# Neighbors' Spillovers on High School Choice*

Juan Matta† Alexis Orellana‡

December 4, 2024

**Abstract**

Do residential neighbors affect each others' schooling choices? We exploit oversubscription lotteries in Chile's centralized school admission system to identify the effect of close neighbors on application and enrollment decisions. A student is 9-12% more likely to rank a high school as their first preference and to attend that school if their closest neighbor attended it the prior year. Lower-achieving applicants are more likely to follow neighbors when their closest neighbor's test scores are higher. A neighbor enrolling in a school with $1\sigma$ higher effectiveness, measured by average test scores, value-added, or school climate induces increases of $0.13$-$0.35\sigma$ in the applicant's attended school. Our findings suggest that targeted policies aimed at increasing information to families have the potential to generate significant multiplier effects.

**Keywords:** spillovers, high school choice, centralized school systems

**JEL Codes:** I21, I24

*We are grateful to the *Ministerio de Educación de Chile* and *Agencia de Calidad de la Educación* for granting access to the administrative records used in this project. We thank Andrew Bacher-Hicks, Andrés Barrios-Fernández, Max García, Josh Goodman, Marcus Winters, and seminar participants at Boston University, Northwestern University, Boston College, Universidad Alberto Hurtado, AEFP, and SOLE for helpful comments and discussions.

†Department of Economics, Universidad Diego Portales. E-mail: juan.matta@udp.cl

‡School of Education and Social Policy, Northwestern University. E-mail: boris.orellana@northwestern.edu

# 1 Introduction

The determinants of school choice behavior have been a focus of policy interest and academic research for several decades. While differences in access to schools can be partially attributed to parental preferences for observable attributes, such as distance or cost, growing evidence suggests that information also plays a significant role in explaining racial and socioeconomic gaps in access to high-quality schools (Hastings and Weinstein, 2008; Andrabi et al., 2017; Corcoran et al., 2018; Ainsworth et al., 2023). School choices not only directly impact children's academic outcomes and future life opportunities, but they may also spillover to future cohorts of students, potentially exacerbating segregation patterns or widening achievement gaps if more disadvantaged families disproportionately attend low-performing schools. Understanding how local environments shape families' preferences for schools is important, given that neighborhoods and communities influence other relevant economic margins, such as the likelihood of working in the same location (Bayer et al., 2008; Hellerstein et al., 2011) or engaging in criminal activity (Billings et al., 2019). However, providing credible evidence on whether residential neighbors influence school decisions can be challenging in most educational contexts due to nonrandom sorting into schools and the presence of unobserved neighborhood attributes.

In this paper, we study the importance of close neighbors on families' high school application and enrollment decisions. Using data from Chile's centralized school assignment mechanism between 2019 and 2022, we link applicants to their closest residential neighbor and show that shocks to neighbors' enrollment decisions spillover to applicants in the next year, affecting their choices of applying to and attending the same schools. From a policy perspective, taking into account these dynamic responses to neighbors' enrollment has important implications for the design and evaluation of school choice interventions.

Identifying spillover effects using observational data is subject to two empirical problems, known in the literature as the reflection problem and the existence of correlated effects (Manski, 1993). To surmount these two challenges, we take advantage of the implementation of a centralized school admission system in Chile. Under this system, student assignment is determined using the Deferred Acceptance mechanism and lottery tie-breakers in oversubscribed schools. Building on earlier work (Abdulkadiroğlu et al., 2011, 2017; Gray-Lobe et al., 2023) these features motivate an instrumental variables strategy to identify the effect of the closest neighbor's enrollment on an applicant's decision. We exploit the exogeneity introduced by the tie-breaking rules in a large number of oversubscribed schools to overcome the correlated effects problem. Regarding the reflection problem, we employ multiple rounds of the centralized system and focus on the effect of the closest neighbor being offered a seat in the *previous* round on the probability of applying to the same school in the *current* round.

We find that close neighbors influence future applicants' behavior. Our main results, based on two-stage least squares (2SLS) estimates, show that an applicant exposed to a neighbor who attends their target school is 2.1 and 1.4 percentage points more likely to include this school in the application list and rank it as the most preferred school, respectively. These estimates represent an increase of 6% and 9% relative to the average for the non-treated group. In terms of school attendance, the presence of a neighbor attending a given school increases by 1.3 percentage points the probability of enrolling in the same school in ninth grade. This estimate corresponds to a 12% increase relative to the average for the non-treated group.

We conduct a series of heterogeneity analyses to investigate how these estimates vary by applicants' and neighbors' observable characteristics. Firstly, in terms of gender, we find these effects are much larger when the applicant is male and the neighbor is female. Secondly, we characterize students' prior academic performance using eighth-grade GPA and standardized tests in prior grades. Our results show that spillovers are larger when applicant and neighbor have GPA scores below the top quintile. For this group, neighbors' enrollment increases the probability of ranking the same school as top choice by 10% and attending the same school by 14%. In a different exercise, we test whether spillovers decay with distance using an extended sample where we link each applicant to a set of ten nearby neighbors. We find that only the nearest has a meaningful effect on applicants' decisions. As the distance order increases, our estimates become imprecise and not statistically distinguishable from zero.

Guided by recent literature studying parental preferences for schools (Burgess et al., 2015; Abdulkadiroğlu et al., 2020; Beuermann et al., 2022; Ainsworth et al., 2023), we supplement our analysis with available data to characterize schools across different dimensions, such as distance, average test scores, school value-added, and school climate. We employ these metrics to analyze heterogeneous effects by school traits. We find that applicants are more likely to rank the same school as their top preference when the distance to the school is shorter and each of the school quality proxies is higher. Neighbors enrolling in schools within the top tercile of the tenth-grade average scores distribution increase the probability that students in the subsequent cohort will list that school in any rank and as their top choice by 6.2 and 4.1 percentage points, respectively. These patterns are qualitatively similar when we examine heterogeneous effects by school climate or school value-added.

Next, we move to the question of whether neighbors' decisions also shape the characteristics of the schools applicants select. We employ our set of school traits to quantify changes in the characteristics of schools applicants attend as a consequence of following neighbors. We find that neighbors enrolling in their most preferred school induce applicants to attend more effective schools. A neighbor enrolling in a school that is one standard deviation ($\sigma$) above the average in the tenth-grade test score distribution induces an increase of $0.13\sigma$ in the school where the applicant enrolls. Similarly, a $1\sigma$ increase in school value-added on high school graduation and college enrollment implies that applicants attend schools with $0.27\sigma$ and $0.35\sigma$ larger school value-added, respectively.

We provide evidence supporting that learning from neighbors' previous choices is one of the mechanisms explaining our results. Applicants scoring below the median in math standardized tests whose closest neighbor scored above the median are 2.1 percentage points more likely to rank the same school as their top choice and attend it. By contrast, we find negligible effects for applicants who scored above the median but their neighbors did not. This pattern suggests that neighbors convey information about school quality or other attributes valued by families and that applicants internalize these signals based on ability differences. While we view these findings as evidence consistent with learning, we lack statistical power in some of our analyses and we cannot completely rule out the relevance of other potential mechanisms, such as reductions in search costs or preferences for unobserved traits.

This paper contributes primarily to the literature studying spillover effects on human capital decisions.[1] Previous work related to the effects of social networks on educational choices has focused mostly on siblings effects at the secondary level (Joensen and Nielsen (2018) for Denmark, Dustan (2018) for Mexico, and Dahl et al. (2023) for Sweden) and at the college level (Goodman et al., 2015; Aguirre and Matta, 2021; Altmejd et al., 2021).[2] By contrast, evidence about neighbors' effects on educational decisions is less common. One recent study is Barrios-Fernández (2022), who estimates neighbors' spillovers on college attendance. Most related to this paper, Bobonis and Finan (2009) and Lalive and Cattaneo (2009) show evidence of neighbors' effects on school enrollment in primary grades leveraging variation from the implementation of the PROGRESA program in Mexican rural communities. This paper differentiates from these studies in two important ways. First, while Bobonis and Finan (2009) and Lalive and Cattaneo (2009) focus on extensive margin changes in enrollment, we are interested in application decisions for students already attending eighth grade. According to the 2021 *Education at a Glance* report, Chile has an attendance rate of 82% for students aged 15-19, similar to the average 84% in OECD countries.[3] Thus, our results are likely to be generalizable to other educational systems in developed and middle-income countries.

[1] A related literature has studied the effects of residential proximity on other economic outcomes and decisions, such as the effects of working on a specific job or establishment (Bayer et al., 2008; Hellerstein et al., 2011), consumption choices (Grinblatt et al., 2008; Angelucci and De Giorgi, 2009; Kuhn et al., 2011; Agarwal et al., 2021), engaging in youth criminal activity (Billings et al., 2019), or perceptions about well-being (Luttmer, 2005).

[2] See Qureshi (2018), Nicoletti and Rabe (2019), and Gurantz et al. (2020) for siblings spillover effects on student achievement.

[3] https://doi.org/10.1787/b35a14e5-en

Second, the school admission system is available to all students enrolling in non-private schools, which account for around 90% of total enrollment in Chile. By merging application records to a rich set of background characteristics, we examine how families respond to their closest neighbors' decisions across several dimensions, such as previous achievement, school characteristics, and residential proximity. We are not aware of such type of analysis in previous work.

We also contribute to the literature examining the indirect effects of centralized school choice mechanisms. Unlike previous literature studying the short-term (Cullen et al., 2006; Hastings et al., 2006; Dobbie and Fryer, 2011; Abdulkadiroğlu et al., 2018; Lincove et al., 2018) and long-term impacts (Deming, 2011; Deming et al., 2014; Dustan et al., 2017; Gray-Lobe et al., 2023) of winning admission to an oversubscribed school under a lottery-based design, this paper focuses on how applicants' decisions spillover to future cohorts. Our heterogeneity analyses show that applicants are significantly more likely to be influenced by close neighbors when they attend more effective and less distant schools. Understanding how these indirect effects vary across observable characteristics is important for at least two reasons. First, it has implications for the design of information interventions (Andrabi et al., 2017; Neilson et al., 2019; Ainsworth et al., 2023) or other policies, such as introducing (or expanding) quotas for specific groups. Second, as pointed out by Angelucci and De Giorgi (2009), to assess the effectiveness of these interventions correctly, it is necessary to consider spillover effects generated by treated units. Our analysis confirms that these spillover effects are meaningful in school choice contexts.

Taken together, our findings suggest that, conditional on their choice sets, neighbors' enrollment decisions propagate to other applicants. These spillover effects vary by applicants' background characteristics and have an influence on the characteristics of schools they attend. As a consequence, residential neighbors' school choices may partly explain the persistence of unequal access to high-quality schools and subsequent achievement gaps.

# 2 Institutional Background and Data

In 2016, Chile started a transition from a decentralized admission system to a centralized system based on the Deferred Acceptance mechanism (Gale and Shapley, 1962). The *Ley de Inclusión Escolar* (School Inclusion Bill), enacted in 2015, introduced stark changes to how parents applied to schools through the implementation of the *Sistema de Admisión Escolar* (School Admission System) for all schools receiving total or partial public funds. Before the law's passing, voucher schools could charge tuition add-ons and run admission processes independently, while public schools faced more restrictions. By 2017, public and voucher schools concentrated 36% and 55% of the nationwide enrollment, respectively. Private schools, which account for 9% of total enrollment, were not included in the reform and do not participate in the centralized assignment mechanism.

