, receiving an offer for a specific major does not translate one-

[33] In principle, optimal bandwidths should be estimated for each admission cutoff independently. However, given the number of cutoffs in our sample, doing this would be impractical. Therefore, we compute optimal bandwidths pooling together all the cutoffs. Appendix Figures B3, B4 and B5 illustrate how sensitive our estimates are to the choice of bandwidth.

[34] Strictly speaking, our estimates represent a weighted average of multiple LATEs. See Section 3.4.4 for additional details. In addition, Appendix Tables B8, B9 and B10 present the results of an additional specification that controls by target major × counterfactual major fixed effect. The effects are very similar to the ones presented in the main section of the paper.

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

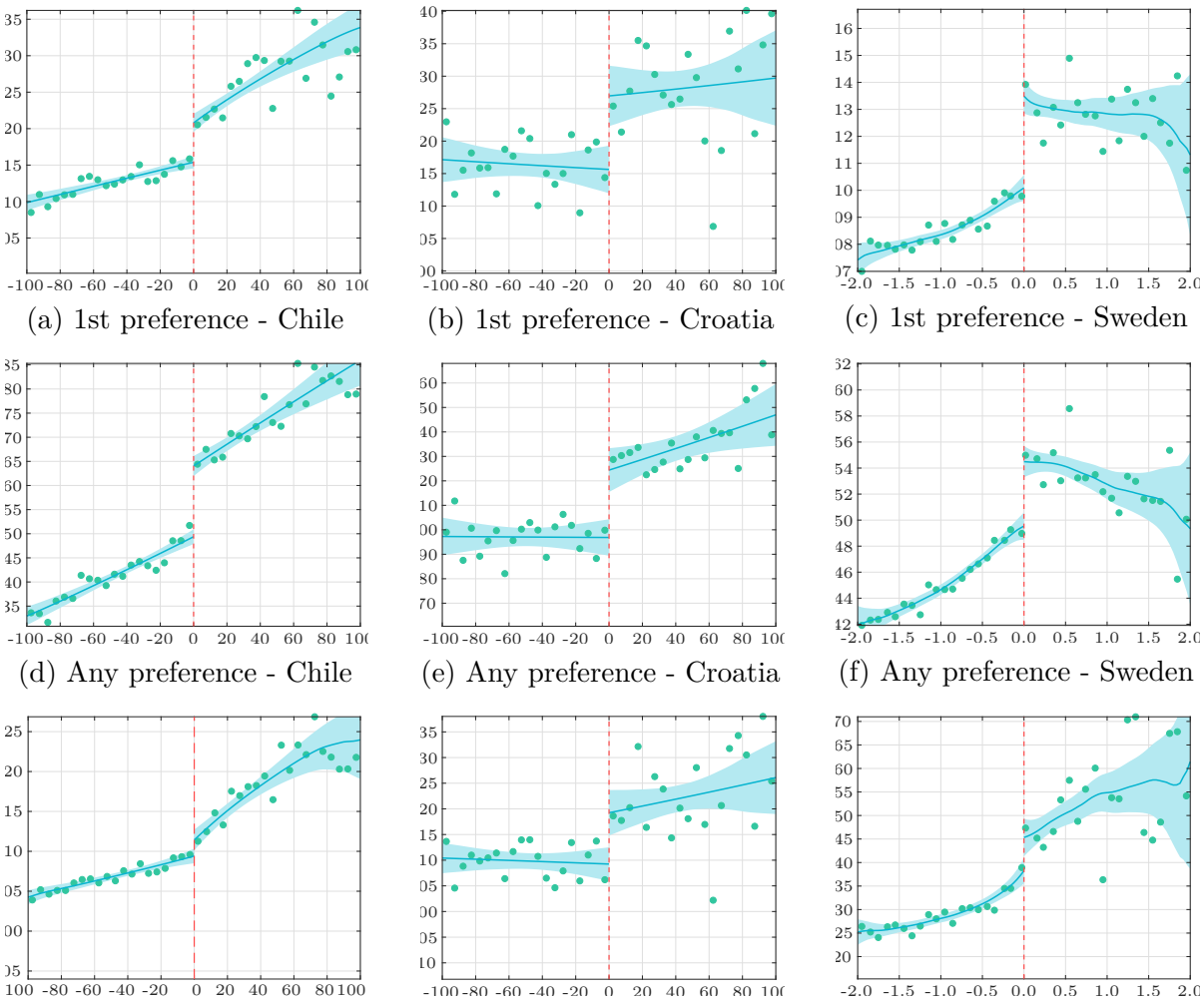
to-one into enrollment in any of the settings that we study. Thus, in order to estimate the effect of older siblings' enrollment on individuals applications and enrollment decisions, we combine the reduced form results discussed in the previous paragraph with the respective first stages illustrated in Figure 3.1, and obtain the fuzzy-RD estimates presented in Table 3.3. Under the identification assumptions discussed in Section 3.4, these fuzzy-RD provide consistent estimates for the effects of interest.

We find that in Chile, having an older sibling "marginally enrolling" [36] in a specific major increases the likelihood of applying to that major in the first preference by 0.8 percentage points (40%) and in any preference by around 2.8 pp (55%). These changes in applications also translate into an increase of around 0.3 pp (30%) in enrollment (although this last figure is not statistically significant). The results for Croatia are very similar. Individuals are 1.4 pp (45%) more likely to apply to their older siblings' target major in the first preference, 3.4 pp (33%) more likely to apply to it in any preference and 1.4 pp (58%) more likely to enroll in it. Finally, in Sweden, the likelihood of ranking older siblings' target major in the first place increases by around 2 pp (180%), while the likelihood of ranking it in any position increases by around 3 pp (63.8%). We also show that enrollment in older siblings' major increases by roughly 0.4 pp (100%).

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

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

Figure 3.2: Probabilities of Applying and Enrolling in the Target Major-College of the Older Siblings



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

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

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

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

older siblings' target major once they learn their application score. In Table 3.4 we present additional results that come from specifications that expand the baseline specification by adding an interaction between older siblings' marginal enrollment and a proxy of younger siblings' eligibility for their older siblings target major.[37]. According to the

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

results presented in columns (1) to (3) of Table 3.4, younger siblings are more likely to apply and enroll in their older siblings' target major if they are eligible for it.[38]

In order to gain a deeper understanding about what is behind this "major following"behavior, in columns (4) to (6) of Table 3.4 we estimate the same specifications just discussed, but this time focusing on the sub-sample of older siblings whose target and counterfactual majors were offered by the same college. For these older siblings, being rejected from their target major does not change the college in which they end up being admitted. Finding that even in this restricted sample younger siblings are more likely to apply to and enroll in their older siblings target major, suggests that the effects discussed in this section are not only driven by an increase in applications and enrollment in the older sibling's target college.

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

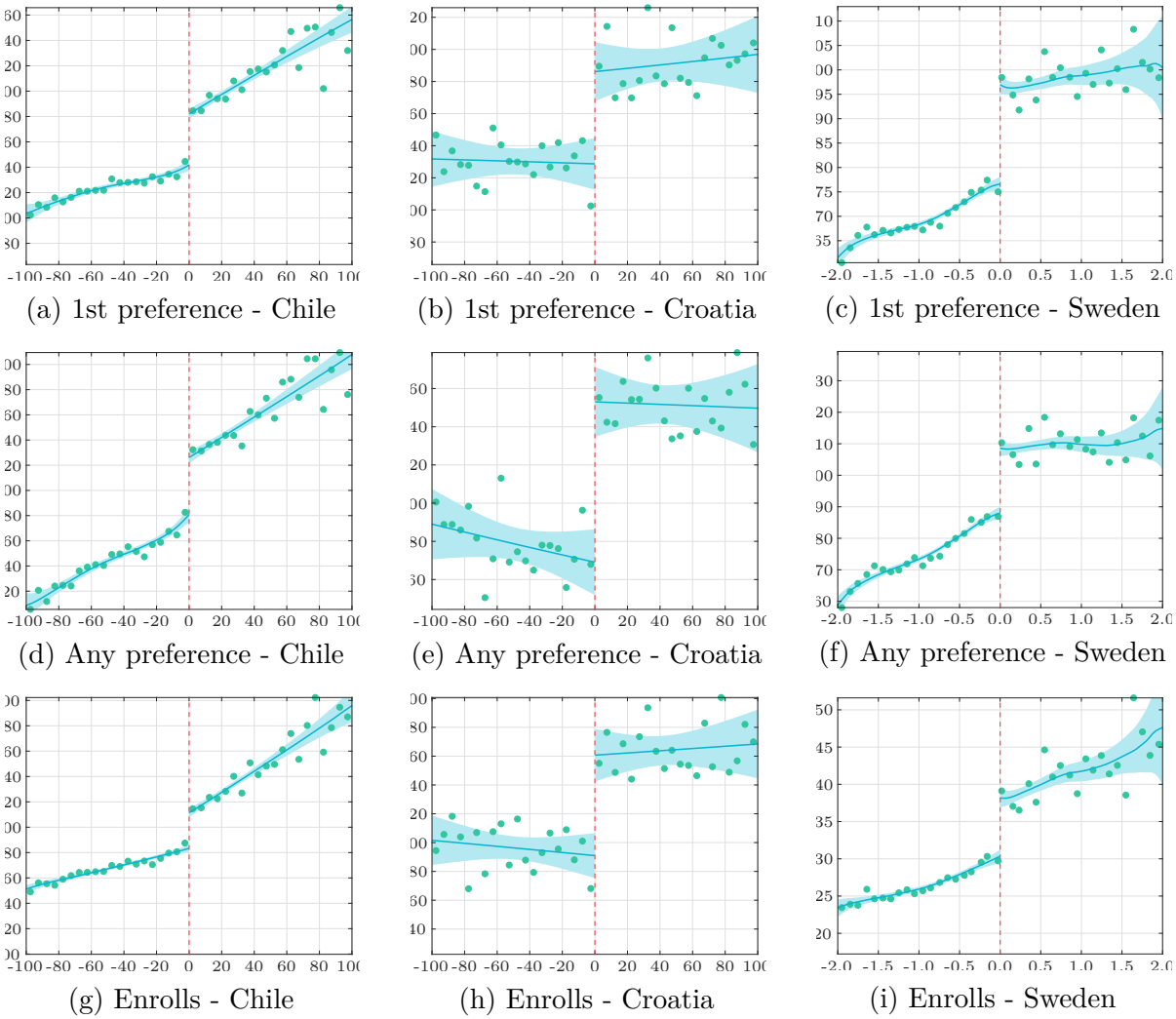
3.5.3 Effects of Older Siblings on College and Field of Study Choices

While the focus of the previous section was on the specific major-college choice, this section independently investigates how younger siblings' choices of college and field of study are affected by older siblings. To study these margins, we slightly modify the baseline specification of the previous section by replacing the outcome for a dummy variable that indicates if the younger sibling applies or enrolls in the target college or in the target field of study of the older sibling.[39] Depending on the margin being investigated, we focus our attention on the College Sample or on the Field Sample defined in Section 3.4.2.[40]

Table 3.5 summarizes the results of siblings' spillovers on the choice of college. In Chile, individuals are 7.2 pp (45%) more likely to rank their older siblings' target college first and 10.1 pp (30%) more likely to apply to it in any preference. They are also 4.4 pp (44%) more likely to enroll in that college. For Croatia, the same figures are 7.5 pp (23%), 10.9 pp (19%) and 8.4 pp (29%), respectively, and for Sweden they are 15 pp (170%), 15.3 pp (79%) and 6.4 pp (188%).

[38] In section 3.5.8, we show that older siblings' enrollment on their target major does not increase younger siblings' academic performance in high school or in the university admission exam. These

Figure 3.3: Probabilities of Applying and Enrolling in the Target College of Older Siblings



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

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

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

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

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

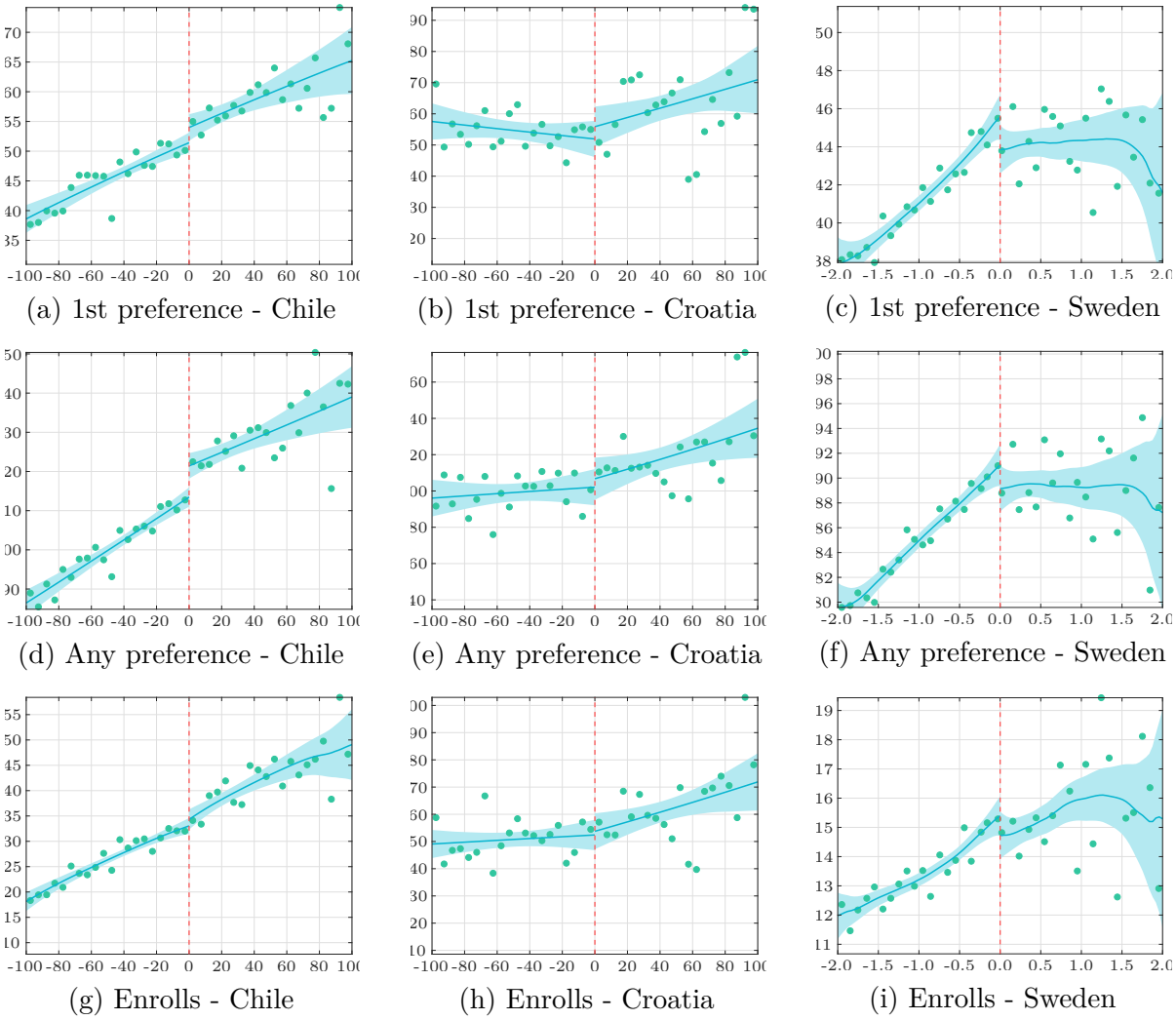
results attenuate selection concerns that could have arisen by adding eligibility into the analisys.

