Older Schoolmate Spillovers on Higher Education Choices*

Nagui Bechichi
Paris School of Economics

Gustave Kenedi
London School of Economics and Sciences Po

Job Market Paper

This Version: November 20, 2024
updated frequently, click here for latest version

Abstract

How do students choose where to apply to higher education? Answering this question is important because the returns to higher education differ greatly across majors and institutions. This paper examines a previously overlooked factor shaping these choices: the high school environment. Specifically, we analyze how students' applications and enrollments are influenced by the higher education trajectory of recent graduates from their high school. Exploiting random admission cutoffs in France's centralized admission system, we compare high schools where older schoolmates were marginally admitted versus rejected from specific degrees (subject-institution combination). Our findings reveal significant older schoolmate spillovers: students are 6 percentage points (+19%) more likely to apply to and 2 percentage points (+45%) more likely to enroll in the same degree as a marginally enrolled older schoolmate. These effects are large, corresponding to roughly 45% of spillovers across siblings estimated in other countries. We find both teacher influence and homophily/role model effects mediate these cross-cohort spillovers. Lastly, we quantify the extent to which exposure to high-achieving older schoolmates affects the socioeconomic gap in applications to very selective degrees and find that equalizing this exposure would narrow the gap by around 10%. These results demonstrate how inequalities in higher education choice can be perpetuated or mitigated through high school peer networks.

*Bechichi: Paris School of Economics, 48 boulevard Jourdan, 75014 Paris France, nagui.bechichi@gmail.com. Kenedi: Centre for Economic Performance, London School of Economics, 32 Lincoln Inn's Field, WC2A 3PH London United Kingdom, and Sciences Po, 28 rue Saints-Pères, 75007 Paris France, gustavekenedi@gmail.com. Gustave gratefully acknowledges the financial support of PSE's Education policy and social mobility chair and the ESRC through the Centre for Economic Performance. This paper greatly benefited from feedback and suggestions from Léonard Bocquet, Cécile Bonneau, Naomi Cohen, Pierre-Philippe Combes, Paul Dutronc-Postel, Etienne Fize, Julien Grenet, Gabrielle Fack, Mylène Feuillade, Manon Garrouste, Marc Gurgand, Nina Guyon, Xavier Jaravel, Camille Hémet, Clément Imbert, Camille Landais, Mélina London, Steve Machin, Lindsey Macmillan, Sandra McNally, Florian Münch, Ekaterina Oparina, Fabien Petit, Roland Rathelot, Adam Soliman, Margaux Suteau and Camille Terrier as well as participants at numerous conferences and seminars.

1 Introduction

Choosing where to apply to higher education is likely the most consequential and complex decision high school students face. Returns to higher education vary widely across fields of study and institutions (Hastings et al., 2013; Altonji et al., 2016; Kirkeboen et al., 2016; Britton et al., 2022; Chetty et al., 2023), making students' decisions critical for their long-term earnings and career trajectories. Students face this pivotal decision while choosing from thousands of distinct options about which they often have limited information. These information frictions lead students to rely heavily on their social networks for guidance, exacerbating inequalities between privileged and disadvantaged students and impeding upward mobility (Hoxby and Avery, 2013; Altmejd et al., 2023; Campbell et al., 2023). Social learning has been widely hypothesized as shaping students' choices, yet the sources of this learning are not fully understood. Notably, the influence of high school actors such as schoolmates and teachers remains unexplored, despite students frequently discussing their higher education plans with these actors (Marbach and Van Zanten, 2023; MESR, 2023).

In this paper, we provide the first causal evidence that the high school environment, one of the main sources of students' social networks, influences students' higher educational choices. We focus on a novel and understudied channel: the impact of older schoolmates—that is, students who graduated from the same high school in the previous year—on younger students' decisions. Existing work has so far only established correlational conjectures regarding the role of older schoolmates.[1] This limitation stems from both the difficulty of finding exogenous variation in older schoolmates' enrollments and the scarcity of comprehensive datasets linking students' higher education choices with their high school affiliations.

We overcome these obstacles by implementing a research design based on admission cutoffs, using rich administrative data on students' applications to higher education programs (subject-institution combination) in France between 2012 and 2017. We show that random shocks to older peers' higher education trajectories substantially influence the subsequent application and enrollment choices of younger students within the same high school. In particular, students are significantly more likely to apply to, and enroll in, the exact same degree in which a schoolmate enrolled the previous year. These findings reveal an overlooked source of social learning—older schoolmates—which contributes to inequalities in higher education choice between students exposed to different high school environments (Hoxby and Avery, 2013).

[1] Hoxby and Avery (2013, p.3) note that high-achieving, low income students who do not apply to selective colleges are predominantly from schools where they "have a negligible probability of meeting a [...] schoolmate from an older cohort who attended a selective college." Similarly, Black et al. (2020, p.232) find using conditional logit regressions that "all students learn something about a college campus through the feeder relationship of their high school peers."

We identify spillovers across cohorts by exploiting admission cutoffs generated by France's centralized higher education allocation mechanism. Unlike systems where test scores determine admissions, French institutions have discretion in ranking applicants. Our unique dataset includes these degree-specific rankings, which applicants themselves cannot observe. We leverage each degree's cutoff rank—the rank of the last admitted student—which creates a sharp discontinuity in admission probabilities. Under this mechanism, students receive a single offer from their highest-ranked preferred degree where they clear the cutoff, while those below are rejected. These cutoffs determine the degrees to which younger students are exposed through their older schoolmates. Crucially, these cutoffs are unpredictable ex-ante, as they result from the interplay of *all* applicants' preferences, *all* degrees' rankings, and capacity constraints.

Based on this institutional setting, we implement a regression discontinuity design comparing application and enrollment outcomes in the subsequent year $(t+1)$ for high schools (technically, high school x high school tracks)[2] with an older peer ranked just above or below a cutoff in year t.[3] Due to the randomness of cutoffs, these high schools are virtually identical, except for having an older schoolmate offered a seat or rejected from this degree. Our approach differs from previous studies exploiting academic thresholds (e.g., Altmejd et al., 2021; Altmejd, 2023) in that we utilize relative rankings rather than absolute scores, providing novel quasi-random variation in degree exposure across high schools.

Our focus on high schools at the margin of admission cutoffs captures the most relevant cases for our analysis. In schools well above cutoffs, students likely already consider these degrees attainable, while in schools well below cutoffs, younger peers may face low admission chances. High schools around a degree's admission cutoff represent a critical middle ground where an older schoolmate's admission could open up new possibilities in the eyes of younger students while maintaining realistic chances of admission. Thus, these marginal cases represent schools where older schoolmates' outcomes are most likely to shape younger students' higher education choices.

We find that students follow their older schoolmates to the same higher education degree. High schools with a marginally enrolled older schoolmate to a specific degree are 6 percentage points (+19% increase relative to the counterfactual mean) more likely to have at least one applicant to that same degree in the following year relative to high schools with a marginally rejected older schoolmate. These spillovers on applications

[2]Our analysis is undertaken at the high school x high school track level because (i) in France, high school tracks are very segregated within high schools, and (ii) higher education programs are often largely track-specific. There are 13 high school tracks. To ease legibility, we use "high school" to refer to "high school x high school track".

[3]In cases where there are multiple applicants from the same high school to the same degree, we follow Estrada et al. (2024) and consider only the highest-ranked applicant from each high school for each higher education degree. This effectively ensures that only high schools whose applicants are all ranked below the last admitted student have no enrolled students in that degree.

translate into actual enrollment: high schools with a marginally enrolled older schoolmate are 2 percentage points (+45%) more likely to have at least one enrolled student in the same degree the following year. We also uncover large impacts on the number of applicants and enrolled students, with increases of 0.19 (+23%) and 0.03 (+49%) students respectively. The stronger relative effect on enrollment compared to applications stems from degrees ranking subsequent applicants from these high schools more favorably, rather than from changes in students' preference rankings. This suggests that degrees may use the performance of students from previously underrepresented high schools as a signal of future applicants' potential.

The magnitude of these effects is substantial. Compared to spillovers across siblings estimated in Chile, Croatia, and Sweden by Altmejd et al. (2021), our observed spillovers across cohorts represent approximately 45% of the magnitude for applications and 55% for enrollments. This indicates that the high school environment's impact on students' higher education choices is about half as strong as that of siblings, a considerable effect. Furthermore, these spillovers persist over time. Two years post-treatment ($t + 2$), the impact remains substantial: spillovers on applications are still 65-75% of the magnitude observed one year after treatment ($t + 1$). This persistence implies that the total impact of exposure to different degrees is even larger than our main estimates suggest. Importantly, we also observe spillovers to similar degrees, that is, not just the older schoolmate's exact degree. These findings have significant implications for educational interventions: one-time initiatives aimed at influencing students' higher education choices (e.g., information campaigns or mentoring programs) are likely to have cascading effects on subsequent cohorts, amplifying their impact.

We investigate two mechanisms underlying our observed spillovers across cohorts: (i) teacher influence, and (ii) student homophily/role model effects. These mechanisms are not mutually exclusive but lead to different policy implications: improving teacher guidance versus facilitating peer connections across cohorts. To examine teacher influence, we assess whether students are more likely to follow an older schoolmate if they share the same "principal" teacher—the teacher responsible for the administrative duties of a class and guidance on higher education applications. We find that students sharing the same principal teacher as the older peer are significantly more likely to emulate the marginally admitted older schoolmate's higher education choices relative to students with a different principal teacher. This suggests a key role for teacher influence in these spillover effects. These stronger spillover effects for students sharing the same principal teacher are unlikely to be explained by increased cross-cohort interactions, as sharing a teacher does not typically facilitate more contact between students from different years in France. From a policy perspective, this finding highlights the importance of providing adequate training and support to teachers in charge of helping their students make their higher education decisions.

Moreover, we assess student homophily/role model effects by analyzing whether spillovers are larger for students who share the same demographic characteristics as the marginally admitted older schoolmate. We find strong support for the homophily/ role model effects. Importantly, the spillover effects for girls are significantly larger when the older schoolmate was a girl compared to boys. Since we compare students of the same gender on both sides of the cutoff, this difference cannot be explained by gender-specific degree preferences. However, we do not find that boys are more likely to follow a boy compared to a girl older schoolmate. Additionally, low socioeconomic status (SES) students are significantly more likely to apply to a degree when the older schoolmate was also from a low SES background. Interestingly, very high SES students are also more likely to follow a very high SES older schoolmate, an intriguing result given very high SES students are likely to be well-informed about their higher education opportunities. These findings suggest that the demographic alignment between younger students and older schoolmates plays a crucial role in shaping the influence of peer outcomes on higher education choices. This homophily effect could be explained by degrees being more salient when students share the same characteristics, or perhaps by a greater likelihood for these students to have direct links.

We find three insightful heterogeneities in spillover effects. First, smaller high schools exhibit larger spillovers across cohorts, potentially due to increased inter-cohort interactions or stronger teacher-student relationships. Second, spillover effects vary based on the selectivity of degrees and high schools' academic level. Notably, students from lower-performing high schools are most influenced by peers entering selective programs, while those from top-performing schools are more affected by peers entering less selective programs. This suggests that peer effects may be expanding students' horizons, encouraging applications to programs they might not otherwise consider. Third, degrees at moderate distances induce the largest spillovers, likely reflecting a balance between awareness of degrees and willingness to apply. This distance effect suggests that peer influence is most potent in exposing students to less familiar yet accessible opportunities. These heterogeneities underscore the complex interplay between school characteristics, degree prestige, and geography in shaping peer influence on higher education choices.

Finally, we quantity the extent to which older schoolmate spillovers could narrow the persistent application gap between high-achieving students of different socioeconomic backgrounds. Among students in the top 10% of the academic ability distribution, those from low SES backgrounds are 27 percentage points less likely to apply to degrees in the top 10% in terms of selectivity compared to their very high SES peers. The same high-achieving low SES students also face a deficit in exposure: they are 20 percentage points less likely to have an older schoolmate who enrolled in such prestigious programs. We use our baseline spillover estimates to conduct a back-of-theenvelope counterfactual analysis that suggests equalizing exposure to high-achieving older schoolmates could narrow the application gap by approximately 10%. This reduction, while not eliminating the disparity, represents a significant improvement.

Our analysis underscores the potential of leveraging peer influences as a complementary strategy to traditional policy approaches. We conclude by discussing several policy options that could harness these spillover effects, including implementing high school quotas for prestigious degrees, reducing residential and within-high school segregation, enhancing mentorship programs, and fostering cross-cohort interactions within schools through alumni meetings. These findings contribute to the broader debate on effective measures to promote educational equity and social mobility.

Related literature. Our paper contributes to several strands of the literature. First, it contributes to the literature on higher education choice. Recent research has highlighted substantial heterogeneity in returns across fields of study and institutions (Altonji et al., 2012; Chetty et al., 2023), emphasizing the critical nature of these choices on individuals' future outcomes. While financial barriers have been a central focus (Dynarski et al., 2023b), a growing body of work examines how informational and behavioral frictions shape students' choices (Lavecchia et al., 2016; Dynarski et al., 2023a). Within this literature, various school-level influences have been documented: teachers affect students' long-term outcomes including college attendance (Chetty et al., 2011, 2014a; Jackson, 2018); Mulhern (2023) shows that more effective high school guidance counselors in Massachusetts positively influence high school graduation, college going and college selectivity; and Mulhern (2021) demonstrates how Naviance, a software providing personalized college admissions information, influences students' college choices based on older peers' admission experiences. We advance this literature by highlighting the important role of the high school environment in shaping students' higher education choices, not just attendance. Our key contribution lies in elucidating the mechanisms through which the high school environment, particularly older schoolmates and teachers, affects students' higher education decisions. By doing so, we provide novel insights into how social dynamics within schools can address informational and behavioral barriers in the college choice process.

Second, our paper contributes to the literature on peer effects in education, particularly in higher education (for a comprehensive review, see Barrios-Fernandez, 2023). Prior work has established that students' higher education choices are strongly influenced by their closest social connections. Altmejd (2023) demonstrates a 50% increase in the likelihood of children graduating from the same major as their parents in Sweden. Aguirre and Matta (2021), Altmejd et al. (2021) and Avdeev et al. (2024) document substantial effects of older siblings' degree enrollment on younger siblings' higher education choices in Chile, Croatia, Sweden, the United States and the Netherlands. Barrios-Fernández (2022) reveals the significant impact of close neighbors on higher education enrollment decisions. We extend this literature by examining a distinct yet understudied social network: older schoolmates. While a recent paper by Estrada et al. (2024) and contemporaneous work by Valdebenito (2023) have begun to explore schoolmate influences in specific contexts—elite secondary schools in Peru for the first, and gender dynamics in higher education field of study choices in Chile for the latter—our analysis provides the first comprehensive evidence on these spillovers. By studying the universe of higher education choices in France and identifying key transmission mechanisms, we show how older schoolmates represent an important channel through which schools can shape students' educational trajectories.

Third, our paper more generally contributes to the literature on social learning (Mobius and Rosenblat, 2014). This literature demonstrates that individuals frequently rely on their social networks when making decisions under imperfect information. For instance, (Conley and Udry, 2010) shows that farmers' technology adoption decisions are significantly influenced by the experiences of neighboring farmers, particularly for new agricultural innovations where returns are initially uncertain. In labor markets, job search processes are heavily impacted by individuals' social connections (e.g., Kramarz and Skans, 2014), with individuals learning not only about job opportunities but also about difficult-to-observe job characteristics (e.g., work atmosphere, team dynamics, salary) through their networks. Social networks similarly shape migration decisions: individuals locate in places where they know more people (Blumenstock et al., 2023). Our study extends this literature by examining social learning in the context of higher education choices, a decision environment characterized by significant informational frictions. We demonstrate that even small changes in exposure to others' choices—in our case, through older schoolmates—can substantially impact students' decisions. These findings suggest that similar social learning mechanisms might operate in other educational and career transitions, such as graduate school selection or first employer choice upon graduation.

Last, our paper contributes to our understanding of intergenerational inequalities. Persistent disparities in labor market outcomes across socioeconomic backgrounds have been extensively documented (Black and Devereux, 2011; Mogstad and Torsvik, 2023). These inequalities stem from multiple sources, including differences in early childhood environments and investments, neighborhood effects, and parental job networks. A significant portion of these disparities can be traced to differential access to and choices within higher education (Chetty et al., 2014b; Bonneau and Grobon, 2022). Given the important heterogeneity in returns across fields of study and institutions, differences in higher education choices between students of varying socioeconomic backgrounds may significantly impede intergenerational mobility. Our research offers new insights for policy interventions by highlighting a specific mechanism—school-level peer effects—through which these inequalities persist. We show that differences in exposure to high-achieving older schoolmates contribute to the socioeconomic gap in applications to selective programs, suggesting that targeted efforts to diversify peer groups or enhance cross-cohort interactions within schools could help promote social mobility.

The rest of the article is organized as follows. Section 2 describes the institutional background and data. The empirical strategy is presented in Section 3, while Section 4 discusses the main results and their robustness. Section 5 presents insightful heterogeneities. In Section 6, we investigate the mechanisms that might explain the observed older schoolmate spillovers. The counterfactual exercise is presented in Section 7, and we discuss policy implications in Section 8. Finally, Section 9 concludes.

2 Institutional Background and Data

This section outlines the most important features of the French higher education system and our data source, with additional details provided in Appendix A. Central to our analysis is France's centralized application platform, where students apply to specific degrees (subject-institution combinations). Notably, there are no nationally standardized admission criteria; each degree independently determines its (undisclosed) criteria based on all the information available in students' application, including both academic (e.g., high school grades) and non-academic (e.g., teacher comments, high school of origin, geographic location) factors.[4] We also explain how the allocation of students to degrees generates the discontinuities we exploit to identify spillovers across cohorts within the same high school.

Our analysis is conducted at the *high school* $\times$ *high school track* level[5] rather than at the *high school* level. This granular approach is motivated by two key considerations: (i) the significant segregation of tracks within French high schools, and (ii) the common restriction of higher education programs to specific high school tracks. To ease legibility, we will continue to refer to "high schools" throughout the paper, although this technically denotes high school $\times$ high school track combinations. Consequently, all statistics pertaining to high schools are computed at this more precise level.

[4] Charousset et al. (2021) attempt to reverse engineer the potential criteria used by elite programs.

[5] There are 13 high school tracks, grouped into three aggregate tracks (general, technological, and professional). The three general tracks are Sciences (S), Social Sciences (ES), and Literature (L). The 8 technological tracks are Management Sciences and Technologies (STMG), Sustainable Development Sciences and Technologies (STI2D), Health Sciences and Technologies (ST2S), Laboratory Sciences and Technologies (STL), Design and Applied Arts Sciences and Technologies (STD2A), Agronomy and Life Sciences and Technologies (STAV), Hospitality (H), and Music and Dance Techniques (TMD). The two broad professional tracks are Professional (P) and Agricultural Professional (PA).

2.1 Institutional Background

From 2009 to 2017, senior high school students in France applied to degrees through a centralized online platform called *Admission Post-Bac*.[6] The platform's coverage expanded progressively, encompassing up to 90% of post-secondary programs (academic and vocational) by the end of this period (Bechichi et al., 2021). In 2017, it managed over 10,000 programs and 800,000 applicants.[7] The allocation of students to degrees occurred in three distinct stages:

1. *Application submission:* students submit a rank-ordered list of up to 36 program choices.[8]

2. *Degree ranking:* higher education programs rank their applicants (without knowing applicants' rank-ordered lists) and specify their capacity constraints (available seats) to the platform.

3. *Offer distribution:* a three-round college-proposing deferred acceptance algorithm allocates offers.[9]

The three-round structure of the allocation algorithm allowed for seat reallocation due to (i) students opting for programs outside the platform, (ii) students entering the labor market directly, and (iii) students failing the *Baccalauréat*, the high school exit exam required for higher education enrollment (see Appendix Figure A.2 for a timeline of the procedure). A wait-list system was embedded in the system to enable this seat reallocation. Throughout the three-round process, two key features stand out. First, students receive only *one* admission offer at a time, specifically from the highest-ranked degree for which they are ranked above the last admitted student in that round. Crucially, this means that a student offered a seat in their top-ranked degree does not know whether they would have received offers from any of the programs in their rank-ordered list. Second, at each round, students face one of two scenarios. If offered admission to a degree, they can either accept the offer or reject it and exit the platform. If wait-listed, they can choose to remain in consideration for future rounds or accept

[6] This platform was replaced in 2017 by *Parcoursup*, which introduced significant changes to the application and admission system. For a detailed comparison between *Admission Post-Bac* and *Parcoursup*, see Frouillou et al. (2019).

[7] Several institutions, including Paris Dauphine University, Institutes of Political Studies (IEP), certain private programs, and nursing schools, maintained separate application and admission processes outside this platform.

[8] The number of program choices was reduced in 2016 to 24. In all years, students could apply to at most 12 programs within each program "type", e.g., vocational degrees (BTS), technical degree (DUT), preparatory classes (CPGE).

[9] The implemented algorithm was a slight variation of the Gale-Shapley algorithm, accommodating joint applications for degrees and degrees' student accommodations. This variation is irrelevant for our study. For more details, see Appendix 2.A.2 in Charousset et al. (2021).

their current admission offer (if applicable). Our analysis focuses on outcomes after the last-round.

2.2 Data

We use comprehensive application-level administrative data from the *Admission Post Bac* application platform spanning 2012-2017, encompassing the entire period for which data from the platform is available. This rich dataset contains students rank-ordered list of applications to degrees, degrees rankings of applicants, and final matching outcomes. While applicants do not have access to degrees' rankings, our use of the platform's data provides this crucial information which we will use to determine admission cutoffs. We do not directly observe *enrollment*, but can track *offer acceptance*. Thus, some students who accept offers may ultimately not enroll. For simplicity, we use offer acceptance to refer to enrollment. The dataset also provides detailed student background characteristics, including high school and specific track within the school, high school exit exam grades, and socioeconomic status based on legal guardian's occupation. Unique identifiers allow us to track both high schools and programs over time.

2.3 Admission Cutoffs

We rely on admission cutoffs generated by France's centralized higher education allocation mechanism. Unlike systems where test scores determine admissions (e.g., Chile, Sweden), French institutions have discretion in ranking applicants. As we observe these degree-specific rankings, we can leverage each degree's cutoff rank, i.e., the rank of the last admitted student, which creates a sharp discontinuity in admission probabilities: students ranked above this cutoff may receive an offer[10], while those below are rejected. We refer to the rank of the last admitted student for a given degree as the degree's *admission cutoff rank*. Figure 1 provides an example of how our admission cutoff ranks are defined for a fictitious degree.

3 Empirical Strategy

Our aim is to identify the extent to which students' higher education choices are influenced by the higher education trajectory of older schoolmates. Specifically, do these students have higher propensities to apply to and enroll in the same degrees as students from their high school's previous graduating cohort. We hypothesize that the

[10] Recall that because the allocation mechanism provides students with a single admission offer, they do not receive an admission offer from all degrees by which they are ranked above the last admitted student.

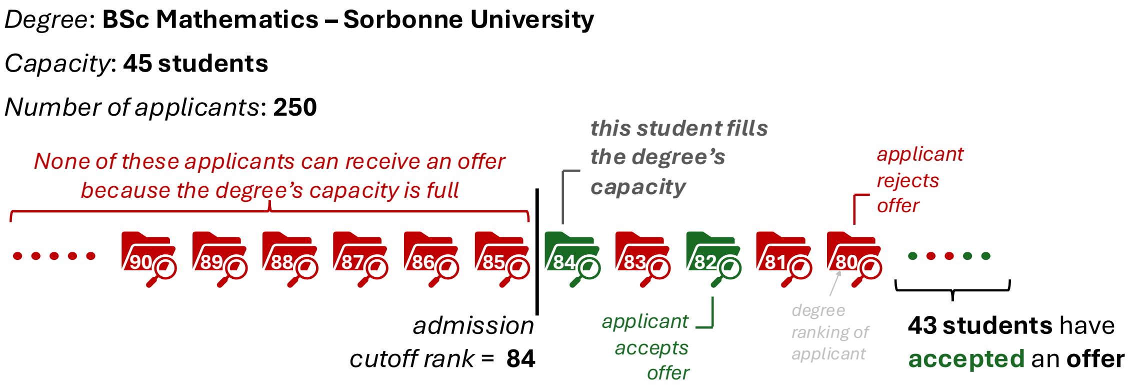


Figure 1: Illustration of Admission Cutoff Rank

Notes: This figure illustrates how we define the admission cutoff rank for a fictitious degree.

experiences of these older students at specific degrees could spillover to younger cohorts, potentially through increased awareness or information about these programs.