The implementation of the policy was staggered across regions and grades. Starting in 2017, every year an additional group of regions was incorporated for pre-K, K, first, seventh, and ninth grades, adding the remaining grades in the following year. For ninth-grade applicants, the reform was fully implemented by 2019. Figure 1 shows the number of applicants observed each year, and Online Appendix Figure A.1 shows the distribution of applicants by grade in the 2019-2022 rounds. Most applications are observed in school transition grades (pre-K, first, and ninth grades). In the Chilean educational system, some secondary flagship schools (*liceos emblemáticos*) start in seventh grade, which explains the high number of applications observed at this level. Online Appendix Figure A.2 shows the number of participating schools by grade. Between 2019 and 2022, around 2,000 high schools offer at least five seats for ninth-grade students, with an average of 69 vacant seats. Online Appendix Figure A.3 summarizes the main stages of the admission process. Each year, families submit their school preferences between September and October. After receiving all applications, the main assignment round is conducted and families observe the outcomes around November. There is a complementary round where unassigned applicants or families who did not participate in the main round can submit a new application. Unassigned students in this complementary round are allocated to the closest tuition-free school with available seats. The process ends in late December when all students have received an assignment. Online Appendix Table A.1 shows the acceptance rates for each round at different school levels. Considering the 2019-2022 rounds, more than 80% of applicants obtain a seat in any of their three most preferred schools. Depending on the school level, between 40% and 60% obtained a seat in their most preferred alternative.

Two important features of this centralized system are worth mentioning. First, some groups of students receive priority in the assignment rule. There are four priority groups served in strict order: (i) students with siblings enrolled at the school, (ii) students with a parent working at the school, (iii) former students previously enrolled at the school, and (iv) all other applicants (Correa et al., 2022). Furthermore, the system includes special quotas for vulnerable students, and some schools can select a fraction of their seats based on admission tests.[4] In the former case, disadvantaged students are given the second highest priority after (i). In the latter case, the system first fills these quotas by assigning students based on their admission test scores, and the remaining seats are assigned following the priority groups (i)-(iv). Online Appendix Figure A.4 summarizes the seat classification within schools and the priorities in each case.[5] Second, ties are broken randomly

[4]Some schools incorporate a quota reserved for special-needs students. We do not incorporate this last group of students in our analysis.

[5]An additional issue relates to the fact that students might receive different priorities depending on which group they are considered. For example, a disadvantaged student with a working parent fits into two different seat categories

within each priority group in oversubscribed schools. Based on these considerations, we simulate the algorithm used to allocate students to schools and compute the assignment propensity score for each applicant (Abdulkadiroğlu et al., 2017). Figure 2 shows the proportion of schools receiving more first-rank applications than vacant seats. This figure shows that in ninth grade, more than 30% of the schools participating in the system are oversubscribed.

Crucially for our purposes, the administrative records include the geocoded location of every applicant. For confidentiality purposes, these locations contain a small amount of noise.[6] Additionally, not all addresses in the data correspond to actual residences. We take on a number of steps to discard unreliable geographic locations. First, we drop all students with imputed addresses.[7] Second, we drop applicants whose registered location indicates one region, but their school enrollment records in the same year indicate a different region. This sample selection drops around 20% of all applicants.

## 2.1 Sample Construction

We build our sample using data for ninth-grade applicants observed in the 2019-2022 application rounds. As Figure 1 shows, 2019 is the first year when we observe applications for each region in the country. We match each applicant to the closest neighbor who applied to ninth grade in the previous year after excluding cases following the criteria described at the end of the previous section. Since our location data contain noise, the closest neighbor we identify needs not be the actual nearest neighbor. To distinguish between close neighbors and members of the same family applying in different years, we employ anonymized parent identifiers and discard siblings or pairs of students associated with the same adult responsible for the application. Unfortunately, we do not observe student names in any of the administrative records we employ in this study. We also drop observations where the (noisy) distance between the applicant and neighbor exceeds 0.5 miles. Finally, we exclude from the estimation sample neighbors who took an admission test in their most preferred school.[8] For all applicants, we observe the outcome of the first round of the assignment process. At this stage, parents can accept the designation, accept it conditionally on not receiving

(disadvantaged and no trait). In these cases, the allocation mechanism assumes that students have preferences over contracts that specify the school and the type of seat to be used, whereas schools have preferences over contracts specifying the student and the type of seat (shown in Online Appendix Figure A.4). See Correa et al. (2022) for additional details.

[6] Observed locations are displaced between 50 and 350 meters from the actual ones, with a median value of 175 meters.

[7] Applicants whose residential address was not accurately captured are assigned the location of the municipality where they live.

[8] In our sample, 4% of all assigned seats correspond to schools authorized to select applicants based on admission tests.

an offer from a more preferred school, or reject it and apply to a private school. We link each applicant to enrollment records in the next year to observe which school they finally enrolled in.

We compute the probability of being offered admission at the top-ranked school by simulating the assignment algorithm 1,000 times for each student, changing the random tie-breaker and obtaining the school allocation in each iteration while holding preferences and priorities fixed. Based on the simulated allocations, we use the fraction of times when students receive an offer at their most preferred school as the assignment propensity score. Online Appendix Figure A.5 presents the distribution of the propensity scores for all ninth-grade applicants between 2019 and 2022. Online Appendix Figure A.6 shows that our simulation almost exactly replicates the observed assignments. The slope of a bivariate regression between an offer indicator and the propensity score is 0.999.

We supplement our analysis sample with two additional sources of information. Firstly, we link students' previous math and language test scores in standardized national exams (SIMCE tests).[9] These administrative records also contain survey information about family characteristics, such as reported income, parents' education, and college expectations. We can merge these records to 80% of applicants and 82% of neighbors in our sample. Online Appendix Table A.2 summarizes the grade from which we can observe these records for each cohort. Secondly, we incorporate available student- and school-level data for the 2015-2018 cohorts of tenth-grade students to construct additional attributes, such as average tenth-grade SIMCE scores, school value-added, and school climate proxies.

## 2.2 Sample Description

Our analysis sample consists of eighth-grade students who apply to a high school using the centralized admission system. Students enrolled in K-12 schools can choose to participate if they want to move to a new school, while students enrolled in K-8 schools necessarily need to participate unless they prefer to switch to a private school. Table 1 compares observable characteristics of applicants relative to the universe of eighth-grade students enrolled in non-private schools. Column (1) shows the average characteristics of all students enrolled in K-12, non-private schools, while column (2) restricts the sample to students enrolled in K-8 non-private schools. Column (3) shows the characteristics of students participating in the centralized system, which correspond to approximately 49% of all eight-grade students in the country. Relative to column (1), the subset of applicants is more disadvantaged, measured by the fraction of low-income students (*prioritario* and *preferente*

[9]SIMCE is an acronym of *Sistema de Medición de la Calidad de la Educación* (National System of Quality Measurement). It was created in 1988 and has been the primary indicator to identify effective schools (Mizala and Urquiola, 2013) or intervene ineffective ones (Chay et al., 2005).

statuses), previous performance, and parents' education.[10] The comparison of the number of students in columns (2) and (3) shows that a small fraction of students enrolled in K-12 schools choose to apply to a different school in ninth grade.

Column (4) presents summary statistics for the subset of students in our estimation sample, defined as applicants who are linked to neighbors with a propensity score strictly between zero and one. The comparison of columns (3) and (4) shows that our estimation sample is representative of the total applicant population. Panel A shows that around 53% of applicants in our estimation sample are girls, 63% have a disadvantaged (*prioritario*) status, and 60% attended a public school in eight grade. The average baseline math and language test scores are -0.26$\sigma$ and -0.19$\sigma$, respectively, primarily reflecting differences in achievement between students enrolled in public and private schools. Around 16% of applicants' mothers have a college degree, and 9% of applicants' families report a monthly income higher than CLP800k ($\approx$ US$1,000 in year 2021). Similarly, Panel B summarizes application metrics for all applicants and randomized applicants. Students in our estimation sample apply on average to 3.6 schools, and around 60% of them submit three schools or less. When comparing these outcomes to the universe of applicants, our estimation sample exhibits only minor differences. Applicants in our sample submit almost the same number of schools on average and the proportion submitting four or more schools is only 0.8 percentage points larger.

Online Appendix Figure A.7 shows the distribution of applications pooling all rounds. We observe that the modal number of applications is three and that less than 25% of families apply to more than five schools. Online Appendix Figure A.8 shows the distribution of the applicant-vacant ratio for schools offering ninth grade. For each school $s$ and year $t$, we compute the number of students applying to this school as their top choice $A_{st}$ and use the reported number of vacant seats offered by the school, $V_{st}$. The ratio $A_{st}/V_{st}$ summarizes the excess demand at each school and year. Online Appendix Figure A.8 shows that around 30% of schools face a ratio $A_{st}/V_{st} > 1$.

Finally, Online Appendix Table A.3 presents differences in application patterns by socioeconomic status and other student characteristics using simple OLS regressions. Overall, disadvantaged applicants are more likely to submit fewer schools and to apply to schools with lower average tenth-grade test scores. To characterize socioeconomic status, we employ the disadvantaged status variable included in the application records. Column (1) shows estimates of a regression of the number of applications onto students' characteristics and priority groups. Conditional on the priority group,

[10]The *prioritario* and *preferente* statuses were introduced in 2008 by the *Ley de Subvención Escolar Preferencial* or SEP bill, which established a new targeted voucher to transfer additional resources to schools receiving these students. Each status is determined based on household economic hardship, income, and mother's education. See the work of Mizala and Torche (2017), Feigenberg et al. (2017), and Neilson (2023) for additional details about the implementation of the bill and its consequences on student outcomes.

gender, and being in the top quintile of the eighth-grade GPA distribution, low-SES applicants submit 0.4 fewer schools on average. Columns (2)-(3) indicate that disadvantaged students apply to schools with significantly lower average scores. We pool all schools included in students' applications and regress the average tenth-grade language and math scores of each school (measured in 2018) onto student characteristics and distance to school, conditioning on priority status and the number of schools included in the application. The estimates imply that low-SES applicants consider schools that perform $0.21\sigma$ less in standardized tests compared to higher-SES applicants with similar characteristics and location. In columns (4) and (5) we repeat the analysis focusing on the top-ranked school. In this case, the gap increases to $0.28\sigma$ and $0.27\sigma$ in language and math school average tenth-grade test scores, respectively. These patterns suggest that low-SES students in Chile are more likely to apply to lower performing schools although they may list as many schools as they want on their application, consistent with prior research about heterogeneous preferences for school attributes across socioeconomic groups (Hastings et al., 2005; Hastings and Weinstein, 2008; Burgess et al., 2015; Neilson, 2023). Based on these descriptive patterns, our objective for the rest of the paper is to analyze whether this differential behavior impacts the application decisions of future cohorts.

# 3 Empirical Strategy

In this section, we describe our empirical strategy to identify the impact of close neighbors' enrollment on applicants' decisions. Our strategy leverages variation in neighbors' enrollment induced by random admission offers in oversubscribed schools, conditional on their choice sets.