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

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

share of the population, but also multiple universities.[41].

Figure 3.4: Probabilities of Applying and Enrolling in the Target Field of Study of Older Siblings



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

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

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

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

Table 3.6 presents the results of an exercise in which we estimate the baseline specification on a sample of Chilean students from Santiago, Valparaíso and Concepción whose older siblings apply to institutions in their hometowns. If the effects documented in Table 3.5 were driven only by geographic preferences, we should not find sibling spillovers on the choice of college for this subsample. However, the coefficients that we obtain in this case are very similar to the main results previously discussed.

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

On the other hand, when investigating how the choice of field of study —defined by the three- digit level code of the ISCED classification— is affected, we only find a marginally significant effect on younger siblings' applications in the case of Chile. In Croatia and Sweden none of the estimated coefficients is statistically significant (Table 3.7). Considering that the comparison of results across samples must be treated with caution, the results discussed so far suggest that individuals' major choice is only affected when younger siblings are likely to be admitted in their older siblings' specific major-college combination.

Since the choices of major and college seem to be the margins more affected by older siblings' higher education decisions, in the rest of the paper we will focus on these margins.[42]

3.5.4 Effects on Applications to Major and College by Gender:

This section explores if the responses in major and college choice documented in the previous sections vary depending on siblings' gender.[43]

The results of this section are summarized in Table 3.8. The first three columns look at differences in applications to majors, while the following three columns look at differences in applications to colleges. To perform these analyses we expand the baseline specification by adding an interaction between the treatment and a dummy variable that indicates whether the gender of both siblings is the same. The main effect of the "same gender"dummy is also included as a control in all these specifications.

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

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

When looking instead at the college choice, we find no significant difference in how male and female applicants respond to the choices of their older brothers or sisters. Being of the same gender as younger siblings does not seem to increase the likelihood of being

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

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

followed by them. However, in this case, independently of their gender, younger siblings seem to be more responsive to older brothers than to older sisters.

Overall, the results discussed in this section indicate that males are more likely to apply to the same major and college of an older brother than of an older sister. However, their applications are also affected by the higher education decisions of their older sisters. In the case of females, the pattern is less clear. They seem to be more responsive to what happens with their older sisters when choosing a major, but the opposite is true when looking at applications to college.

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

In this section we investigate how the applications to major and college change depending on how close siblings are in terms of age and academic potential.[44] To investigate differential effects by age, we expand the baseline specification with an interaction between the treatment and a dummy variable indicating whether siblings were born 5 or more years apart. To investigate if the effects change depending on differences in academic potential, we proceed in a similar way by adding an interaction with the absolute difference in siblings' high school GPA.[45] In Croatia, we only observe high school GPA for students completing their secondary education before 2015; this explains the smaller sample used in this part of the analysis for Croatia.

Table 3.9 summarizes the results of this section. The first two columns look at the choice of major, while the last two at the choice of college. In Chile and Croatia, the effects do not significantly decrease with the age difference between siblings. In the case of Sweden, the effects are stronger for siblings who are closer in age. However, even for those who are more than 5 years apart the effects are significant both statistically and economically.

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

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

[45] Note that if younger siblings are still in high school when their older siblings apply to higher education, their high school GPA could be an outcome of the treatment. However, as shown in Section 3.5.8 "marginal enrollment"of older siblings in their target major does not seem to affect individuals' academic performance.

choices depending on siblings' academic potential.

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

This section studies how the effects documented in Section 3.5.2 change depending on the quality of the target major of the older sibling.[46] We measure quality in terms of admitted students' academic potential, first-year dropout rates and graduates' earnings.[47]

Student quality is the only variable in this section that we observe for the three countries. We define the quality of the students in a program in a given year using the average performance of admitted students in the college admission exams in Chile and Croatia, and as the average high school GPA of admitted students in Sweden. We are able to compute dropout rates and graduates earnings only for Chile and Sweden. We compute dropout rates for each major using individual level data provided by the Ministry of Education (Chile) and by the Council for Higher Education (Sweden). The data from Chile allow us to compute dropout rates for all college cohorts beginning in 2006;[48] in Sweden we observe dropout rates for the entire sample period. Variables measuring the labor market performance of former students in Chile are available at the major-college level. They are computed by the Ministry of Education with the support of the National Tax Authority.[49] In the case of Sweden, information on earnings comes from Statistics Sweden.

The main results of this section are summarized in Table 3.10. All variables, except for dropout rates, are standardized to facilitate the interpretation of the results. The first three columns of the table investigate heterogeneous effects on applications to majors, while the last three on applications to colleges.

When looking at heterogeneous effects on the major choice by the quality of the students admitted to that major, we only find a significant difference in Sweden. In this country, a difference of 1σ in the quality of the applicants admitted to the older siblings' major increases the younger sibling's applications to that major by 1.2 pp. Differences are more

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

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

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

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

clear when looking instead at the college choice. In this case, an increase in the quality of the students admitted to the older siblings major increases younger siblings' applications to the college offering that major by 2.4 pp in Chile, 2.7 pp in Croatia and 3.6 pp in Sweden. [50]

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

Finally, when looking at heterogeneity by graduates' labor market outcomes we find that younger siblings are more likely to apply to their older siblings' major when past graduates' earnings are higher. A similar pattern arises when focusing on the college choice, but in this case the coefficients are unprecisely estimated.

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

Table 3.11 presents results of a similar exercise, but in which we study heterogeneous effects by the difference in the quality indexes of older siblings' target and counterfactual majors (counterfactual major is the major in which they would have been admitted in the event of being rejected from their target choice).[51] This forces us to restrict the sample to older siblings for whom it is possible to identify a counterfactual alternative. Therefore, those not admitted to any program are not part of this analysis. We find no heterogeneous effects by differences in any of the quality measures we use. In part, this could be due to the smaller sample size used for this exercise and to the fact that on average there is no significant difference between the quality of the target program and the quality of the next best option.

3.5.7 Effects on Application and Enrollment by the College Experience of Older Siblings

This section investigates whether the effects on the choice of major and college depend on the experience of older siblings in their target major. Table 3.12 provides evidence consistent with the hypothesis that individuals learn from their older siblings' experience if a specific major or college would be a good match for them. Siblings are similar in many dimensions, and therefore if an older sibling has a negative experience in a specific major or college, their younger siblings may infer that applying and enrolling in that

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

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

alternative is not necessarily good for them. In our data, the best available proxy for older siblings' experience in college is dropout. We are only able to compute dropout for Chile and Sweden, and therefore this section only presents results for these countries.

We add to the baseline specification an interaction between the treatment and a dummy that indicates whether the older sibling drops out from the major or college in which she first enrolls,[52] and the main effect of older siblings' dropout.[53] The results of this exercise should be interpreted with caution. Dropping out from college is not random, and although controlling by dropout helps to capture some of the differences that may exist between individuals who remain at and leave a particular college, there could still be differences that we are not able to control for.[54] In addition, the dropout variable can only be built for older siblings who actually enroll in some major. Appendix Table B4 shows that in Chile and Sweden, marginal admission does not translate into relevant increases in older siblings' total enrollment. However, only focusing on applicants whose older siblings enroll in a program affects the composition of the sample used in this analysis.

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

3.5.8 Effects on Academic Performance

In this section we study if the increase in the likelihood of applying and enrolling in the major attended by an older sibling could be driven by an improvement in younger siblings' academic performance. To study this we use the same fuzzy-RD strategy discussed in Section 3.4, but this time we look at younger siblings' high school GPA and at their scores in the admission exams. Since not all potential applicants take the admission exam, we replace missing values by zero. This means that when looking at effects on exams scores, our estimates capture differences in performance, but also differences in the probability of taking the exam. The bandwidths used in this section are the same as those used in Section 3.5.2.

Table 3.13 summarizes these results. We show that, having an older sibling "marginally

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

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

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

enrolling" in her target major does not seem to generate significant changes in younger siblings' high school performance or in their performance in the university admission exams.

These results hold for the three countries in our study, and suggest that the effects documented on the choice of program are not driven by an improvement in the academic performance of younger siblings.[55]

3.6 Discussion

The results presented in Section 3.5 show that the path followed by older siblings in higher education affects the major and college choice of their younger siblings. Although documenting the existence of sibling spillovers in the choice of major and college in three settings as different as Chile, Croatia and Sweden is interesting in itself, from a policy perspective it is also relevant to understand the mechanisms behind these responses. In the rest of this section, we discuss three broad classes of mechanisms that could drive our results using a simple framework of discrete choice and utility maximization.

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

$$\max_{m \in M_i} U_i(\vec{m}), \quad m = (x_{1m}, ..., x_{nm})$$
$$\text{s.t. } P_m \leq B_i$$

With this simple framework in mind, the first way in which older siblings could affect the decision of applying and enrolling in a specific major or college is by affecting the costs of that option. For instance, by attending the same college as an older sibling, individuals might save in commuting and living costs. However, we find that the effects persist even among siblings who, due to age differences, are unlikely to attend college at the same time. This result, and the fact that the effects look very similar when we focus on a group of individuals whose older siblings apply to majors offered in their hometown, suggest that this convenience channel is not the main driver of our results. [56]

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

[56] In some settings, the admission systems give an advantage to siblings of current or former students. This, however, is not a concern in our case. In Chile, Croatia and Sweden universities use centralized

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

The persistence of the effects among siblings with large age differences suggests that our results are not driven by them enjoying each other's company. In addition, finding no heterogeneous effects by differences in the quality of target and counterfactual majors of older siblings and finding no effects on younger siblings' academic performance, suggests that individuals' aspirations are not affected. If this were the case, we would expect to see them exerting additional effort in preparation for college, something that is not reflected in their applications, nor in their high school and college admission exam performance.

Joensen and Nielsen (2018) argue that the fact that their results are driven by brothers who are close in age and in academic performance is evidence in favor of competition being the main driver of their results. As previously discussed, in our case the results persist even among siblings born more than 5 years apart, and also among sisters and different-gender siblings, suggesting that if competition mostly arises between brothers close in age, it cannot be the only driver of our results.

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

Finally, older siblings' enrollment in a specific major-college could affect the choice set of their younger siblings by making some options more salient or by providing information about relevant attributes of the available options.[57] Considering the amount of major-college combinations from which applicants can choose, both hypothesis could play a relevant role. However, we find stronger effects when older siblings' majors are of higher quality, which goes against a pure salience story. If salience were the main driver of

admission systems that select students based only on their academic performance in high school and on a national level admission exam. Although in Chile some colleges offer discounts in tuition fees when many siblings simultaneously attend the same program, finding that the effect persists even when looking at siblings born 5 or more years apart makes this an unlikely driver of our results. In Croatia, students do not pay tuition fees if they accept the offer they receive the first time that they apply and in Sweden all higher education institutions are free.

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

our results, we should see individuals following their older siblings independently of the quality of their majors. On the other hand, we show that the effects are driven by older siblings who enroll in majors that are better in terms of student quality, retention and graduates' labor market performance. In addition, the difference found on the effects depending on older siblings' dropout suggest that the experience that they have in higher education matters, and that younger siblings are more likely to follow their older siblings when they have a good experience in higher education.

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

3.7 Conclusions

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

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

Despite the differences that exist between the three countries, we consistently find statistically and economically significant spillovers. In the three settings studied, we show that individuals are more likely to apply and enroll in the same major-college combination as their older siblings. In Chile, we document an increase of 2.8 pp (55%) in applications and 0.3 pp (30%) in enrollment; the same figures for Croatia are 3.4 pp (33%) and 1.4 pp (58%); and 3 pp (63.8%) and 0.4 pp (100%) for Sweden. These effects are stronger when individuals are more likely to be admitted in their older siblings' target major and persist even for individuals whose target and next best majors are offered by the same

institution. This suggests that the spillovers we find in the specific major-college choice are not only driven by increased preferences for older siblings' colleges.