Identifying such cross-cohort spillovers is challenging because (i) students and their older schoolmates share high school characteristics (e.g., teachers, location, resources) that independently influence higher education choices regardless of older schoolmates' higher education trajectories, (ii) simple comparisons between high schools with admitted versus rejected students may overstate spillover effects due to inherent differences between these schools (e.g., number of applicants, academic achievement levels), and (iii) degrees may show preferences for high schools for which they regularly admit students.

To overcome these challenges, we employ a fuzzy regression discontinuity design, leveraging admission cutoff ranks—i.e., the degree's last admitted student rank—described above. The fuzzy nature of the design stems from the fact that we define (conceptually) high schools as being treated if an older schoolmate *enrolled* in the degree, not only if he or she was *admitted*. Admission cutoff ranks induce a discontinuity in the likelihood that a high school has an older student who is admitted and enrolls in the degree. This approach allows us to compare high schools that are essentially identical except for having an older student just above or below the rank of the last admitted student for a given degree. The key distinction between these high schools is that an older student above the cutoff has a higher probability of receiving an offer, potentially enrolling, and subsequently influencing younger cohorts' choices.

Crucially for the validity of our research design, these cutoff ranks cannot be anticipated ex-ante by applicants or degrees. The rank of the last admitted student depends on *all* applicants' rank-ordered lists, which are unknown to programs when reporting capacities and rankings, and degrees' rankings, and capacity constraints, which are unknown to applicants (and other degrees). Additionally, the evolution of admission cutoff ranks across rounds is determined by applicants' decisions to accept or decline offers, making it unpredictable during the application phase. This design ensures that high schools on either side of the cutoff are essentially identical, ruling out the possibility that results are driven by differences in student body composition, size, or location. Moreover, it addresses the reflection problem (Manski, 1993), as the older cohort's enrollment decisions cannot be impacted by younger schoolmates' future choices.

3.1 Running Variable

Our running variable is defined for each degree as the high school's rank relative to the admission cutoff rank. To avoid misclassification of a high school's treatment status, we follow Estrada et al. (2024) and consider only each high school's best-ranked applicant for each degree. This approach ensures that (i) we have one observation per high school, per degree, per year—specifically, that of the best-ranked applicant from that high school to that degree in that year, and (ii) if a high school has (at least) one rejected and (at least) one admitted student to a degree it is classified as "treated" by keeping only the rank of the (best-ranked) admitted student.

We then center these high schools' best ranks around the degree's admission cutoff rank—the rank of the last admitted student. Formally, for each high school s with student(s) $i(s)$ applying to degree j in year t, we define the running variable as:

$$\textit{distance to last admitted student}_{sjt} = \textit{rank of last admitted student}_{jt} - \max_{sjt}\{\textit{rank}_{i(s)jt}\}.$$

A positive value of *distance to last admitted student*$_{sjt}$ indicates that high school s is potentially treated by degree j in year t, meaning the best-ranked applicant might have received an offer and enrolled in degree j. A strictly negative value indicates the high school is not treated. Our running variable for a given degree may not have actual observations at each rank. This is because it's very common for degrees to have several applicants from the same high school—in fact, 98% of degrees have at least two applicants from the same high school. Additional details on the running variable are provided in Appendix B. Figure 2 uses the same example as in Figure 1 to illustrate our running variable.

3.2 Estimation Sample

We restrict our sample to (i) high schools and degrees present in both the treatment and the following year (95% of all high schools-year and degrees-year, respectively); (ii) degrees with at least one applicant ranked after the admission cutoff rank (86% of all degrees-year), ensuring the potential for rejected applicants; (iii) degrees reporting the same capacity constraint across all three admission rounds (84% of all degrees-year), preventing manipulation of the admission cutoff rank; and (iv) degrees with at

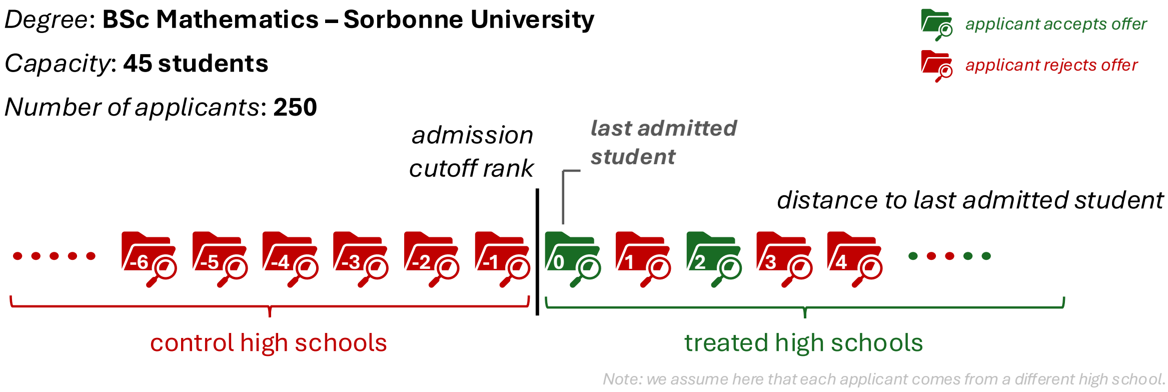


Figure 2: Illustration of the Running Variable

Notes: This figure illustrates how we construct our running variable for a fictitious degree.

least 30 high schools in their application pool, mitigating autocorrelation of the running variable over time for degrees with few applicants.

To further enhance high schools' comparability around the admission cutoff rank, we implement additional conditions: (i) we "symmetrize" the running variable such that, for example, if a degree has n applicants ranked after the admission cutoff rank, we retain n applicants above the admission cutoff rank; (ii) we drop the last admitted student; and (iii) we keep only degrees with at least two observations on both sides of the cutoff within the chosen regression discontinuity bandwidth, after symmetrization and exclusion of the last admitted student. Our final sample, within the bandwidth, contains 19,577 degree-years and 56,763 high school-years, totaling 375,566 observations.

Table 1 presents descriptive statistics for high schools and degrees in both the raw sample[11] (*Full sample*) and the sample used for the analysis (*RD sample*). Comparing these samples reveals the extent of our results' external validity, as differences between our analysis sample and the general applicant population reflect how our sample restrictions may focus on a potentially unrepresentative subset of high schools and degrees.

Overall, the two samples exhibit reasonable similarity in terms of high school characteristics, including size, student body composition, and academic achievement. However, some notable differences exist. The analysis sample slightly over-represents larger high schools and technological high school tracks, while under-representing professional tracks. Regarding degree characteristics, the analysis sample contains degrees with significantly more applicants, though they have very similar average academic records. These degrees also have more high schools within their application pools. Perhaps most importantly, our analysis sample contains a higher proportion of programs

[11]Our raw sample is restricted to applications from high school students with a non-missing high school identifier.

Table 1: Descriptive Statistics

	Full sample *All high school applicants* *2012-2016* (1)	RD sample *Distance to cutoff rank* $\in [-20, 20]$ (2)
Panel A. High school characteristics		
Number of high schools	66,586	56,763
Mean number of students	41.6	47.0
Female (%)	53.5	54.3
Mean Bac grade	12.0	12.1
Very high SES (%)	28.3	28.3
High SES (%)	13.3	13.8
Middle SES (%)	30.2	30.2
Low SES (%)	24.6	24.8
Missing SES (%)	3.5	2.9
Scientific high-school track (%)	19.3	20.9
Social science high-school track (%)	17.4	18.3
Literature high-school track (%)	15.2	15.1
Technological high-school track (%)	23.6	25.2
Professional high-school track (%)	24.5	20.5
Panel B. Degree characteristics		
Number of degrees	50,423	19,577
Number of applicants	391.6	542.9
Mean Bac grade of applicants	12.1	12.1
Number of admitted students	44.6	54.2
Mean Bac grade of admitted students	12.2	12.4
Number of high-schools within application pool	148.3	204.1
Public university (%)	24.1	33.2
Vocational degree (STS) (%)	49.9	44.2
Technical degree (IUT) (%)	8.6	8.0
Academic preparatory classes (CPGE) (%)	8.3	6.7
Other institutions (%)	9.1	7.8
Number of observations	7,376,434	375,566

Notes: This table shows descriptive statistics for two samples: (1) the full sample, i.e., all high school applicants between 2012-16, and (2) our regression discontinuity (RD) sample. *High school* refers to high school $\times$ high school tracks-year, while *degrees* refer to degrees-year. The Bac is the French high school exit exam.

in public universities (33.2% in the analysis sample vs 24.1% in the full sample).[12] Conversely, there is a slight under-representation of other types of programs, particularly vocational degrees (44.2% vs 49.9%) and preparatory classes (6.7% vs 8.3%).

[12]One reason why public university degrees may be over-represented in our analysis sample is that these degrees rank *all* applicants while other degrees rank as many applicants as they wish.

3.3 Empirical Specification

First-stage. To illustrate the fuzzy nature of our regression discontinuity design, we stack all degrees in our sample, centering them around the rank of the last admitted student. Figure 3 depicts the likelihood of a high school having at least one enrolled student (panel A) and the number of enrolled students (panel B) as a function of the high school's distance to the admission cutoff rank. The probability of treatment increases from essentially 0%[13] to the left of the cutoff to 22% just to the right. This discontinuity in enrollment probability is less than 100% because students receive offers only from the highest-ranked degree in their preference list for which they are above the admission cutoff rank. Consequently, a student may exceed the rank threshold for multiple degrees yet receive no offer from them due to an offer from a higher-ranked preference.

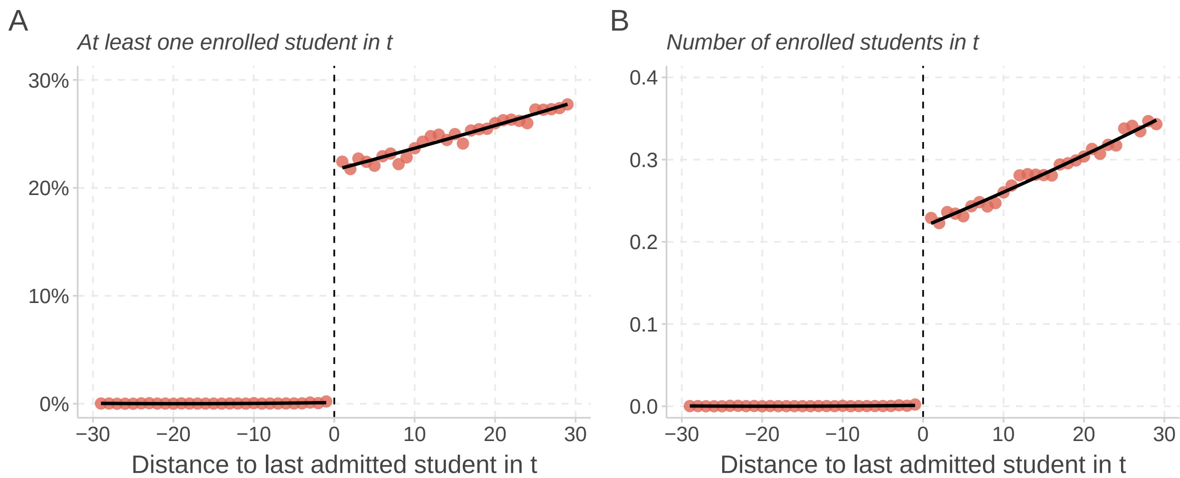


Figure 3: Probability of Older Schoolmate Enrolling in Degree Around Admission Cutoff Rank

Notes: This figure shows the probability that a high school has at least one enrolled student (panel A) and its number of enrolled students (panel B) in the degree as a function of the distance to the last admitted student in the same year. The distance to the last admitted student is defined as the rank of the high school's best ranked applicant by the degree centered around the rank of the degree's last admitted student.

Around the admission cutoff, the treatment *intensity* is essentially equal to one student, as evidenced by the close alignment of panels A and B. Appendix Figure E.1 provides a detailed breakdown of cases with one, two, or more enrolled students at each distance from the admission cutoff rank. Our current analysis focuses on spillovers for high schools transitioning from zero to one enrolled student in a degree.

[13] We observe rare cases (0.04% of our sample) where students accept offers despite being ranked below the last admitted student.

Main specification. Our approach involves stacking thousands of degree-year specific regression discontinuities, such that the admission cutoff ranks is zero for all degree-years. The following equation describes our baseline specification:

$$y_{sjt+1} = \beta \textit{distance to last admitted student}_{sjt} + \gamma(\textit{distance to last admitted student} \geq 0)_{sjt} + \delta \textit{distance to last admitted student}_{sjt} \times (\textit{distance to last admitted student} \geq 0)_{sjt} + \mu_{jt} + \epsilon_{sjt+1}. \tag{1}$$

y_{sjt+1} indicates whether high school s, with a marginally admitted student to degree j in year t, experiences more applications and enrollments in degree j in year $t+1$. *distance to last admitted student* is the running variable described above, and *(distance to last admitted student* $\geq 0)$ is a dummy variable indicating whether an older schoolmate was ranked above the last admitted student by degree j in year t. The interaction term allows for different slopes around the cutoff. Following Fort et al. (2022), we include degree-year fixed effects (μ_{jt}), thus our identifying variation comes from differences in *exposure* to a given degree across high schools. ϵ_{sjt} is the error term. γ is our coefficient of interest. Figure 4 illustrates how we stack thousands of degree-year specific regression discontinuities.

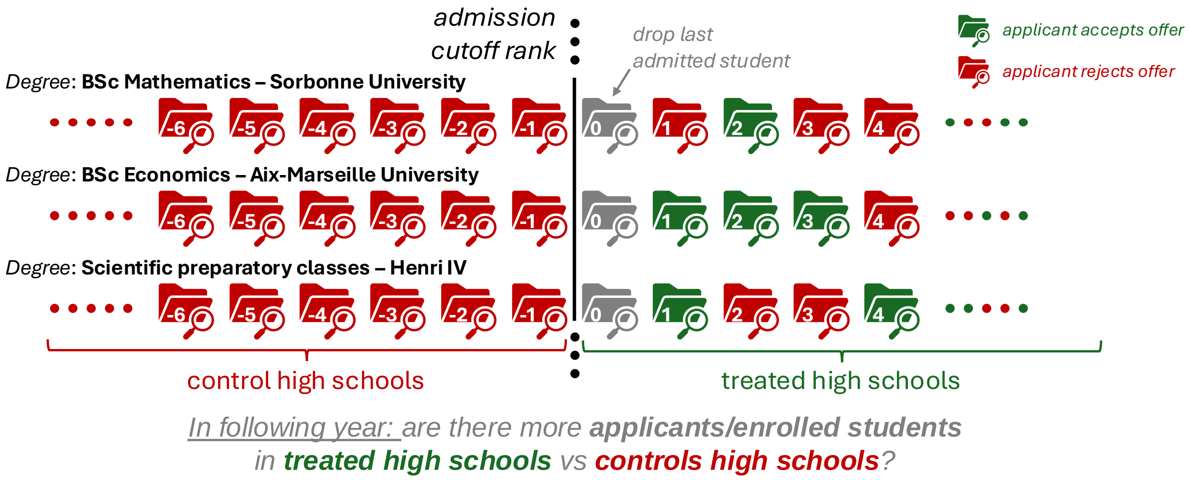


Figure 4: Illustration of Stacking of Degree-Specific Regression Discontinuities

Notes: This figure illustrates how we stack degree-specific regression discontinuities for fictitious degrees.

This specification estimates intent-to-treat effects—i.e., the effect of having an older schoolmate ranked above the last admitted student to the degree but not necessarily enrolling in it. To estimate the effect of an older schoolmate's actual enrollment, we employ a 2SLS approach, instrumenting enrollment with an indicator for being ranked above the last admitted student. This 2SLS estimate corresponds to the ratio between the intent-to-treat and the first-stage estimates. As noted in Section 2.2, we observe degree offer acceptance rather than actual enrollment. This leaves our intent-to-treat estimates unchanged but potentially underestimates our 2SLS coefficients. Since the first-stage for actual enrollment is necessarily smaller than the first-stage for offer acceptance, if anything the 2SLS coefficients would be larger, and thus the reported instrumented estimates are lower bounds on the effect of an older schoolmate's actual enrollment.

Following Cattaneo et al. (2019)'s guidelines, we estimate the coefficient of interest nonparametrically using local linear regressions. Specifically, linear regressions are fit on both sides of the threshold using a triangular kernel which gives more weight to observations near the threshold. We compute MSE-optimal bandwidths following Calonico et al. (2014) for our four main outcomes: (i) having at least one applicant to the same degree as an older schoolmate, (ii) the number of applicants to this degree, (iii) having at least one enrolled student in this degree, and (iv) the number of enrolled students in this degree. To ensure a consistent sample across estimates, we use a common bandwidth of 21 ranks, corresponding to the smallest bandwidth for these main outcomes, as in Altmejd et al. (2021). Given that high schools may have students at the admission margin for multiple degrees, we cluster standard errors at the high school-year level.

Robustness. We ensure our main results are robust to (i) varying the bandwidth used, (ii) estimating equation (1) using a second-order polynomial of the running variable, (iii) including time-varying high school covariates, (iv) including high school-year fixed effects, and (v) allowing the slopes to vary for each degree-year. Moreover, we show that significant discontinuities are observed exclusively at the true admission cutoff rank, with no such effect at placebo cutoffs.

3.4 Identifying Assumptions

The central identifying assumption in our analysis is the exogeneity of the last admitted students' rank to a degree with respect to that students' high school. In other words, we require that applicants' high school around the admission cutoff rank are as-good-as-random. This assumption implies two conditions common to regression discontinuity designs: (i) the absence of strategic manipulation of applicant rankings by degrees or applicants themselves, and (ii) the continuity of potential confounders around the admission cutoff ranks.

First, as highlighted in Section 2.1, degrees cannot predict ex-ante which student will be the last admitted, due to their lack of knowledge about students' rank-ordered lists.[14] While degrees may use applicants' high schools to rank them, they cannot

[14] Technically, for oversubscribed public university programs, students' rank-ordered lists were used to randomly allocate students to these programs. However, these programs still could not predict the high school of the last admitted student because they could not observe the rank-ordered lists of *all*

predict the high school of the eventual last admitted student. Figure 5 presents the distribution of our running variable, showing no clear evidence of manipulation.[15]

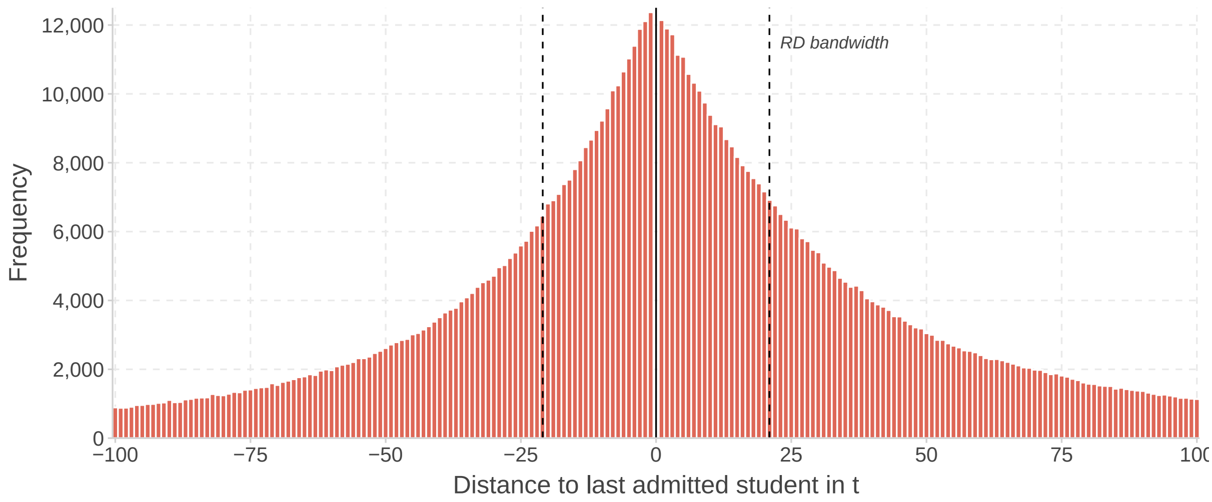


Figure 5: Distribution of Running Variable

Notes: This figure shows the distribution of the running variable, which corresponds to the rank of each high school's best-ranked applicant for a given degree, centered around the rank of that degree's last admitted student. The dashed lines represent the regression discontinuity (RD) bandwidth used in the analysis.

Second, we examine whether high school characteristics are identical around the admission cutoff rank. Figure 6 displays the estimated discontinuities in numerous high school characteristics. All discontinuities are negligible in magnitude and predominantly statistically insignificant. Crucially, there are no differences in the number of applicants in the treatment year.[16] The underlying figures in Appendix Figures E.3 visually confirm the absence of discontinuities in high school characteristics around the cutoff. Additionally, in Appendix Figures E.4 we show there are no differences in degree characteristics around the cutoff. While such differences would be absorbed by the degree-year fixed effects in estimating equation (1), their absence strengthens the validity of the visual evidence presented in the main results.

applicants on the platform, as well as *all* degrees ranking and capacity constraints. See Bechichi et al. (2021) for further details on this lottery system.

[15] Appendix Figure E.2 displays the composition of program types along the running variable.

[16] By construction, high schools on both sides of the cutoff have at least one ranked applicant in the treatment year, as this is a necessary condition for having an applicant rank for the degree.

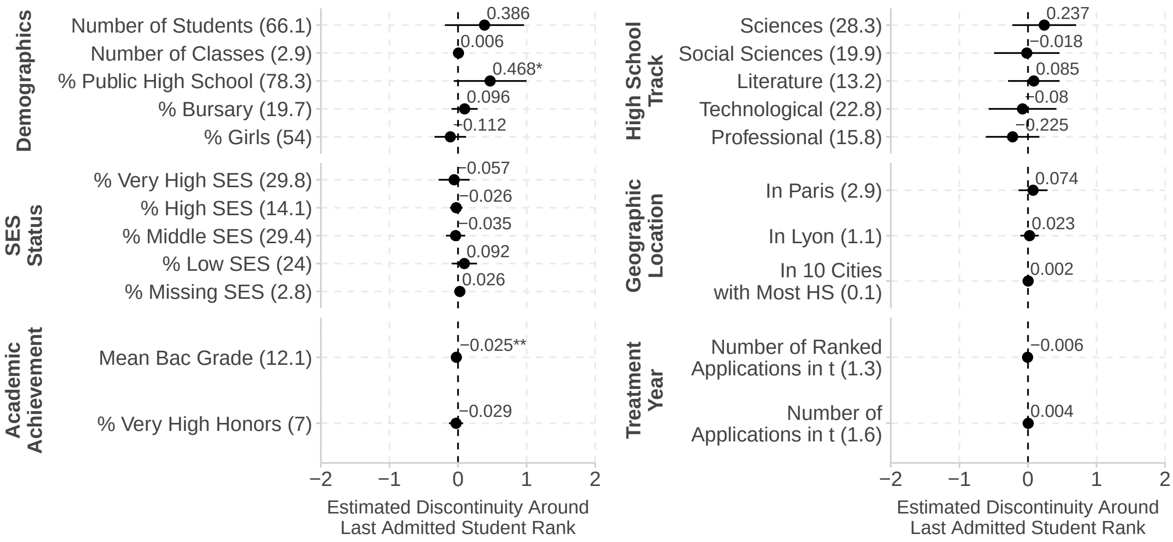


Figure 6: Discontinuity in High School Characteristics

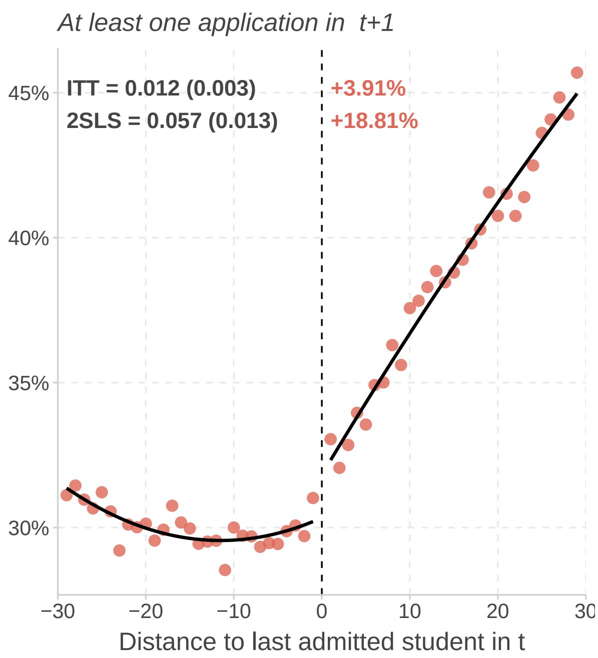
Notes: This figure shows the estimates of discontinuities in high school characteristics around the rank of degrees' last admitted student. High school characteristics are reported on the y-axis, with the mean value just below the rank of the last admitted student (e.g., $[-5, -1]$) shown in parenthesis. All specifications correspond to local linear regressions with a triangular kernel and include degree-year fixed effects. The bandwidth (20.96) is the smallest MSE-optimal bandwidth (Calonico et al., 2014) for the main outcomes. Standard errors are clustered at the high school-year level. 95% confidence intervals are reported. Significance levels: *** 1%, ** 5%, * 10%.