## 3.1 Assignment Propensity Score

As discussed in Section 2, all applicants within a priority group who rank the same oversubscribed school as their first option have the same probability of receiving an admission offer to that school. The features of the Chile's centralized system motivate the use of the assignment propensity score, defined as the probability of obtaining a school offer conditional on each applicant's preferences and priorities (Abdulkadiroğlu et al., 2017; Angrist et al., 2024). Formally, we characterize an applicant $i$ by a set of preferences over schools $\succ_i$ and a vector of priorities at each of the $M$ schools submitted to the centralized system, $\rho_i = (\rho_{i1}, \ldots, \rho_{iM})$. Following the notation of Angrist et al. (2024), student $i$'s propensity score for assignment to their most preferred school is defined as:

$$P_i = \Pr(z_i = 1 | \succ_i, \rho_i) \tag{1}$$

Where $z_i$ is a binary indicator equal to one if $i$ was offered admission to their top-ranked school.

## 3.2 Neighbors' Spillovers

Consider an individual $i$ applying to school in year $t$, and let $n$ be $i$'s nearest neighbor among year $t-1$ applicants. We are interested in studying how $n$'s school enrollment affects the decisions of applicant $i$. We relate both by estimating the following set of equations using two-stage least squares (2SLS):

$$y_{in} = \alpha + \beta x_n + \sum_p \phi_p \mathbb{1}\{P_n = p\} + \varepsilon_{in} \tag{2}$$

$$x_n = \gamma + \delta z_n + \sum_p \pi_p \mathbb{1}\{P_n = p\} + \nu_n \tag{3}$$

where $y_{in}$ is a binary indicator that equals one if $i$ applied to (or enrolled in) school $s_n$ (i.e., the school ranked first by $n$), $x_n$ is a binary indicator that equals one if neighbor $n$ enrolled in $s_n$, and the instrument $z_n$ is a binary indicator for whether $n$ was offered admission to $s_n$. Both equations include a saturated set of indicators for values of the propensity score $P_n$. Our parameter of interest is $\beta$, which captures the causal effect of $n$'s enrollment in school $s_n$ on the probability that $i$ applies to (or enrolls in) the same school.

In our estimation, we use the sample of applicants whose nearest neighbors participate in oversubscribed lotteries, that is, neighbors $n$ such that $0 < P_n < 1$. In all of our specifications, we cluster standard errors at the neighbor level to account for the fact that one neighbor can be linked to multiple applicants.

## 3.3 Identifying Assumptions

Identification of spillovers requires admission offers to affect $n$'s school enrollment (i.e., $\delta \neq 0$), as well as the following conditional independence assumption:

$$z_n \mid P_n \perp \varepsilon_{in}, \nu_n \tag{4}$$

This assumption means that conditional on the neighbor's propensity score $P_n$, admission offers made to $n$ must be independent of unobserved factors affecting $i$'s and $n$'s enrollment. Independence with respect to $\nu_n$ is guaranteed by the fact that all neighbors in the same lottery have the same probability of being offered admission to $s_n$. Independence with respect to $\varepsilon_{in}$ further relies on an exclusion restriction, i.e., we need $z_n$ to affect $y_{in}$ exclusively through its effect on $x_n$. In other words, we need to assume that admission offers made to the neighbor do not affect the applicant's choices unless they affect the neighbor's actual enrollment. This assumption rules out the possibility that an offer can affect the applicant even if the neighbor does not attend that school. In section 4 we present evidence that supports the exclusion restriction.

Our framework can be extended to accommodate the possibility of heterogeneous effects. Under additional assumptions, our 2SLS estimate of $\beta$ can be interpreted as a weighted average of local average treatment effects (Imbens and Angrist, 1994) for compliers who attend their top-ranked school in response to an enrollment offer received by the neighbor at each value of the propensity score $P_n = p$. Following the discussion in Gray-Lobe et al. (2023), let $\beta_p$ be the LATE for compliers with propensity score $p$. Our estimate of $\beta$ will put more weight on values of $\beta_p$ associated with a higher compliance rate ($\delta_p$), a larger fraction of applicants linked to neighbors with propensity score $p$, and higher variance of $z_n$ conditional on $p$ (i.e., closer to $p = 0.5$).[11] It is important to remark that the counterfactual school enrollment depends on the remaining set of ranked schools, whether or not they are also oversubscribed, and the priorities of other applicants to those schools. Therefore, the counterfactual enrollment for compliers is not necessarily the second-ranked school. To better understand the counterfactual scenario, we estimate the effect of receiving an offer on four characteristics of the school attended in ninth grade for the full sample of randomized applicants between 2018 and 2022.[12] Online Appendix Table A.4 shows that receiving an offer to the highest ranked school improves substantially the quality of the school attended in ninth grade. Column (2) shows an increase of $0.4\sigma$ in school average tenth-grade scores and increases of around $0.02\sigma$ in two measures of school value-added.[13] Applicants who receive an offer to their highest ranked school also attend schools with better environments, measured by an increase of $0.38\sigma$ in a school climate index developed by the Ministry of Education.

[11] See Section III.A in Gray-Lobe et al. (2023) for more details and Appendix C of Walters (2018) for a derivation of the formula.

[12] Specifically, we estimate the following set of equations using 2SLS:

$$w_{s(i)} = \alpha + \beta x_i + \sum_p \phi_p \mathbb{1}\{P_i = p\} + \varepsilon_i \tag{5}$$

$$x_i = \gamma + \delta z_i + \sum_p \pi_p \mathbb{1}\{P_i = p\} + \nu_i \tag{6}$$

Where $w_{s(i)}$ denotes a characteristic of the school attended by applicant $i$.

[13] We describe the construction of each school value-added measure in our heterogeneity analysis in section 4.3.3.

Our strategy allows us to identify the causal effect of neighbor $n$'s enrollment on applicant $i$'s decisions. This effect should be interpreted as a reduced form parameter capturing both the direct influence of $n$ over $i$ and any indirect effects of $n$ operating through other applicants who might be affected by $n$'s enrollment and affect $i$'s decisions (Barrios-Fernández, 2022). In the next section we present evidence ruling out contemporaneous effects, which favors an interpretation of $\beta$ as the direct effect of the nearest neighbor applying in the previous round.

# 4 Results

## 4.1 Balance Tests

Before presenting our main results, we examine the validity of our empirical strategy. Under the exclusion restriction, admission offers to each neighbor should be uncorrelated with other determinants of applicants' school attendance conditional on each value of the propensity score $P_n$. Panel A of Table 2 shows that applicants' observable characteristics are balanced based on neighbors' offers. We test differences across several individual and family characteristics. Specifically, we consider gender, socioeconomic status, high-achieving status (defined as being in the top quintile of the eight-grade GPA distribution), baseline test scores, parents' education, college expectations, and family income for each applicant. Columns (1) and (2) show estimates of a separate OLS regression of the observable characteristic onto an offer indicator, conditioning on a full set of propensity score indicators. All but two of the estimates are not statistically significant at the 10% level. The only imbalanced covariates are gender and baseline math test scores. For the latter, we find that applicants whose neighbors were admitted to their target school scored $0.015\sigma$ below those linked to neighbors who did not received an offer. The remaining covariates display small differences not statistically different from zero at the 10% level. We also conduct a joint significance test where we regress the offer indicator onto all background variables listed above and test the hypothesis that all coefficients are jointly zero. The $p$-value of 0.349 provides further evidence that the likelihood of a neighbor receiving an offer is exogenous to applicants' observable characteristics.

Analogously, we test whether neighbors' observable characteristics are balanced between offered and non-offered individuals. Panel B of Table 2 shows the estimates of regressions on the same set of observable characteristics as well as the $p$-value from a joint significance test. As expected, the estimates show that student attributes do not explain seat assignment after conditioning on the assignment propensity score. Finally, the last row shows no statistically significant differences in the geographic distance between each applicant and their closest neighbor.

## 4.2 Neighbors' Spillovers on School Applications and Enrollment

Table 3 shows our intent-to-treat (ITT) and 2SLS estimates of the influence of neighbors on applicants' behavior. Column (1) shows the estimate of the first-stage coefficient $\delta$ in equation (3). This estimate shows that an offer at the top-ranked school increases the probability of attending it in ninth grade by 67 percentage points. Column (2) shows that the probability of including this school in the application list increases by 1.4 percentage points on average if the closest neighbor receives an offer. To contextualize the magnitude of each estimate, we use the estimate of the average outcome in the untreated state for the group of compliers, following Abadie (2002). Relative to the mean for non-treated compliers (i.e., applicants whose closest neighbor did not get an offer), this estimate represents an increase of 4%. Column (3) shows that the probability of applying to the same school as top choice increases by 0.9 percentage points (or 6% relative to the mean for non-treated compliers). Column (4) shows the ITT estimate on school attendance. We find an increase of 0.86 percentage points in the probability of attending the same school as the neighbor's most preferred alternative, corresponding to a 8% increase.

Columns (5)-(7) show our 2SLS estimates using the neighbor's offer receipt as an instrument for attendance. The probability of applying to a school in any preference increases by 2.1 percentage point and the probability of ranking this school as the top alternative increases by 1.4 percentage points. These estimates represent increases of 6% and 9% relative to the mean for non-treated compliers, respectively. Finally, column (7) shows that the closest neighbor's enrollment in their most preferred school also increases the probability of an applicant attending it by 1.3 percentage points. This estimate is equivalent to an increase of 12% relative to the mean for non-treated compliers. We report the Kleibergen-Paap $F$-statistic in columns (5)-(7).

*Standard errors:* In recent work, Lee et al. (2022) show that conducting inference based on $t$-ratios in IV studies might lead to over-rejection and under-covered confidence intervals. They propose using an adjusted $t$-ratio depending on the value of the first-stage $F$-statistic and 2SLS estimates ($tF$ critical values). We examine whether our estimates are robust to this correction by employing their adjustment method for tests with a significance level of 0.05 and 0.01.[14] Considering the large values of our reported $F$-statistics in Table 3, standard errors and confidence intervals remain unchanged.

*Comparison to OLS estimates:* We report OLS estimates from specifications not adjusting for the saturated set of assignment propensity scores in Online Appendix Table A.5. Using the same estimation sample, we find an increase of 6 percentage points in the probability of applicants mimicking

[14] See pages 3271 and 3272 in Lee et al. (2022).

their closest neighbor's top-ranked school. This number is almost four times larger than the 2SLS estimate we report. Similarly, the OLS estimate for enrollment is 9.4 percentage points, around seven times larger than our 2SLS estimate. The upshot of these comparisons is that not properly accounting for correlated effects vastly overstates the magnitude of the spillover effects.