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

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

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

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

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

Table 3.3: Probability of Applying and Enrolling in the Target Major-College of Older Siblings

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

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

Table 3.4: Probability of Applying and Enrolling in the Target Major-College of Older Siblings by Younger Siblings' Eligibility

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

Notes: These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 3.3. In addition, they have an interaction between the treatment and a proxy of younger siblings' eligibility for their older siblings' target program. Columns (1) to (3) focus on the major sample, while columns (4) to (6) on the subset of individuals whose older siblings target and counterfactual major are offered by the same college. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 3.5: Probability of Applying and Enrolling in the Target College of Older Siblings

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

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

Table 3.6: Probability of Applying and Enrolling in the Target College of Older Siblings: Large Cities Sample

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

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target college on younger siblings' probabilities of applying to and enrolling in the same college. The controls and bandwidths used in these specifications are the same described in Table 3.5. The sample only includes pairs of siblings who live in cities with at least 10 colleges and in which the older sibling target college is located in the same city. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 3.7: Probability of Applying and Enrolling in the Target Field of Study of Older Siblings

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

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

Table 3.8: Probability of Applying to the Target Major and Target College of Older Siblings by Older Siblings' Gender

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

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major and college by siblings' gender. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Tables 3.3 and 3.5. Specifications also control by a dummy variable that indicates if the siblings are of the same gender. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 3.9: Probability of Applying in the Target Major and College of Older Siblings by Siblings' Similarity

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

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major and college by siblings' similarity. Columns (1) and (3) investigate heterogeneous effects by age difference, while columns (2) and (4) by difference in high school GPA. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Tables 3.3 and 3.5. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 3.10: Probability of Applying in the Target Major and Target College of Older Siblings by Quality

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

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major or college by different quality measures of their target majors. Columns (1) and (4) investigate heterogeneous effects by the average quality of admitted students, columns (2) and (5) by first year dropout rates and columns (3) and (6) by graduates average earnings. Students' quality is measured by the average scores of admitted students in the admission exam. The measure of students quality and graduates average earnings are standardized. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Tables 3.3 and 3.5. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 3.11: Probability of Applying and Enrolling in the Target Major-College of Older Siblings by Quality Difference with respect to Counterfactual Alternative

	Major			College		
	Δ Admitted students quality (1)	Δ Dropout (2)	Δ Earnings (3)	Δ Admitted students quality (4)	Δ Dropout (5)	Δ Earnings (6)
Panel A - Chile						
Older sibling enrolls	0.028***	0.028***	0.025***	0.108***	0.101***	0.103***
	(0.006)	(0.005)	(0.005)	(0.017)	(0.016)	(0.016)
Interaction	0.000	-0.003	0.006*	-0.005	-0.165	-0.013
	(0.005)	(0.037)	(0.003)	(0.015)	(0.105)	(0.021)
Observations	99652	90784	90082	45082	41229	40836
Outcome mean	0.062	0.062	0.062	0.319	0.322	0.323
Bandwidth	20.000	20.000	20.000	15.000	15.000	15.000
F-statistics	7674.012	7397.956	7219.418	3153.688	2959.387	2908.442
Panel B - Croatia						
Older sibling enrolls	0.034***			0.107***		
	(0.009)			(0.021)		
Interaction	-0.003			0.007		
	(0.005)			(0.010)		
Observations	34510			10693		
Mean y	0.130			0.537		
Bandwidth	80.000			80.000		
F-statistics	6854.732			2607.328		
Panel C - Sweden						
Older sibling enrolls	0.033***	0.017**	0.233***	0.185***	0.116***	0.142***
	(0.006)	(0.006)	(0.008)	(0.015)	(0.014)	(0.020)
Interaction	-0.015***	-0.002	-0.004	-0.053***	-0.009	-0.021**
	(0.003)	(0.002)	(0.003)	(0.010)	(0.007)	(0.009)
Observations	472966	309934	210261	262275	172027	117555
Mean y	0.054	0.053	0.051	0.200	0.196	0.201
Bandwidth	0.510	0.510	0.510	0.367	0.367	0.367
F-statistics	4439.812	4419.105	2264.171	4439.812	4419.105	1125.23

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major and college by the gap between older siblings' target and counterfactual major in different quality measures. Columns (1) and (4) investigate heterogeneous effects by the difference in the average quality of admitted students, columns (2) and (5) by the difference in first year dropout rates and columns (3) and (6) by the difference in graduates average earnings. Students' quality is measured by the average scores of admitted students in the admission exam. The measure of students quality and graduates average earnings are standardized. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 3.3. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. In this table, the sample is restricted to older siblings with counterfactual programs in their application lists. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 3.12: Probability of Applying and Enrolling in the Target Major and Target College of Older Siblings by Older Siblings' Dropout

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

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major on younger siblings' probability of applying to and enrolling in that major. The specifications include the same controls and use the same bandwidths described in Tables 3.3 and 3.5. They also control for a dummy variable that indicates if older siblings dropout from the major in which they initially enroll. The samples used in these last columns only include individuals whose older siblings enroll in a major. In parenthesis, standard errors clustered at family level.*p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table 3.13: Effect of Older Siblings' Enrollment in the Target Major-College on Academic Performance (Major Sample)

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

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major on younger siblings' probability of taking the admission exam and applying to college (columns 1 and 2), and on different measures of academic performance: high school GPA (column 3), reading and math sections of the admission exam (columns 4 and 5) and average performance on the admission exam (column 6). While in Chile and Croatia we only observe applications to college degrees, in Sweden we also observe applications to other higher education programs. These analyses focus on the Major Sample. This means that in this case, marginal admission or rejection from their target major, changes the major, but not necessarily the college or field in which older siblings are admitted. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 3.5. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

.1 Identification Strategy: Further Discussion

This section discusses the assumptions under which our identification strategy provides us with a consistent estimator of the effects of interest. As discussed in Section 3.4.4, a fuzzy RD can be thought of of as an IV. In what follows, and for ease of notation, we drop time and individual indices t, i, τ and focus our analysis on a specific major-college u. Following this notation, the treatment in which we are interested is:

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

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

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

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

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

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

1. Independence of the instrument:

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

2. Exclusion restriction:

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

3. First stage:

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

4. Monotonicity:

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

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

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

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

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

Proposition 3. *Under assumptions* $1 - 4$*:*

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

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

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

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

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

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

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

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

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

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

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

Taken together, 3 and 4 imply:

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

□

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

.2 Robustness Checks

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

.2.1 Manipulation of the Running Variable

A first condition for the validity of our RD estimates is that individuals should not be able to manipulate their older siblings' application scores around the admission cutoff. The structures of the admission systems in Chile, Croatia and Sweden make the violation of this assumption unlikely. However, to confirm this we study whether the distribution of the running variable (i.e. older sibling's application score centered around the relevant cutoff) is continuous at the cutoff. We do this by implementing the test suggested by Cattaneo et al. (2018), the results of which are presented in Figure B1. As expected, we do not detect discontinuities in the distribution of the running variable at the cutoff for any of the three countries.[58] In Sweden, Figure B1 only focuses on the distribution of the high school GPA. As discussed in Section 3.2, the admission exam is voluntary in Sweden, and institutions select their students using two independent pools that consider either the applicants' high school GPA or the applicants' scores in the admission exam. Considering that the distribution of admission exam scores is coarser, to investigate manipulation of these scores we present histograms of these variables.

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

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

.2.2 Discontinuities in Potential Confounders

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

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

Figure B2 summarizes this result. It plots the estimated discontinuities at the cutoff and their 95% confidence intervals. To estimate these discontinuities we control for a linear polynomial of the running variable and allow for the slope to change at the cutoff. Using the same bandwidths reported for linear specifications in Section 3.5, we find no statistically significant jump at the cutoff for any of the potential confounders being investigated.

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

.2.3 Different Bandwidths

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

Figures B3, B4 and B5 show how the estimated coefficients change when reducing the bandwidth used in the estimations. Although the standard errors increase as the sample size is reduced, the coefficients remain stable.

.2.4 Placebo Exercises

This setting allows us to perform two types of placebo exercises. First, in Figures B9, B10 and B11 we study if younger siblings' enrollment affect the application decisions of their older siblings. Since younger siblings apply to college after their older siblings, being marginally admitted or rejected from a major or college should not affect what happens with older siblings. These figures show that this is indeed the case. Even though when looking at the placebo on college choice in Sweden we find some discontinuities at the cutoff, their size is considerably smaller than the ones documented in the main body of the paper. In addition, in Figures B6, B7 and B8 we show that only at the real cutoff we observe a discontinuity on younger siblings' outcomes This is not surprising since these

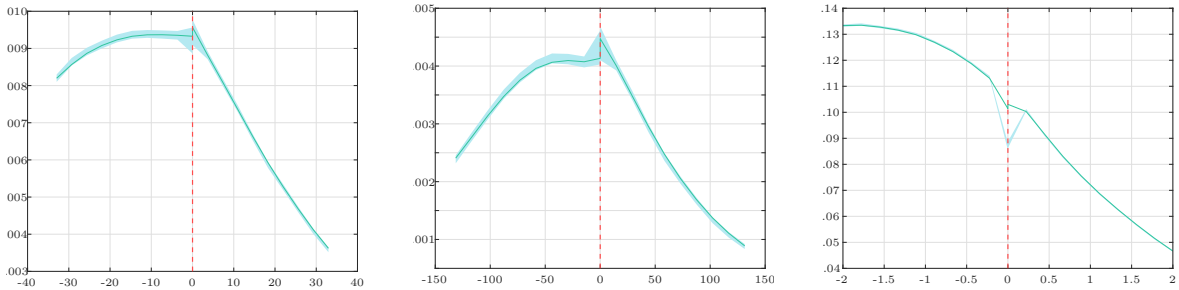
fake cutoffs do not generate any increase in older siblings' admission.

.2.5 Alternative Specifications and Total Enrollment

Figures B12, B13 and B14 and Tables B1, B2, B3, B5, B6 and B7 study how robust our estimates are to the degree of the local polynomial used, to re-weighting the observations by the inverse of the total number of applicants in the proximity of each cutoff and to allowing the running variable to have different slopes for each cutoff-major. In addition, Tables B8, B9 and B10 present results in which target $\times$ counterfactual major fixed effects are used. The results are robust to these changes, and although the magnitude of the coefficients is smaller when re-weighting the observations, the general picture remains unchanged.

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

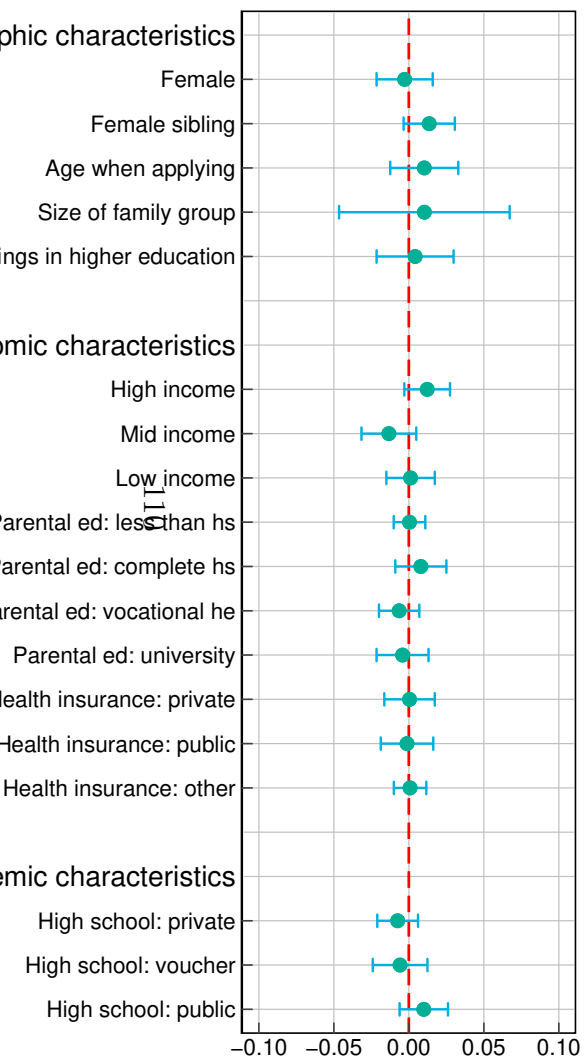
Figure B1: Density of Older Siblings' Application Scores at the Target Major-College Admission Cutoff