4 Main Results

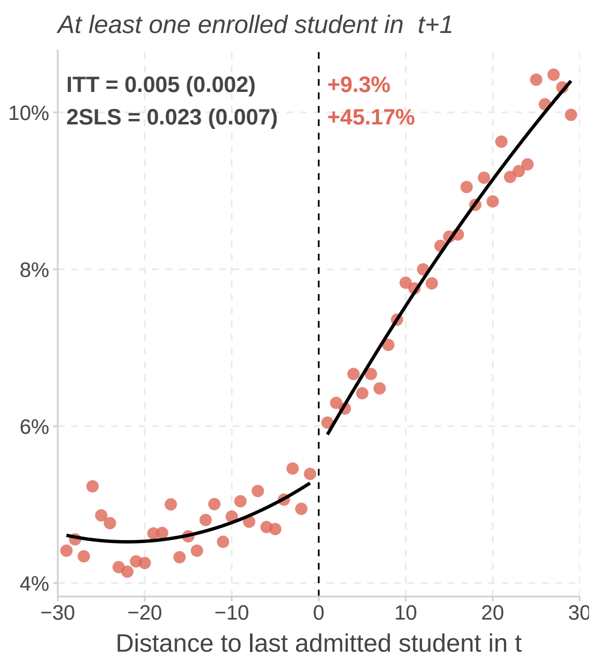
This section presents results on spillovers across cohorts on higher education choices. We show that students follow in their older schoolmates' footsteps. Specifically, the marginal admission of an older schoolmate to a specific degree (subject-institution combination) significantly influences the application and enrollment patterns of younger cohorts from the same high school, compared to instances where these older peers were marginally rejected.

4.1 Older Schoolmate Spillovers on Applications and Enrollment

Applications. Students are more likely to apply to the exact same degree in which a student from the previous cohort within the same high school was admitted. Figure 7a illustrates this phenomenon. These figures show the reduced-form relationship, across all degrees and years, between a high school's applications in $t + 1$ and the rank of this high school's best-ranked student relative to the rank of the last admitted student to the degree in year t. Panel A displays the extensive margin of applications—i.e., whether there is at least one applicant to the degree—while panel B shows the intensive margin—i.e., the number of applicants to the degree.



(a) Spillovers on Applications



(b) Spillovers on Enrollment

Figure 7: Older Schoolmate Spillovers on Applications to and Enrollment in Marginally Admitted Older Schoolmate's Degree

Notes: This figure presents non-parametric binned scatter plots of the relationship between high schools' application and enrollment outcomes for a degree in $t+1$ and these high schools' distance to the degrees' last admitted student in t. The specific application and enrollment outcomes are reported in each facet's title. Each point represents the average outcome value for high schools at a given distance from the admission cutoff rank. The fitted lines correspond to second-order polynomial fits through the conditional expectations. Both the intent-to-treat (ITT) and instrumented (2SLS) estimates are reported, using local linear regressions with a triangular kernel, and including degree-year fixed effects. The bandwidth (20.96) corresponds to the smallest MSE-optimal bandwidth (Calonico et al., 2014) for these main outcomes. Standard errors, clustered at the high school-year level, are reported in parentheses.

There is a clear discontinuity in both the likelihood of having at least one applicant and the number of applicants to degrees for which there was a marginally admitted older schoolmate. The first row of Table 2 reports the intent-to-treat ("Older schoolmate above cutoff (ITT)") estimates. These suggest that a high school with a marginally admitted older peer to a degree experiences a 1.2 percentage points increase in its likelihood of having at least one applicant to the same degree in the following year, and a 0.038 increase in its number of applicants. These correspond to a 3.9% and 4.7% increase relative to the mean outcome just below the cutoff, respectively, indicating slightly larger effects for the intensive margin of applications.

The second row of Table 2 presents 2SLS estimates, corresponding to the ratio between these reduced-form and the first-stage estimates ("Older schoolmate enrolls (2SLS)"). These 2SLS coefficients represent the effect of a high school's older schoolmate enrolling in a given degree on the application and enrollment decisions of the following cohort of students from the same high school. Due to our moderate first-stage estimates (a 22 percentage point increase in the probability of treatment), the 2SLS estimates are significantly larger than the intent-to-treat estimates. An older schoolmates' enrollment in a degree leads to a 5.7 percentage points increase in the probability of having at least one applicant from the same high school to the same degree the following year, and a 0.19 increase in the number of applicants. These effects are substantial, representing a 18.8% and 23.5% increase relative to the baseline mean, respectively.

To contextualize these magnitudes, we compare them with spillovers across siblings on higher education choices. Altmejd et al. (2021) find that an older siblings' enrollment in a given degree increases their younger siblings' application to the same degree by 36% in Chile, 32% in Croatia, and 58% in Sweden. Our estimated older schoolmate spillovers correspond to between 32% and 58% of these siblings spillovers.[17] In other words, the high school environment's influence on students higher education choices is about half as strong as that of siblings. This comparison underscores the considerable impact of high school environments on students' higher education choices, especially considering the more diffuse nature of spillovers across cohorts compared to sibling influences.

Enrollment. Students are also more likely to enrol in the same degree as marginally admitted students from the same high school's previous cohort. Figure 7b panels A and B shows the same relationship as for applications except with enrollment as the outcome. As with applications, there is a clear discontinuity in a high school's likelihood of having at least one student enroll and in the number of enrolled students in the same degree as a marginally admitted student from the previous cohort. Table 2

[17] Since no siblings spillovers have yet been estimated for France, we can only compare our older schoolmate spillovers with the siblings spillovers estimated in other countries.

reports the reduced-form and 2SLS estimates of the discontinuity.

Having a marginally admitted student in the previous cohort increases a high school's probability of having at least one enrollment in the program by 0.5 percentage points and its number of enrolled students by 0.006. These coefficients are small in absolute terms but considering how unlikely it is for a high school to have a student enroll in any given degree, these represent 9% of the counterfactual means. The 2SLS estimates are larger by construction, yielding a 2.3 percentage point increase in the probability of having at least one enrolled student and a 0.032 increase in the number of enrolled students. Relative to the baseline mean, these correspond respectively to 45% and 49% increases. Compared to sibling spillovers, these older schoolmate spillovers on enrollment are between 27% and 90% of those found in Sweden (167%) and Chile (50%) respectively.

Degree and applicant response. The observed difference in spillover magnitude between spillovers on applications (+19%) and enrollment (+45%) can be explained by two factors:

1. *Degree reponse:* when evaluating applicants, degrees consider all available information, including applicants' high school of origin. If a degree enrolls a student from a previously unrepresented high school, it may reassess that school's quality, potentially influencing how future applicants from the same institution are ranked.

2. *Applicant behavior:* younger cohorts might rank the degree higher in their rank-ordered list, increasing their likelihood of enrollment if they meet or exceed the admission cutoff rank.

Our results suggest that degree response is the primary driver of the difference between application and enrollment spillovers. Indeed, as shown in Appendix Table F.2, we find that, conditional on applying, degrees tend to rank applicants from high schools with a marginally admitted older schoolmate higher than those from high schools with a marginally rejected older schoolmate. While the estimates are consistently positive, they are imprecisely estimated and generally not statistically significant at conventional levels. Conversely, in Appendix Table F.3, we find no significant difference in how students rank the degree in their rank-ordered lists, regardless of whether they come from high schools with marginally admitted or rejected older schoolmates. This suggests that the observed spillover effects on enrollment are not driven by changes in applicant preferences or ranking strategies.

Table 2: Older Schoolmate Spillovers on Applications to and Enrollment in Degree, HE Institution and Field of Study of Marginally Admitted Older Schoolmate

	Degree Spillovers				HE Institution Spillovers				Field of Study Spillovers			
	Applications		Enrollment		Applications		Enrollment		Applications		Enrollment	
	At least one	Number	At least one	Number	At least one	Number	At least one	Number	At least one	Number	At least one	Number
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Older schoolmate above cutoff (ITT)	0.012***	0.038***	0.005***	0.006***	0.011***	0.316***	0.011***	0.087***	0.003	-0.007	0	-0.01
	(0.003)	(0.013)	(0.002)	(0.002)	(0.003)	(0.063)	(0.003)	(0.021)	(0.003)	(0.043)	(0.003)	(0.013)
% of counterfactual mean	3.91	4.65	9.3	9.52	1.92	7.18	5.56	11.05	0.47	-0.16	-0.13	-0.99
Older schoolmate enrols (2SLS)	0.057***	0.19***	0.023***	0.032***	0.063***	1.956***	0.068***	0.547***	0.026	0.009	0.001	-0.066
	(0.013)	(0.06)	(0.007)	(0.011)	(0.017)	(0.357)	(0.014)	(0.118)	(0.02)	(0.341)	(0.021)	(0.104)
% of counterfactual mean	18.81	23.49	45.17	48.54	11.42	44.48	33.69	69.5	4.18	0.2	0.22	-6.57
Degree-year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Obs. (right)	189,361	189,361	189,361	189,361	189,361	189,361	189,361	189,361	189,361	189,361	189,361	189,361
Obs. (left)	186,205	186,205	186,205	186,205	186,205	186,205	186,205	186,205	186,205	186,205	186,205	186,205
Counterfactual mean [-5,-1]	0.3	0.81	0.051	0.065	0.55	4.397	0.203	0.787	0.614	4.513	0.337	1.012
Bandwidth	20.96	20.96	20.96	20.96	20.96	20.96	20.96	20.96	20.96	20.96	20.96	20.96
First stage	0.217***	0.217***	0.217***	0.217***	0.172***	0.172***	0.172***	0.172***	0.127***	0.127***	0.127***	0.127***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
First stage F-stat	19,506	19,506	19,506	19,506	8,161	8,161	8,161	8,161	3,357	3,357	3,357	3,357

Notes: This table reports estimates of within-high school spillovers. "At least one application" refers to the probability that a high school has at least one application to the degree, HE institution or field of study in $t+1$ of a marginally admitted older schoolmate in t; "number of applications" refers to the number of applications to the marginally admitted students' degree, HE institution or field of study. The same applies for the enrollment outcomes. The first row for each outcome presents intent-to-treat estimates, while the second row presents 2SLS estimates in which older schoolmates enrollment is instrumented with them being ranked above the admission rank cutoff. All the specifications in the table correspond to local linear regressions using a triangular kernel, and include degree-year fixed effects. The bandwidth (21.32) corresponds to the smallest MSE-optimal bandwidth (Calonico et al., 2014) for degree outcomes. Standard errors clustered at the high school-year level are reported in parentheses. ***, **, and * indicate significance at 1, 5, and 10%, respectively.

Treatment intensity. We examine whether spillover effects are amplified when multiple older schoolmates are admitted to the same degree. Using cases where high schools have exactly two ranked applicants in the treatment year (33% of our high school Œ degree Œ year observations), we analyze how effects differ between first and second-ranked applicants being marginally admitted.

When examining second-ranked applicants, control high schools start from a higher baseline: they have on average 0.20 students admitted in the treatment year, likely because their best-ranked applicant was above the cutoff. Nevertheless, being just above the cutoff for the second-ranked applicant leads to a significant increase of 0.195 in the average number of admitted students, compared to 0.27 for first-ranked applicants. Table **??** presents the results. For the extensive margin of applications, we find similar effects whether looking at first or second-ranked applicants (2.7 and 2.4 percentage points respectively, both statistically significant). However, while first-ranked applicants generate significant effects on the number of applications (0.091 additional applications), we find no such effect for second-ranked applicants (-0.009, not significant). Enrollment effects are smaller and statistically insignificant for both first and second-ranked applicants.

These results suggest that while having an additional older schoolmate admitted doesn't diminish the basic spillover effect, it also doesn't substantially amplify it. This pattern aligns with our teacher mechanism findings: if teachers are key conduits of information about admission possibilities, the admission of a single student may be sufficient to make them aware of specific degree programs and their accessibility to future cohorts.

4.2 Robustness

Our results are qualitatively robust to standard regression discontinuity robustness checks. Figure 8 visually presents estimates from six different specifications, including our baseline, illustrating the broad stability of our results.

Bandwidth size. Appendix Figure D.1 shows that our baseline estimates for degree-specific older schoolmate spillovers remain stable across various bandwidth choices. However, institution-specific spillovers are very sensitive to the bandwidth chosen, shrinking to zero for small bandwidths, suggesting they are null and statistically insignificant. Field of study spillovers consistently show very small and insignificant effects across bandwidths.

Alternative specifications. We evaluate our baseline estimates' stability under several alternative specifications: (i) including a second-order polynomial of the running vari-

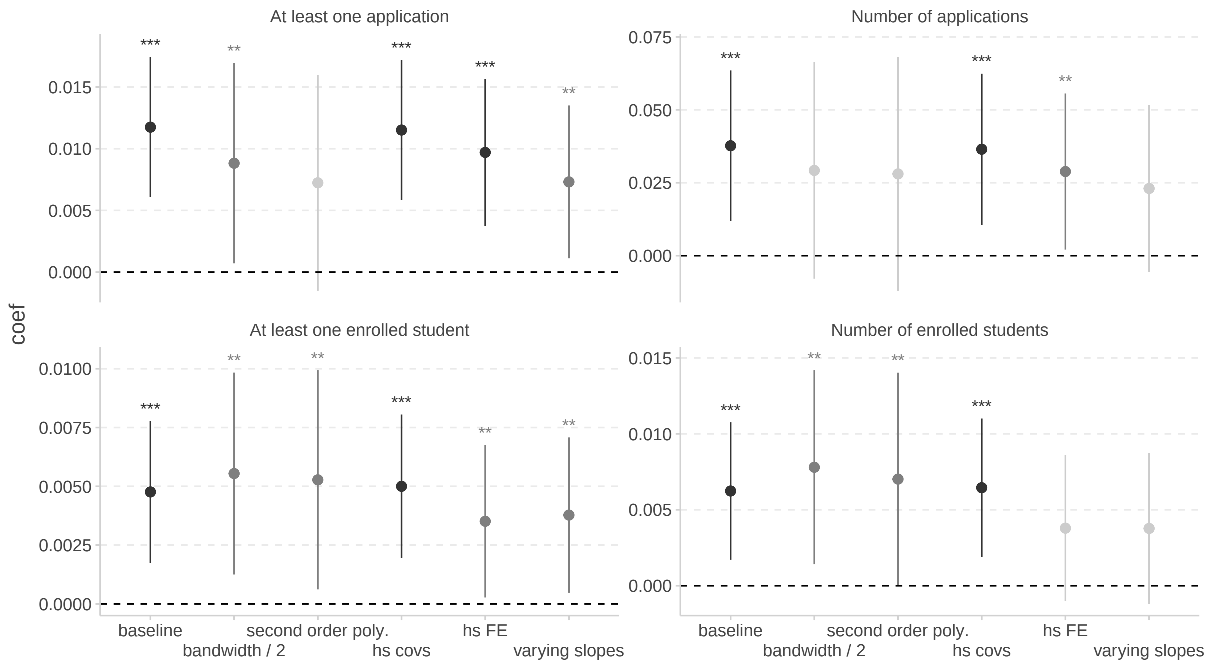


Figure 8: Robustness of Baseline Older Schoolmate Spillovers to Alternative Specifications

Notes: This figure presents estimates from six different specifications, including our baseline.

able, (ii) including time-varying high school characteristics as controls, (iii) including high school-year fixed effects, and (iv) allowing slopes to vary at each degree admission cutoff rank. Figure 8 presents these results. While some estimates lose precision in certain specifications, the coefficient magnitudes remain largely consistent across specifications.

Randomization inference. To ensure our results are not driven by chance, we implement a randomization inference approach. Specifically, we randomize the rank of each high school for degrees with at least one older schoolmate applicant. Figure E.6 displays the results, showing the randomized estimates are centered around zero, as expected, with our baseline estimates being far from this distribution.

Placebo cutoffs. Appendix Figure D.2 demonstrates our estimates' robustness to placebo thresholds. We estimate discontinuities at various points along the running variable, finding significant effects only at the actual cutoff.

Placebo outcomes. We estimate discontinuities in our four main outcomes measured in the year prior to treatment. We should observe no spillover effects in the pre-treatment year because by construction these students' choices occur before the treatment takes place. Figure 9 presents these results, focusing on treatment years 2013-

2015. This range allows us to measure outcomes in $t-1$ (pre-treament year), t (treatment year), $t+1$ (baseline outcomes) and $t+2$ (snowball effect) for a constant sample of high schools and degrees. All coefficients in $t-1$ are insignificant, aligning with our expectations.

4.3 Older Schoolmate Spillovers Persist Over Time

Our analysis thus far has demonstrated that older peers' enrollment decisions influence the higher education choices of the following cohort of students from the same high school. However, the impact of these decisions may extend beyond a single year. Figure 9 provides compelling evidence that older schoolmate spillovers persist over time, creating what we term "snowball effects."

The coefficients in $t+2$, representing spillover effects two years after the initial treatment, remain substantial and statistically significant, albeit slightly smaller than those in $t+1$. For the extensive margin of applications, the $t+2$ spillovers are 75% as large as those in $t+1$, while for the intensive margin, they are 65% as large. This persistence underscores the path-dependent nature of higher education choices within high schools, demonstrating that seemingly minor shocks can continue to influence

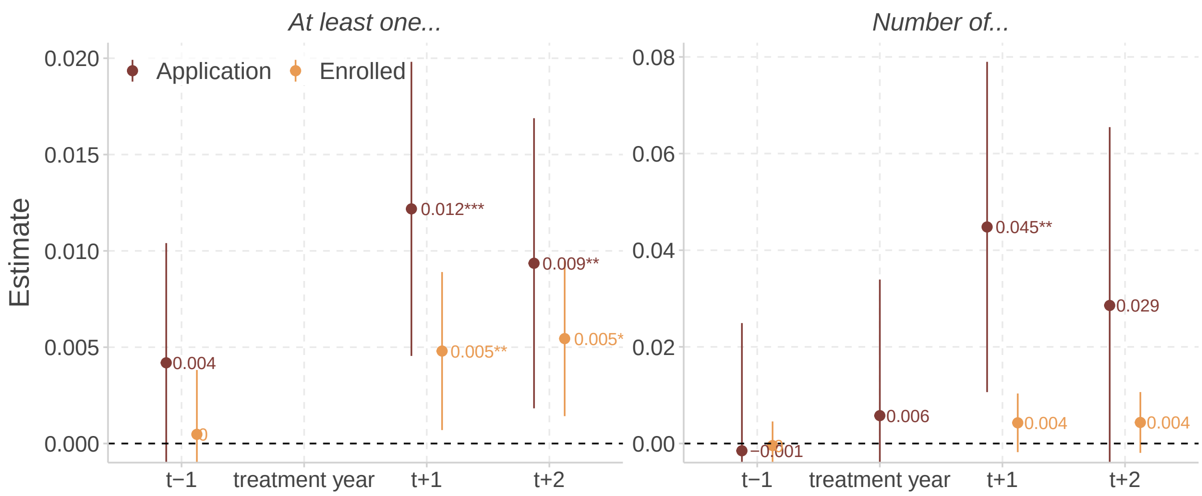


Figure 9: Placebo and Snowball Effects

Notes: This figure shows estimates of within-high school spillovers in applications and enrollments for outcomes measured in different years. "$t-1$" refers to the year prior to the older schoolmate's marginal admission, "treatment year" refers to the year of an older schoolmate's marginal admission, and $t+1$ and $t+2$ correspond, respectively to high schools' application and enrollment outcomes one and two years following the marginal admission of one of its students. The sample is restricted to treatment years 2014 and 2015 to ensure the sample is constant across estimates. All the specifications in the figure correspond to local linear regressions using a triangular kernel, and include college-major - year fixed effects. The bandwidth (20.96) corresponds to the smallest MSE-optimal bandwidth (Calonico et al., 2014) for the main college-major outcomes. Statistically significance is based on standard errors clustered at the high school - year level. Confidence intervals correspond to 95% confidence intervals. ***, **, and * indicate significance at 1, 5, and 10%, respectively.

students at least two years after the initial event. Intriguingly, the coefficients for enrollment do not decrease in $t + 2$. This observation lends support to the hypothesis that programs may adjust their perceptions of student quality from high schools that have previously sent few students. Such adjustments could explain the sustained enrollment effects over time.

These findings have significant implications for understanding the dynamics of higher education choices and the long-term impacts of individual enrollment decisions. They suggest that interventions or changes in enrollment patterns could have cascading effects on future cohorts, potentially amplifying initial impacts over time. This persistence highlights the importance of considering long-term consequences when designing policies or interventions in higher education admissions processes.

4.4 Older Schoolmate Spillovers on HE Institution, Field of Study and Similar Degrees

HE Institution and Field of Study. Students' enrollment decisions could also have broader spillovers on the higher education institution or field of study applications and enrollments of students in the subsequent cohort of the same high school. So far we have shown there are sizable degree-specific spillovers from one high school cohort to the next. Each degree corresponds to a higher education institution and a field of study and thus either component could influence subsequent students' higher education decisions.

We undertake the exact same analysis as for degrees, this time for institutions, and for subjects separately. Due to the specificity of the French higher education system, fields of study are partly institution-specific because some tracks are only offered in some types of institutions. For example, technical degrees are only offered in University Institutes of Technology *(Instituts Universitaires de Technologie (IUT))* and vocational degrees are only offered in Sections of Superior Technicians *(Sections de Techniciens Supérieurs (STS))*.

Appendix Figures E.7a and E.7b and Appendix Table 2 report the results. We find positive older schoolmate spillovers on applications and enrollments in a given institution, though these are very sensitive to the chosen bandwidth (see Appendix Figure D.1) and therefore are unreliable, and no spillovers on fields of study.

Similar degrees. Our analysis has shown that an older schoolmate's enrollment in a specific degree increases the likelihood of younger students applying to that same degree. This finding suggests a potential learning effect about the degree. However, this influence may extend beyond the exact degree to similar programs. For instance, if a student's high school experiences a "shock" with a peer enrolling in a Law degree at Sorbonne University, it might induce applications not only to that specific program but also to law degrees at other Paris universities or to related programs like Law and Economics at the same institution.

To test this hypothesis, we estimate spillover effects on similar degrees. We characterize degree similarity using the top three ranked applications of admitted students for each degree. We then group degrees based on their relative frequency and categorize them into quartiles, ranging from the 25% most similar to the 25% least similar. Additionally, we estimate spillovers for unrelated degrees by randomly selecting 10 degrees not present in the top three ranked applications of admitted students.

Figure 10 presents the results of this analysis. To facilitate comparison across different spillovers, we present the estimates as a percentage of the baseline mean. Our findings reveal significant spillover effects for the most similar degrees, with a magnitude of approximately 50% of the baseline effect. In contrast, estimates for less similar degrees are small in magnitude and statistically insignificant. These results suggest that the influence of an older schoolmate's enrollment is relatively targeted. Students affected by a peer's enrollment in a particular degree are more likely to apply not only to that specific degree but also to very similar programs. However, this effect does not seem to extend to less related degrees.

This pattern of spillovers provides insights into how information or influence flows

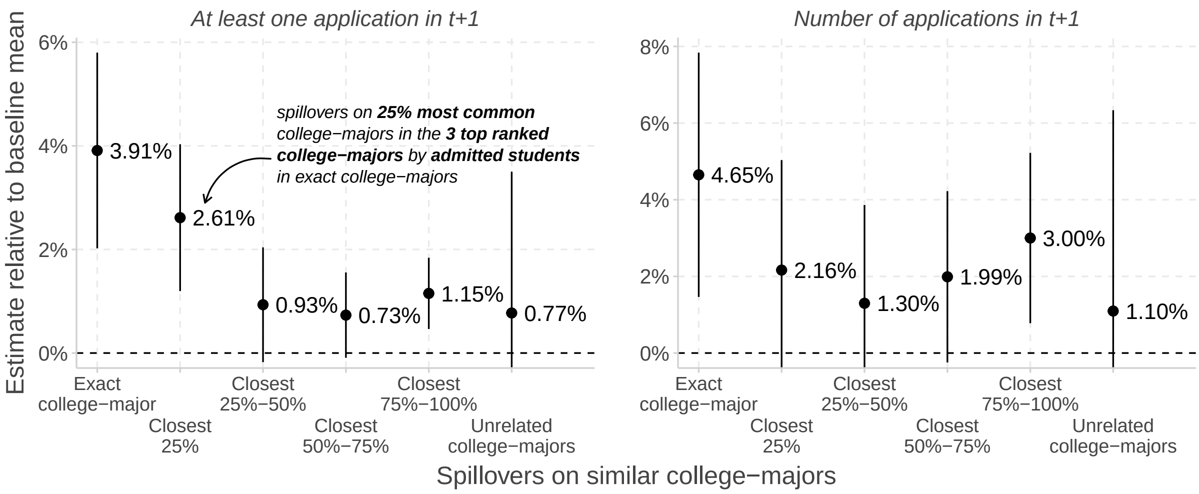


Figure 10: Older Schoolmate Spillovers on Similar Higher Education Degrees

Notes: This figure shows estimates of within-high school spillovers in applications for similar college-majors. "Closest 25%" refers to college-majors that account for the 25% most commonly ranked among the top 3 ranked college-majors of admitted students. "Unrelated college-majors" refers to 10 randomly selected college-majors. All the specifications in the figure correspond to local linear regressions using a triangular kernel, and include college-major - year fixed effects. The bandwidth (20.96) corresponds to the smallest MSE-optimal bandwidth (Calonico et al., 2014) for the main college-major outcomes. Statistically significance is based on standard errors clustered at the high school - year level. Confidence intervals correspond to 95% confidence intervals. ***, **, and * indicate significance at 1, 5, and 10%, respectively.

within high schools. It suggests that students may be gaining specific knowledge about certain academic paths or career options, rather than experiencing a general increase in higher education aspirations.