*Comparison to previous literature:* Previous research on neighbors' spillovers in school enrollment decisions (Bobonis and Finan, 2009; Lalive and Cattaneo, 2009) has documented the relevance of peers living in the same community.[15] However, our results are not directly comparable to these estimates. First, these studies report the change in the likelihood of attending a school when the peer group's enrollment rate increases by 1 percentage point, while our treatment variable is defined only by the closest neighbor's enrollment. In addition, our sample is not restricted to a particular subpopulation (such as the villages participating in the PROGRESA program) and includes applicants from different backgrounds. For these reasons, we also consider how our estimates relate to siblings' effects on school choices at the secondary level. Overall, our estimates align with the effects reported by other work in this literature.[16] These orders of magnitude are also observed for siblings' effects on college major choices. For example, Altmejd et al. (2021) show that the probability of a younger sibling applying to the same college in first preference increases by 3.3 to 6.3 percentage points and by 0.6 to 1.2 percentage points by applying to the same college-major combination in the first preference. Similarly, Aguirre and Matta (2021) find an increase of 1.9 percentage points in the probability of choosing the same college-major combination.

*Placebo Tests:* In addition to the balance tests presented in Table 2, a second test exploits the fact that applicants should be influenced only by neighbors' previous choices. If neighbors' influence drives our results, future choices should not affect current behavior. To conduct this falsification exercise, we first match each applicant in year $t$ to their closest neighbor in $t+1$ and test whether there is an effect of the offer received by this neighbor on applications observed in the *previous* year. Table 4 show ITT estimates of the offer indicator in $t+1$ on outcomes observed in $t$. The estimates are of smaller magnitude than our main estimates and not statistically different from zero at the 10% level. In addition, we can exploit the timing of the allocation to conduct a second placebo test, now linking applicants and neighbors who participate in the *same* year. Considering the timeline of the assignment process presented in Online Appendix Figure A.3, by the time neighbors receive

[15] Bobonis and Finan (2009) find an increase in secondary school enrollment rate of 5 percentage points in ineligible households of treated villages in the PROGRESA program, relative to ineligible households in control villages. Lalive and Cattaneo (2009) find that an increase of 10 percentage points in peer group school attendance leads to a 5 percentage points increase in individual attendance.

[16] Joensen and Nielsen (2018) find an increase of 7 percentage points in the likelihood of applying to the same math-science major as the older sibling from a pilot program in Denmark. Dustan (2018) finds an increase of 7 percentage points in the likelihood of applying to the same school in Mexico. Dahl et al. (2023) find that younger siblings are 2.4 percentage points more likely to choose the same high-school major as their older sibling in Sweden.

an offer in November, whether or not they accept it should not determine which schools applicants submit before October. Therefore, there should be no causal link between neighbors' offers and applicants' decisions for applicant-neighbor pairs in the same year. Columns (1) and (2) in Table 5 present similar results to the previous placebo test, showing no evidence of neighbors' offers on applicants' choices. However, it could be possible that neighbors' offers still affect attendance to the same school by inducing applicants to reject their initial assignment in the main round and try to obtain a seat in the complementary round, which starts in December. We assess this potential concern by estimating whether a neighbor's offer causes an applicant to accept their assignment or to participate in the complementary round. Columns (3) and (4) in Table 5 show that this behavior is not observed. Consequently, column (5) reports no effect of same-year neighbors' offers on school enrollment. In conclusion, both placebo tests provide additional support to our identification strategy.

*Additional Robustness Checks:* Our main results are robust to several checks of the estimation sample and our preferred empirical specification. First, one might be concerned that our sample excludes students enrolled in K-12 schools who choose to participate in the assignment process. To check the robustness of our results to this type of selection, we include all applicants enrolled in K-12 schools in our estimation sample. Panel A in Online Appendix Table A.6 shows that our results remain almost similar to the main results reported in Table 3. Second, our definition of school attendance in ninth grade is based on enrollment data. These records span the first months of the year and do not include information about other schools attended if a student transfers or moves during the year. To account for potential misclassification of the school where an applicant effectively enrolled, we also use student-level academic achievement data, which lists all schools where a student was registered, allowing us to see if a student transferred to another school during the year. We repeat our main analysis defining the school attended as the school where each student was enrolled the largest number of days in ninth grade. Panel B in Online Appendix Table A.6 shows very small differences in our estimates when we consider this alternative definition of school attendance. Third, we exclude from the sample applicants who have a sibling priority. This sample selection helps to assuage concerns related to the misclassification of siblings registered in different locations as neighbors. Panel C in Online Appendix Table A.6 shows that our results remain practically unchanged. Finally, Online Appendix Table A.7 presents ITT and 2SLS estimates using a linear control for values of the assignment propensity score rather than the saturated set of indicators in equations (2) and (3).

*Persistence:* We investigate the persistence of spillover effects by estimating equations (2) and (3) in a different sample where we link each applicant to the closest neighbor who participated in the centralized assignment process two years before. Figure 3 shows these estimates alongside our main results and placebo tests. Panel A shows that the effect of applying to the same school in any rank two years later is 1.2 p.p. ($p$-value>0.1). Although the estimate is statistically indistinguishable from zero we cannot reject the null hypothesis of equal effects between $t$=1 and $t$=2. Our test of equal effects has a $p$-value of 0.25. In contrast, separate tests reject the hypothesis of equality of effects between $t$=1 and $t$=0 ($p$-value=0.018) and $t$=1 and $t$=-1 ($p$-value=0.014). For spillovers on enrollment, our results are similar. Panel B shows a spillover effect of 0.6 p.p. ($p$-value>0.1) two years later. For this outcome we also cannot rule out the equality of effects between $t$=1 and $t$=2 ($p$-value=0.22), but we reject the equality of effects between $t$=1 and $t$=0 ($p$-value=0.003) and $t$=1 and $t$=-1 ($p$-value=0.039).

To summarize, our estimates show economically important effects relative to the baseline levels. On average, neighbors' enrollment affect applicants' behavior in the next admission process. Applicants are more likely to rank a school as their top choice and enroll in it when the closest neighbor is also enrolled. We next turn to examining differences in both margins by applicants, neighbors, and school characteristics to investigate the potential channels driving our results.

## 4.3 Heterogeneous Spillovers

In this section, we study whether neighbors' influence varies according to the characteristics of applicants, neighbors, and schools. To conduct these analyses, we split the sample in different subgroups and estimate equations (2) and (3) separately for each of them.

### 4.3.1 Heterogeneity by Applicant and Neighbor Characteristics

We begin by considering heterogeneous effects by gender. Table 6 presents estimates of spillover effects across four subgroups, defined by the genders of applicants and their neighbors. Panel A shows that spillover effects are small and not statistically different from zero when applicants and neighbors are boys. In contrast, panel B highlights larger spillovers when the applicant is male and the neighbor is female. For this group, neighbors' enrollment in their top-ranked school increases the probability of submitting the same school as the highest-ranked choice by 1.9 p.p. (14%). Furthermore, the effect on attendance corresponds to an increase of 2 p.p. (21%). Panel C presents a statistically significant spillover effect when the applicant is female and the neighbor is male, although this is observed only for the probability of applying to the same school in any rank. Here, our estimates indicate an increase of 2.6 p.p. (8%), while the impacts on the remaining outcomes are smaller and imprecise. Finally, panel D shows that neighbor spillovers are imprecisely measured when the applicant and neighbor are females. In this subgroup, we cannot reject that the effects are different from zero for any of the outcomes considered. The last row reports $p$-values from tests of the hypothesis that effects are equal across sample splits. For each outcome, we cannot reject the null hypothesis of no heterogeneity across gender groups.

In addition to examining differences by gender, we evaluate how spillover effects vary in terms of two measures of academic performance, eighth-grade GPA and previous standardized test scores. In doing so, we investigate the possibility that our results can be explained by neighbors conveying information about school quality. For instance, applicants with relatively lower academic performance could be more likely to follow the choices of others, believing that a school attended by a neighbor with better performance would be a good fit for them. Conversely, applicants with better GPA or test scores might be less inclined to consider the school where a lower-performing neighbor is enrolled, as they could infer that the quality of that school is low. Table 7 presents estimates of spillover effects based on applicants' and neighbors' eighth-grade GPA. We categorize students using an indicator of high-achieving status, defined as those ranked in the top 20% of their eighth-grade GPA distribution within their cohort.[17] We evaluate how spillovers vary depending on the similarity in GPA ranking. Column (1) indicates that the spillover effects on the probability of including the same school in any rank are very similar across groups. A joint test cannot reject the null hypothesis of no heterogeneity among the four groups ($p$-value=0.978). Columns (2) and (3) show that spillovers are stronger for applicants with a GPA below the top quintile. For column (2), the comparison of panels A and B suggests that applicants with GPA below the top quintile are more inclined towards ranking the same school as their most preferred choice, regardless of their neighbors' GPA. In contrast, high-achieving applicants are less likely to consider the choices of their residential neighbors when applying to schools. Similarly, column (3) also suggests larger spillovers on attending the same school when applicants have GPA below the top quintile, irrespective of their neighbors' grades. Although we find differences in our estimates across different groups, the $p$-value of the last row indicates that they are not statistically distinguishable.

Although the GPA in Chile uses a common scale from 1 to 7, identical grades from different schools do not necessarily indicate comparable academic performance. This variation can stem from diverse grading practices among teachers or differences in evaluation methods across schools. For this reason, we also utilize data from prior math and language test scores to assess whether neighbors' influence depends on relative differences in academic performance. To conduct this analysis, we construct indicators equal to one if applicants scored above the median in the corresponding test score distribution, and similarly for neighbors. As before, we estimate our baseline equations (2) and (3) separately for each of the four subgroups defined by these indicators. Our sample is restricted by the availability of test scores for applicants and neighbors. We are able to use 69% of

[17]This classification is calculated by the Ministry of Education and included in the applicants' data.

the observations included in our main results presented in Table 3.

Table 8 presents our estimates using prior math standardized tests. Panel A presents our results for the subgroup of applicants and neighbors who scored below the median in their respective cohorts. Column (1) shows that within this subgroup the probability of considering the same school in any rank increases by 1.5 percentage points. Columns (2) and (3) display increases of 0.5 and 0.9 percentage points in the probability of ranking the same school as top choice and attending the same school, although these estimates are imprecise and not statistically significant at the 10% level.

Panel B restricts the sample to applicants scoring below the median linked to a neighbor who scored above the median. In this case, we find larger spillovers. The probability of including the same school in any rank and as the top choice increases by 2.8 p.p. and 2.1 p.p., respectively. Both estimates are statistically significant at the 5% level. The probability of enrolling in the same school also increases by 2.1 p.p. and it is statistically significant at the 5% level. In contrast, panel C shows small and insignificant estimates when the applicant scored above the median but the neighbor did not, suggesting that neighbors' enrollment is irrelevant when applicants locate relatively higher in the math test score distribution. We interpret this asymmetry with respect to panel B as suggestive evidence that neighbors' enrollment provides a positive signal of school quality to applicants when the former have better academic performance than the latter.

Finally, panel D presents our results for the subgroup of applicants and neighbors who scored above the median. We find positive spillover effects for this group but the estimates are imprecise and not statistically significant at the 10% level. Online Appendix Table A.8 reports our findings when we employ language tests to measure prior academic performance. In this case, most of our estimates are not statistically distinguishable from zero at conventional levels.

### 4.3.2 Heterogeneity by Distance

Our main results indicate that the closest neighbor plays an important role in applicants' decisions. However, the closest neighbor is one among potentially multiple members of each applicant's social network. To analyze whether neighbors' influence varies with distance we augment our sample to include the ten closest neighbors within 0.5 miles for each applicant. Using the pooled sample, we use our baseline specification to estimate the effect of each neighbor, $\beta_n$, where $n = \{1, \ldots, 10\}$, separately by distance order.