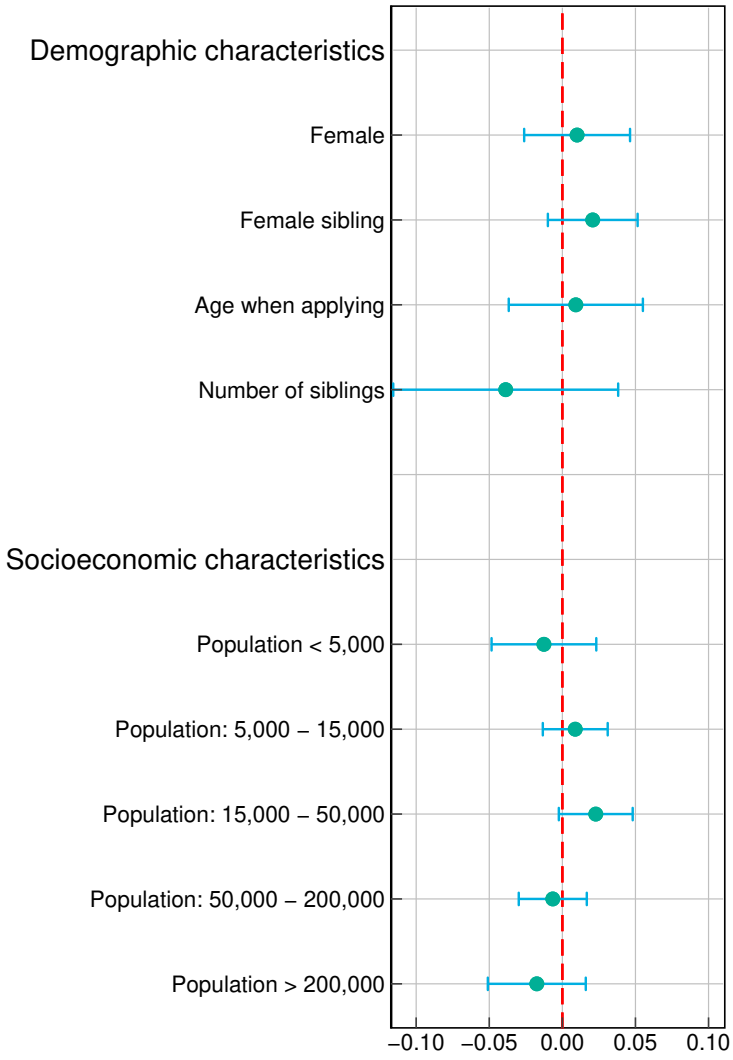
(a) 1st preference - Chile (b) 1st preference - Croatia (c) 1st preference - Sweden

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

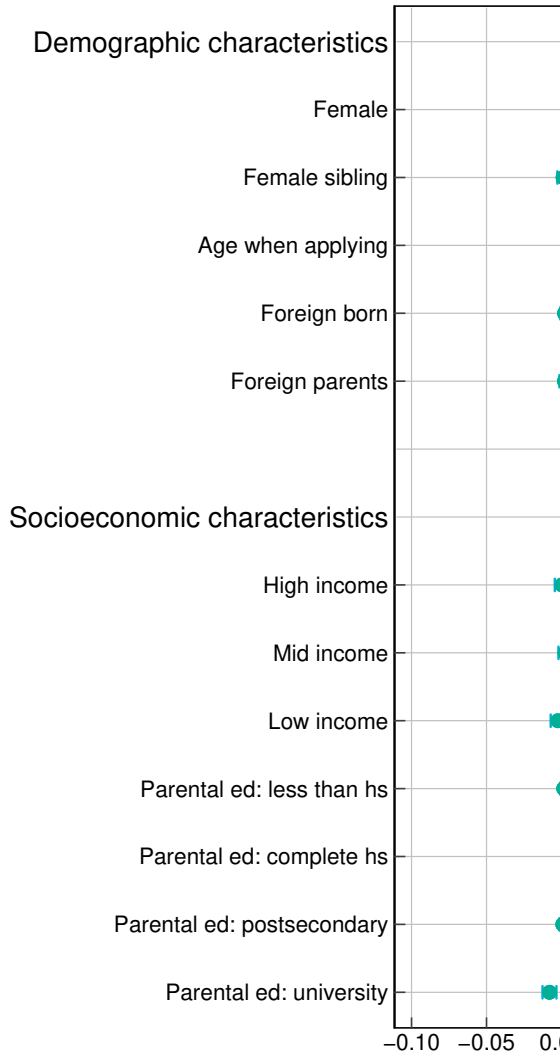
Figure B2: Discontinuities in other Covariates at the Cutoff



(a) Chile



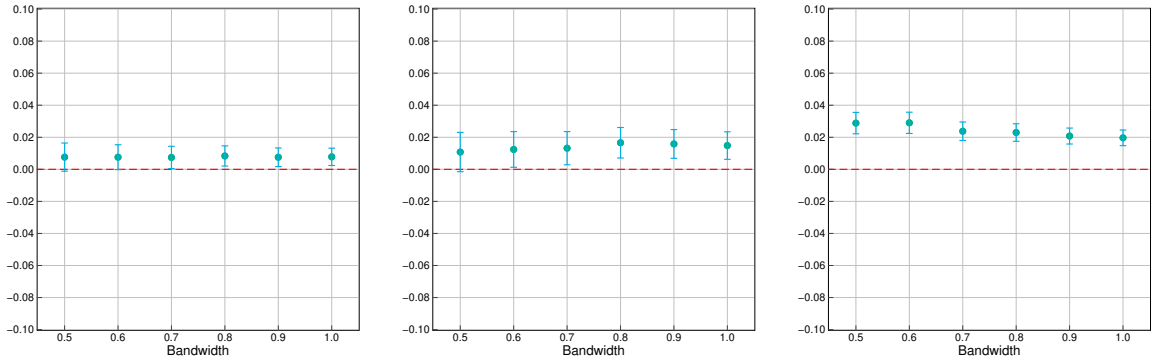
(b) Croatia



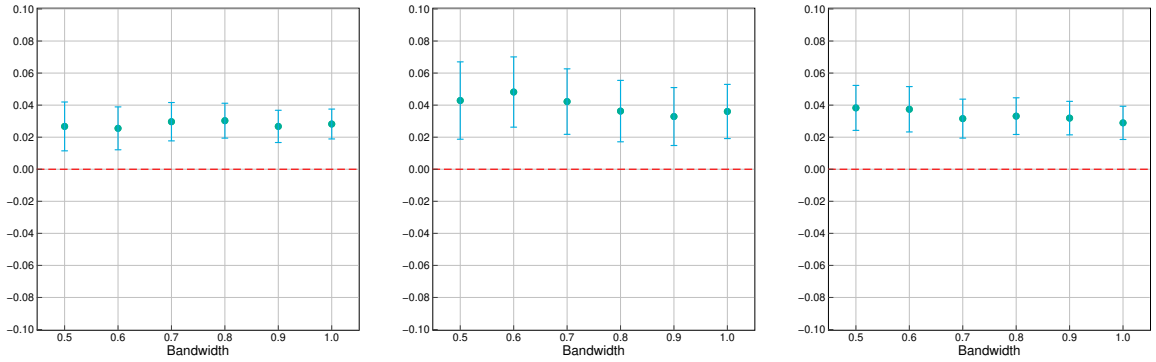
(c) Sweden

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

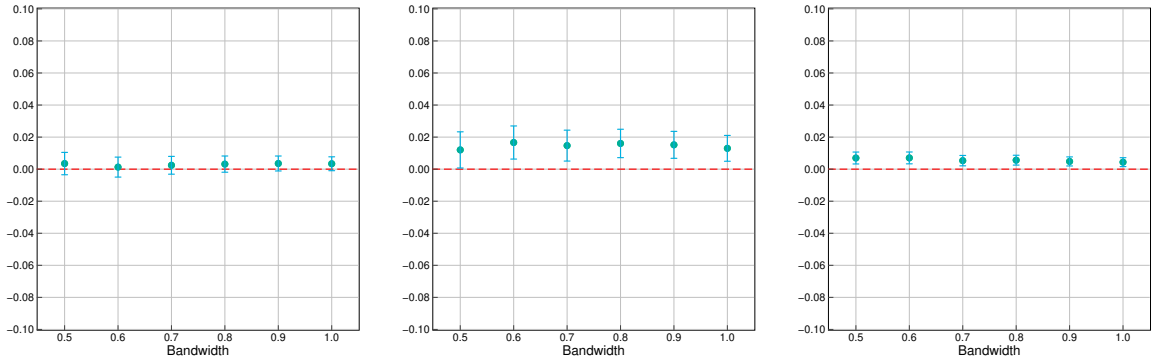
Figure B3: Probabilities of Applying and Enrolling in the Target Major-College of Older Siblings - Different Bandwidths



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



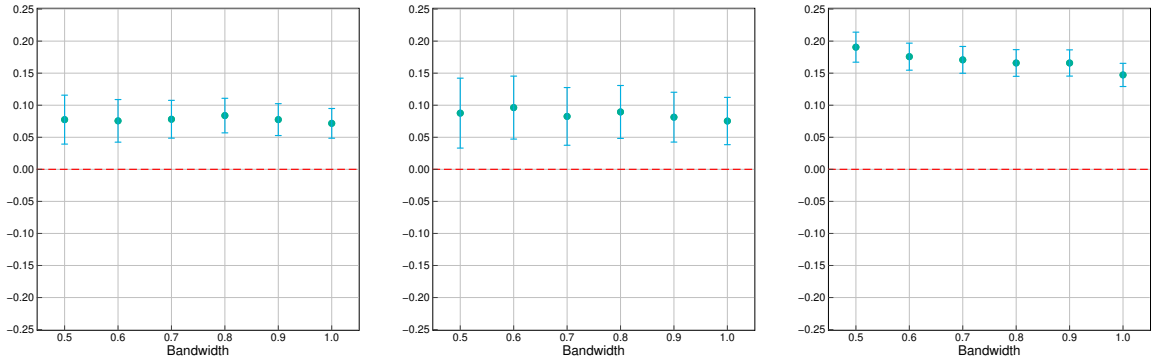
(d) Any preference - Chile (e) Any preference - Croatia (f) Any preference - Sweden



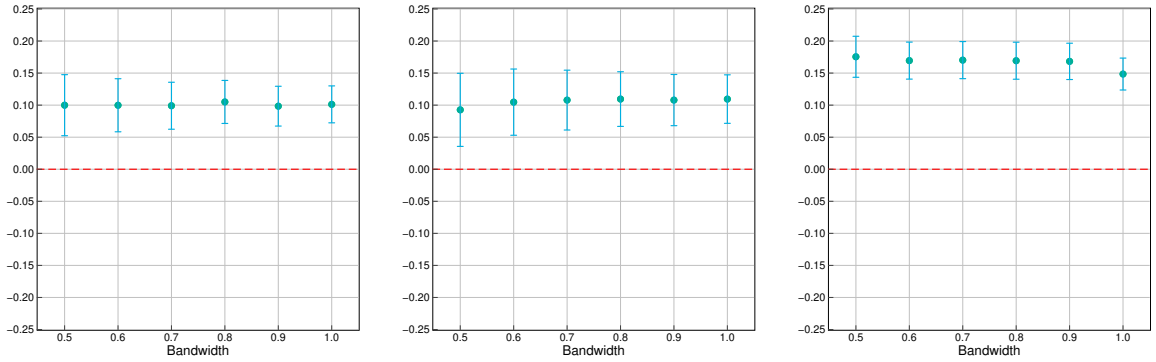
(g) Enrolls - Chile (h) Enrolls - Croatia (i) Enrolls - Sweden

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

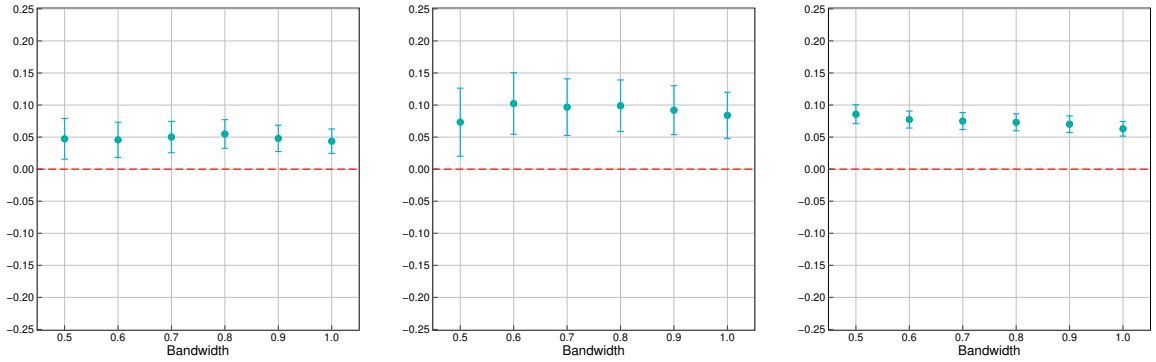
Figure B4: Probabilities of Applying and Enrolling in the Target College of Older Siblings - Different Bandwidths



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



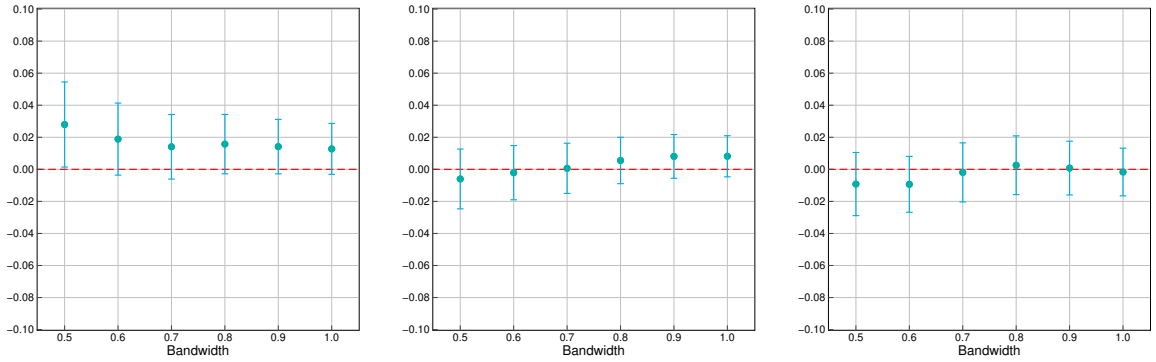
(d) Any preference - Chile (e) Any preference - Croatia (f) Any preference - Sweden