5 Heterogeneity

Having established the presence of significant spillovers across cohorts, we now examine how these effects vary across different dimensions to better understand their nature and implications. We find that the estimates are remarkably stable across the five years of our sample period, despite not always reaching conventional significance levels. This temporal consistency suggests that these spillovers represent a structural rather than transitory influence on students' higher education choices. To further characterize these effects, we examine variation along four key dimensions: student characteristics, high school attributes, program characteristics, and the interaction between high school and program features. This comprehensive analysis of heterogeneity not only helps identify which students are most influenced by their older schoolmates, but also illuminates the conditions under which these influences are strongest. We discuss each dimension in turn below.

5.1 Student Characteristics

We begin by examining how spillover effects vary across student characteristics to understand which students are most susceptible to be influenced by marginally admitted older schoolmates. Figure 11 presents these results. We find no significant differences in either application or enrollment effects between male and female students, suggesting that older schoolmate influence operates similarly across genders. Using the French Ministry of Education's classification based on students' legal guardians' occupations, we categorize students into low (24% of applicants), middle (28%), high (13%), and very high (34%) socioeconomic status (SES) groups. Surprisingly, we find relatively similar levels of responsiveness between low and very high SES students. This pattern is notable given the conventional wisdom that low SES students, who typically have less information about higher education options, might rely more heavily on peer experiences. This finding suggests that the information transmission mechanism we document may operate similarly across socioeconomic groups.

5.2 High School Characteristics

The results in Figure 12 reveal notable variation across high school characteristics. While we observe similar magnitude spillovers across most academic tracks (with

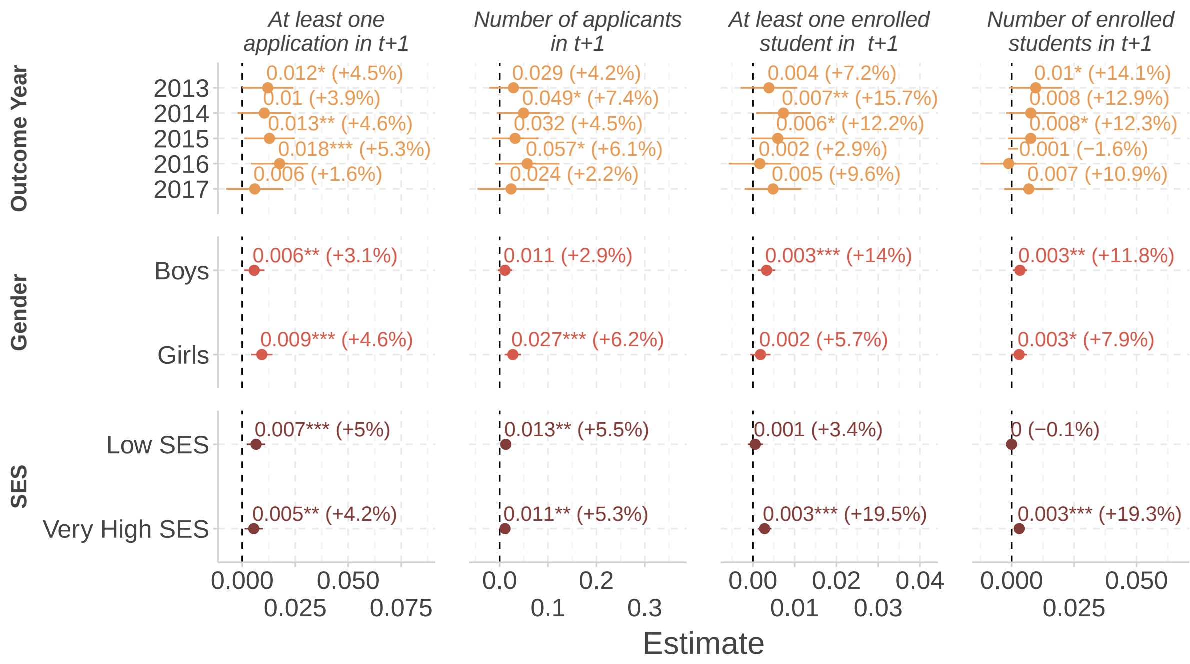


Figure 11: Heterogeneity by Student Characteristics

Notes: This figure shows estimates of within-high school spillovers in applications and enrollments for subsamples of students. The application and enrollment outcomes are reported in the figure facet titles, while the subsamples are reported on the y-axis. Socioeconomic status (SES) is based on students' legal guardian's occupation. All the specifications in the figure correspond to local linear regressions using a triangular kernel, and include college-major - year fixed effects. The bandwidth (20.96) corresponds to the smallest MSE-optimal bandwidth (Calonico et al., 2014) for the main college-major outcomes. Statistically significance is based on standard errors clustered at the high school - year level. Confidence intervals correspond to 95% confidence intervals. ***, **, and * indicate significance at 1, 5, and 10%, respectively. The percentage change relative to the counterfactual mean ($[-5, -1]$) is shown in parenthesis next to the estimates.

slightly larger effects in social sciences and literature), school size emerges as a key differentiating factor. Spillover effects are strongest in small high schools (less than 30 students) and decrease with school size. This pattern could reflect stronger student-teacher relationships in small schools, better alumni networks, or faster information diffusion in smaller school communities. Examining variation by school academic level (measured by median high school exit exam—*Baccalauréat*—scores and grouped into quintiles), we find no systematic relationship between school performance and spillover intensity. While some effects appear in the second quintile and top quintile for the extensive margin of applications, the overall pattern suggests that a school's academic standing is not a primary determinant of cross-cohort influence.

5.3 Degree Characteristics

We next examine how older schoolmate spillovers vary across degree characteristics. Our analysis reveals a clear pattern: spillover effects are strongest for public university,

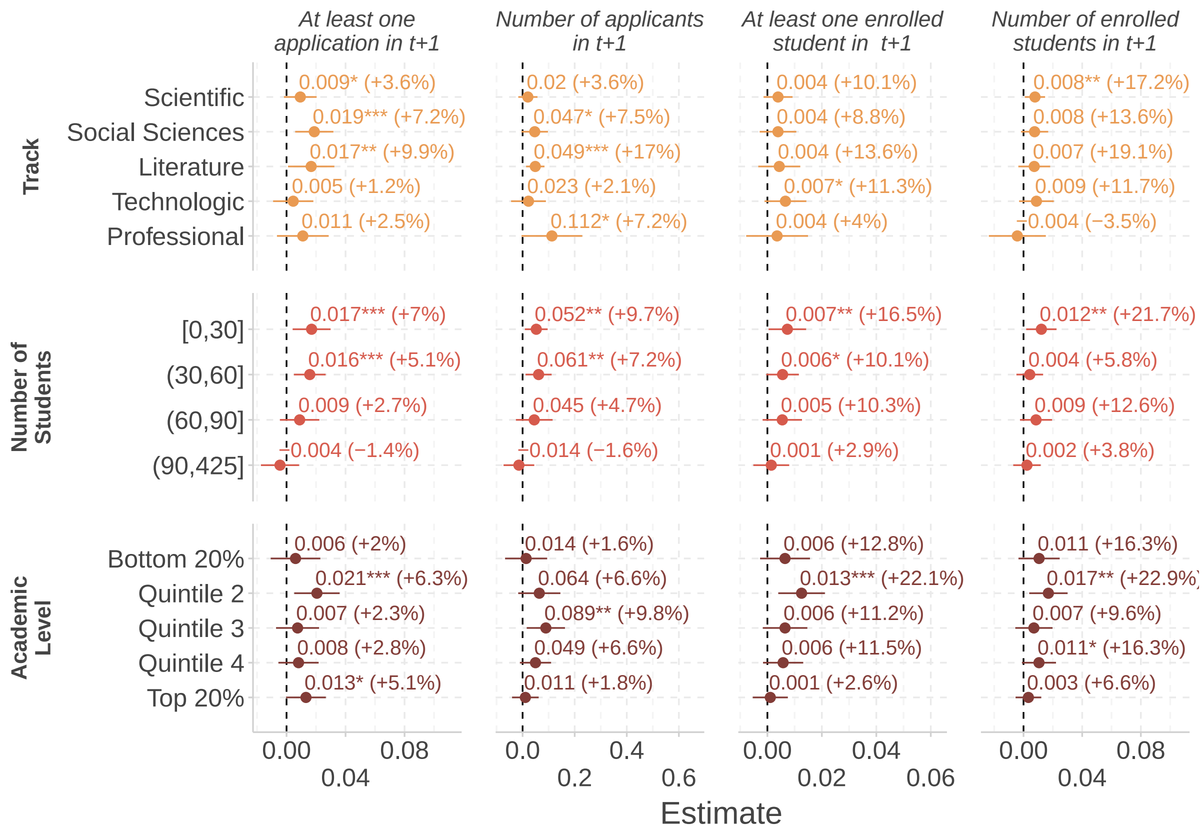


Figure 12: Heterogeneity by High School Characteristics

Notes: This figure shows estimates of within-high school spillovers in applications and enrollments for subsamples of high schools. The application and enrollment outcomes are reported in the figure facet titles, while the subsamples are reported on the y-axis. High schools' academic level is defined as the median of its students' end of high school exam (Bac) grades. The quintiles of academic level are calculated over *all* high schools in the full sample, not only among high schools in the regression discontinuity sample. All the specifications in the figure correspond to local linear regressions using a triangular kernel, and include college-major - year fixed effects. The bandwidth (20.96) corresponds to the smallest MSE-optimal bandwidth (Calonico et al., 2014) for the main college-major outcomes. Statistically significance is based on standard errors clustered at the high school - year level. Confidence intervals correspond to 95% confidence intervals. ***, **, and * indicate significance at 1, 5, and 10%, respectively. The percentage change relative to the counterfactual mean ($[-5, -1]$) is shown in parenthesis next to the estimates.

technical, and professional programs, while we find no statistically significant effects for the prestigious "preparatory classes" (two-year intensive programs to enter *Grandes Écoles*, elite graduate schools) or other degrees. This pattern is reinforced when examining program selectivity: spillovers are concentrated among programs in the lower deciles of selectivity, as measured by enrolled students' median high school exit exam scores. At first glance, this pattern might suggest that informational barriers play a larger role than aspirational barriers in our setting. However, as we show below, the relationship between program selectivity and spillover effects becomes more complex when considering interactions with high school characteristics.

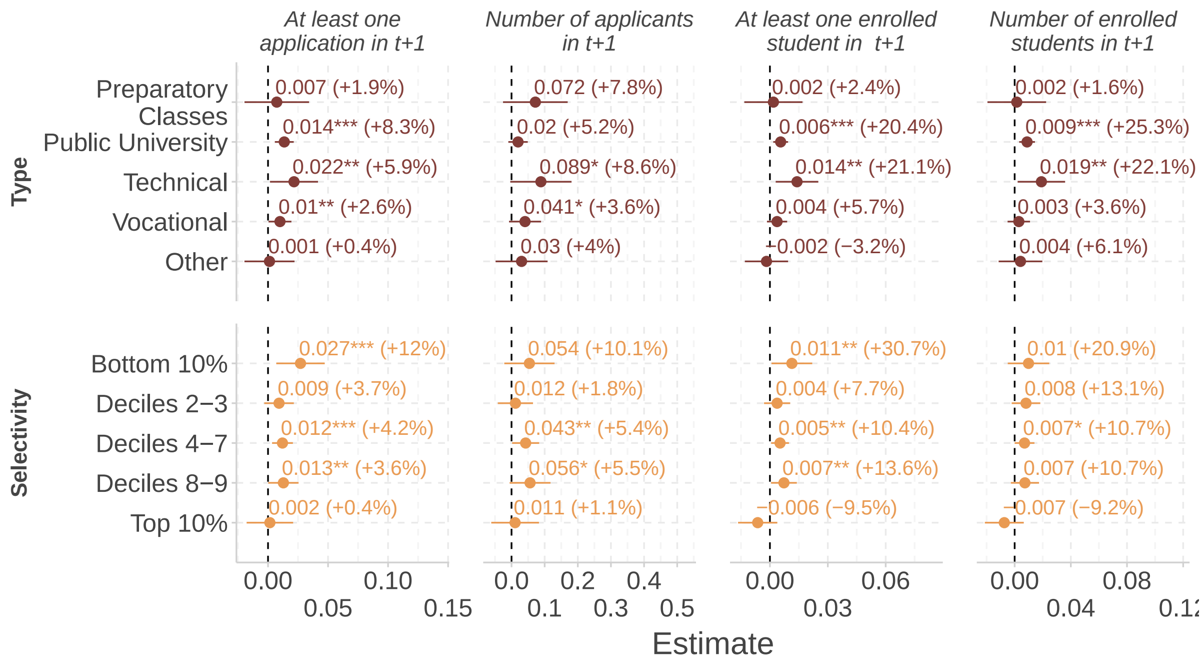


Figure 13: Heterogeneity by College-Major Characteristics

Notes: This figure shows estimates of within-high school spillovers in applications and enrollments for subsamples of college-majors. The application and enrollment outcomes are reported in the figure facet titles, while the subsamples are reported on the y-axis. College-majors' selectivity is measured as median end-of-high school exam grade of enrolled students. The deciles of selectivity are calculated over *all* college-majors in the full sample, not only among college-majors in the regression discontinuity sample. All the specifications in the figure correspond to local linear regressions using a triangular kernel, and include college-major - year fixed effects. The bandwidth (20.96) corresponds to the smallest MSE-optimal bandwidth (Calonico et al., 2014) for the main college-major outcomes. Statistically significance is based on standard errors clustered at the high school - year level. Confidence intervals correspond to 95% confidence intervals. ***, **, and * indicate significance at 1, 5, and 10%, respectively. The percentage change relative to the counterfactual mean ($[-5, -1]$) is shown in parenthesis next to the estimates.

5.4 Interaction Between High School and Degrees Characteristics

Lastly, we examine interactions between high school and degree characteristics. Using our measures of high school academic level and degree selectivity (both based on median high school exit exam scores), we focus on high schools in the bottom and top quartiles and examine their responses to degrees across the selectivity spectrum. Two striking patterns emerge: students from lower-performing high schools show stronger responses when an older schoolmate is marginally admitted to a highly selective program (top 25%), potentially reflecting increased aspirations among these schools' top students. This suggests that exposure to successful older peers can expand students' horizons, making selective programs feel more attainable. Conversely, students from top-performing high schools show large responses to older schoolmates admitted to less selective programs (bottom 25%), indicating that peer influences can broaden the set of programs students consider beyond their school's typical trajectory.

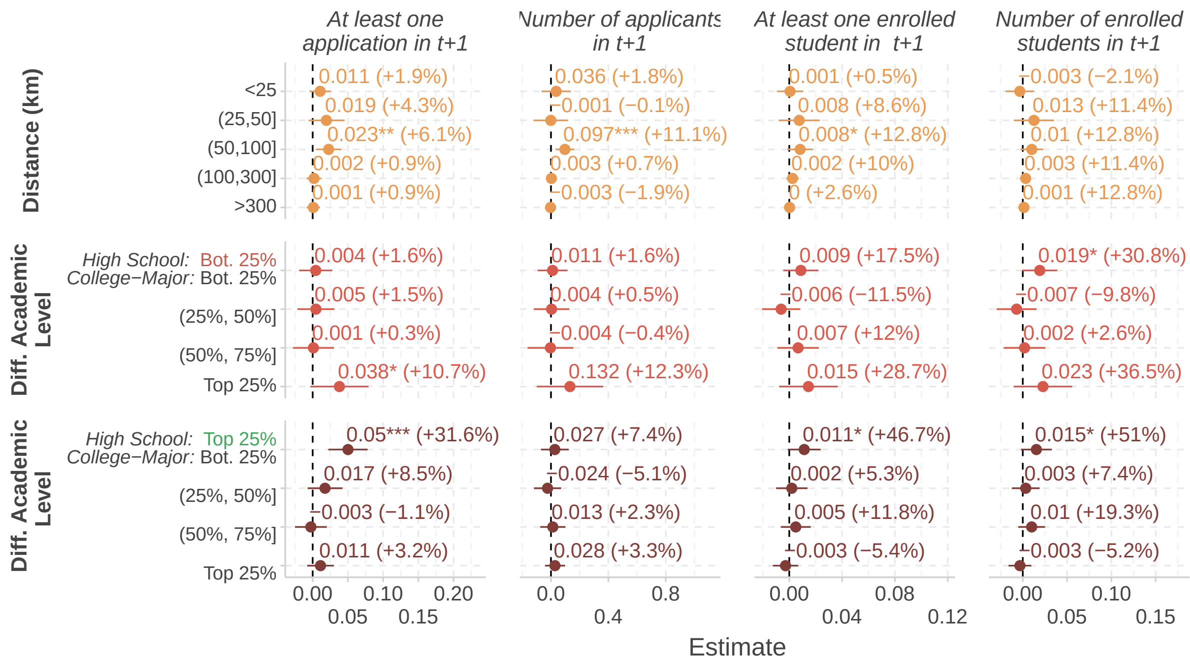


Figure 14: Heterogeneity by High School *and* College-Major Characteristics

Notes: This figure shows estimates of within-high school spillovers in applications and enrollments for subsamples of high schools and college-majors. The application and enrollment outcomes are reported in the figure facet titles, while the subsamples are reported on the y-axis. See Figures 12 and 13's notes for details on the definitions of high schools academic level and college-major selectivity, used for the *Diff. Academic Level* results. All the specifications in the figure correspond to local linear regressions using a triangular kernel, and include college-major - year fixed effects. The bandwidth (20.96) corresponds to the smallest MSE-optimal bandwidth (Calonico et al., 2014) for the main college-major outcomes. Statistically significance is based on standard errors clustered at the high school - year level. Confidence intervals correspond to 95% confidence intervals. ***, **, and * indicate significance at 1, 5, and 10%, respectively. The percentage change relative to the counterfactual mean ($[-5, -1]$) is shown in parenthesis next to the estimates.

Geographic distance between the high school and the degree also shapes spillover effects.[18] While nearby programs might be more familiar to students and distant ones involve higher transport costs, we find that programs at moderate distances (50-100 km) generate the strongest effects, both in applications and enrollment. This suggests that older schoolmates' experiences are particularly valuable for programs that are beyond immediate local knowledge but still within a manageable distance.

6 Mechanisms

Having established the presence and heterogeneity of older schoolmate spillovers, we now examine two non-mutually exclusive mechanisms that might drive these effects: (i) teacher influence, and (ii) student homophily/role model effects. Teachers may

[18]We obtain high schools' and degrees' precise geographic location (longitude and latitude) from available open data. Exact geographic location is missing for 7% of high schools and 4% of degrees.

recommend specific degrees based on their past students' higher education choices, thereby serving as conduits of information across cohorts. Students may also be more responsive to older schoolmates who share similar characteristics such as gender or socioeconomic background, suggesting a role for homophily and role models. Understanding these mechanisms helps inform potential policy interventions that could replicate these spillover effects.

6.1 Teacher Influence

To examine the role of teachers, we leverage data on France's "principal teachers"—i.e., teachers who, in addition to teaching their subject-area, are in charge of all administrative duties for their assigned class and, in particular, of helping their students with higher education applications. Survey evidence from the French Ministry of Education and Higher Education suggests that over 60% of students discuss their higher education plans with these principal teachers, who are more likely to be well-informed about their students' higher education choices relative to non-principal teachers (MESR, 2023). This institutional feature allows us to test the teacher mechanism: we examine whether students who share the same principal teacher with the marginally admitted older schoolmate are more likely to apply to the older schoolmate's degree than students who do not share the same principal teacher. We restrict the analysis to high schools with at least two classes, with different principal teachers and at least one in common with older schoolmates, to ensure the subsample is consistent across estimation samples.

Table 3 displays the results of this analysis. The estimates for students sharing the same principal teacher as the marginally admitted older schoolmate are significantly larger than those for students not sharing the same principal teachers. These findings complement a growing literature on teachers' long-run effects on students' educational trajectories. Chetty et al. (2011) show that kindergarten teachers affect their students' college attendance, while Jackson (2018) demonstrates that teachers influence non-academic outcomes crucial for educational attainment. In higher education specifically, Carrell et al. (2010) and Lim and Meer (2020) find that teachers shape students' college major choices, while Mulhern (2023) shows substantial effects of school counselors on college enrollment. We interpret these results as suggesting teachers can play an important role in mediating the within-high school spillovers we documented above.

6.2 Student Homophily/Role Model

Next, we examine the role played by homophily and role models in shaping older schoolmate spillovers. The literature on role models has shown substantial impacts

Table 3: Effect of Principal Teacher on Older Schoolmate Spillovers

	Same principal teacher as older schoolmate		Different principal teacher as older schoolmate	
	At least one application (1)	Number of applications (2)	At least one application (3)	Number of applications (4)
Older schoolmate above cutoff (ITT)	0.01**	0.026**	-0.001	0.001
	(0.004)	(0.011)	(0.005)	(0.015)
% of counterfactual mean	5.32	7.5	-0.37	0.2
Degree-year FE	✓	✓	✓	✓
Obs. (right)	73,706	73,706	73,706	73,706
Obs. (left)	71,935	71,935	71,935	71,935
Counterfactual mean [-5,-1]	0.182	0.348	0.223	0.484
Bandwidth	20.96	20.96	20.96	20.96

Notes: This table shows estimates of older schoolmate spillovers in applications for students sharing the same *principal* teacher and class number as the marginally admitted older schoolmate and for students who do not. The application outcomes are reported in the figure facet titles. All the specifications in the figure correspond to local linear regressions using a triangular kernel, and include degree-year fixed effects. The bandwidth (20.96) corresponds to the smallest MSE-optimal bandwidth (Calonico et al., 2014) for the main degree outcomes. Statistically significance is based on standard errors clustered at the high school-year level. Confidence intervals correspond to 95% confidence intervals. ***, **, and * indicate significance at 1, 5, and 10%, respectively. The percentage change relative to the counterfactual mean ($[-5, -1]$) is shown in parenthesis next to the estimates.

of demographic matching: female professors significantly increase girls' likelihood of pursuing STEM degrees (Carrell et al., 2010) and economics degrees (Canaan and Mouganie, 2021). Similarly, exposure to Black teachers substantially increases college enrollment for Black students (Gershenson et al., 2022), while interactions with female scientists significantly affect girls' STEM applications (Breda et al., 2023).

In the spirit of these studies, we examine whether students are more responsive to older schoolmates who share their demographic characteristics. Specifically, we test for gender matching effects (whether girls follow female older schoolmates) and socioeconomic status matching (whether low-SES students follow low-SES older schoolmates). Since all students within a high school likely have similar information about previous cohorts' outcomes, evidence of differential effects by demographic matching would suggest that older schoolmates serve as some form of role models rather than mere information sources. Importantly, because our regression discontinuity design compares students of the same gender or SES background on both sides of the admission cutoff, any differences in effects cannot be explained by gender- or SES-specific program preferences. Figure 15 presents these results, with underlying visual evidence in Appendix Figures E.9-E.12.

We find strong evidence for gender-specific role model effects, but only for girls. Female students are 1.6 percentage points (+8%) more likely to apply to the same degree as a marginally admitted female older schoolmate, while showing no significant response to male older schoolmates. This pattern holds for both the extensive and intensive margins of applications. Male students, however, show no differential response based on the gender of the older schoolmate.