We present the set of estimates $\{\beta_n\}$ in Figure 4. The horizontal axis shows the distance order of each neighbor $n$ and the y-axis plots the corresponding estimate $\beta_n$ for each outcome. Overall, we find that when pooling across nearby neighbors only the first one influences applicants' decisions while the remaining have a smaller and imprecisely estimated effect. This pattern suggests that neighbors' influence works at a very local level, similar to what previous literature has documented. For example, Barrios-Fernández (2022) shows that only neighbors located at distances closer than 200 meters matter in the context of college enrollment.

#### 4.3.3 Heterogeneity by School Characteristics

In this section, we investigate whether spillovers also depend on the characteristics of the school neighbors attend. Following recent evidence about parental preferences in the school choice literature (Burgess et al., 2015; Abdulkadiroğlu et al., 2017; Beuermann et al., 2022; Ainsworth et al., 2023), we study heterogeneity along the following dimensions: (i) distance to school, (ii) average tenth-grade test scores, (iii) school value-added on tenth-grade test scores, and (iv) school climate.

To characterize each of these attributes, we employ data between 2015 and 2018. For average tenth-grade test scores, we use the school-level average for math and language tests. We construct school value-added using information from the same 2015-2018 cohorts of tenth-graders.[18] Finally, we employ a school climate index reported by the Ministry of Education. This index uses parental surveys from tenth-grade students in the same years, capturing attitudes and perceptions about non-academic dimensions of schools.[19] We pool information across surveys in each year and standardize the index to have mean zero and unit variance. All the previous variables consider public and private schools so that these proxies of school quality capture differences across all high schools in the country. Online Appendix Figure A.9 shows that schools with higher indexes are more demanded. Each index strongly associates with the first-rank submissions to vacancies ratio.

Then we classify each school attribute into terciles and estimate equations (2) and (3) separately for each of these groups. Table 9 summarizes our findings. Columns (1)-(3) present estimates of spillover effects on the likelihood of applying to the same school in any rank and columns (4)-(6) present our results for the likelihood of ranking the same school as top choice. We find that applicants are more likely to apply to the school where the neighbor enrolled if the school is closer and has better average test scores and school climate. Panel A shows that the probability of applying to the same school increases by 9.2 p.p. when the school is located less than 1 mile to the applicant (bottom tercile). As distance to the school increases, this probability decreases. Considering

[18] Additionally, we link information about high school graduation and college enrollment to estimate school value-added on high school graduation and college enrollment. See Online Appendix A.2 we discuss our estimation approach and the distribution of school value-added.

[19] Parents are asked multiple questions about relationships between school members, episodes of discrimination, conflict or violence incidents, and school responses to conflict situations.

our results for ranking the school as the most preferred alternative, the probability is 5.6 p.p. for closer schools. A joint test rejects the null hypothesis of equal effects across terciles at the 1% level for both outcomes. Panel B explores heterogeneity by school average tenth-grade scores. The differences across terciles show that neighbors' spillovers are larger for schools with better average performance on standardized tests. For schools in the top tercile, the spillover effects are 6.2 and 4.1 p.p. for each outcome. Similar to distance, a joint test rejects the hypothesis of equal effects across subgroups at the 5% level.

Panel C shows that applicants are similarly responsive to school value-added on test scores, although we find some differences in the probability of ranking the same school as the most preferred option. Our estimates in columns (4)-(6) show spillover effects of similar magnitude between terciles 2 and 3 and a joint test cannot reject the equality of effects across terciles. This pattern suggests that applicants are less responsive to a dimension of school quality that is harder to observe, compared to distance or school-level average performance. Finally, panel D shows that spillovers are also larger for schools with better climates. Neighbors enrolled in schools in the top tercile of the distribution increase by 4.8 and 3.6 p.p. the likelihood that applicants submit the same school in any rank and as their top choice, respectively. These differences are statistically significant across subgroups as indicated by the reported $p$-value of a joint test of equal effects. Since this index is constructed using parents' perceptions about their children's school experiences, the differences across terciles suggest that neighbors can transmit this kind of information related to school experiences that is not available to parents and which might influence their choices in the next application rounds. Unlike value-added, school climate is a dimension that parents can assess based on their children's experiences and therefore easier to transmit to future cohorts.

These results are somewhat different from those reported in the siblings effects literature. For example, Altmejd et al. (2021) find that an older sibling's admission to their target college-major increases the probability that the younger sibling applies to the same college, independent of the quality of the older sibling's target. In contrast, our analysis of different school attributes shows that applicants respond differently to neighbors' enrollment based on school characteristics. Specifically, they are less likely to follow neighbors if this school is distant and more likely to follow them if it offers a better learning environment or has better performance on tenth-grade standardized tests.

## 4.4 Effects on Schools Attended by Applicants

Finally, in this section, we examine whether neighbors' spillover effects impact the characteristics of schools chosen by applicants. Specifically, we estimate the causal relationship between neighbor's and applicant's school characteristics. We employ the same source of variation used to estimate equations (2) and (3) but focus on school attributes rather than binary decisions as the outcomes of interest. Formally, we estimate the following equations using 2SLS:

$$w_{s(i)} = \alpha + \beta w_{s(n)} + \sum_{p} \phi_p \mathbb{1}\{P_n = p\} + \epsilon_i \tag{7}$$

$$w_{s(n)} = \kappa + \rho w_{s(n)}^{offer} + \sum_{p} \pi_p \mathbb{1}\{P_n = p\} + \eta_n \tag{8}$$

Where the indexes $s(i)$ and $s(n)$ refer to the schools where $i$ and $n$ enroll, respectively. We instrument the characteristics of the school where neighbors go, $w_{s(n)}$, using the same attributes of the school where they received an offer, $w_{s(n)}^{offer}$. The outcome $w_{s(\cdot)}$ corresponds to each of the following school attributes: average tenth-grade scores, school value-added on high school graduation and college enrollment, and school climate. As before, the estimate of interest is $\beta$, which represents the effect of an increase of one standard deviation in the attribute of the school where $n$ enrolled on the value of the same attribute in the school ranked as top choice and attended by the applicant.

Table 10 shows our results. Column (1) reports estimates of the first-stage coefficient $\rho$ in equation (8). Since the receipt of an offer increases attendance by a large fraction (67 percentage points according to Table 3), we also observe a strong relationship between the characteristics of the school where each neighbor received an offer and enrolled. Columns (3) and (4) show the ITT and 2SLS estimates of spillover effects on school characteristics for each applicant's target. These estimates suggest that neighbors have a positive influence on the type of schools applicants consider. For average test scores, our 2SLS estimates show that a neighbor who attends a school $1\sigma$ higher than the average induces the applicant to rank as top choice a school located $0.125\sigma$ higher in the school-level test scores distribution. For school value-added on high school graduation and college enrollment, we also find that following neighbors lead applicants to choose schools with better attributes. Neighbors attending a school with value-added $1\sigma$ above the average in each of these dimensions implies that applicants will rank as their top choice schools with $0.24\sigma$ and $0.36\sigma$ above the average, respectively. Finally, we also find positive spillovers on school climate. Our estimates indicate a spillover effect of $0.25\sigma$ on the most preferred school. Columns (6) and (7) display similar effects on the characteristics of schools applicants attend.

We also explore whether neighbors' enrollment change the characteristics of other schools included in the application. Online Appendix Table A.9 presents the effect on the same set of characteristics for the second- and third-ranked schools. Conditional on submitting more than one preference, we observe a positive effect not only on the top-ranked school's characteristics but also on the other alternatives considered by the applicant. For example, columns (5) and (8) show that a neighbor receiving an offer in their most preferred school the average tenth-grade test scores of the schools ranked second and third by $0.087\sigma$ and $0.081\sigma$, respectively. Overall, the magnitudes for the remaining school characteristics are similar to those of the top-ranked school, suggesting that neighbors' enrollment shapes the complete application profile of students in the next round, inducing them to consider schools with better attributes.

These findings are noteworthy since they suggest that, by means of spillover effects, information interventions conducted in a given period might induce changes in school preferences, allowing applicants to enroll in schools with higher effectiveness. We next turn to an analysis of potential mechanisms to explain these patterns.

# 5 Discussing Mechanisms

This section investigates the mechanisms behind the spillover effects we document. It is worth remarking that we employ exogenous variation in the likelihood of receiving an offer for one of potentially multiple members of each applicant's network.[20] Furthermore, we do not observe school preferences before exposure to the neighbors' influence, so we cannot separately identify effects on increasing awareness of alternative options from changes in preferences. One analogy to our setting corresponds to work in the job search literature related to the importance of residential neighbors (Bayer et al., 2008; Hellerstein et al., 2011; Schmutte, 2015). Following this work, we assume the closest neighbor acts as an indirect proxy of each applicant's network. Considering these data limitations, our results could be capturing multiple causal channels. For these reasons, we discuss the plausibility of three explanations: learning from neighbors' choices, reducing decision-making costs, and utility gains.

## 5.1 Learning from Neighbors

We start by exploring the possibility that neighbors convey information internalized by applicants depending on relative academic performance. Under this hypothesis, applicants with lower relative academic performance will be more likely to mimic previous choices because they might infer this school would also be a good fit for them. By contrast, applicants with relatively better performance will be less willing to consider the school where the neighbor is enrolled because they infer that school quality (or other attribute they value) is low. Our results from Table 8 are consistent

[20] An alternative treatment margin corresponds to the share of close neighbors who obtain a seat in their most preferred school. We leave this task for future research.

with this explanation, although we lack statistical power to make any stronger conclusions. The $p$-values from tests of the null hypothesis of equal effects across subgroups show that the estimates are not statistically distinguishable. However, our analysis of heterogeneous effects by school characteristics presented in Table 9 supports the learning hypothesis. Overall, we find that families are more likely to consider neighbors' choices when schools are closer and have higher school effectiveness and better learning environments. Since value-added or school climate are not reported in the application platform these findings can be interpreted as evidence of parents reacting to information from neighbors' previous experiences. Unfortunately, we do not observe families perceptions about schools in our analysis sample. Such information would be useful to test whether applicants who follow neighbors are more likely to have a positive valuation of the school they enroll.

## 5.2 Search Costs

If searching for schools is more costly for disadvantaged families or there are information frictions, households could primarily rely on social networks and other informal sources to choose where to apply. Parents lacking information about school attributes has been extensively documented in the school choice literature.[21] Our results in Table 7 showing that spillover effects concentrate on applicants with lower academic performance might be explained by these or additional factors, such as the admission system's complexity or residential segregation, all of them motivating the use of informal networks. Consistent with this hypothesis, Arteaga et al. (2022) show for the Chilean context that: (i) the search process is costly in terms of the steps required to acquire information about schools and (ii) families have limited knowledge about the options they submit.[22] In addition, previous research documents small changes in school segregation levels and the proportion of vulnerable students across schools before and after implementing the reform (Kutscher et al., 2023; Honey and Carrasco, 2023). Therefore, we cannot rule out that search costs or other frictions stemming from the structural characteristics of the Chilean educational system are the main mechanism explain our results.