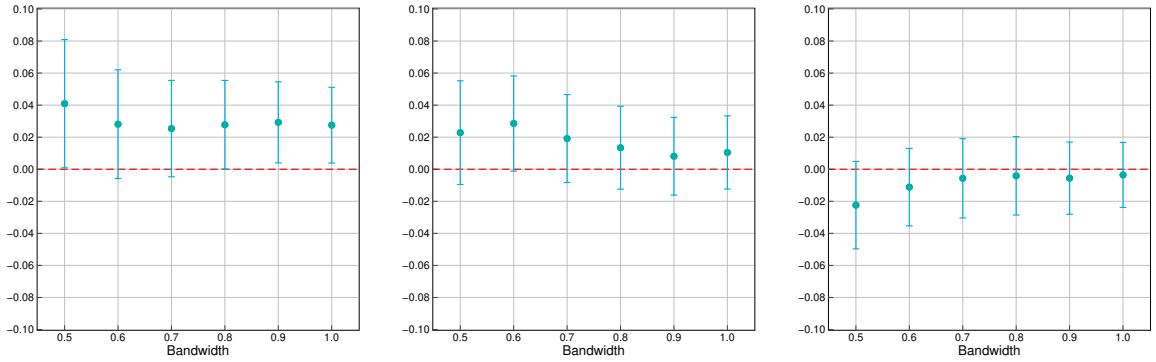
(g) Enrolls - Chile (h) Enrolls - Croatia (i) Enrolls - Sweden

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

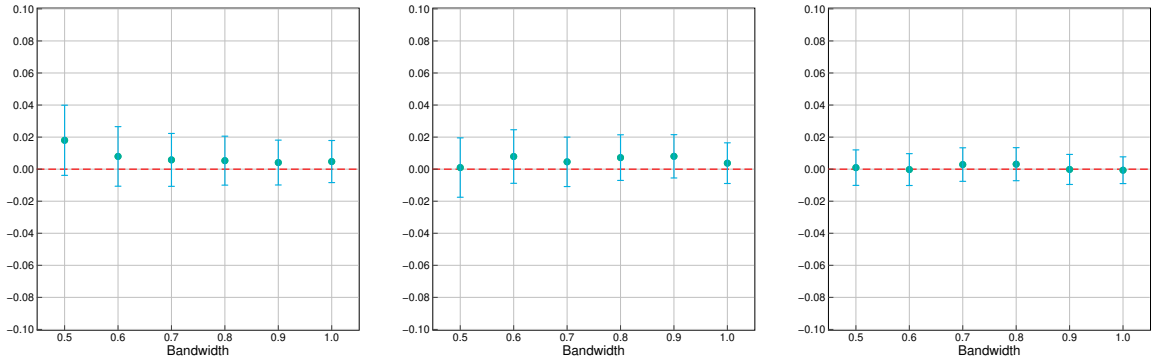
Figure B5: Probabilities of Applying and Enrolling in the Target Field of Study of Older Siblings - Different Bandwidths



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



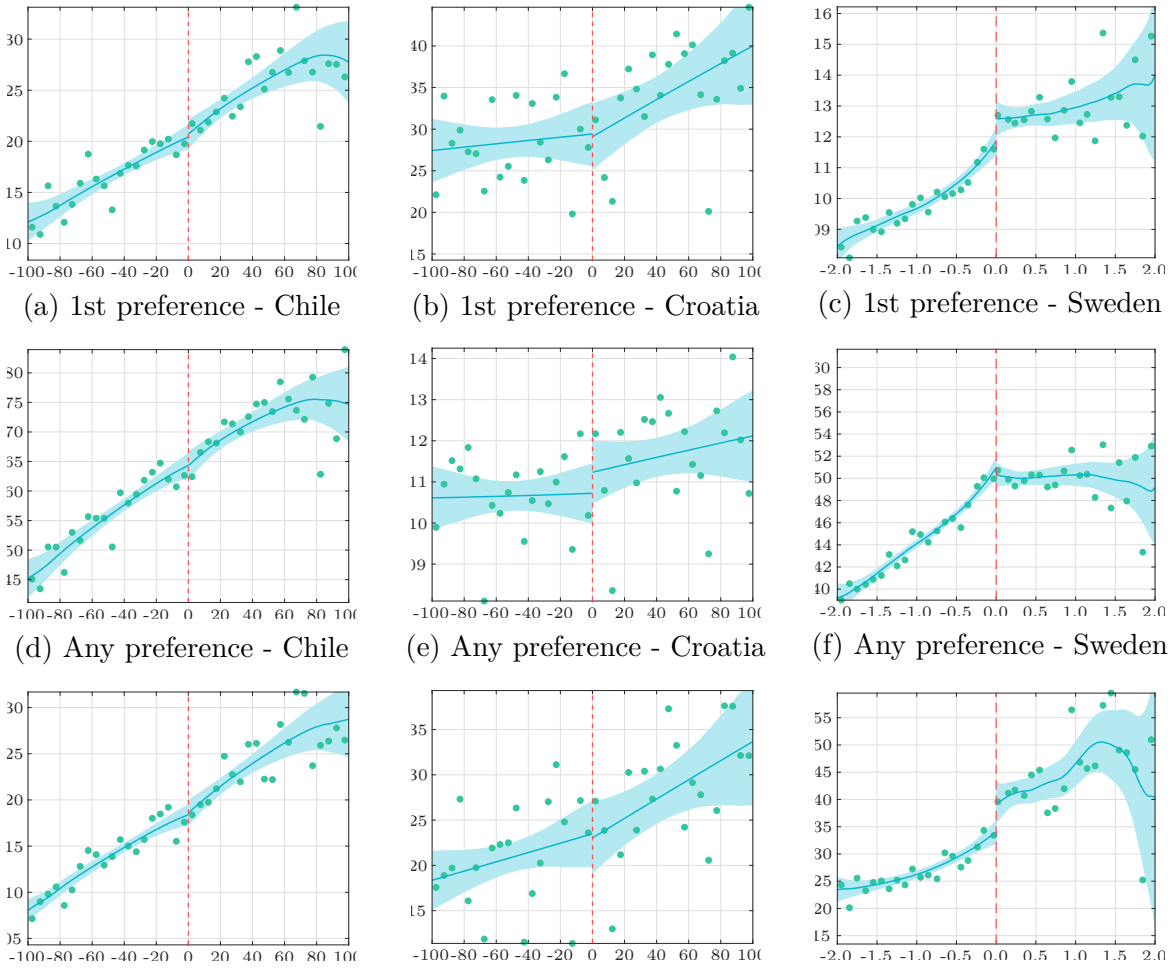
(d) Any preference - Chile (e) Any preference - Croatia (f) Any preference - Sweden



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

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

Figure B6: Placebo - Probabilities of Applying and Enrolling in the Target Major-College of Younger Siblings



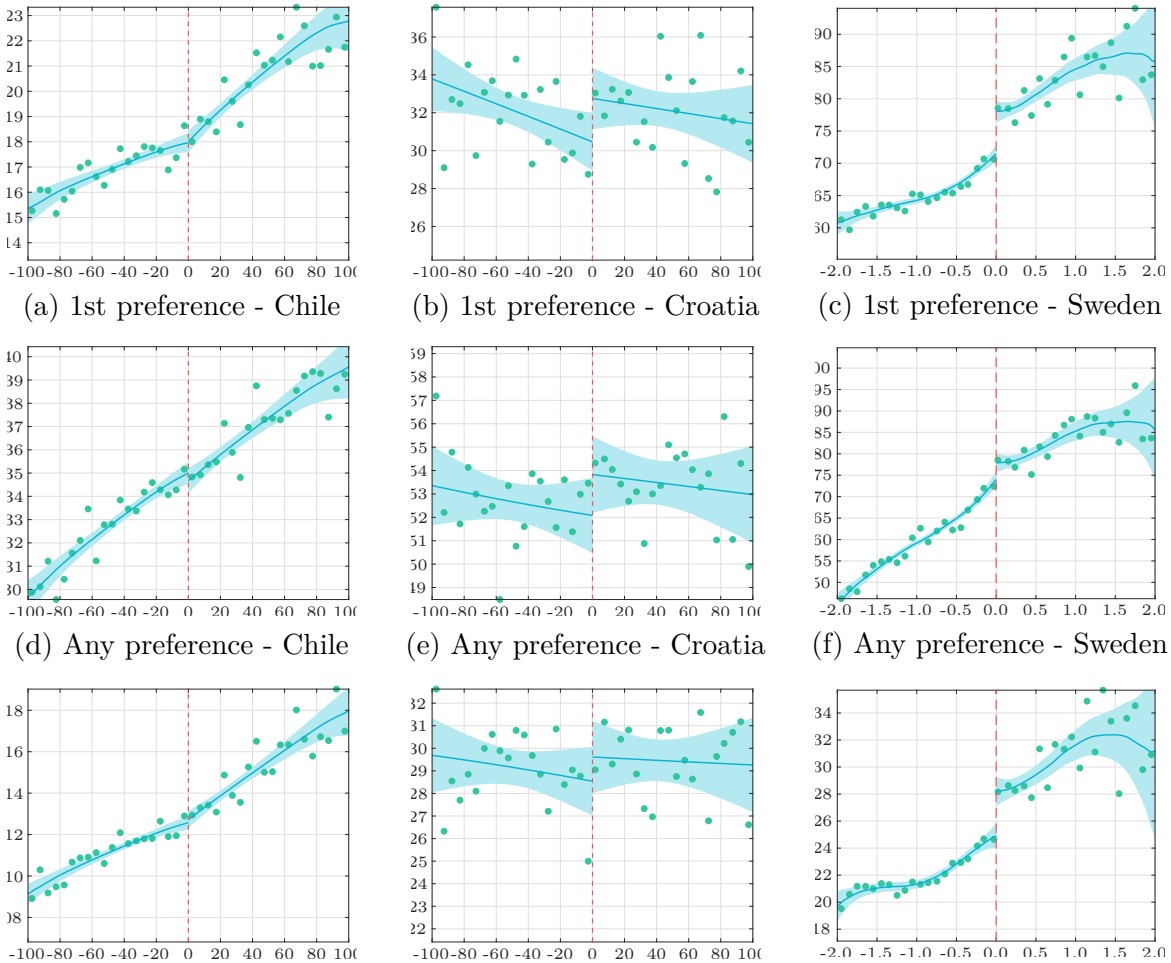
(a) 1st preference - Chile (b) 1st preference - Croatia (c) 1st preference - Sweden

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

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

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

Figure B7: Placebo - Probabilities of Applying and Enrolling in the Target College of Younger Siblings



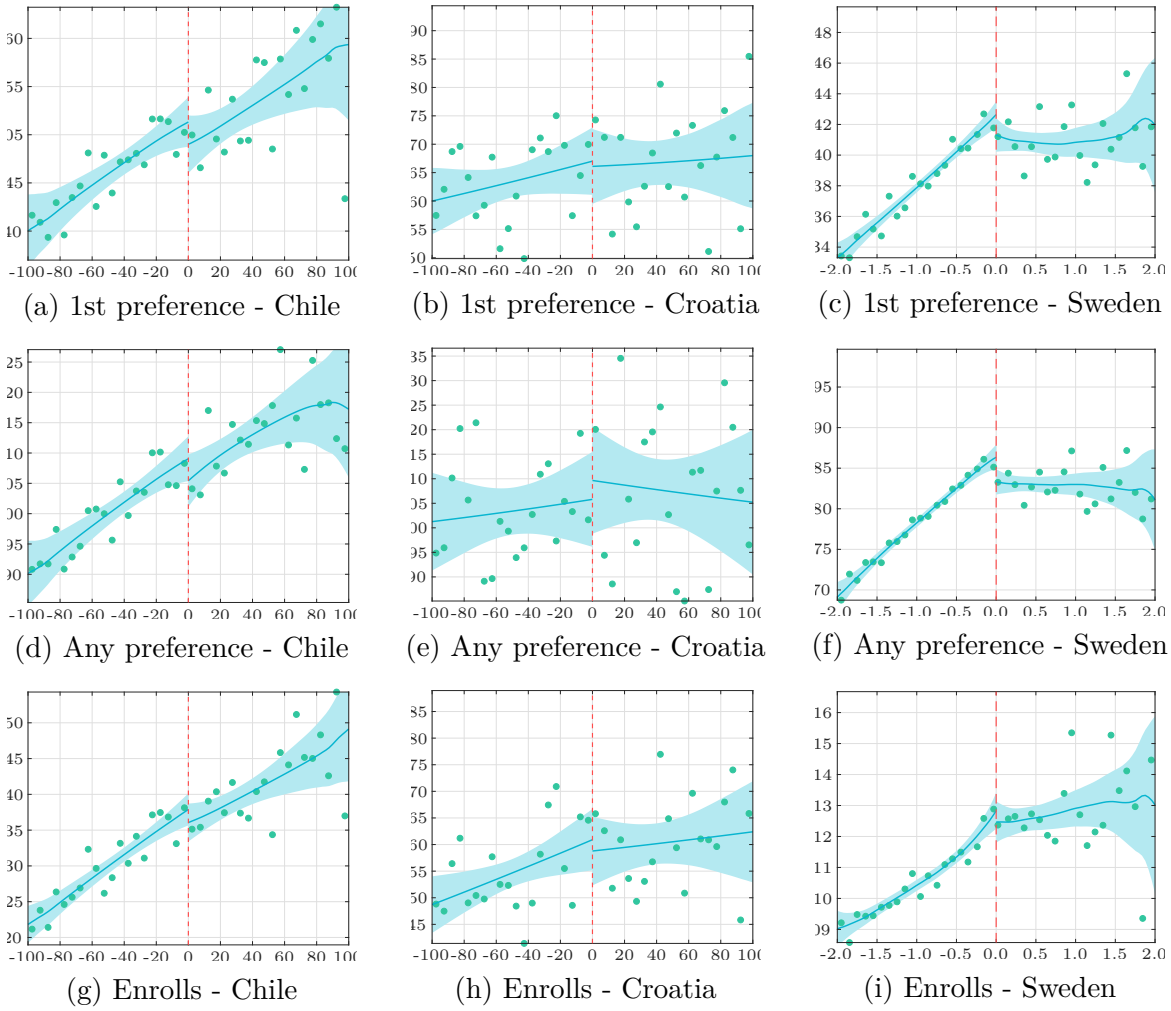
(a) 1st preference - Chile (b) 1st preference - Croatia (c) 1st preference - Sweden