The socioeconomic dimension reveals symmetric role model effects. Low-SES students are 1.2 percentage points (+10%) more likely to follow a marginally admitted

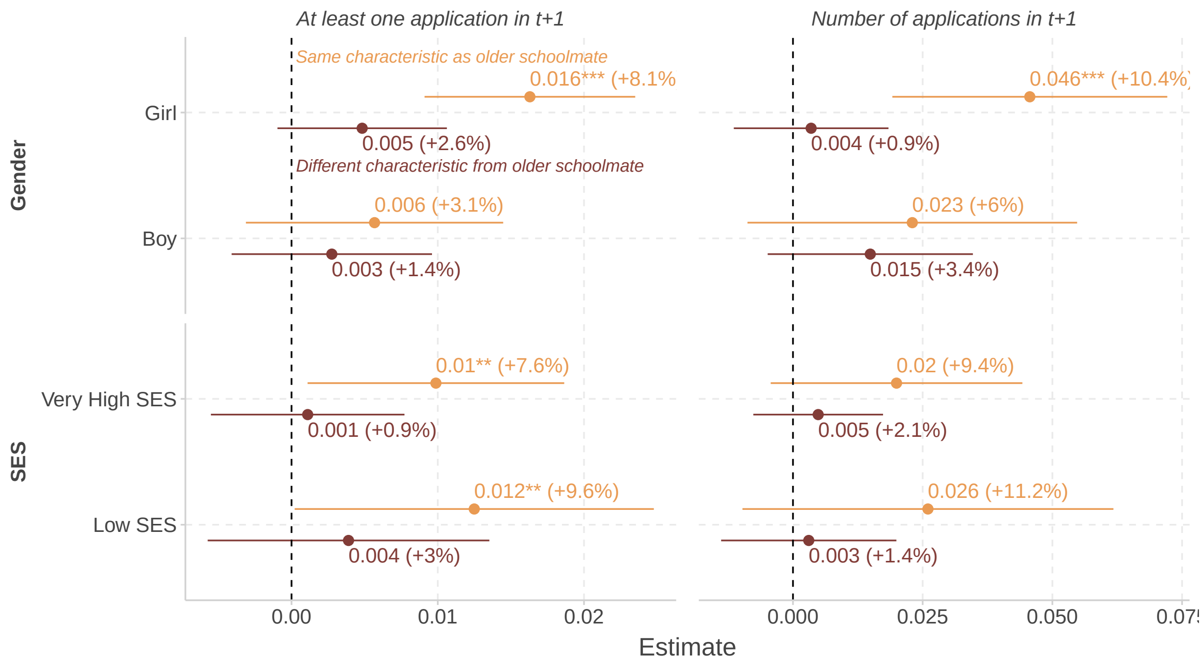


Figure 15: Older Schoolmate Gender- and SES-Specific Spillovers

Notes: This figure shows estimates of within-high school spillovers in applications for students sharing the same gender and SES as the marginally admitted older schoolmate and for students who do not. The application outcomes are reported in the figure facet titles, while the marginally admitted older schoolmate's characteristic are reported on the y-axis. Socioeconomic status (SES) is based on students' legal guardian's occupation. All the specifications in the figure correspond to local linear regressions using a triangular kernel, and include degree-year fixed effects. The bandwidth (20.96) corresponds to the smallest MSE-optimal bandwidth (Calonico et al., 2014) for the main college-major outcomes. Statistically significance is based on standard errors clustered at the high school-year level. Confidence intervals correspond to 95% confidence intervals. ***, **, and * indicate significance at 1, 5, and 10%, respectively. The percentage change relative to the counterfactual mean ($[-5, -1]$) is shown in parenthesis next to the estimates.

low-SES older schoolmate but show minimal response (+0.4 percentage points) to very high-SES older schoolmates. Similarly, very high-SES students are 1 percentage point (+8%) more likely to follow very high-SES older schoolmates while being unresponsive (+0.1 percentage points) to low-SES older schoolmates. This symmetric pattern aligns with our earlier finding that average responsiveness to older schoolmates does not vary substantially by student SES.

The asymmetric gender effects we find—with girls responding strongly to female older schoolmates while boys show no gender-specific response—echo similar patterns documented in other educational contexts. Breda et al. (2023) find that female role models increase girls' interest in math-intensive fields with no effects on boys, while Porter and Serra (2020) show female role models increase women's propensity to major in economics without affecting men's choices. This aligns with research in social psychology finding that female students draw greater inspiration from and more strongly identify with female than male role models, while male students' responses appear less influenced by role model gender (Lockwood, 2006).

More broadly, our results connect to the literature on identity and educational choice. Akerlof and Kranton (2000) theorize that students make educational choices that align with their social identity, while Ray (2006) argues that individuals form aspirations through social learning from "attainable" role models—those similar enough to make their achievements feel possible. Our findings suggest that seeing a demographically similar peer enrol in a specific degree makes that path feel more viable, perhaps by making abstract possibilities more concrete and attainable. This interpretation is particularly relevant for understanding why both low and very high-SES students are more responsive to same-SES peers, as it suggests students across the socioeconomic spectrum look to those with similar backgrounds when forming their educational aspirations.

7 Implications of Older Schoolmate Spillovers

To assess the broader implications of older schoolmate influences on higher education inequalities, we conduct a counterfactual analysis examining how low SES students' applications would change if they were exposed to the same set of older schoolmates as their very high SES peers.[19] We first focus on high-achieving students, i.e., those in the top 10% of the academic ability distribution, who have been shown to disproportionately apply to less selective institutions/degrees than their academic credentials would warrant (Hoxby and Avery, 2013; Campbell et al., 2023). We then generalize this analysis to the entire academic ability distribution.

We measure students' academic ability using their percentile rank in the national distribution of high school exit exam scores (*Baccalauréat*) by academic year, and we define degree selectivity based on median exam score of admitted students in the previous year in each degree-year combination, which we then rank into percentiles by academic year. Among students in the top decile of the exam score distribution, low SES students are 27 percentage points less likely to apply to degrees in the top selectivity decile compared to their very high SES counterparts of similar academic ability (the full student academic decile - degree selectivity decile application matrix can be found in Appendix Figure E.13). This stark disparity in application behavior might partly reflect differential exposure to high-achieving older peers: while 95% of high-achieving, very high SES students have at least one older schoolmate who enrolled in a top-decile degree, only 75% of similarly academically able low SES students do (the full older schoolmate exposure matrix can be found in Appendix Figure E.14).

[19] We continue to follow the socioeconomic status (SES) categories constructed by the French Ministry of Education based on students' legal guardian's occupation: low SES (XX% of all applicants), middle SES (XX%), high SES (XX%), and very high SES (XX%).

To quantify how older schoolmate exposure shapes these application gaps, we conduct a counterfactual analysis that simulates high-achieving, low SES students' applications if they were exposed to the same older schoolmates as their high-achieving, very high SES peers. Our methodology proceeds in three steps. First, we identify the subset of high-achieving, low SES students with no older schoolmates enrolled in top-decile degrees, which comprises 25% of all high-achieving, low SES students. Second, we calculate how many of these students' applications should be adjusted using our baseline 2SLS estimate and the observed exposure gap. Our baseline estimate suggests that having a marginally enrolled older schoolmate increases the number of applicants to a degree by 23.5% (Table 2, column 2). We aim to close the exposure gap between low and very high SES students: while 95% of high-achieving, very high SES students have at least one older schoolmate who enrolled in a top-decile degree, only 75% of high-achieving, low SES students do. To equalize exposure, we need to expose an additional 20 percentage points of high-achieving, low SES students to top-decile degrees. Since these 20 percentage points represent 80% (20/25) of students with no current exposure, and the spillover effect is 23.5%, we calculate that 18.8% (23.5% Œ 80%) of students without current exposure should be affected. Finally, we randomly select this calculated number of students and elevate their most ambitious application to the median rank within the top decile of degree selectivity (95.5), but only if their current most ambitious application is below this threshold. We run this counterfactual simulation exercise 100 times and compute low SES' average counterfactual most ambitious application at each academic ability rank.

This counterfactual exercise suggests that equalizing older schoolmate exposure would increase the share of high-achieving, low SES students applying to top-decile degrees from 47.4% to 50.3% (Table 4, Panel A). While a 3 percentage point increase may appear modest, it represents an 11% reduction in the raw application gap, a sizable reduction. Using our baseline spillover estimates may perhaps not accurately reflect how high-achieving, low SES students would react to older schoolmates enrolling in a top decile degree. Moreover, these baseline estimates tell us the additional number of applicants to the *exact* degree to which a high school is exposed, but not the additional number of applicants to *any* degree in the top decile if a student is exposed to an older schoolmate enrolling in a specific top decile degree. This would allow us to capture, for example, how exposure to a marginally enrolled older schoolmate in a top 10% degree affects the likelihood of low SES students in the top 10% of academic achievement to apply to any top 10% degree. Thus, we estimate spillovers at the SES x academic ability decile x degree decile level, and use these estimates to recompute the counterfactual analysis. The results, reported in Table 4's Panel B, are quantitatively similar to the ones obtained using our baseline spillover effects: a roughly 10% in the application gap to top decile degrees.

Table 4: Results of Counterfactual Analysis for Top 10% Students

	Very High SES *observed*	Low SES *observed*	Low SES *counterfactual*
Panel A. Homogeneous treatment effects (baseline estimates)			
Applications to top 10% degrees (%)	74.2	47.4	50.3
Difference relative to very high SES (p.p.)	-	26.8	23.9
Change in difference (%)	-	-	10.8
Panel B. Heterogeneous treatment effects (SES x academic decile x degree decile estimates)			
Applications to top 10% degrees (%)	74.2	47.4	50.0
Difference relative to very high SES (p.p.)	-	26.8	24.2
Change in difference (%)	-	-	9.7

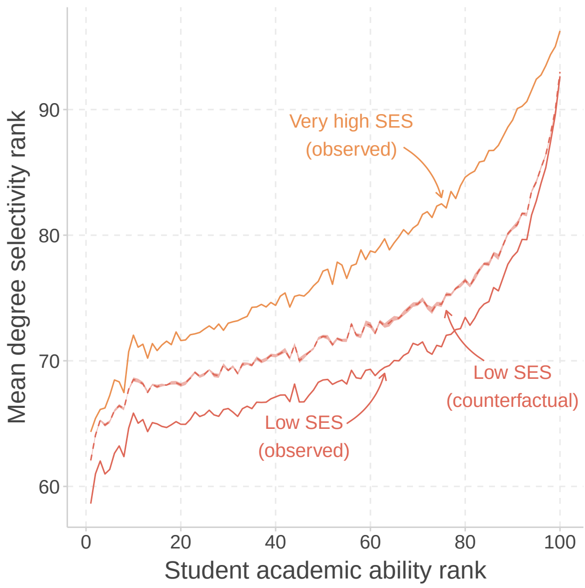
Notes: This table displays of the counterfactual analysis for two sets of older schoolmate spillovers: using baseline estimates (*Panel A*), and more non-parametric estimates at the SES - student academic decile - degree selectivity decile level (*Panel B*).

We generalize this analysis to students across the ability distribution, and for degrees along the selectivity distribution, using two sets of estimates: our baseline spillover effects and heterogeneous effects that vary by student SES, academic decile, and degree selectivity decile, which we just discussed. For each student academic ability decile decile combination, we apply the same simulation approach, adjusting applications upward or downward based on the sign of the estimated spillover and the older schoolmate exposure gap. Figure 16 plots the relationship between students' high school exit exam rank and the selectivity rank of their most ambitious application, comparing observed patterns (solid lines) to counterfactual scenarios (dashed lines). Under both simulation scenarios, equalizing older schoolmate exposure consistently narrows the SES gap in application ambition across the academic ability distribution. The magnitude varies by specification: homogeneous treatment effects suggest particularly strong impacts at lower academic ranks, while heterogeneous estimates indicate more modest improvements.

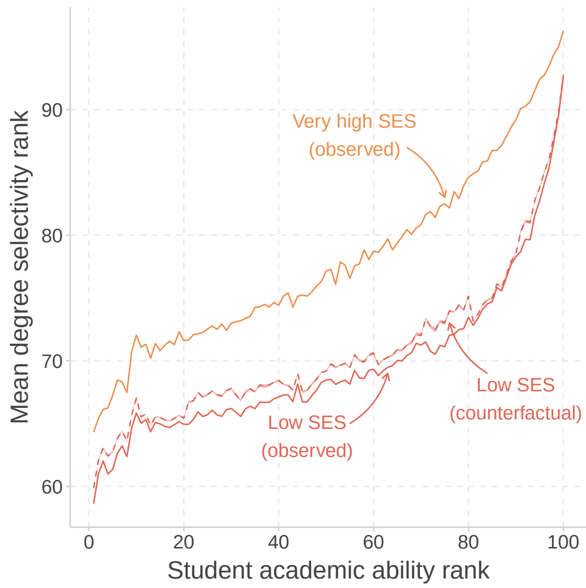
8 Discussion and Policy Implications

Our findings on older schoolmate spillovers on higher education choices have several potential policy implications. These implications touch on issues of educational equity, school segregation, and the role of information in college choice.

First, our results suggest that implementing high school quotas for elite degrees could have cascading effects on educational equity. While potentially controversial, such quotas could ensure a more diverse representation of high schools in top programs, triggering positive spillovers across multiple cohorts and potentially broad-



(a) Homogeneous Treatment Effects



(b) Heterogeneous Treatment Effects

Figure 16: Application Mismatch With and Without Counterfactual Spillovers

Notes: This presents the results from the counterfactual analysis.

ening access to elite education. The need for such measures is underscored by the stark concentration of current admission patterns. In the UK, Montacute and Cullinane (2018) found that a mere eight high schools account for 50% of admitted students to Oxford University and Cambridge University. Similarly, in the French context, Bonneau et al. (2021) reveal that just 8% of high schools contribute 50% of students enrolled in the most selective elite graduate institutions.

Second, our findings underscore the importance of reducing residential and within-high school segregation. The strong influence of older schoolmates on higher education choices implies that segregation may perpetuate inequalities in higher education access. Policies aimed at increasing socioeconomic and racial diversity within high schools could lead to more diverse peer influences and, consequently, more diverse college applications.

Third, our research underscores the crucial role of high-quality higher education guidance provided by schools. This guidance could potentially be significantly improved through enhanced training for principal teachers, who may also struggle the complexity of the French higher education landscape. Furthermore, recent evidence by Mulhern (2023) demonstrates that school counselors have substantial impacts on students' educational outcomes, suggesting that broader policies to improve and expand school counseling services could be highly beneficial. These improvements in guidance could amplify the spillover effects we observe, by ensuring that students are better informed about their options and more capable of leveraging the experiences of their older schoolmates.

Lastly, our results suggest that encouraging high schools to organize alumni forums, facilitate feedback from older schoolmates, and establish mentoring programs across cohorts could be beneficial. However, care must be taken to ensure these initiatives do not exacerbate existing inequalities between high schools. Schools with fewer resources or less diverse alumni networks might struggle to implement such programs effectively.

9 Conclusion

This paper provides evidence that the high school environment, particularly the influence of older schoolmates, plays a crucial role in shaping students' higher education choices. By leveraging a novel research design based on degree-specific admission cutoffs in the French higher education system, we uncover significant spillover effects across cohorts within the same high school.

We find that students are substantially more likely to apply to and enroll in the same degrees as their older schoolmates from the previous year. These effects are both statistically significant and economically meaningful, with a 19% increase in applications and a 45% increase in enrollments to specific programs following the admission of an older peer. Importantly, these spillover effects persist over time and extend beyond exact programs to similar degrees, indicating a broader influence on students' educational trajectories.

The mechanisms underlying these spillovers appear to be twofold. First, we find evidence of significant teacher influence, with students sharing the same principal teacher as the older peer being more likely to emulate their choices. Second, we observe strong homophily effects, where spillovers are larger when students share demographic characteristics with the older schoolmate. These findings have important implications for policy design, suggesting that interventions should consider both the role of teachers and the importance of relatable role models.

The potential for these spillover effects to narrow application gaps between high-achieving students of different socioeconomic backgrounds is particularly noteworthy. Our counterfactual analysis suggests that equalizing exposure to high-achieving older schoolmates could reduce the application gap by approximately 11%. While this would not eliminate disparities entirely, it represents a significant step towards greater educational equity.

These findings open up several avenues for policy interventions. Implementing high school quotas for prestigious degrees, enhancing mentorship programs, and fostering cross-cohort interactions within schools through alumni meetings are all potential strategies that could harness these spillover effects. By leveraging peer influences, policymakers and educators may be able to complement traditional approaches to promoting educational equity and social mobility.

References

Aguirre, Josefa and Juan Matta, "Walking in Your Footsteps: Sibling Spillovers in Higher Education Choices," *Economics of Education Review*, February 2021, *80*, 102062.

Akerlof, George A. and Rachel E. Kranton, "Economics and Identity," *Quarterly Journal of Economics*, August 2000, *115* (3), 715–753.

Altmejd, Adam, "Inheritance of Fields of Study," 2023.

_, Andres Barrios-Fernandez, Aspasia Bizopoulou, Martti Kaila, Rigissa Megalokonomou, Jose Montalban, Christopher Neilson, Sebastian Otero, and Xiaoyang Ye, "Inequality in College Applications: Evidence from Three Continents," 2023.

_, Andrés Barrios-Fernández, Marin Drlje, Joshua Goodman, Michael Hurwitz, Dejan Kovac, Christine Mulhern, Christopher Neilson, and Jonathan Smith, "O Brother, Where Start Thou? Sibling Spillovers on College and Major Choice in Four Countries," *The Quarterly Journal of Economics*, June 2021, *136* (3), 1831–1886.

Altonji, J.G., P. Arcidiacono, and A. Maurel, "The Analysis of Field Choice in College and Graduate School," in "Handbook of the Economics of Education," Vol. 5, Elsevier, 2016, pp. 305–396.

Altonji, Joseph G., Erica Blom, and Costas Meghir, "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers," *Annual Review of Economics*, September 2012, *4* (1), 185–223.

Avdeev, Stanislav, Nadine Ketel, Hessel Oosterbeek, and Bas Van Der Klaauw, "Spillovers in Fields of Study: Siblings, Cousins, and Neighbors," *Journal of Public Economics*, October 2024, *238*, 105193.

Barrios-Fernández, Andrés, "Neighbors' Effects on University Enrollment," *American Economic Journal: Applied Economics*, July 2022, *14* (3), 30–60.

Barrios-Fernandez, Andrés, *Peer Effects in Education*, Oxford University Press, 2023.

Bechichi, Nagui and Georgia Thebault, "Students' Preferences, Capacity Constraints and Post-Secondary Achievements in a Non-Selective System," Technical Report N° G2021/01, INSEE, Paris January 2021.

_, Julien Grenet, and Georgia Thebault, "D'Admission Post-Bac à Parcoursup : Quels Effets Sur La Répartition Des Néo-Bacheliers Dans Les Formations d'enseignement Supérieur ?," Insee Références Édition 2021, INSEE, Paris 2021.

Black, Sandra E. and Paul J. Devereux, "Recent Developments in Intergenerational Mobility," in "Handbook of Labor Economics," Vol. 4, Elsevier, 2011, pp. 1487–1541.

_, Kalena E. Cortes, and Jane Arnold Lincove, "Apply Yourself: Racial and Ethnic Differences in College Application," *Education Finance and Policy*, March 2020, *15* (2), 209–240.

Blumenstock, Joshua E, Guanghua Chi, and Xu Tan, "Migration and the Value of Social Networks," *Review of Economic Studies*, December 2023, p. rdad113.

Bonneau, Cécile and Sébastien Grobon, "Unequal Access to Higher Education Based on Parental Income: Evidence from France," Technical Report 2022/01, World Inequality Lab, Paris 2022.

_, Pauline Charousset, Julien Grenet, and Georgia Thebault, "Quelle démocratisation des grandes écoles depuis le milieu des années 2000?," Technical Report 30, Institut des Politiques Publiques, Paris 2021.

Breda, Thomas, Julien Grenet, Marion Monnet, and Clémentine Van Effenterre, "How Effective Are Female Role Models in Steering Girls Towards STEM? Evidence from French High Schools," *The Economic Journal*, June 2023, *133* (653), 1773–1809.

Britton, Jack, Laura Van Der Erve, Chris Belfield, Anna Vignoles, Matt Dickson, Yu Zhu, Ian Walker, Lorraine Dearden, Luke Sibieta, and Franz Buscha, "How Much Does Degree Choice Matter?," *Labour Economics*, December 2022, *79*, 102268.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs: Robust Nonparametric Confidence Intervals," *Econometrica*, November 2014, *82* (6), 2295–2326.

Campbell, Stuart, Lindsey Macmillan, Richard Murphy, and Gill Wyness, "Matching in the Dark? Inequalities in Student to Degree Match," *Journal of Labor Economics*, 2023.

Canaan, Serena and Pierre Mouganie, "Does Advisor Gender Affect Women's Persistence in Economics?," *AEA Papers and Proceedings*, May 2021, *111*, 112–116.

Carrell, Scott E, Marianne E Page, and James E West, "Sex and Science: How Professor Gender Perpetuates the Gender Gap," *Quarterly Journal of Economics*, 2010, *125* (3), 1101–1144.

Cattaneo, Matias D., Nicolas Idrobo, and Rocio Titiunik, *A Practical Introduction to Regression Discontinuity Designs: Foundations* November 2019.

Charousset, Pauline, Fack Gabrielle, Julien Grenet, and He YingHua, "How Do Elite Post-Secondary Programs Select Their Students?," 2021.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan, "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star," *The Quarterly Journal of Economics*, November 2011, *126* (4), 1593–1660.

Chetty, Raj, David J Deming, and John N Friedman, "Diversifying Society's Leaders? The Determinants and Causal Effects of Admission to Highly Selective Private Colleges," NBER Working Paper 31492, National Bureau of Economic Research 2023.

_, John N. Friedman, and Jonah E. Rockoff, "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood," *American Economic Review*, September 2014, *104* (9), 2633–2679.

— , **Nathaniel Hendren, Patrick Kline, and Emmanuel Saez**, "Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States," NBER Working Paper 19843, National Bureau of Economic Research, Cambridge, MA 2014.

Conley, Timothy G and Christopher R Udry, "Learning about a New Technology: Pineapple in Ghana," *American Economic Review*, March 2010, *100* (1), 35–69.

Dynarski, Susan, Aizat Nurshatayeva, Lindsay C. Page, and Judith Scott-Clayton, "Addressing Nonfinancial Barriers to College Access and Success: Evidence and Policy Implications," in "Handbook of the Economics of Education," Vol. 6, Elsevier, 2023, pp. 319–403.

— , **Lindsay Page, and Judith Scott-Clayton**, "College Costs, Financial Aid, and Student Decisions," in "Handbook of the Economics of Education," Vol. 7, Elsevier, 2023, pp. 227–285.

Estrada, Ricardo, Jeremie Gignoux, and Agustina Hatrick, "Learning About Opportunity: Spillovers of Elite School Admissions in Peru," 2024.

Fack, Gabrielle and Julien Grenet, "Improving College Access and Success for Low-Income Students: Evidence from a Large Need-Based Grant Program," *American Economic Journal: Applied Economics*, April 2015, *7* (2), 1–34.

Fort, Margherita, Andrea Ichino, Enrico Rettore, and Giulio Zanella, "Multi-Cutoff Rd Designs with Observations Located at Each Cutoff: Problems and Solutions," *SSRN Electronic Journal*, 2022.

Frouillou, Leïla, Clément Pin, and Agnès Van Zanten, "Le rôle des instruments dans la sélection des bacheliers dans l'enseignement supérieur. La nouvelle gouvernance des affectations par les algorithmes," *Sociologie*, 2019, *10* (2), 209.

Gale, D. and L. S. Shapley, "College Admissions and the Stability of Marriage," *The American Mathematical Monthly*, January 1962, *69* (1), 9.

Gershenson, Seth, Cassandra M. D. Hart, Joshua Hyman, Constance A. Lindsay, and Nicholas W. Papageorge, "The Long-Run Impacts of Same-Race Teachers," *American Economic Journal: Economic Policy*, November 2022, *14* (4), 300–342.

Hastings, Justine, Christopher Neilson, and Seth Zimmerman, "Are Some Degrees Worth More than Others? Evidence from College Admission Cutoffs in Chile," Technical Report w19241, National Bureau of Economic Research, Cambridge, MA July 2013.

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

Jackson, C Kirabo, "What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes," *Journal of Political Economy*, 2018, *126* (5).

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad, "Field of Study, Earnings, and Self-Selection," *The Quarterly Journal of Economics*, August 2016, *131* (3), 1057–1111.

Kramarz, F. and O. N. Skans, "When Strong Ties Are Strong: Networks and Youth Labour Market Entry," *The Review of Economic Studies*, July 2014, *81* (3), 1164–1200.

Lavecchia, A.M., H. Liu, and P. Oreopoulos, "Behavioral Economics of Education," in "Handbook of the Economics of Education," Vol. 5, Elsevier, 2016, pp. 1–74.