[21] Hastings and Weinstein (2008) show evidence of parents lacking information about schools and their characteristics in the Charlotte-Mecklenburg school choice program, while Jensen (2010) shows evidence that families underestimate the returns to secondary school from an experimental intervention in the Dominican Republic. Using surveys from applicants in New Haven, Kapor et al. (2020) find that families' beliefs about their admission chances are off by 30 percentage points on average.

[22] In one question asking parents about what they needed to know about a school to feel that they knew it well, 79% of applicants answered that "asking for references from current families" is a relevant step to know a school. In addition, when asked about how much they knew about the schools submitted, 64% of applicants declared that they "knew well" their target school.

### 5.3 Preferences

Finally, there is the possibility that applicants and neighbors simply have the same preferences for school traits unavailable in our data or that applicants derive utility from sharing the same environment with residential neighbors. For example, families could apply to the same schools where other residential neighbors attend to improve school-parent communication, avoid exposure to crime if the routes to school are unsafe, or simply be part of the same community. Findings from the siblings effects literature (Goodman et al., 2015; Altmejd et al., 2021; Aguirre and Matta, 2021), suggesting that there could be intrinsic value in following the path of a sibling who enrolls in a particular college or major, might also be relevant in our case to explain why applicants are more likely to follow neighbors, particularly when schools are closer. In addition, recent work by Ainsworth et al. (2023) shows that, after providing information about school value-added, families are responsive but a significant fraction still leave value-added "on the table". Thus, preferences or other unobserved school traits might be a relevant factor in explaining neighbors' spillovers.

To summarize, there are a number of reasons that could explain a causal impact of a residential neighbor impacting applicants' future decisions. Although we show evidence consistent with the idea of applicants learning about schools characteristics based on who enrolls in them, other alternative explanations might rationalize our results. Although recent work has shown that information and preferences are relevant to understanding families' choices, further research is required to quantify the importance of these different mechanisms in explaining how choices spillover to future cohorts.

## 6 Conclusion

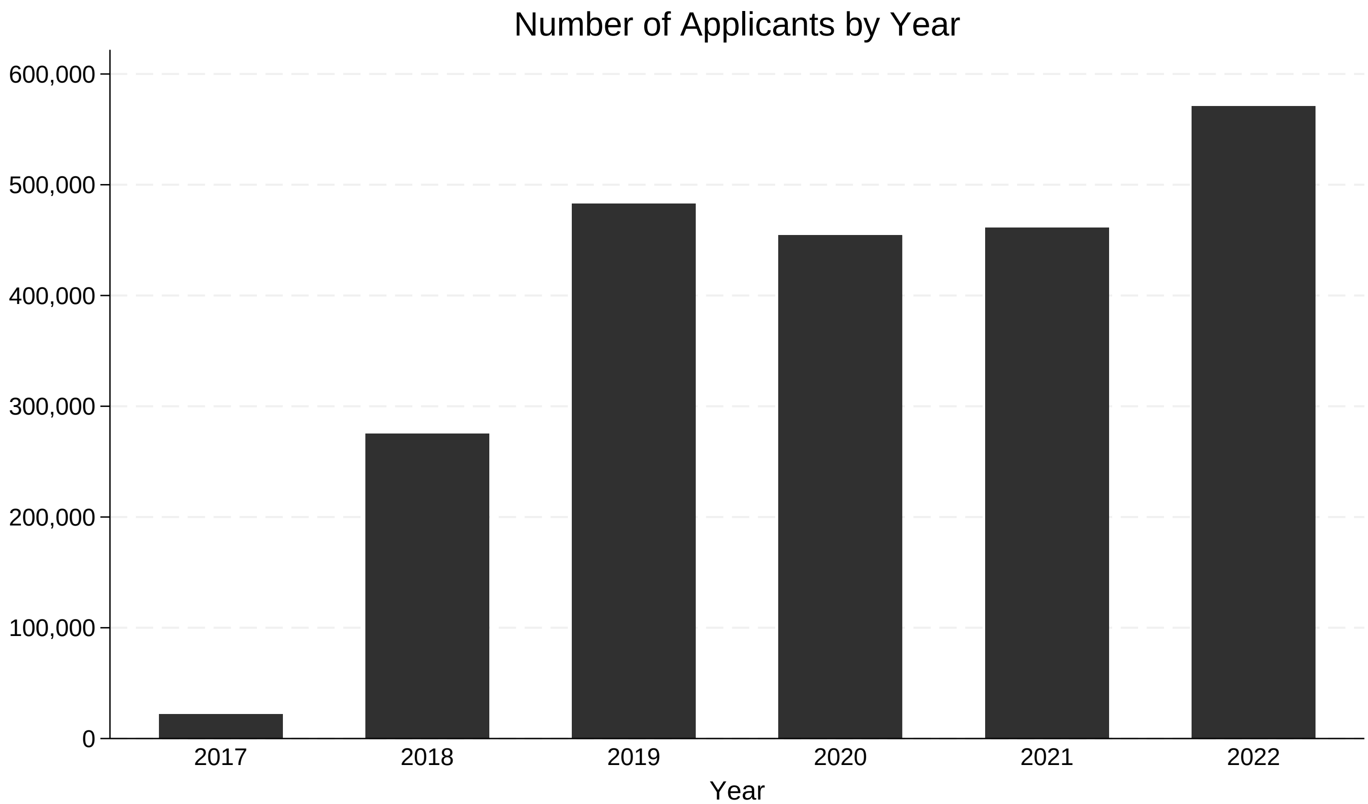
In this paper, we investigate the influence of nearby neighbors on school application and enrollment decisions. We utilize data from the Chilean centralized school admission system spanning the years 2019 to 2022. The large number of oversubscribed schools in ninth grade and the use of lottery-based tie-breakers to determine assignments motivate an instrumental variables design, enabling us to identify causal effects. To our knowledge, no prior research has examined this type of spillover effects within centralized school systems.

Our findings indicate meaningful spillover effects on school applications and enrollment. On average, having a close neighbor enrolled in their most preferred school in the previous round increases the likelihood that an applicant ranks that school as their first preference and subsequently attends it by 1.4 and 1.3 percentage points, respectively. These estimates represent increases of 9% and 12% relative to the non-treated rates. Our heterogeneity analysis reveals that these effects are more pronounced when both applicant and neighbor do not belong to the top quintile of the eighth-grade GPA distribution. However, we do not have sufficient statistical power to rule out equal effects with respect to other achievement groups. Neighbors attending schools with higher average tenth-grade test scores, school value-added, and school climate increase the probability of applicants including these schools in their rankings the following year. We also show that the influence of neighbors extends to the quality of schools applicants consider and ultimately attend. When neighbors enroll in schools with better proxies of school quality, this positively affects the characteristics of schools that applicants rank as their first, second, and third choices. On average, neighbors' enrollment in a school $1\sigma$ above the average in the distribution of tenth-grade test scores, school value-added on twelfth-grade outcomes, and school climate leads to increases of 0.13-0.35$\sigma$ in the characteristics of schools attended by the next cohort of applicants.

Our analysis indicates that information transmission could be one mechanism influencing our findings. Unfortunately, our data do not enable us to differentiate between changes in applicants' choice sets and shifts in their preferences. Future research could explore which of these two channels is more significant in explaining spillover effects among residential neighbors. Such information would be valuable in evaluating the indirect effects of targeted interventions designed to improve the allocation of educational investments.