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

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

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

Figure B8: Placebo - Probabilities of Applying and Enrolling in the Target Field of Study of Younger Siblings



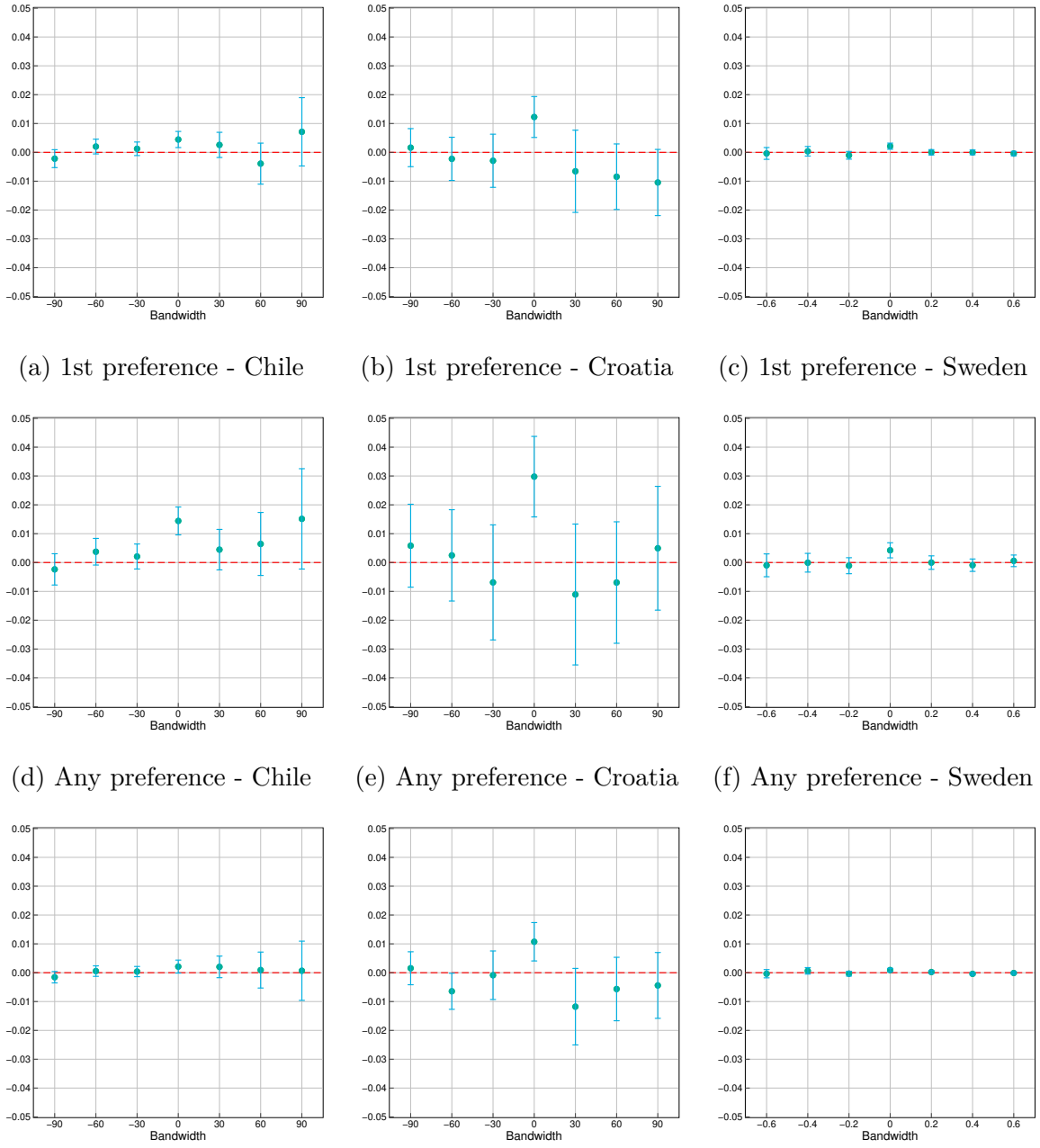
(a) 1st preference - Chile (b) 1st preference - Croatia (c) 1st preference - Sweden

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

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

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

Figure B9: Placebo Cutoffs - Probabilities of Applying and Enrolling in the Target Major-College of Older Siblings



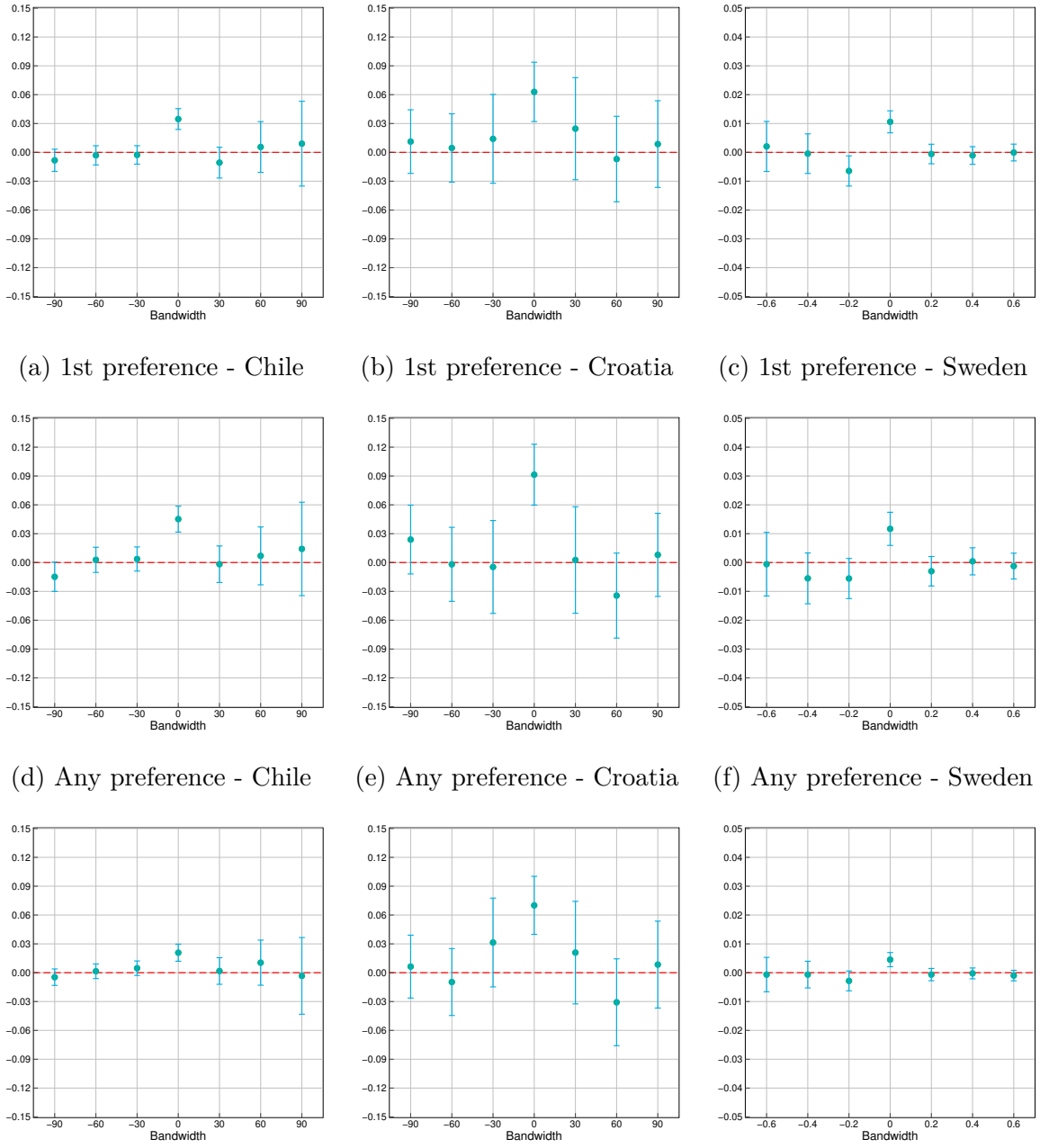
(a) 1st preference - Chile (b) 1st preference - Croatia (c) 1st preference - Sweden

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

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

This figure illustrates the results of a placebo exercise that investigates if effects similar to the ones documented in figure 3.2 arise at different values of the running variable. Therefore, the x-axis corresponds to different (hypothetical) values of cutoffs - 0 corresponds to the actual cutoff used in the main body of the paper. The other values correspond to points where older siblings' probability of being admitted to their target major is continuous. Blue points illustrate estimated effect, and the blue bars denote the 95% confidence intervals. Figures (a), (d) and (g) illustrate the case of Chile, figures (b), (e) and (h) the case of Croatia, while figures (c), (f) and (i) the case of Sweden.

Figure B10: Placebo Cutoffs - Probabilities of Applying and Enrolling in the Target College of Older Siblings



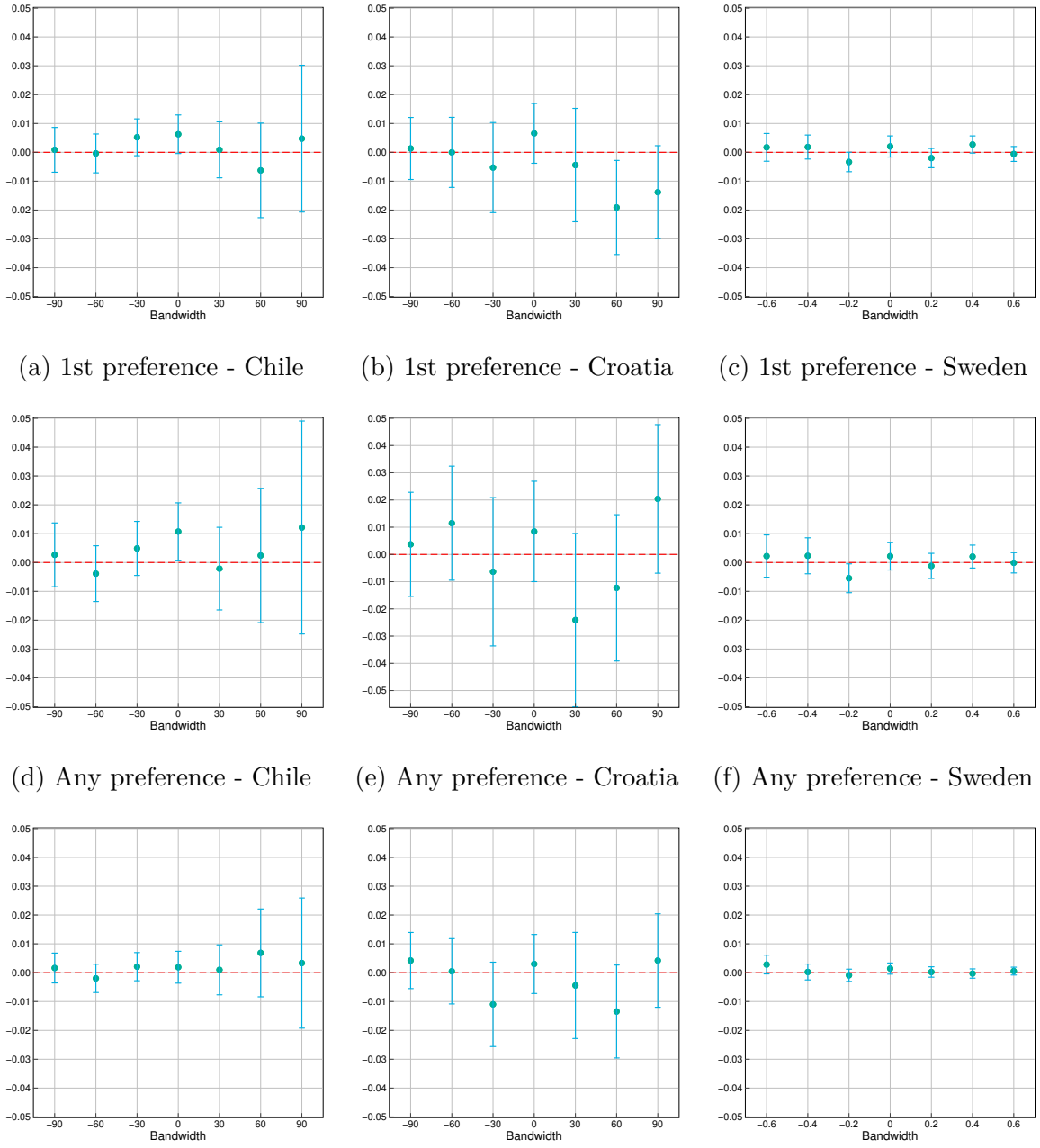
(a) 1st preference - Chile (b) 1st preference - Croatia (c) 1st preference - Sweden

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

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

This figure illustrates the results of a placebo exercise that investigates if effects similar to the ones documented in figure 3.3 arise at different values of the running variable. Therefore, the x-axis corresponds to different (hypothetical) values of cutoffs - 0 corresponds to the actual cutoff used in the main body of the paper. The other values correspond to points where older siblings' probability of being admitted to their target majors is continuous. Blue points illustrate estimated effect, and the blue bars denote the 95% confidence intervals. Figures (a), (d) and (g) illustrate the case of Chile, figures (b), (e) and (h) the case of Croatia, while figures (c), (f) and (i) the case of Sweden.