Lim, Jaegeum and Jonathan Meer, "Persistent Effects of Teacher–Student Gender Matches," *Journal of Human Resources*, 2020, *55* (3), 809–835.

Lockwood, Penelope, ""Someone Like Me Can Be Successful": Do College Students Need Same-Gender Role Models?," *Psychology of Women Quarterly*, March 2006, *30* (1), 36–46.

Manski, Charles F., "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 1993, *60* (3), 531–542.

Marbach, Léon and Agnès Van Zanten, "With a Little Help from My Family and Friends: Social Class and Contextual Variations in the Role of Personal Networks in Students' Higher Education Plans," *British Journal of Sociology of Education*, October 2023, pp. 1–22.

MESR, "L'état de l'Enseignement Supérieur et de La Recherche," Technical Report, Ministère de l'Enseignement Supérieur et de la Recherche, Paris 2019.

_ , "Baromètre Parcoursup 2023," Technical Report, MESR September 2023.

Mobius, Markus and Tanya Rosenblat, "Social Learning in Economics," *Annual Review of Economics*, August 2014, *6* (1), 827–847.

Mogstad, Magne and Gaute Torsvik, "Family Background, Neighborhoods, and Intergenerational Mobility," in "Handbook of the Economics of the Family," Vol. 1, Elsevier, 2023, pp. 327–387.

Montacute, Rebecca and Carl Cullinane, "Access to Advantage," Technical Report, The Sutton Trust, London 2018.

Mulhern, Christine, "Changing College Choices with Personalized Admissions Information at Scale: Evidence on Naviance," *Journal of Labor Economics*, January 2021, *39* (1), 219–262.

_ , "Beyond Teachers: Estimating Individual School Counselors' Effects on Educational Attainment," *American Economic Review*, November 2023, *113* (11), 2846–2893.

Porter, Catherine and Danila Serra, "Gender Differences in the Choice of Major: The Importance of Female Role Models," *American Economic Journal: Applied Economics*, July 2020, *12* (3), 226–254.

Ray, Debraj, *Aspirations, Poverty, and Economic Change*, 1 ed., Oxford University Press-New York, May

Valdebenito, Rocío, "Peer Influence and College Major Choices in Male-Dominated Fields," 2023.

Online Appendix

A Institutional Background: Additional Details

A very clear overview of the French higher education landscape and its costs can be found in Fack and Grenet (2015). We summarise some of the important features of France's higher education below. Figure A.1 provides an (somewhat simplified) illustration.

A.1 Access to Higher Education

In France, the only requirement to enter higher education is to obtain the end of high school exam, the *Baccalauréat* (hereafter Bac). Over 2013-2016, roughly 88% of students who took the Bac obtained it. Three types of Bac can be prepared by high school students, all of them corresponding to the type of high school track they are enrolled in. They are categories, in their more aggregate versions, as *general* (academic;), *technological* (technical), and *professional* (vocational). In 2021, half of Bac holders obtained a general Bac, the remaining half were divided between technological tracks (20%) and professional tracks. About three out of four high school students who obtained the Bac continued into tertiary education MESR (2019). This share is much higher for students from general and technical high school tracks as compared to students from vocational tracks.

A.2 Types of Higher Education Programs

The French higher education system is composed of five types of programs: non-selective public universities, (ii) selective *vocationally*-oriented post-secondary schools (Sections of Superior Technicians *(Sections de Techniciens Supérieurs (STS))*), (iii) selective *technically*-oriented institutes (University Institutes of Technology *(Instituts Universitaires de Technologie (IUT))*), (iv) selective *academically*-oriented preparatory classes (Preparatory Classes for the *Grandes Écoles (Classes Préparatoires aux Grandes Écoles (CPGE))*), and (v) other private schools (mostly engineering, business, art, and paramedical and social schools).

Selective institutions are free to select their applicants according to their own (undisclosed) criteria. Non-selective programs could not select its students. If capacity constraints were bindings, they distinguished applicants based on non-academic priority rules such as whether the student was from the same academic region as the institution and how applicants ranked the degree in their rank-ordered list. Lotteries were implemented to break ties should capacity constraints continue to bind despite these priority criteria. See Bechichi and Thebault (2021) for more details, and analysis of these lotteries.

A.3 Cost of Higher Education

The cost of higher education depends exclusively on whether the institution is public or private. 82% of students are enrolled in a public institution. Public institutions charge annual tuition fees of slightly under 200 euros. There is no limit on the tuition fees private institutions can charge.

A.4 Admission Post-Bac (APB)

From 2009 to 2017, students seeking admission to higher education programs were required to go through a centralised national platform called *Admission Post-Bac* (APB), where they could apply to both non-selective and selective programs. The APB system gathered roughly 12,000 programs and 800,000 applicants each year. Candidates submitting applications were asked to provide a rank-ordered list (ROL) of programs from January to March. Following this application phase, program administrators rank applicants. Each selective program produces its own

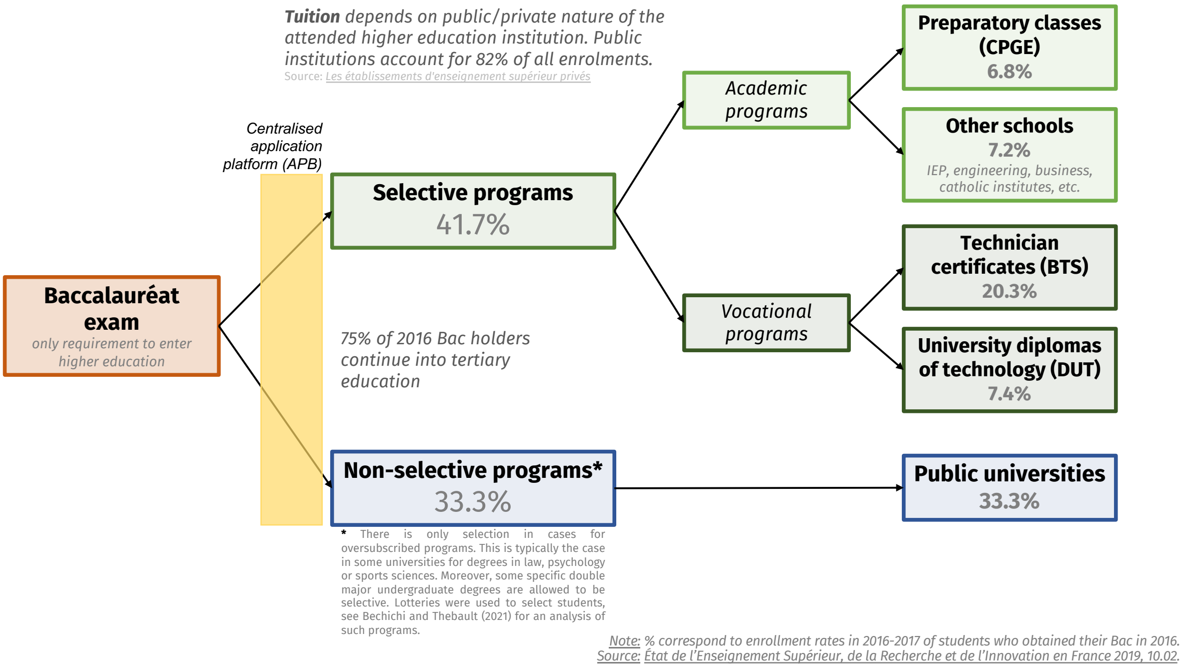


Figure A.1: Higher Education Landscape in France

specific ranking based on discretionary criteria and without any legal constraints. Selective programs are not required to rank all their applicants. The ranking for a non selective program is produced automatically by the centralised platform on the basis of applicants' non-academic priorities. In contrast to selective programs, a rank is assigned to all the applicants to a given non-selective program. It is important to note that the local decision rules or algorithms used by selective programs to rank applicants are not public information, neither for applicants nor for the centralised platform. The only information that the platform collects is the rank of each applicant, the outcome produced by local algorithms.

Taking into account programs' capacities in addition to applicants' ROL and programs rankings, applicants get offered a seat to their best feasible option through a three-rounds college-proposing deferred acceptance (DA) algorithm (Gale and Shapley, 1962) taking place from June to July (Figure A.2). For each program j and each admission round k, applicant i gets offered a seat only if (i) the rank $r^k_{i,j}$ is above the cutoff c^k_j which corresponds to the rank of the last applicant receiving an offer; (ii) there is no higher-ranked program j' where applicant i is ranked above the cutoff $c^k_{j'}$. Applicants could accept the offer, turn it down or conditionally accept placement while waiting for applicants selected by higher-ranked programs to withdraw from the selection process in subsequent admission rounds. This sequential procedure implies that programs cutoffs could evolve from round 1 to round 3, always observing the following rule : $c^1_j \leq c^2_j \leq c^3_j$. The final results of the Bac were published between the second and the third rounds of the procedure. Students who failed the exam were not able to compete for a seat anymore, and their seats were re-offered in the third round. Finally, applicants could participate in supplementary rounds, which took place between June and September, and helped students to apply to programs with remaining seats. Figure A.2 summarises the timeline of this process.

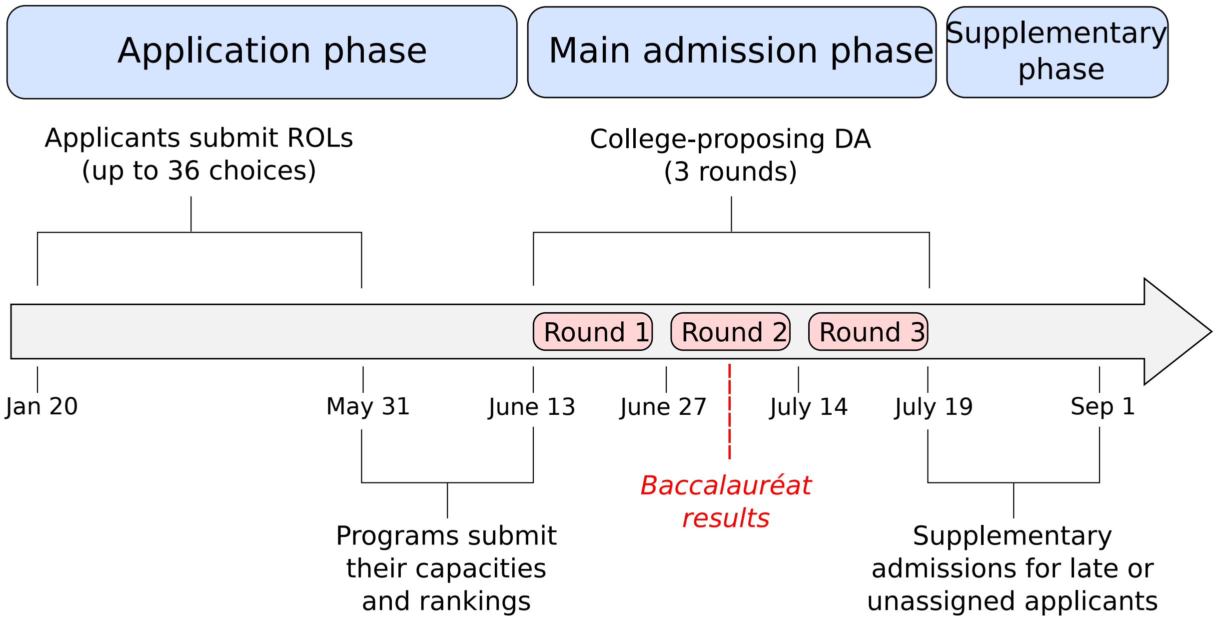


Figure A.2: Timeline of the Application and Admission Procedure into Higher Education Programs in France

B Running Variable: Additional Details

In the main text we made a very slight simplification. In practice, degrees rank their applicants within *"ranking groups"*. These ranking groups typically relate to applicants high school track, though they can be even more specific. This implies that the same degree might have different rankings for different types of applicants. In our setting what matters for high school students is whether an older schoolmate was marginally admitted to a degree or not, regardless of the ranking group to which they belonged. Thus, we overcome this minor issue (only 0.05% of high school $\times$ high school tracks, the level at which we conduct our analysis, have students in several ranking groups) by defining the high school's best ranked applicant as the best ranked applicant to the degree across all ranking groups. In the extremely few cases where the best rank is tied, we keep one at random.

C Spillover effects and older schoolmates' experience

Are high school students more likely to be influenced by older peers who successfully graduate from a program after enrolling? Admission to a program does not necessarily lead to satisfaction or success for all students once enrolled. The question of how the educational trajectories of older peers influence the decisions of new generations of high school students is crucial. Encouraging students to apply to colleges where former peers had negative experiences may not be an effective strategy for improving higher education outcomes. This disconnect may arise due to insufficient skills for the program or mismatches related to factors such as course content or class composition. In France, there is a notably high proportion of students who drop out of university without obtaining a degree or change majors early in their studies.

When high school students apply to French universities through centralized platforms in January, they have limited information about the experiences of their older peers who were admitted the previous year and began their studies in September. Furthermore, no administrative data is available to track the academic performance of these peers during their first semester.

To approximate the negative experiences of older peers, we identify those who reapplied to new programs shortly after the start of the academic term. This is the earliest indicator we can observe. Early program changes by older peers could serve as a valid proxy for a negative experience, as this information becomes available in January, coinciding with the period when new students are making their college decisions.

When re-estimating the ITT effect on the subset of cases where the older peer was identified as an "early mover" (almost 30% of our observations), we find that spillover effects on applications are not significantly different from the full sample estimates. However, spillover effects on program rankings and final admissions decisions are slightly lower when the older peer was an early mover. Since applications reflect only students' choices, while program rankings and final admissions are influenced by both student preferences and colleges screening of applications, these results suggest that new students do not heavily consider older peers' experiences when applying to higher education programs. Additionally, college programs seem to update their assessment of student potential based on older peers' experiences.

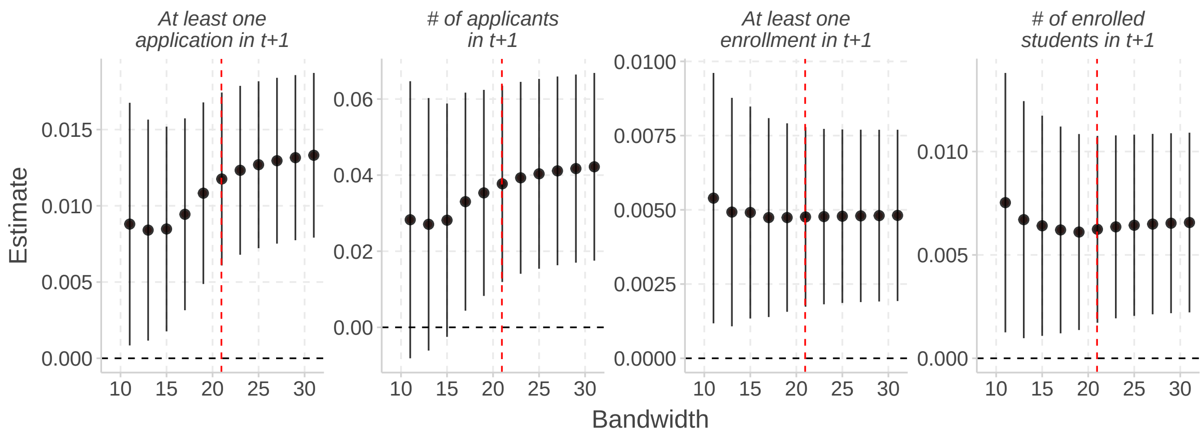
	Full sample	No early change for the older schoolmate	Early change for the older schoolmate
Sample size	375,566	267,841	107,725
Spillovers on applications			
Baseline mean	0.79	0.828	0.701
ITT	0.039***	0.039**	0.035
	(0.0131)	(0.0163)	(0.0231)
Spillovers on programs' ranking			
Baseline mean	0.503	0.522	0.455
ITT	0.02**	0.022*	0.015
	(0.0092)	(0.0115)	(0.0162)
Spillovers on final admission			
Baseline mean	0.064	0.066	0.059
ITT	0.006**	0.006**	0.001
	(0.0023)	(0.0028)	(0.004)

Table C.1: Spillover effects and older schoolmates' experience

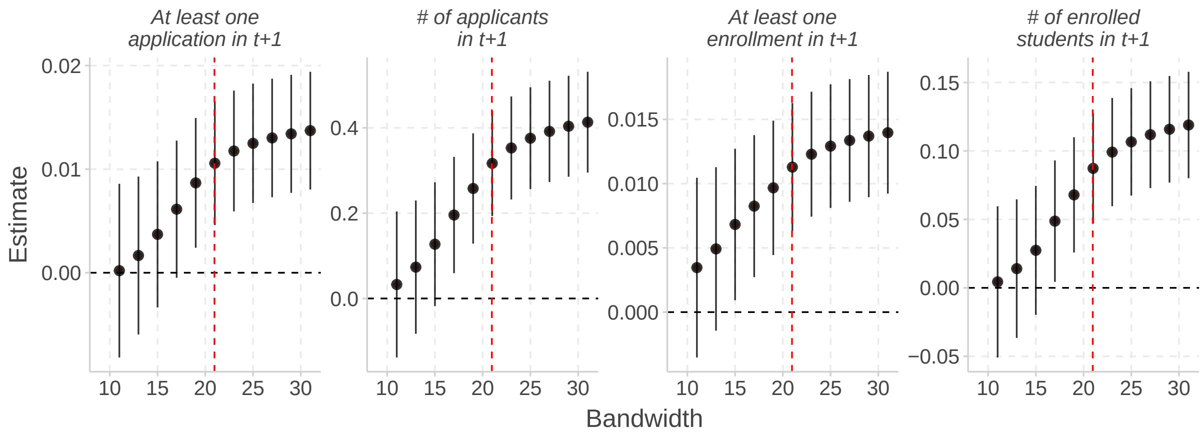
Notes: This table shows spillover effects based on the experience of older schoolmates. The first column presents results for the full sample, while the second and third columns distinguish between cases where the older schoolmate did or did not experience early changes in their college-major. The estimates account for applications, program rankings, and final admissions outcomes.
***, **, and * indicate significance at 1, 5, and 10%, respectively.

D Robustness Checks

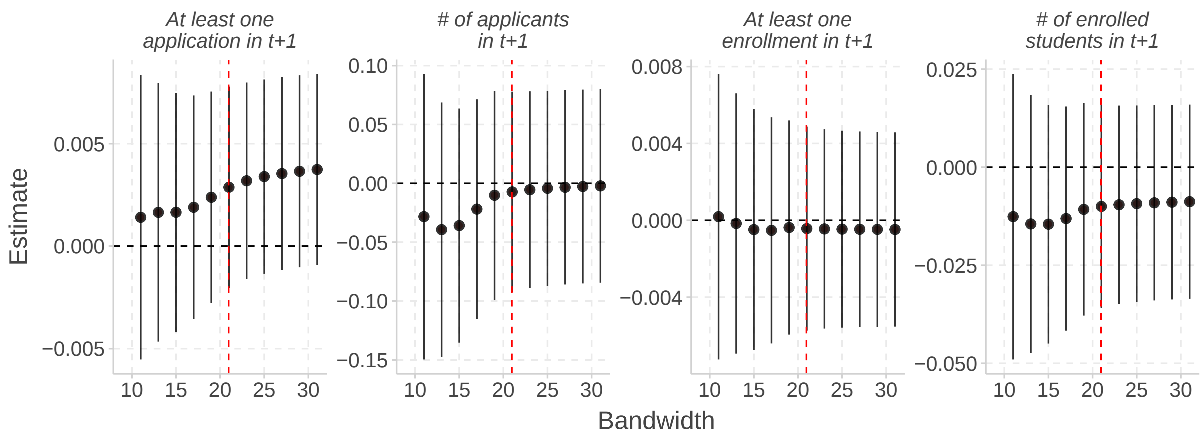
We assess the robustness of our baseline results to (i) varying the bandwidth over which the estimates are computed (Appendix Figure D.1, and (ii) estimating older schoolmate spillovers at placebo admission rank cutoffs (Appendix Figure D.2).



(a) Degree Spillovers



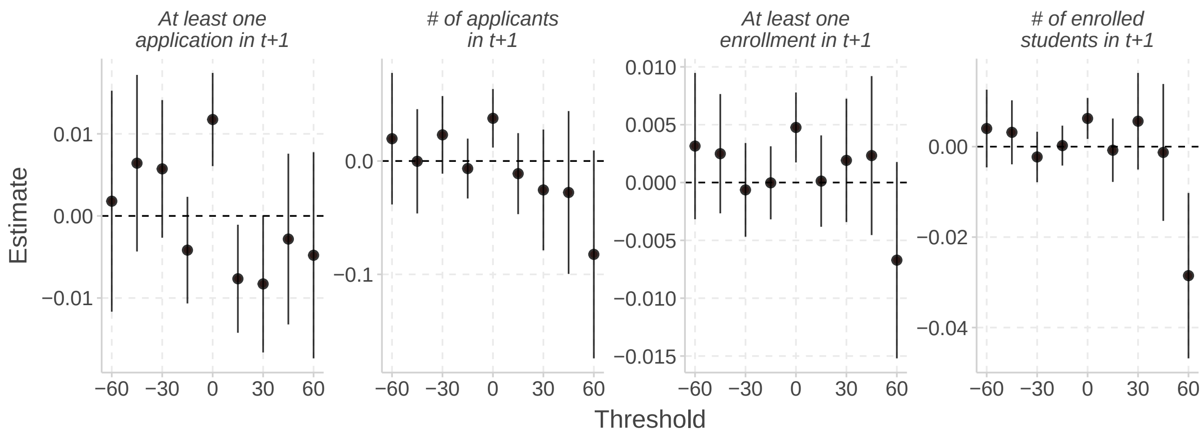
(b) Higher Education Institution Spillovers



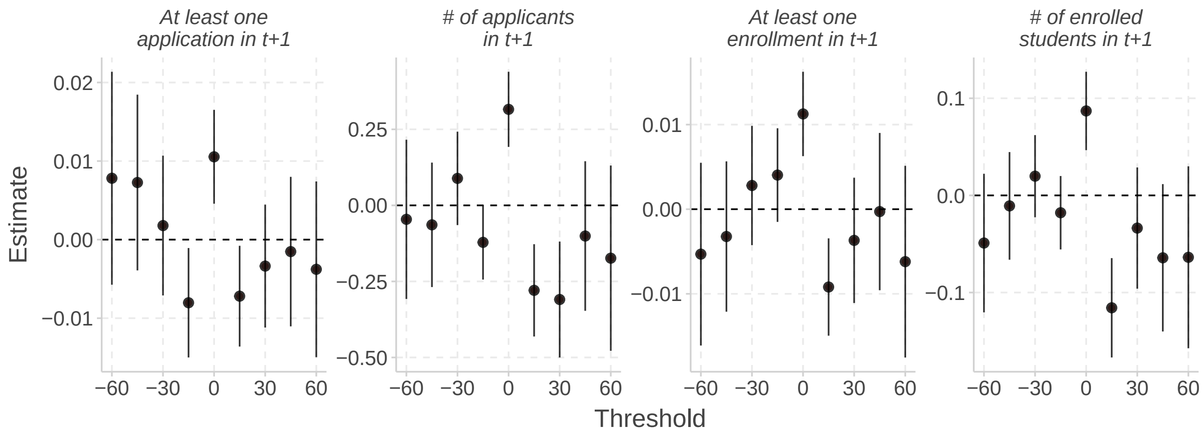
(c) Field of Study Spillovers

Figure D.1: Robustness of Baseline Older Schoolmate Spillovers to Using Different Bandwidths

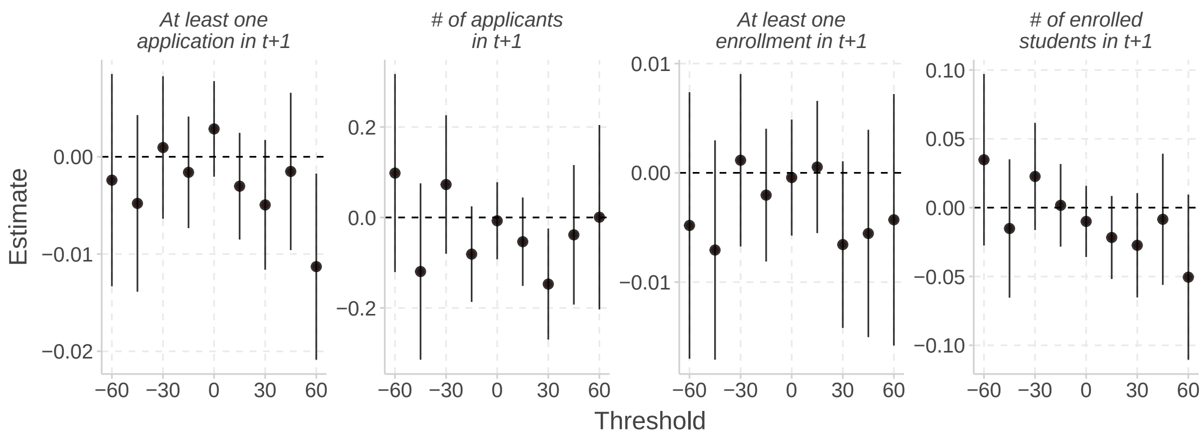
Notes: This figure shows estimates of older schoolmate spillovers, varying the bandwidth over which the estimates are obtained. The application and enrollment outcomes are reported in the figure facet titles, while the type of spillover (degree, HE institution, or field of study) is indicated by the subfigure caption. The baseline bandwidth is denoted by the vertical dashed line. All the specifications in the figure correspond to local linear regressions using a triangular kernel, and include degree-year fixed effects. Statistically significance is based on standard errors clustered at the high school-year level. Confidence intervals correspond to 95% confidence intervals.