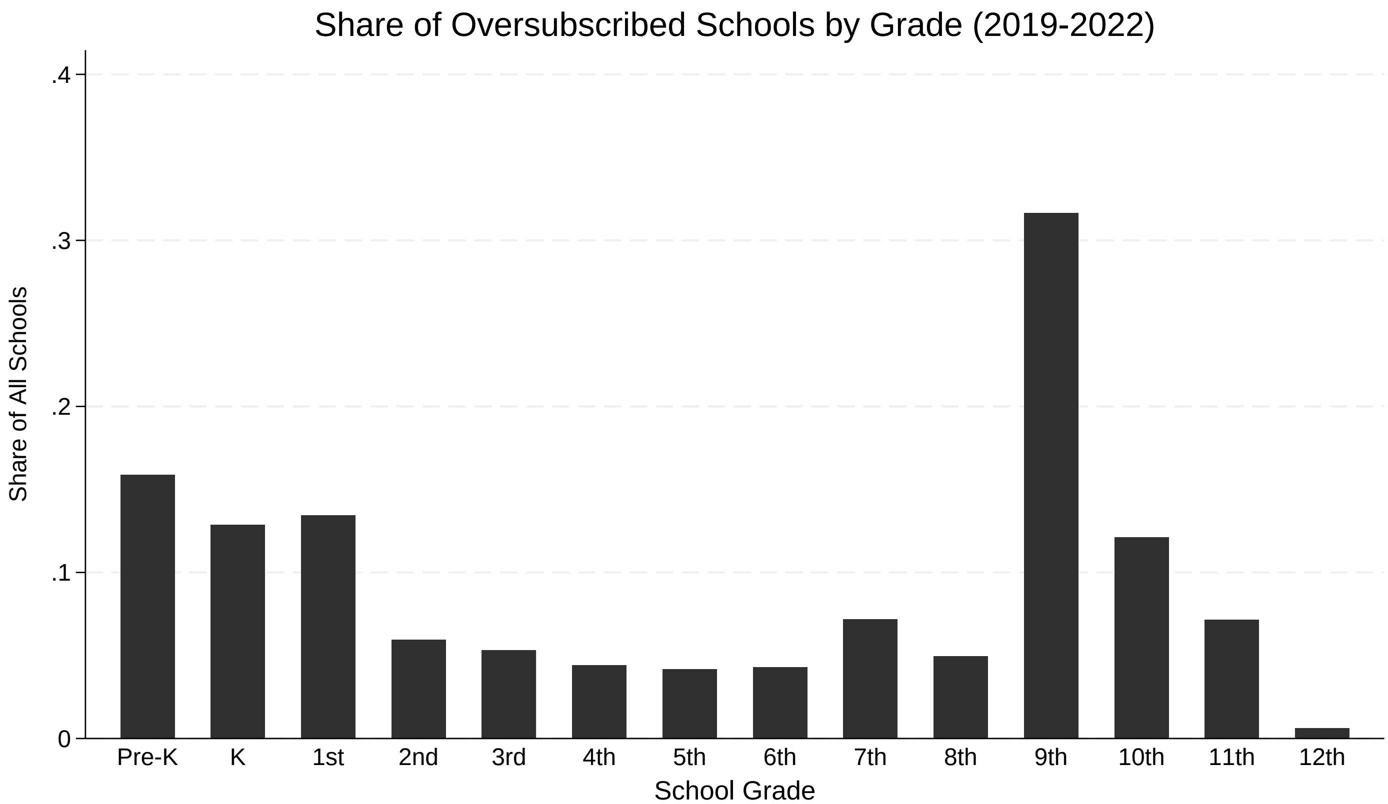
# 7 Figures and Tables

Figure 1: Implementation of the Centralized School Choice System



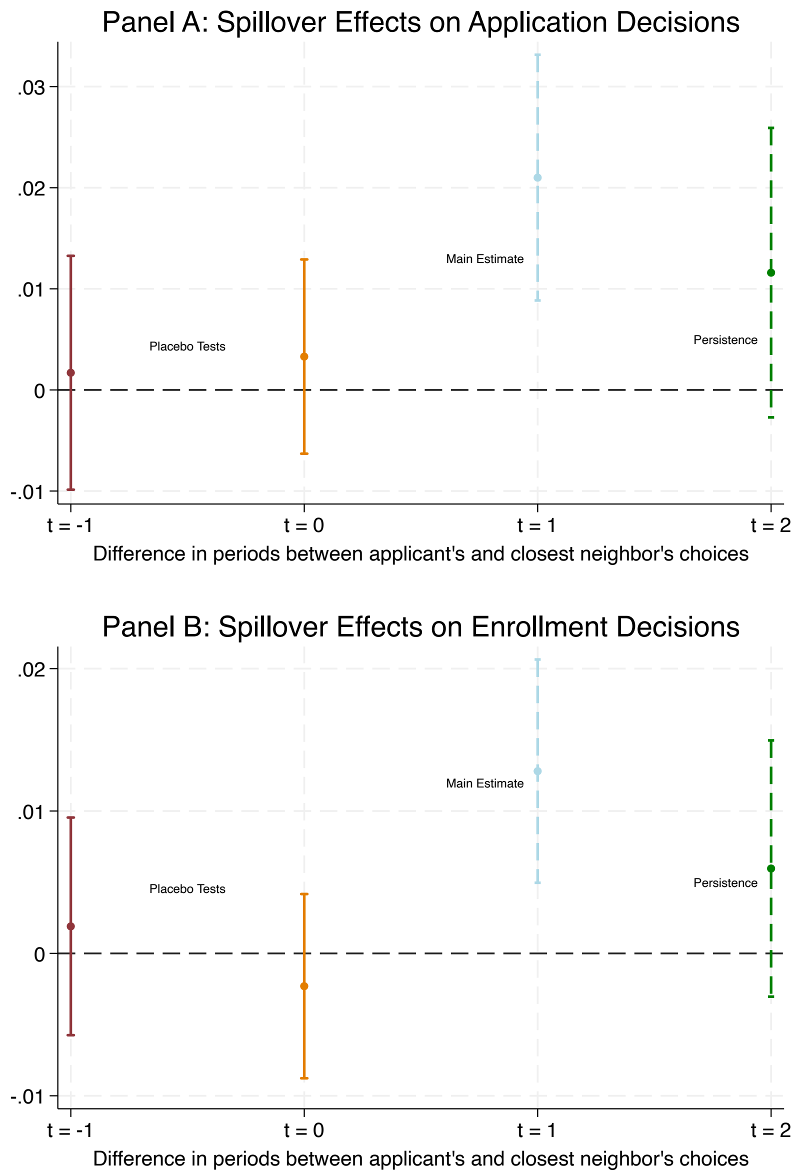

Notes: This plot shows the number of applicants observed in the centralized system between 2017 and 2022. Rollout was staggered across regions and grades. Starting in 2017, each year a new set of regions was incorporated to the system. By 2019, the centralized admission system is used for admission to ninth grade in all public and private voucher schools.

Figure 2: Oversubscribed Schools



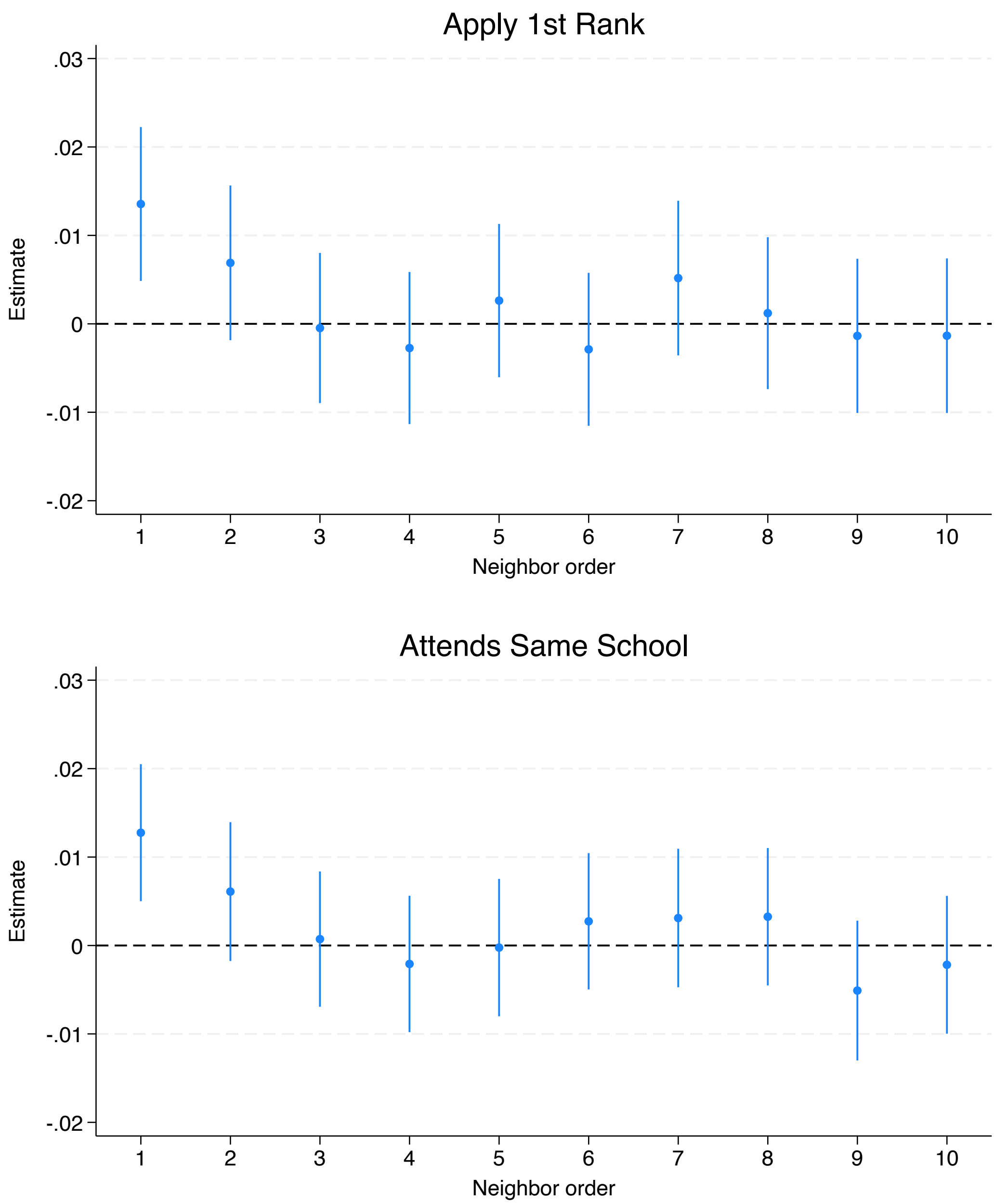

Notes: This plot shows the share of schools where the number of applicants submitting the school as first option surmounts the number of vacant seats in the corresponding grade. The share is computed pooling the 2019-22 application rounds.

Figure 3: Impacts of Neighbors From Different Time Horizons: Separate 2SLS Estimates




Notes: This figure shows how spillover effects vary depending on the number of periods used to link each applicant with their closest neighbor. Each plot reports separate 2SLS estimates for each sample and outcome using our baseline equations (2) and (3). Panel A uses as outcome an indicator equal to one if the applicant ranks the same school attended by the closest neighbor between in any preference while panel B uses as outcome an indicator equal to one if the applicant attends the same school. Neighbor's enrollment is instrumented with an indicator equal to one if the neighbor received an offer in their highest ranked school. We cannot reject the null hypothesis of equality of effects between $t = 1$ and $t = 2$ for both outcomes ($p$-value=0.254 for panel A and $p$-value=0.219 for panel B) while we reject the hypothesis of equal effects between $t = 1$ and the remaining periods for both outcomes. All models include a saturated set of propensity scores and standard errors are clustered at the neighbor level.

Figure 4: Heterogeneous Neighbor Spillovers: by Distance




Notes: This figure shows how spillover effects vary with the distance between each applicant and their ten nearest neighbors. We estimate equations (2) and (3) separately by each neighbor. Neighbor's enrollment is instrumented with an indicator equal to one if the neighbor received an offer in their most preferred school. Standard errors are clustered at the neighbor level.

Table 1: Summary Statistics

| | All 8th-Grade Students (1) | All 8th-Grade in K-8 Schools (2) | All Applicants (3) | Estimation Sample (4) |
|---|---|---|---|---|
| *Panel A: Background Characteristics* | | | | |
| Girl | 0.512 | 0.532 | 0.504 | 0.530 |
| Prioritario | 0.529 | 0.639 | 0.610 | 0.631 |
| Preferente | 0.294 | 0.250 | 0.265 | 0.266 |
| Public School | 0.382 | 0.599 | 0.545 | 0.600 |
| SIMCE (Math) | -0.037 | -0.266 | -0.210 | -0.256 |
| SIMCE (Language) | -0.022 | -0.194 | -0.152 | -0.185 |
| Missed SIMCE | 0.156 | 0.193 | 0.178 | 0.180 |
| Father Education: College | 0.278 | 0.159 | 0.187 | 0.162 |
| Father Education: Less than HS | 0.673 | 0.784 | 0.758 | 0.781 |
| Mother Education: College | 0.308 | 0.186 | 0.215 | 0.186 |
| Mother Education: Less than HS | 0.673 | 0.791 | 0.763 | 0.792 |
| College Expectations | 0.756 | 0.671 | 0.692 | 0.678 |
| Family Income >CLP800,000 | 0.185 | 0.088 | 0.111 | 0.088 |
| *Panel B: Application Characteristics* | | | | |
| Number of applications | | | 3.530 | 3.583 |
| Submits one school | | | 0.026 | 0.016 |
| Submits two schools | | | 0.280 | 0.257 |
| Submits three schools | | | 0.318 | 0.343 |
| Submits four schools or more | | | 0.376 | 0.384 |
| Observations | 840,755 | 367,045 | 410,412 | 128,085 |

Notes: This table presents average characteristics of the estimation sample relative to eighth-grade students who participate in the centralized system between 2019 and 2022 and all students enrolled in K-12 non-private schools. Column (1) shows average characteristics for all eighth-grade students; column (2) restricts the sample to students enrolled in K-8 schools. Column (3) displays average characteristics for all applicants with a valid (non-imputed) geographic location. Column (4) shows average values after restricting the sample to applicants whose closest neighbor's top-choice was an oversubscribed school (i.e., seat offer was determined by the tie-breaking rules).