Figure B11: Placebo Cutoffs - Probabilities of Applying and Enrolling in the Target Field of Study of Older Siblings



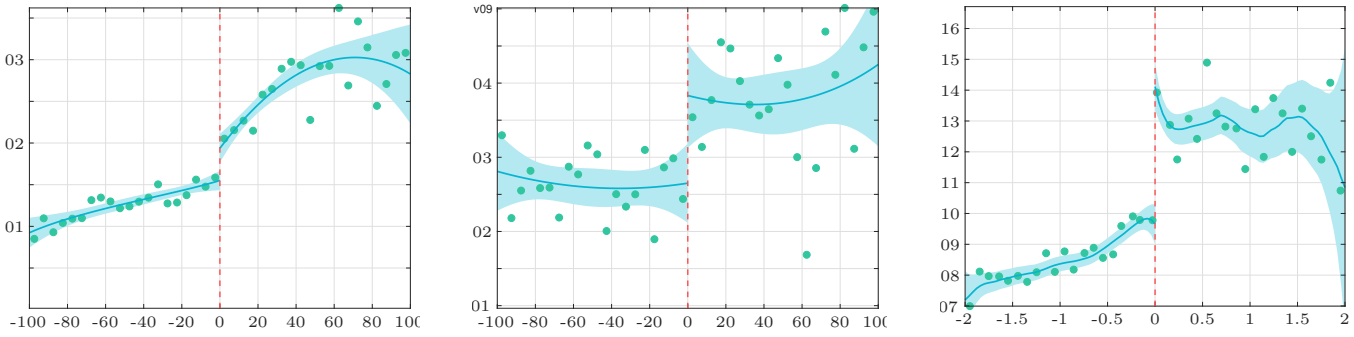
(a) 1st preference - Chile (b) 1st preference - Croatia (c) 1st preference - Sweden

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

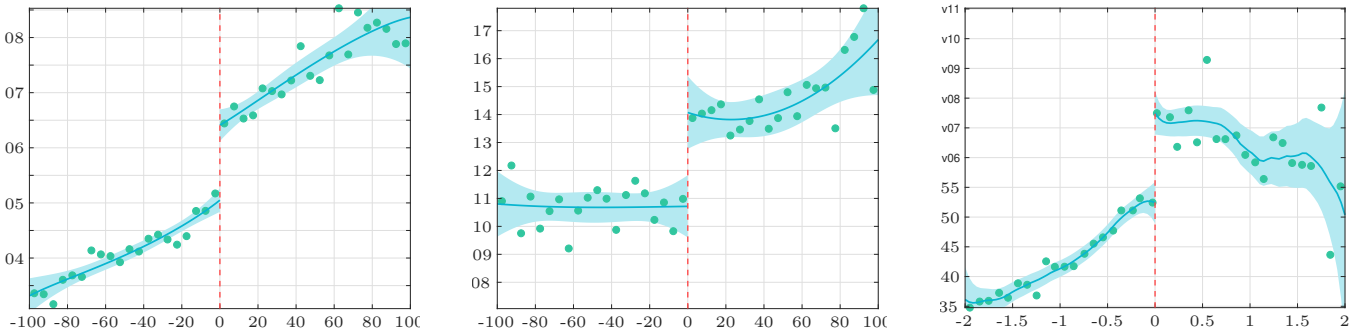
(g) Enrolls - Chile (h) Enrolls - Croatia (i) Enrolls - Sweden

This figure illustrates the results of a placebo exercise that investigates if effects similar to the ones documented in figure 3.4 arise at different values of the running variable. Therefore, the x-axis corresponds to different (hypothetical) values of cutoffs - 0 corresponds to the actual cutoff used in the main body of the paper. The other values correspond to points where older siblings' probability of being admitted to their target major is continuous. Blue points illustrate estimated effect, and the blue bars denote the 95% confidence intervals. Figures (a), (d) and (g) illustrate the case of Chile, figures (b), (e) and (h) the case of Croatia, while figures (c), (f) and (i) the case of Sweden.

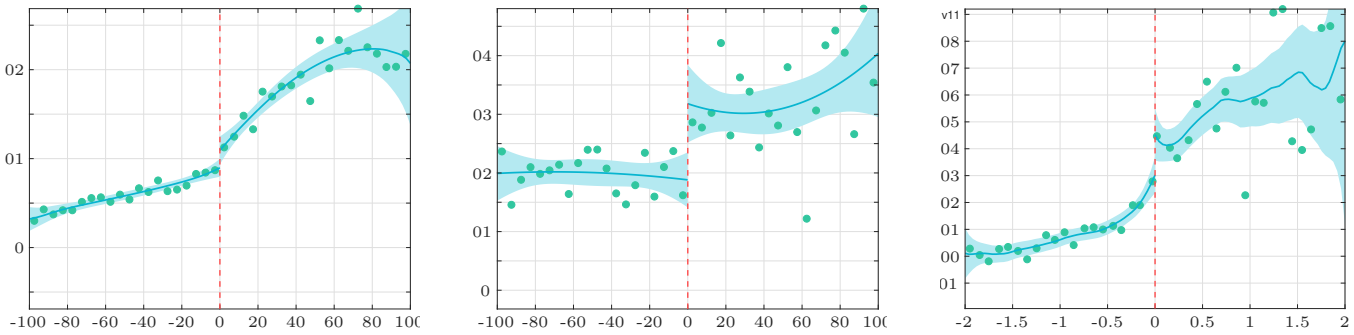
Figure B12: Probabilities of Applying and Enrolling in the Target Major-College of Older Siblings (Polynomial of degree 2)



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



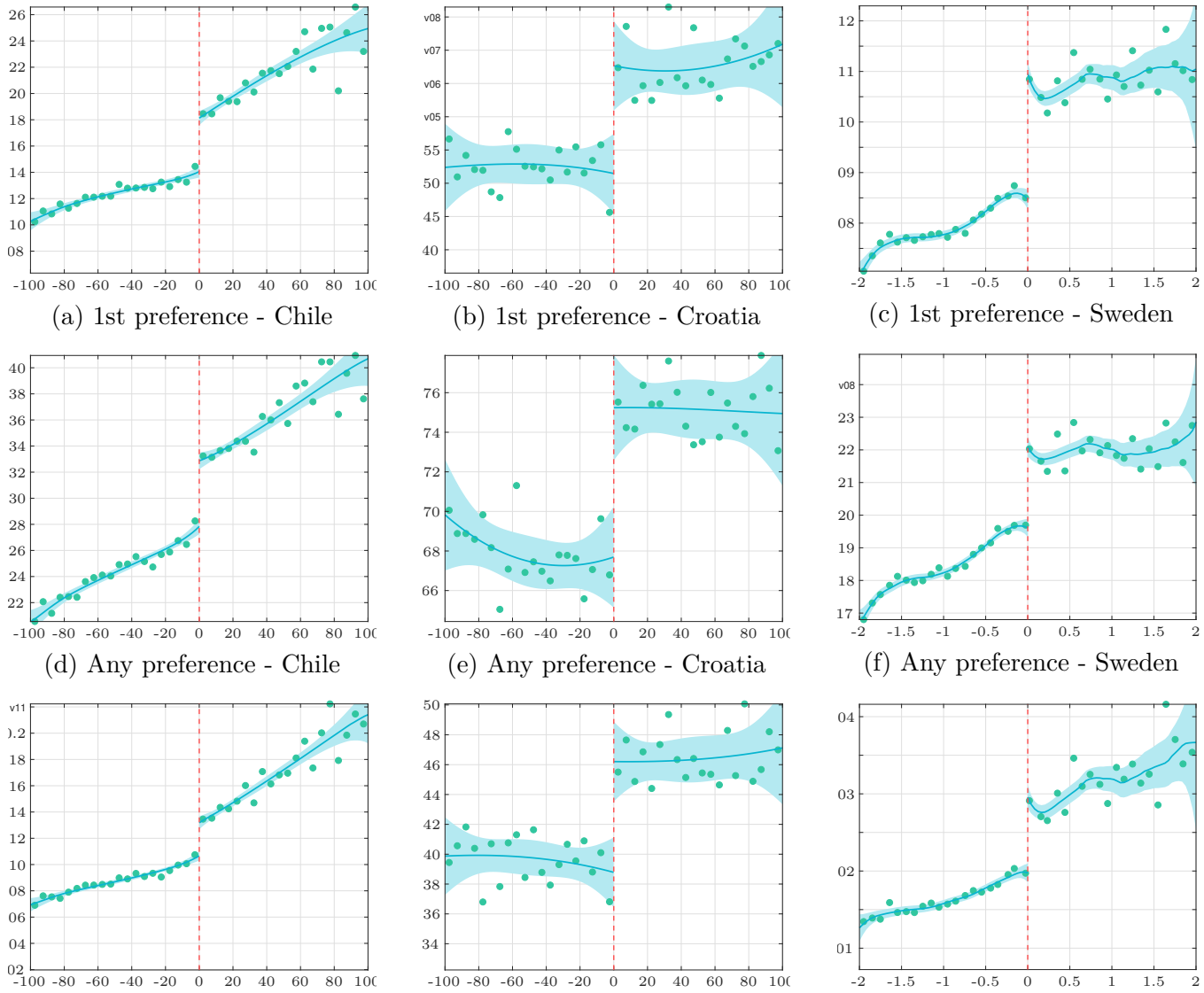
(d) Any preference - Chile (e) Any preference - Croatia (f) Any preference - Sweden



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

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

Figure B13: Probabilities of Applying and Enrolling in the Target College of Older Siblings (Polynomial of degree 2)



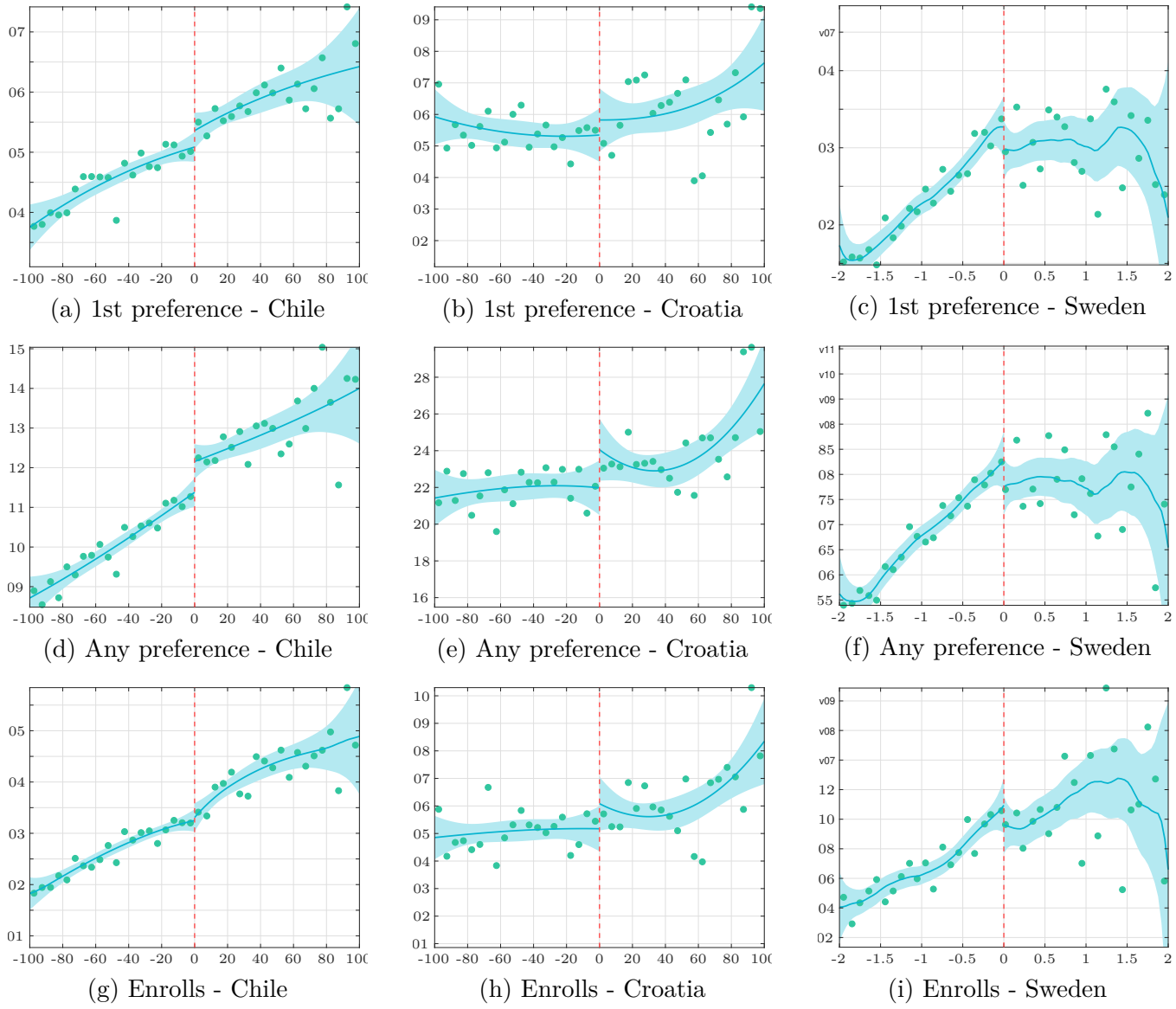
(a) 1st preference - Chile (b) 1st preference - Croatia (c) 1st preference - Sweden