(a) Degree Spillovers



(b) Higher Education Institution Spillovers



(c) Field of Study Spillovers

Figure D.2: Robustness of Baseline Older Schoolmate Spillovers to Placebo Admission Cutoffs

Notes: This figure shows estimates of older schoolmate spillovers, varying the admission cutoff at which the estimates are obtained. The application and enrollment outcomes are reported in the figure facet titles, while the type of spillover (degree, HE institution, or field of study) is indicated by the subfigure caption. All the specifications in the figure correspond to local linear regressions using a triangular kernel, and include degree-year fixed effects. The bandwidth (21.32) corresponds to the smallest MSE-optimal bandwidth (Calonico et al., 2014) for the main college-major outcomes. Statistically significance is based on standard errors clustered at the high school-year level. Confidence intervals correspond to 95% confidence intervals.

E Appendix Figures

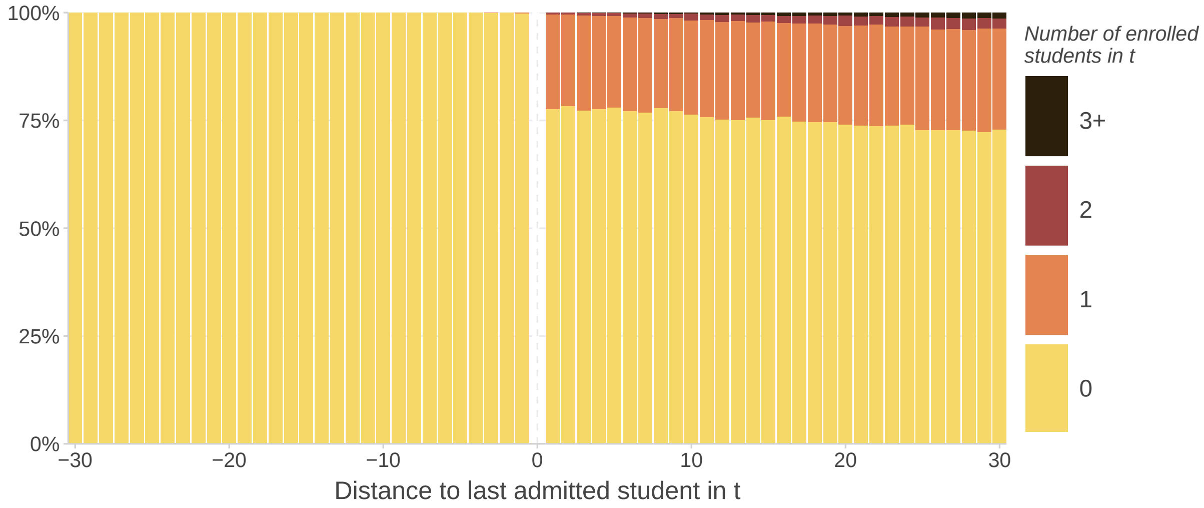


Figure E.1: Exact Number of Enrolled Students in Treatment Year

Notes: This figure shows the number of enrolled students for a given high school and college-major as a function of the high school's distance to the last admitted student.

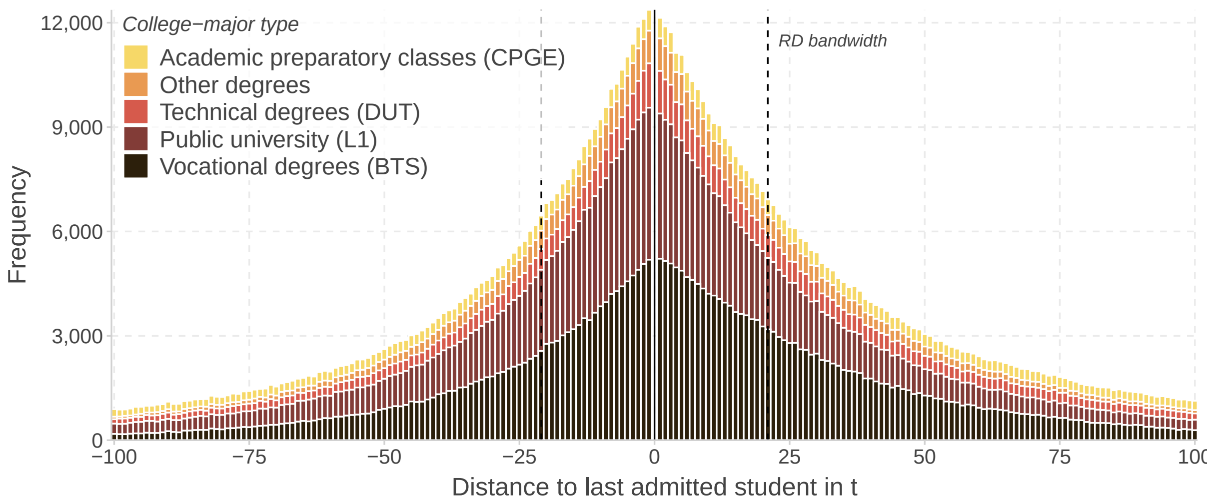


Figure E.2: Running Variable With Degree Type Composition

Notes: This figure shows the composition in terms of degree type of the running variable which corresponds to the rank of the high school's best ranked applicant by the college-major centered around the rank of the college-major's last admitted student. The dashed lines represent the the regression discontinuity (RD) bandwidth used in the analysis.

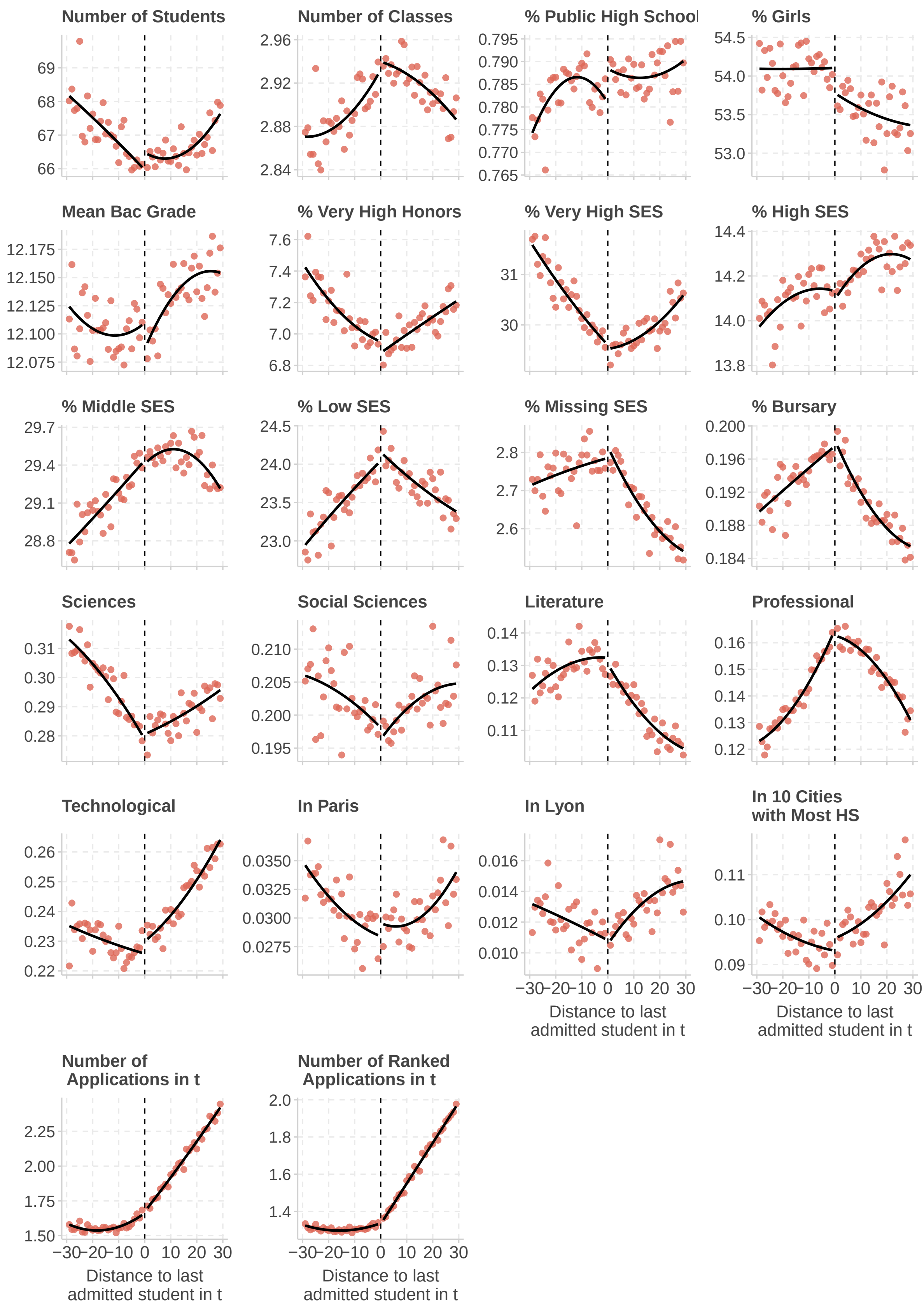


Figure E.3: Discontinuity in High School Characteristics

Notes: This figure shows non-parametric binned scatter plots of the relationship between various high school characteristics and distance to the last admitted student. The high school characteristics are reported in the figure facet titles. Each point corresponds to the average high school characteristic for high schools with distance to the last admitted student equal to the value on the x-axis. The fitted lines correspond to second-order polynomial fits through the conditional expectation.

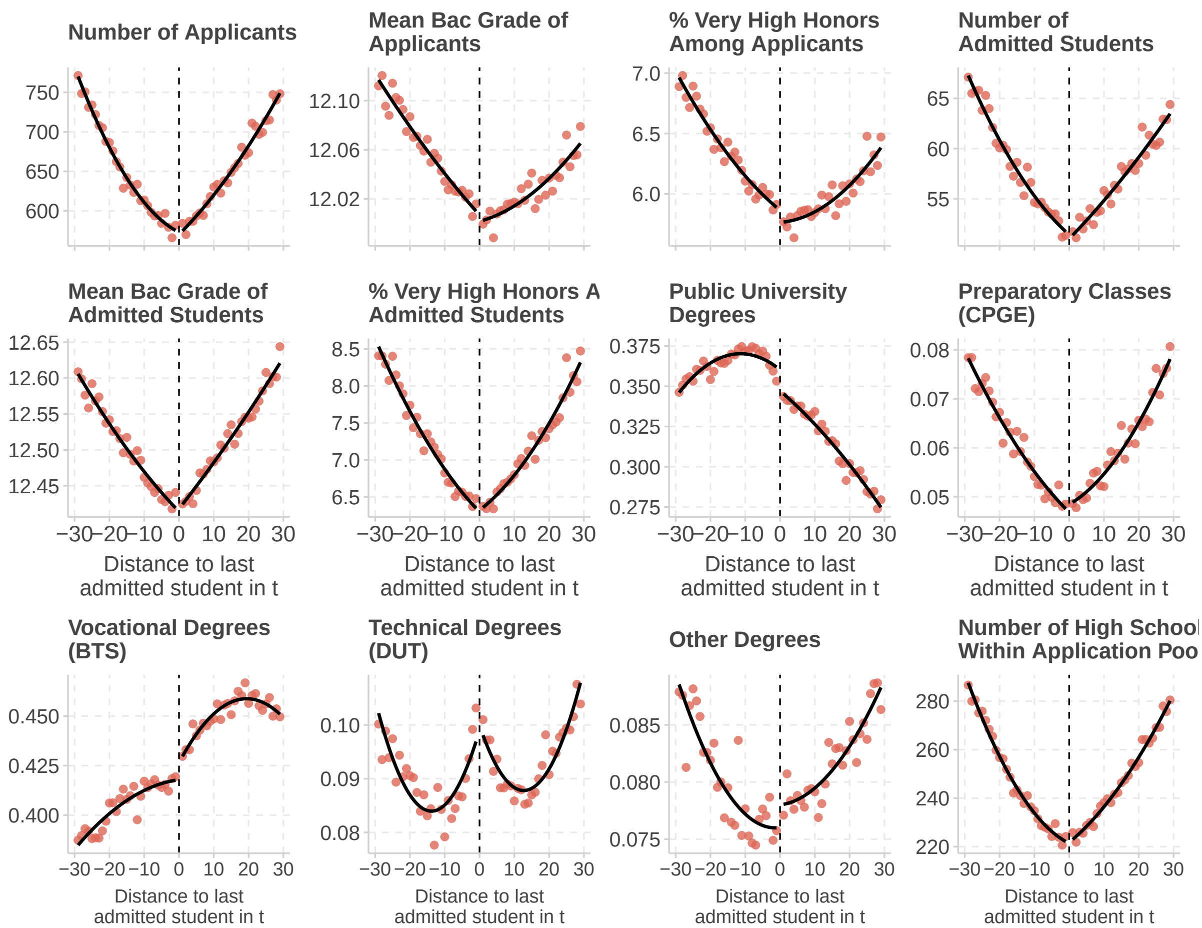


Figure E.4: Discontinuity in College-Major Characteristics

Notes: This figure shows non-parametric binned scatter plots of the relationship between various college-major characteristics and distance to the last admitted student. The college-major characteristics are reported in the figure facet titles. Each point corresponds to the average college-major characteristic for high schools with distance to the last admitted student equal to the value on the x-axis. The fitted lines correspond to second-order polynomial fits through the conditional expectation.

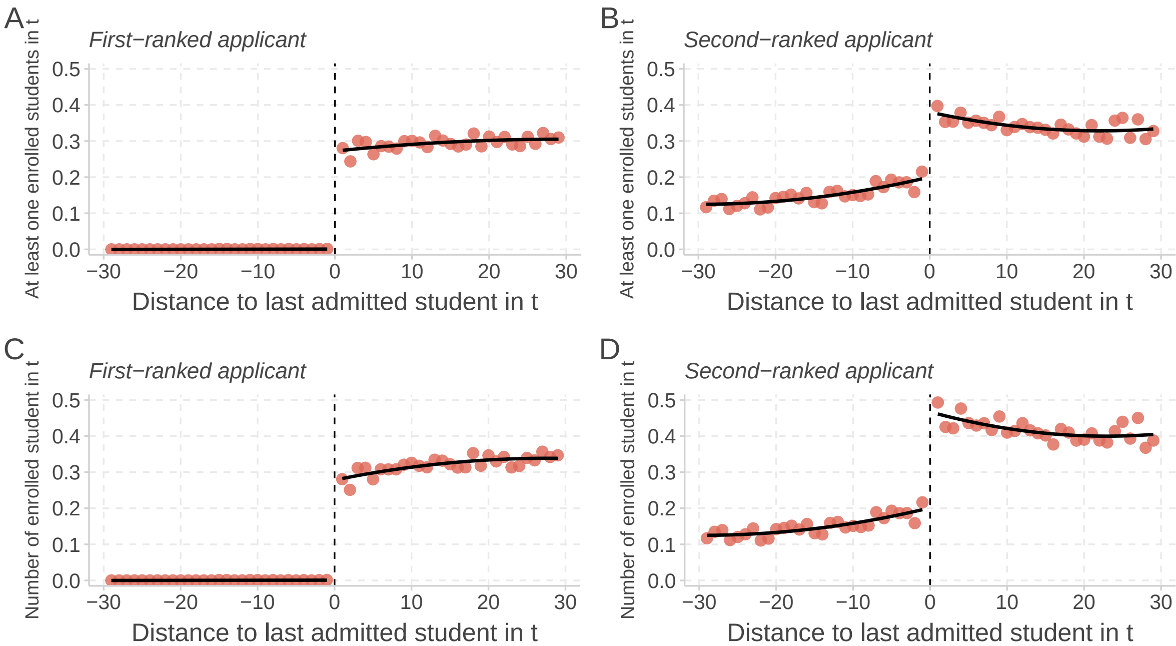


Figure E.5: Probability and Number of Older Schoolmates Enrolling in Degree Around Admission Cutoff Rank - For High School First- and Second-Ranked Applicant

Notes: Only high schools with exactly two applicants to a given degree in a given year. This figure shows, for high-schools with two students ranked by the college-major, the number of enrolled students in the college-major as a function of the distance to the last admitted student in the same year. The figures compare two alternative definitions of the distance to the last admitted student. The first one is the rank of the high school's best ranked applicant by the college-major centered around the rank of the college-major's last admitted student, the second one is the rank of the high school's second-best ranked applicant by the college-major centered around the rank of the college-major's last admitted student.

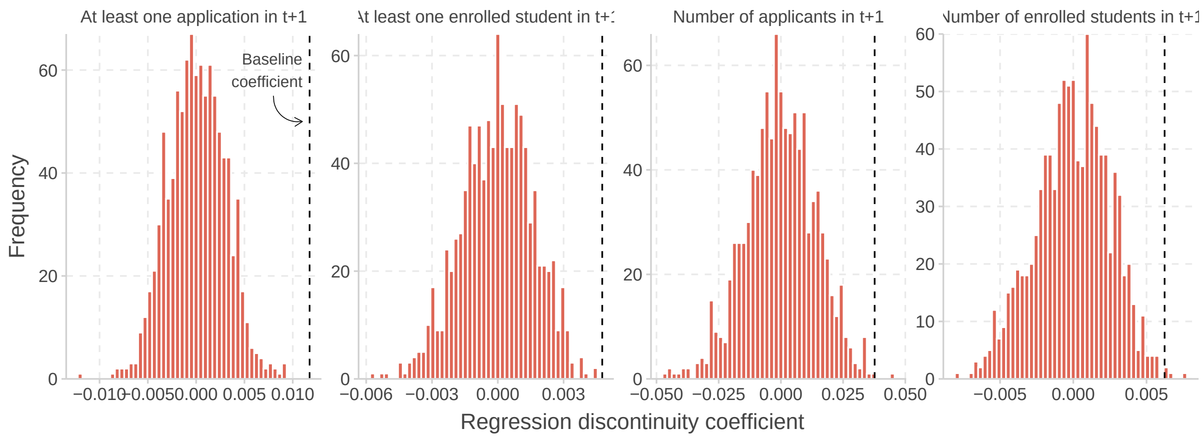
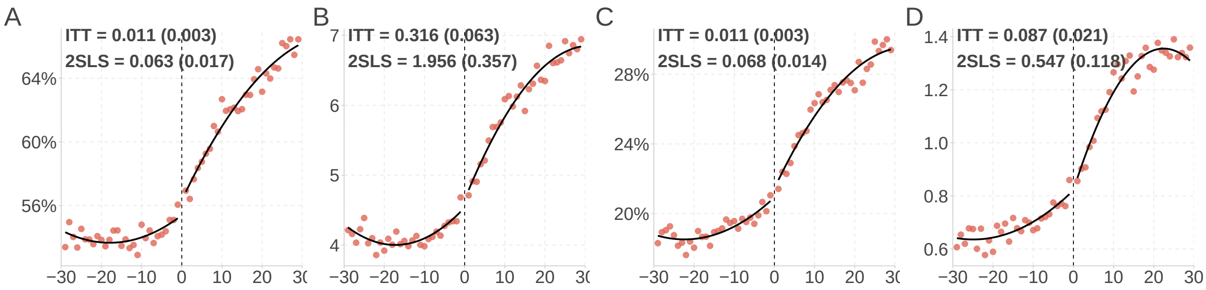
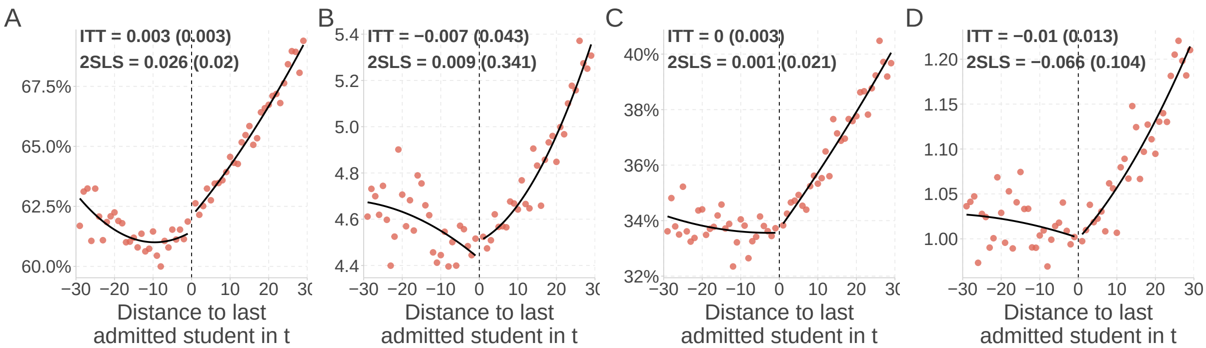


Figure E.6: Randomization Inference

Notes: This figure shows add figure explanation.



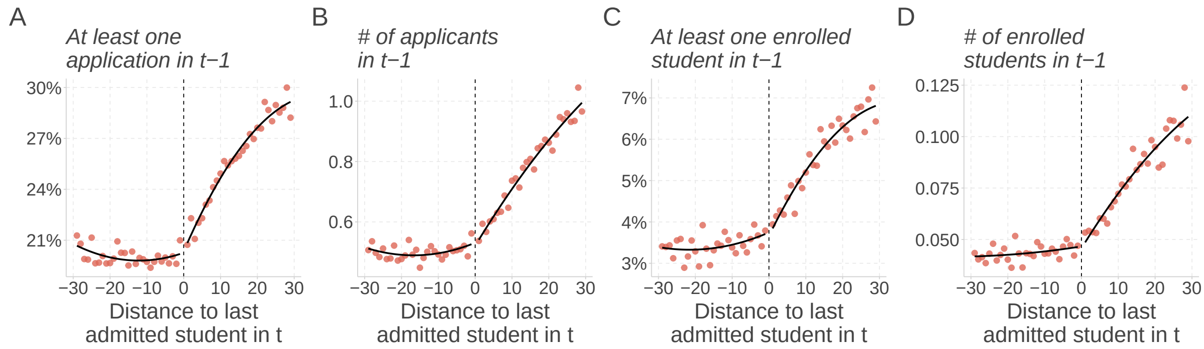
(a) HE Institution



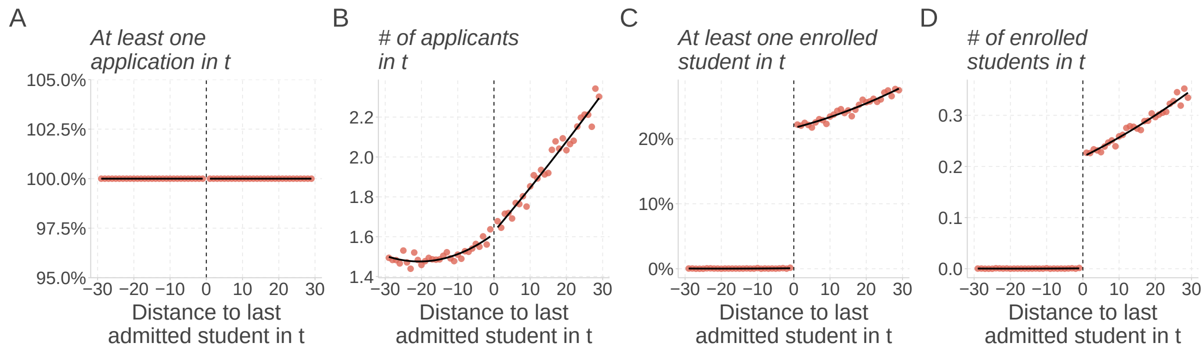
(b) Spillovers on Enrollment

Figure E.7: Older Schoolmate Spillovers on Applications to and Enrollment in Marginally Admitted Older Schoolmate's HE Institution and Field of Study

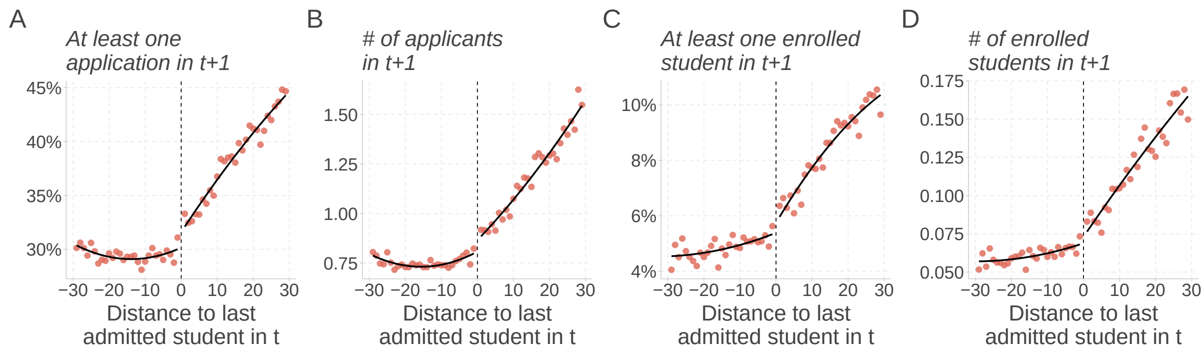
Notes: This figure presents non-parametric binned scatter plots of the relationship between high schools' application and enrollment outcomes for a degree in $t+1$ and these high schools' distance to the degrees' last admitted student in t. The specific application and enrollment outcomes are reported in each facet's title. Each point represents the average outcome value for high schools at a given distance from the admission cutoff rank. The fitted lines correspond to second-order polynomial fits through the conditional expectations. Both the intent-to-treat (ITT) and instrumented (2SLS) estimates are reported, using local linear regressions with a triangular kernel, and including degree-year fixed effects. The bandwidth (20.96) corresponds to the smallest MSE-optimal bandwidth (Calonico et al., 2014) for these main outcomes. Standard errors, clustered at the high school-year level, are reported in parentheses.