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

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

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

Figure B14: Probabilities of Applying and Enrolling in the Target Field of Study of Older Siblings (Polynomial of degree 2)



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

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

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

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

Table B1: Probability of Applying and Enrolling in the Target Major of Older Siblings - Reweighting

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

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

Table B2: Probability of Applying and Enrolling in the Target College of Older Siblings - Reweighting

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

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

Table B3: Probability of Applying and Enrolling in the Target Field of Older Siblings - Reweighting

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

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

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

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

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

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

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

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

Table B6: Probability of Applying and Enrolling in the Target College of Older Siblings - Different Slope for each Admission Cutoff

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

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

Table B7: Probability of Applying and Enrolling in the Target Field of Older Siblings - Different Slope for each Admission Cutoff

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

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

Table B8: Probability of Applying and Enrolling in the Target Major-College of Older Siblings - Target × Counterfactual Major Fixed Effects

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

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

Table B9: Probability of Applying and Enrolling in the Target College of Older Siblings - Target × Counterfactual Major Fixed Effects

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

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

Table B10: Probability of Applying and Enrolling in the Target Field of Older Siblings - Target × Counterfactual Major Fixed Effects

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

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

.3 Additional Results

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

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

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

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

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major and college by siblings' gender. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Tables 3.3 and 3.5. Specifications also control by a dummy variable that indicates if the siblings are of the same gender. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

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

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

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target field of study by siblings' gender. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 3.5. Specifications also control by a dummy variable that indicates if the siblings are of the same gender. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C3: Probability of Enrolling in the Target Major and Target College of Older Siblings by Siblings' Similarity

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

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major and college by siblings' similarity. Columns (1) and (3) investigate heterogeneous effects by age difference, while columns (2) and (4) by difference in high school GPA. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Tables 3.3 and 3.5. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

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

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

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target field of study by siblings' similarity. Columns (1) and (3) investigate heterogeneous effects by age difference, while columns (2) and (4) by difference in high school GPA. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 3.5. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C5: Probability of Enrolling in the Target Major and College of Older Siblings by Quality

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

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major and college by different quality measures of their target majors. Columns (1) and (4) investigate heterogeneous effects by the average quality of admitted students, columns (2) and (5) by first year dropout rates and columns (3) and (6) by graduates average earnings. Students' quality is measured by the average scores of admitted students in the admission exam. The measure of students quality and graduates average earnings are standardized. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Tables 3.3 and 3.5. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

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

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

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target field by different quality measures of their target programs. Columns (1) and (4) investigate heterogeneous effects by the average quality of admitted students, columns (2) and (5) by first year dropout rates and columns (3) and (6) by graduates average earnings. Students' quality is measured by the average scores of admitted students in the admission exam. The measure of students quality and graduates average earnings are standardized. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 3.7. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C7: Probability of Enrolling in the Target Major and College of Older Siblings by Quality Difference respect to Counterfactual Alternative

	Major			College		
	Δ Admitted students quality (1)	Δ Dropout (2)	Δ Earnings (3)	Δ Admitted students quality (4)	Δ Dropout (5)	Δ Earnings (6)
Panel A - Chile						
Older sibling enrolls	0.005	0.006*	0.005	0.044***	0.042***	0.042***
	(0.003)	(0.002)	(0.002)	(0.011)	(0.011)	(0.011)
Interaction	-0.001	0.017	0.000	-0.002	-0.120	-0.016
	(0.002)	(0.016))	(0.001)	(0.010)	(0.066)	(0.013)
Observations	99652	90784	90082	45082	41229	40836
Outcome mean	.013	0.013	0.013	0.105	0.106	0.106
Bandwidth	20.000	20.000	20.000	15.000	15.000	15.000
F-statistics	7674.012	7397.956	7219.418	3153.688	2959.387	2908.442
Panel B - Croatia						
Older sibling enrolls	0.013**			0.101***		
	(0.004)			(0.020)		
Interaction	0.002			0.007		
	(0.002)			(0.010)		
Observations	34510			10693		
Outcome mean	0.024			0.268		
Bandwidth	80.000			80.000		
F-statistics	6854.732			2607.328		
Panel C - Sweden						
Older sibling enrolls	0.006***	0.004**	0.005**	0.071***	0.049***	0.056***
	(0.002)	(0.002)	(0.002)	(0.007)	(0.007)	(0.009)
Interaction	-0.002	0.000	0.000	-0.016***	-0.005	-0.000
	(0.001)	(0.001)	(0.001)	(0.005)	(0.004)	(0.005)
Observations	472966	309934	210261	262275	172027	117555
Outcome mean 0.004	0.005	0.004	0.032	0.036	0.036	
Bandwidth	0.510	0.510	0.510	0.367	0.367	0.367
F-statistics	4439.812	4419.105	2264.171	2282.347	2063.087	1125.23

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major and college by the gap between older siblings' target and counterfactual major in different quality measures. Columns (1) and (4) investigate heterogeneous effects by the difference in the average quality of admitted students, columns (2) and (5) by the difference in first year dropout rates and columns (3) and (6) by the difference in graduates average earnings. Students quality is measured by the average scores of admitted students in the admission exam. The measure of students quality and graduates average earnings are standardized. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Tables 3.3 and 3.5. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. In this table, the sample is restricted to older siblings with counterfactual programs in their application lists. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Table C8: Probability of Applying and Enrolling in the Target Field of Study of Older Siblings by Difference in Quality respect Counterfactual Alternative

	Applies			Enrolls		
	Δ Admitted students quality (1)	Δ Dropout (2)	Δ Earnings (3)	Δ Admitted students quality (4)	Δ Dropout (5)	Δ Earnings (6)
Panel A - Chile						
Older sibling enrolls	0.012	0.013	0.012	-0.002	-0.006	-0.006
	(0.013)	(0.012)	(0.012)	(0.007)	(0.007)	(0.007)
Interaction	0.006	0.022	0.001	0.000	0.059	-0.001
	(0.012)	(0.077)	(0.005)	(0.006)	(0.040)	(0.003)
Observations	45591	40142	39660	45591	40142	39660
Outcome mean	0.122	0.124	0.125	0.034	0.035	0.035
Bandwidth	15.000	15.000	15.000	15.000	15.000	15.000
F-statistics	2608.326	2397.713	2325.023	2608.326	2397.713	2325.023
Panel B - Croatia						
Older sibling enrolls	0.005			0.000		
	(0.012)			(0.007)		
Interaction	0.010			0.005		
	(0.006)			(0.004)		
Observations	29466			29466		
Outcome mean	0.218			0.053		
Bandwidth	80.000			80.000		
F-statistics	4707.803			4707.803		
Panel C - Sweden						
Older sibling enrolls	0.012	-0.006	0.009	0.006	0.003	0.009
	(0.014)	(0.013)	(0.018)	(0.006)	(0.006)	(0.008)
Interaction	-0.023***	-0.005	-0.005	-0.004	0.001	-0.002
	(0.007)	(0.004)	(0.005)	(0.003)	(0.002)	(0.002)
Observations	207042	126204	85936	207042	126204	85936
Outcome mean	0.094	0.090	0.091	0.015	0.016	0.016
Bandwidth	0.390	0.390	0.390	0.390	0.390	0.390
F-statistics	1746.185	1454.422	746.375	1746.185	1454.422	746.375

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target field of study by the gap between older siblings' target and counterfactual program in different quality measures. Columns (1) and (4) investigate heterogeneous effects by the difference in average quality of admitted students, columns (2) and (5) by the difference in first year dropout rates and columns (3) and (6) by the difference in graduates average earnings. Students' quality is measured by the average scores of admitted students in the admission exam. The measure of students quality and graduates average earnings are standardized. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 3.5. In addition, we add as control the main effect of the interaction used in each column. In parenthesis, standard errors clustered at family level. In this table, the sample is restricted to older siblings with counterfactual programs in their application lists. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

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

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

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major on younger siblings' probability of taking the admission exam and applying to college (columns 1 and 2), and on different measures of academic performance: high school GPA (column 3), reading and math sections of the admission exam (columns 4 and 5) and average performance on the admission exam (column 6). While in Chile and Croatia we only observe applications to college degrees, in Sweden we also observe applications to other higher education programs. These analyses focus on the College Sample. This means that in this case, marginal admission or rejection from their target major, changes the college in which older siblings are admitted. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 3.5. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

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

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

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target field on younger siblings' probability of taking the admission exam and applying to university (columns 1 and 2), and on different measures of academic performance: high school GPA (column 3), reading and math sections of the admission exam (columns 4 and 5) and average performance on the admission exam (column 6). While in Chile and Croatia we only observe applications to university degrees, in Sweden we also observe applications to other higher education programs. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 3.7. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

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

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

Notes: The table presents 2SLS estimates for the effect of older siblings' marginal enrollment in their target major on high school GPA (column 1) and on average performance on the admission exam (column 2). The effect is allowed to vary with age difference between siblings. These specifications use the same set of controls and bandwidths used in the 2SLS specifications described in Table 3.3. Age difference between siblings is added as control. In parenthesis, standard errors clustered at family level. *p-value<0.1 **p-value<0.05 ***p-value<0.01.

Bibliography

Abdulkadiroglu, A., J. Angrist, Y. Narita, and P. Pathak (2019). Breaking Ties: Regression Discontinuity Design Meets Market Design. *Working Papers 2019-024, Human Capital and Economic Opportunity Working Group*.

Abdulkadiroglu, A., J. Angrist, and P. Pathak (2014, January). The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. *Econometrica*, 137–196.

Abdulkadiroglu, A., J. D. Angrist, Y. Narita, and P. A. Pathak (2017, October). Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation. *Econometrica 85*.

Agarwal, N. and P. Somain (2018). Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism. *Econometrica*.

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

Altmejd, A., A. B. Fernandez, M. Drlje, D. Kovac, and C. Neilson (2019a). Siblings Effects on University and Major Choice: Evidence from Chile and Croatia. *Princeton Working paper*.

Altmejd, A., A. B. Fernandez, M. Drlje, D. Kovac, and C. Neilson (2019b, 11). Siblings Spillover Effects on College and Major Choice: Evidence from Chile, Croatia and Sweden. *Princeton University Industrial Relations Section Working Paper Series*.

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

Angrist, J. D., S. R. Cohodes, S. M. Dynarski, P. A. Pathak, and C. R. Walters (2016, January). Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry and Choice. *Journal of Labor Economics*.

Angrist, J. D. and A. B. Krueger (1991). Does Compulsory School Attendance Affect Schooling and Earnings. *The Quarterly Journal of Economics 106*, 979–1014.

Angrist, J. D. and M. Rokkane (2016, January). Wanna Get Away? Regression Discontinuity Estimation of Exam School Effects Away From the Cutoff. *Journal of the American Statistical Association 110*, 41–55.

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

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

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

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

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

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

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

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

Cattaneo, M., M. Jansson, and X. Ma (2019). Manipulation Testing based on Density Discontinuity. *Stata Journal 18*, 234–261.

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

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

de Chaisemartin, C. and L. Behaghel (2020). Estimating the Effect of Treatments Allocated by Randomized Waiting Lists. *Econometrica 88*, 1453–1477.

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

Dustan, A. (2018). Family Networks and School Choice. *Journal of Development Economics*.

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

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

Fack, G., J. Grenet, and Y. He (2019). Beyond Truth-Telling: Preference Estimation with Centralized School Choice and College Admissions. *American Economic Review 109*, 1485–1529.

Fernandez, A. B. (2019). Should I Stay or Should I Go? Neighbor's Effects on University Enrollment. *LSE Working paper*.

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

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

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

Goodman, J., M. Hurwitz, and J. Smith (2020). The Economic Impact of Access to Public Four-Year Colleges. *NBER Working Paper No. w27177 35*.

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

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

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

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

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

Hastings, J. S., C. A. Neilson, and S. D. Zimmerman (2014, July). Are Some Degrees Worth More than Others? Evidence From College Admission Cutoffs in Chile . *NBER Working Paper Series*.

Heckman, J. (1997). Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations. *Journal of Human Resources* (32), 441–462.

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

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

Imbens, G. and K. Kalyanaraman (2012, July). Breaking Ties: Regression Discontinuity Design Meets Market Design. *The Review of Economic Studies 79*.

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

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

Jones, D. (2015). The Economics of Exclusion Restrictions in IV Models. *NBER Working Paper*.

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

Kaufmann, K. M., M. Messner, and A. Solis (2013). Returns to Elite Higher Education in the Marriage Market: Evidence from Chile . *Working paper*.

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

Kirkeboen, L. J., E. Leuven, and M. Mogstad (2016, August). Field of Study, Earnings and Self-selection. *Quarterly Journal of Econometrics 131*, 1057–1111.

Kozakowski, W. (2020). Are Public Four-year Colleges Engines for Mobility? Evidence from Statewide Admissions Thresholds. *Working Paper*.

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

Lee, D. S. (2008). Randomized Experiments from Non-random Selection in U.S. House Elections. *Journal of Econometrics 142*, 675–697.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies* (76), 1071–1102.

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

Lucas, A. M. and I. M. Mbiti (2014, July). Effects of School Quality on Student Achievement: Discontinuity Evidence from Kenya . *American Economic Journal: Applied Economics*, 234–263.

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

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

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

Rosenbaum, P. R. and D. B. Rubin (1983, April). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 41–55.

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

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

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

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

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

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

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

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

Zimmerman, S. D. (2014). The Returns to College Admissions for Academically Marginal Students. *Journal of Labor Economics 32*.