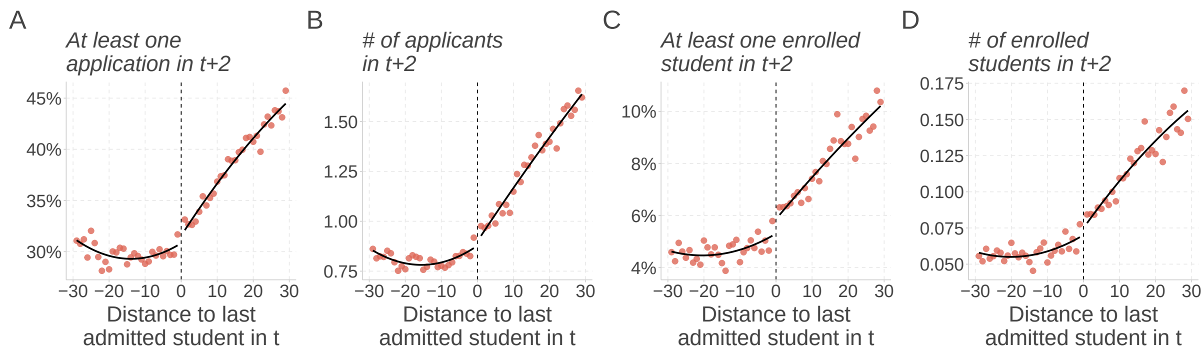
(a) Treatment Year - 1



(b) Treatment Year



(c) Treatment Year + 1



(d) Treatment Year + 2

Figure E.8: Event-Study Analysis Graphs

Notes: This figure shows non-parametric binned scatter plots of the relationship between high schools' application and enrollment outcomes for a college-major in different years and high schools' distance to the college-majors' last admitted student in t. "Treatment Year - 1" refers to the year prior to the older schoolmate's marginal admission, "Treatment Year" refers to the year of an older schoolmate's marginal admission, and "Treatment Year + 1" and "Treatment Year + 2" correspond, respectively to high schools' application and enrollment outcomes one and two years following the marginal admis-

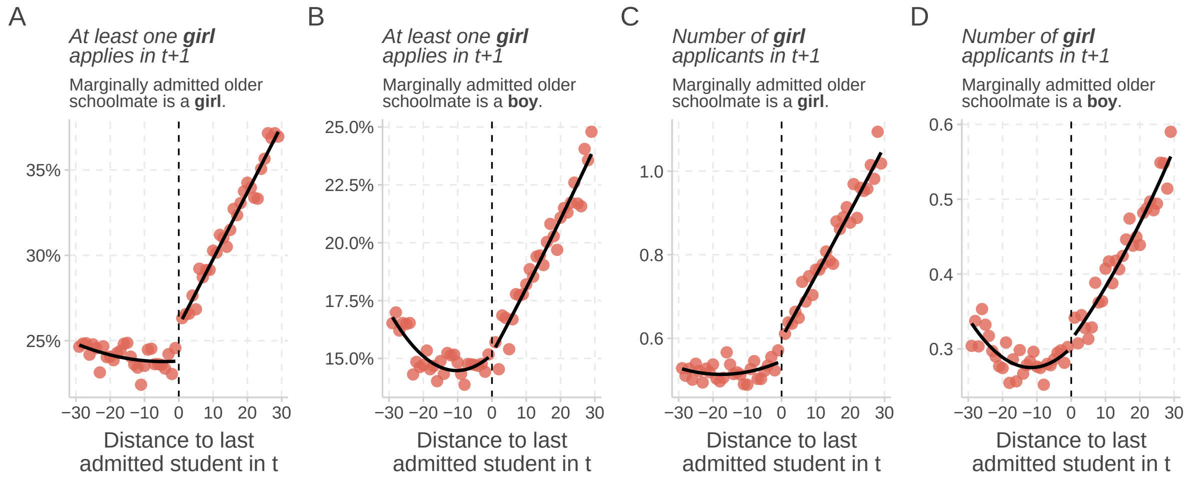


Figure E.9: Role Model Effects for Girls

Notes: This figure shows non-parametric binned scatter plots of the relationship between girls' application and enrollment outcomes for a college-major in $t+1$ and high schools' distance to the college-majors' last admitted girl or boy in t. The application and enrollment outcomes are reported in the figure facet title. Each point corresponds to the average outcome value for girls in high schools with distance to the last admitted student equal to the value on the x-axis. The fitted lines correspond to second-order polynomial fits through the conditional expectation.

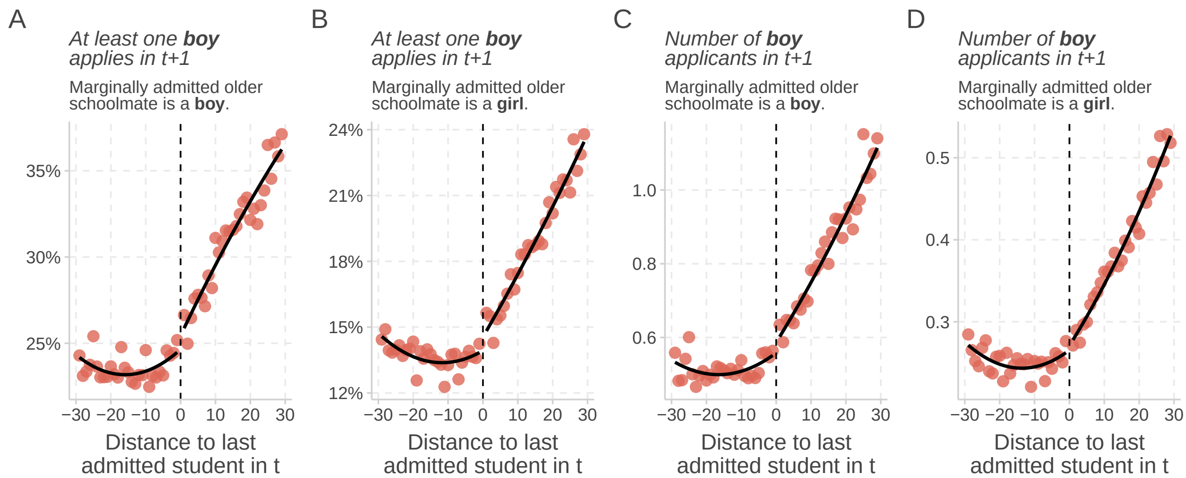


Figure E.10: Role Model Effects for Boys

Notes: This figure shows non-parametric binned scatter plots of the relationship between boys' application and enrollment outcomes for a college-major in $t+1$ and high schools' distance to the college-majors' last admitted boy or girl in t. The application and enrollment outcomes are reported in the figure facet title. Each point corresponds to the average outcome value for boys in high schools with distance to the last admitted student equal to the value on the x-axis. The fitted lines correspond to second-order polynomial fits through the conditional expectation.

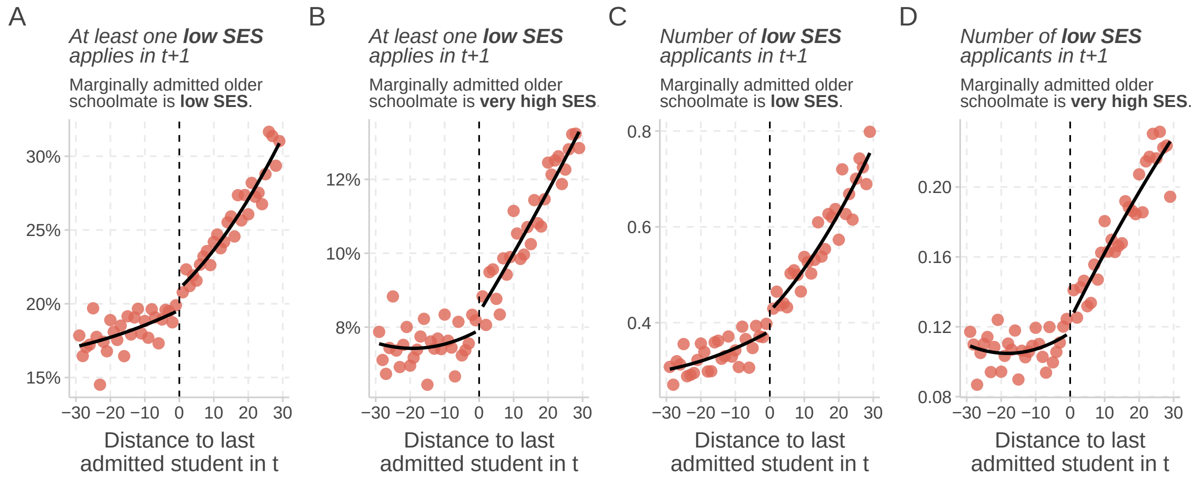


Figure E.11: Role Model Effects for Low SES Students

Notes: This figure shows non-parametric binned scatter plots of the relationship between low SES students' application and enrollment outcomes for a college-major in $t+1$ and high schools' distance to the college-majors' last admitted low SES or very high SES student in t. The application and enrollment outcomes are reported in the figure facet title. Each point corresponds to the average outcome value for low SES students in high schools with distance to the last admitted student equal to the value on the x-axis. The fitted lines correspond to second-order polynomial fits through the conditional expectation.

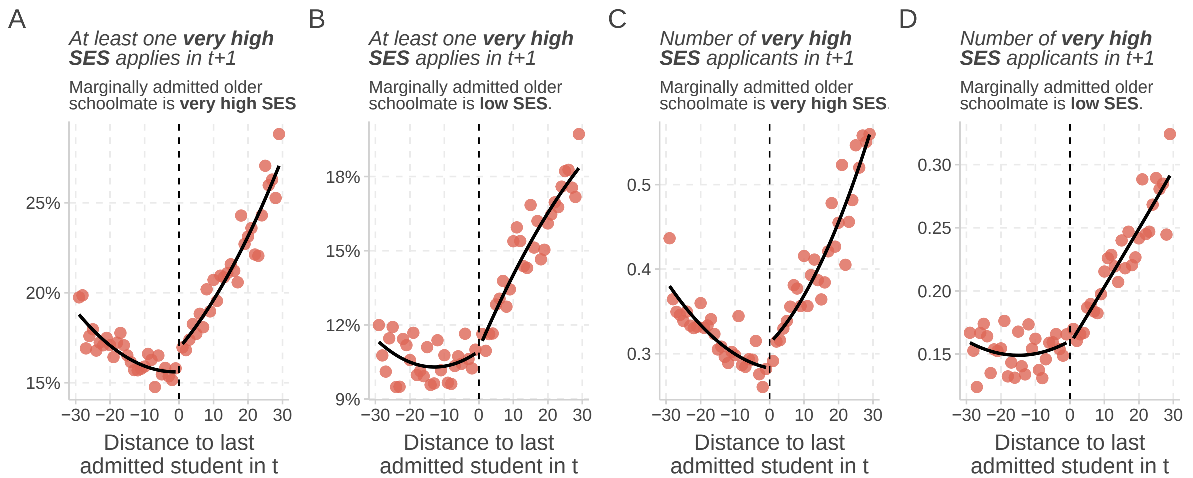


Figure E.12: Role Model Effects for Very High SES Students

Notes: This figure shows non-parametric binned scatter plots of the relationship between very high SES students' application and enrollment outcomes for a college-major in $t+1$ and high schools' distance to the college-majors' last admitted very high SES or low SES student in t. The application and enrollment outcomes are reported in the figure facet title. Each point corresponds to the average outcome value for very high SES students in high schools with distance to the last admitted student equal to the value on the x-axis. The fitted lines correspond to second-order polynomial fits through the conditional expectation.

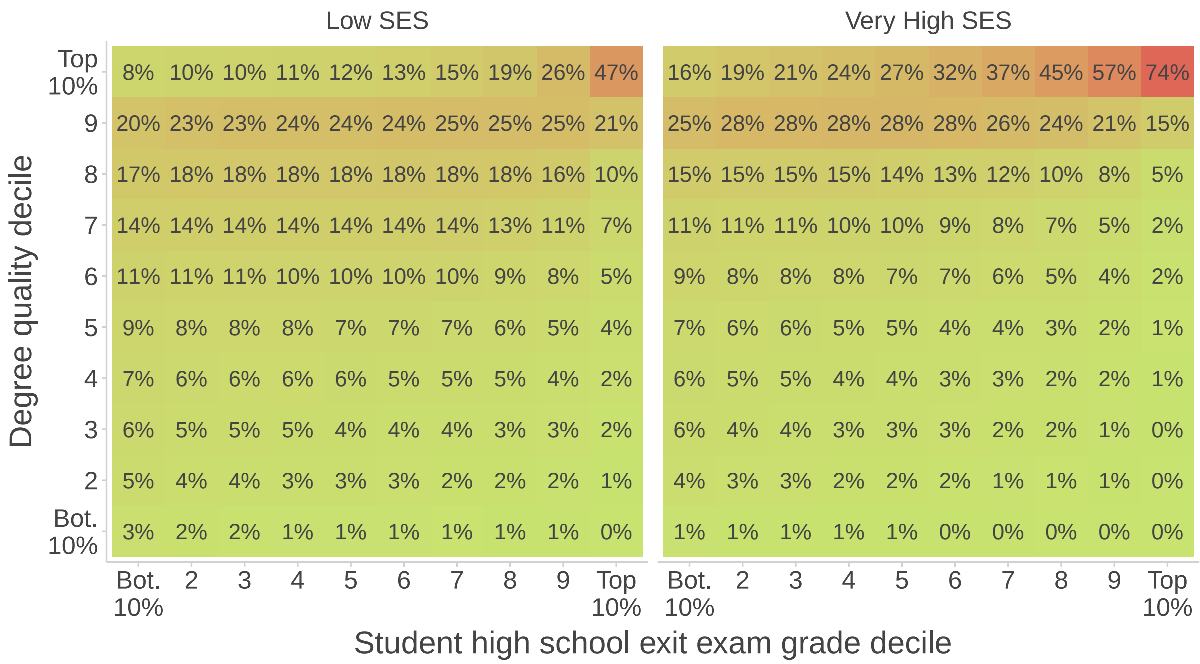


Figure E.13: Most Ambitious Degree by Decile

Notes: This figure reports the likelihood of applying to a degree quality decile for low and very high SES students of different academic ability deciles.

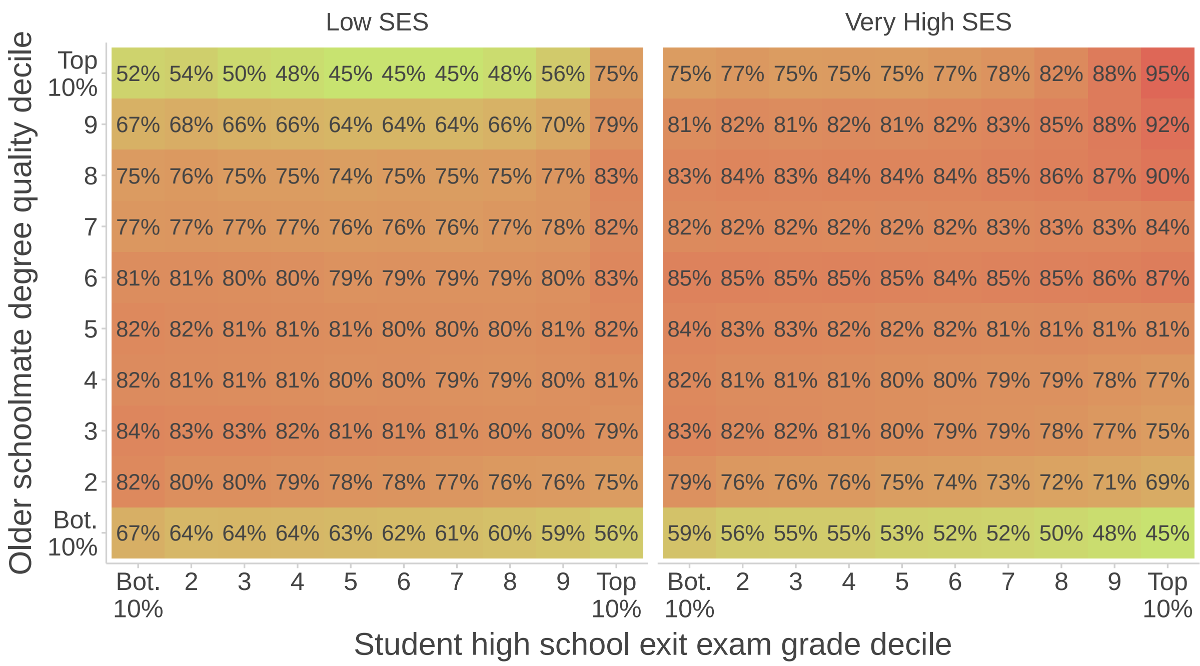


Figure E.14: Older Schoolmate Exposure Matrix

Notes: Write up figure notes.

F Appendix Tables

Table F.1: Number of Observations at Each Sample Restriction

Restriction	Nb. College-Majors	% Change	Nb. High Schools	% Change	Nb. Total Obs.	% Change
Raw number	50,359	100%	66,157	100%	5,541,916	100%
+ College-majors present in treatment and following year	47,991	95.3%	65,956	99.7%	5,381,382	97.1%
+ At least one applicant ranked after last admitted student	41,105	85.65%	65,811	99.78%	4,951,709	92.02%
+ High schools with at least one applicant in two consecutive years	41,105	100%	63,824	96.98%	4,918,338	99.33%
+ No change in reported capacity between admission rounds	34,089	82.93%	63,549	99.57%	3,682,664	74.88%
+ At least 30 high schools within applicant pool	27,497	80.66%	63,364	99.71%	3,526,706	95.77%
+ Symmetrization of running variable	25,701	93.47%	61,275	96.7%	1,143,975	32.44%
+ Drop marginal student	25,419	98.9%	61,208	99.89%	1,130,891	98.86%
+ At least 2 obs. on both sides of cutoff within bandwidth	19,577	77.02%	61,002	99.66%	1,067,809	94.42%

Notes: This table shows the number of college-major - years, high school - years, and observations at each sample restriction mentioned in Section 3.2. High schools refer to high school x tracks.

Table F.2: Discontinuity in Degree Ranking of Applicants at Admission Cutoff Rank

	Ranked		Ranked in Top 10%		Ranked in Top 25%		Ranked in Top 50%		Offer	
	At least one (1)	Number (2)	At least one (3)	Number (4)	At least one (5)	Number (6)	At least one (7)	Number (8)	At least one (9)	Number (10)
Older schoolmate above cutoff (ITT)	0.007	0.017	0.003	0.014*	0.006	0.022	0.006	0.022	0.009	0.01
	(0.005)	(0.027)	(0.004)	(0.008)	(0.006)	(0.014)	(0.006)	(0.021)	(0.006)	(0.015)
% of counterfactual mean	0.93	0.99	2.8	7.52	2.45	4.77	1.33	2.33	1.74	1.12
Older schoolmate enrols (2SLS)	0.028*	0.08	0.014	0.054*	0.025	0.089*	0.024	0.098	0.036*	0.048
	(0.016)	(0.094)	(0.015)	(0.029)	(0.019)	(0.048)	(0.021)	(0.075)	(0.02)	(0.053)
% of counterfactual mean	3.57	4.66	10.99	28.56	9.52	19.12	5.47	10.27	6.78	5.41
Degree-year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Obs. (right)	69,081	69,081	69,081	69,081	69,081	69,081	69,081	69,081	69,081	69,081
Obs. (left)	55,483	55,483	55,483	55,483	55,483	55,483	55,483	55,483	55,483	55,483
Counterfactual mean	0.782	1.71	0.124	0.188	0.261	0.466	0.445	0.957	0.524	0.879
Bandwidth	20.96	20.96	20.96	20.96	20.96	20.96	20.96	20.96	20.96	20.96
First stage	0.286***	0.286***	0.286***	0.286***	0.286***	0.286***	0.286***	0.286***	0.286***	0.286***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
First stage F-stat	7,818	7,818	7,818	7,818	7,818	7,818	7,818	7,818	7,818	7,818

Notes:
Includes degree-year fixed effects. Standard errors clustered at the high school x track-year level. ***, **, and * indicate significance at 1, 5, and 10%, respectively.

Table F.3: Discontinuity in Student Preferences at Admission Cutoff Rank

	Number of Top Ranked Applications	Median Application Rank
	(1)	(2)
Older schoolmate above cutoff (ITT)	0.008	-0.046
	(0.014)	(0.047)
% of counterfactual mean	1.81	-0.97
Older schoolmate enrols (2SLS)	0.036	-0.16
	(0.048)	(0.163)
% of counterfactual mean	7.75	-3.33
Degree-year FE	✓	✓
Obs. (right)	69,081	69,081
Obs. (left)	55,483	55,483
Counterfactual mean	0.47	4.808
Bandwidth	20.96	20.96
First stage	0.286***	0.286***
	(0.004)	(0.004)
First stage F-stat	7,818	7,818

Notes:
Includes degree-year fixed effects. Standard errors clustered at the high school x track-year level. ***, **, and * indicate significance at 1, 5, and 10%, respectively.

Table F.4: Older Schoolmate Spillovers on Applications to and Enrollment in Degree of Marginally Admitted Older Schoolmate - For High School First- and Second-Ranked Applicant

	Applications				Enrollment			
	At least one		Number		At least one		Number	
	First-ranked	Second-ranked	First-ranked	Second-ranked	First-ranked	Second-ranked	First-ranked	Second-ranked
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Older schoolmate above cutoff (ITT)	0.027**	0.024*	0.091*	-0.009	0.007	0.002	0.011	0.003
	(0.011)	(0.014)	(0.055)	(0.065)	(0.007)	(0.009)	(0.01)	(0.013)
% of counterfactual mean	5.12	4.52	6.19	-0.58	7	1.52	8.75	2.27
Older schoolmate enrols (2SLS)	0.1**	0.123*	0.334	0.008	0.025	0.017	0.039	0.026
	(0.042)	(0.069)	(0.205)	(0.331)	(0.026)	(0.045)	(0.038)	(0.066)
% of counterfactual mean	19.15	23.28	22.65	0.48	25.55	14.39	31.54	17.51
College-major-year FE	✓	✓	✓	✓	✓	✓	✓	✓
Obs. (right)	26,798	15,182	26,798	15,182	26,798	15,182	26,798	15,182
Obs. (left)	20,483	23,441	20,483	23,441	20,483	23,441	20,483	23,441
Counterfactual mean [-5,-1]	0.523	0.53	1.476	1.592	0.096	0.115	0.124	0.151
Bandwidth	20.96	20.96	20.96	20.96	20.96	20.96	20.96	20.96
First stage	0.27***	0.195***	0.27***	0.195***	0.27***	0.195***	0.27***	0.195***
	(0.008)	(0.012)	(0.008)	(0.012)	(0.008)	(0.012)	(0.008)	(0.012)
First stage F-stat	2,144	581	2,144	581	2,144	581	2,144	581

Notes:
Includes college-major-year fixed effects. Standard errors clustered at the high school-year level. ***, **, and * indicate significance at 1, 5, and 10%, respectively.