Table 2: Balance Tests

| Variable | Average | | Difference | p-value | Observations | |
|---|---|---|---|---|---|---|
| | Offered | Non-offered | | | Offered | Non-offered |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Panel A: Applicant Covariates* | | | | | | |
| Girl | 0.474 | 0.467 | 0.006* | 0.080 | 58,225 | 69,860 |
| Prioritario | 0.654 | 0.657 | −0.003 | 0.387 | 58,225 | 69,860 |
| High Achiever | 0.301 | 0.301 | −0.000 | 0.945 | 58,225 | 69,860 |
| SIMCE (Math) | -0.263 | -0.248 | −0.015* | 0.056 | 46,716 | 55,538 |
| SIMCE (Language) | -0.184 | -0.183 | −0.001 | 0.939 | 46,457 | 55,340 |
| Father Education: College | 0.161 | 0.164 | −0.003 | 0.282 | 44,612 | 52,464 |
| Father Education: Less than HS | 0.782 | 0.780 | 0.003 | 0.445 | 44,612 | 52,464 |
| Mother Education: College | 0.186 | 0.186 | −0.000 | 0.937 | 44,889 | 52,844 |
| Mother Education: Less than HS | 0.792 | 0.791 | 0.001 | 0.762 | 44,889 | 52,844 |
| College Expectations | 0.682 | 0.675 | 0.006 | 0.112 | 44,570 | 52,441 |
| Family Income > CLP800,000 | 0.088 | 0.089 | −0.002 | 0.516 | 45,050 | 53,020 |
| Joint orthogonality F-test | | | | 0.349 | | |
| *Panel B: Neighbor Covariates* | | | | | | |
| Girl | 0.488 | 0.489 | −0.001 | 0.802 | 32,750 | 40,313 |
| Prioritario | 0.595 | 0.589 | 0.006 | 0.233 | 32,750 | 40,313 |
| High Achiever | 0.314 | 0.316 | −0.002 | 0.684 | 32,750 | 40,313 |
| SIMCE (Math) | -0.197 | -0.181 | −0.016 | 0.100 | 26,690 | 33,440 |
| SIMCE (Language) | -0.114 | -0.110 | −0.004 | 0.693 | 26,575 | 33,256 |
| Father Education: College | 0.175 | 0.181 | −0.006 | 0.114 | 25,088 | 31,074 |
| Father Education: Less than HS | 0.768 | 0.758 | 0.009** | 0.040 | 25088 | 31,074 |
| Mother Education: College | 0.203 | 0.210 | −0.007 | 0.127 | 25,266 | 31,264 |
| Mother Education: Less than HS | 0.773 | 0.766 | 0.007 | 0.125 | 25,266 | 31,264 |
| College Expectations | 0.700 | 0.700 | 0.000 | 0.956 | 25,100 | 31,019 |
| Family Income > CLP800,000 | 0.099 | 0.101 | −0.002 | 0.490 | 25,301 | 31,299 |
| Joint orthogonality F-test | | | | 0.501 | | |
| Distance between neighbors | 0.063 | 0.065 | −0.002 | 0.124 | 58,225 | 69,860 |

Notes: This table presents balance tests of observable characteristics between applicants and neighbors depending on neighbors' offer status. Each row shows the estimate of a regression of the corresponding covariate onto an indicator equals to one if the closest neighbor received an offer in her most preferred school, controlling for a saturated set of indicators for the assignment propensity score. Panel A displays estimates of applicants characteristics, while panel B shows estimates of neighbors characteristics. Joint orthogonality shows the $p$-value of a F-test of joint significance of all covariates listed in the corresponding panel. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 3: ITT and 2SLS Estimates of Neighbor Spillovers

| | First Stage | ITT | | | 2SLS | | |
|---|---|---|---|---|---|---|---|
| | In $t-1$, Neighbor: | In $t$, Applicant: | | | In $t$, Applicant: | | |
| In $t-1$, Neighbor: | Enrolled in 1st Choice (1) | Ranks School Any (2) | Ranks School 1st (3) | Attends School (4) | Ranks School Any (5) | Ranks School 1st (6) | Attends School (7) |
| Admitted to 1st Choice | 0.6736*** | 0.0141*** | 0.0091*** | 0.0086*** | | | |
| | (0.0036) | (0.0042) | (0.0030) | (0.0027) | | | |
| Enrolled in 1st Choice | | | | | 0.0210*** | 0.0136*** | 0.0128*** |
| | | | | | (0.0062) | (0.0044) | (0.0040) |
| Mean (Not Enrolled) | 0.1663*** | 0.3462*** | 0.1458*** | 0.1148*** | 0.3680*** | 0.1502*** | 0.1100*** |
| | (0.0022) | (0.0025) | (0.0018) | (0.0016) | (0.0041) | (0.0027) | (0.0025) |
| $F$-Statistic | | | | | 23,381 | 23,381 | 23,381 |
| N-Obs | 73,063 | 128,085 | 128,085 | 128,085 | 128,085 | 128,085 | 128,085 |
| N-Clusters | | 73,146 | 73,146 | 73,146 | 73,146 | 73,146 | 73,146 |

Notes: This table reports intent-to-treat (ITT) and two-stage least squares (2SLS) estimates of neighbors' spillovers on applicants' decisions. Column (1) presents the OLS estimate from a regression where the dependent variable is an indicator equal to one if the neighbor enrolled in ninth grade in the same school where they received an offer and the explanatory variable is the offer receipt. Columns (2)-(4) display OLS estimates of regressions where the variable of interest is an indicator equal to one if the closest neighbor received an offer in their most preferred school. Columns (5)-(7) report 2SLS coefficients instrumenting neighbors' enrollment with the offer. All models control for a saturated set of indicators for the assignment propensity score. Standard errors are clustered at the neighbor level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 4: Placebo Test: Effect of Neighbor's Next Year Choices

| | In $t$, Applicant: | | |
|---|---|---|---|
| | Ranks School Any | Ranks School 1st | Attends School |
| In $t+1$, Neighbor: | (1) | (2) | (3) |
| Admitted to 1st Choice | 0.001 | 0.004 | 0.001 |
| | (0.004) | (0.003) | (0.003) |
| Mean (Not enrolled) | 0.373*** | 0.156*** | 0.124*** |
| | (0.004) | (0.003) | (0.002) |
| N-Obs | 141,992 | 141,992 | 141,992 |
| N-Clusters | 82,269 | 82,269 | 82,269 |

Notes: This table presents a placebo test where we regress the outcome of an applicant in period $t$ onto an indicator equal to one if the closest neighbor applying in round $t+1$ receives a seat offer at their most preferred school. Clustered standard errors at the neighbor level are reported in parenthesis. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 5: Placebo Test: Effect of Neighbor's Contemporaneous Choices

| | In $t$, Applicant: | | | | |
|---|---|---|---|---|---|
| | Ranks School Any | Ranks School 1st | Accepts Same | Complementary Round | Attends School |
| In $t$, Neighbor: | (1) | (2) | (3) | (4) | (5) |
| Admitted to 1st Choice | 0.002 | $-0.000$ | $-0.000$ | $-0.002$ | $-0.002$ |
| | (0.003) | (0.002) | (0.002) | (0.002) | (0.002) |
| Mean (Not enrolled) | 0.373*** | 0.164*** | 0.079*** | 0.125*** | 0.120*** |
| | (0.003) | (0.002) | (0.002) | (0.002) | (0.002) |
| N-Obs | 174,508 | 174,508 | 174,508 | 174,508 | 174,508 |
| N-Clusters | 125,835 | 125,835 | 125,835 | 125,835 | 125,835 |

Notes: This table shows a placebo test where we regress the outcome of an applicant in period $t$ onto an indicator equal to one if the closest neighbor applying in round $t$ receives a seat offer at their most preferred school. Clustered standard errors at the neighbor level are reported in parenthesis. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 6: Heterogeneous Neighbor Spillovers: by Gender

| | In $t$, Applicant: | | |
|---|---|---|---|
| | Ranks School Any (1) | Ranks School 1st (2) | Attends School (3) |
| *Panel A: Both male* | | | |
| Enrolled in 1st Choice | 0.006 | 0.009 | 0.013* |
| | (0.011) | (0.008) | (0.007) |
| Mean (Not enrolled) | 0.399*** | 0.170*** | 0.122*** |
| | (0.007) | (0.005) | (0.005) |
| $F$-Statistic | 9,882 | 9,882 | 9,882 |
| N-Obs | 34,716 | 34,716 | 34,716 |
| *Panel B: Male applicant, female neighbor* | | | |
| Enrolled in 1st Choice | 0.036*** | 0.019** | 0.020*** |
| | (0.011) | (0.008) | (0.007) |
| Mean (Not enrolled) | 0.350*** | 0.138*** | 0.096*** |
| | (0.007) | (0.005) | (0.004) |
| $F$-Statistic | 8,589 | 8,589 | 8,589 |
| N-Obs | 33,083 | 33,083 | 33,083 |
| *Panel C: Female applicant, male neighbor* | | | |
| Enrolled in 1st Choice | 0.026** | 0.011 | 0.005 |
| | (0.011) | (0.008) | (0.007) |
| Mean (Not enrolled) | 0.340*** | 0.135*** | 0.104*** |
| | (0.007) | (0.005) | (0.005) |
| $F$-Statistic | 8,971 | 8,971 | 8,971 |
| N-Obs | 30,764 | 30,764 | 30,764 |
| *Panel D: Both female* | | | |
| Enrolled in 1st Choice | 0.020 | 0.014 | 0.010 |
| | (0.012) | (0.009) | (0.008) |
| Mean (Not enrolled) | 0.377*** | 0.157*** | 0.116*** |
| | (0.008) | (0.006) | (0.005) |
| $F$-Statistic | 8,204 | 8,204 | 8,204 |
| N-Obs | 29,418 | 29,418 | 29,418 |
| $p$-value for equal effects | 0.272 | 0.833 | 0.495 |

Notes: This table reports 2SLS estimates from equations (2) and (3) of the effects of applicants' nearest neighbor attending their most preferred school depending on applicants' and neighbors' gender. Enrollment is instrumented with an indicator equal to one if the nearest neighbor received a seat offer. All models control for a saturated set of indicators for the assignment propensity score. Clustered standard errors at the neighbor level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 7: Heterogeneous Neighbor Spillovers: by High Achiever Status

| | In $t$, Applicant: | | |
|---|---|---|---|
| | Ranks School Any | Ranks School 1st | Attends School |
| | (1) | (2) | (3) |
| *Panel A: None is high achiever* | | | |
| Enrolled in 1st Choice | 0.022*** | 0.016*** | 0.016*** |
| | (0.008) | (0.006) | (0.005) |
| Mean (Not enrolled) | 0.376*** | 0.153*** | 0.112*** |
| | (0.006) | (0.004) | (0.004) |
| $F$-Statistic | 14,231 | 14,231 | 14,231 |
| N-Obs | 61,381 | 61,381 | 61,381 |
| *Panel B: Only neighbor is high achiever* | | | |
| Enrolled in 1st Choice | 0.017 | 0.015 | 0.012 |
| | (0.013) | (0.009) | (0.008) |
| Mean (Not enrolled) | 0.347*** | 0.131*** | 0.093*** |
| | (0.008) | (0.005) | (0.005) |
| $F$-Statistic | 6,999 | 6,999 | 6,999 |
| N-Obs | 28,138 | 28,138 | 28,138 |
| *Panel C: Only applicant is high achiever* | | | |
| Enrolled in 1st Choice | 0.018 | 0.005 | 0.004 |
| | (0.012) | (0.009) | (0.008) |
| Mean (Not enrolled) | 0.374*** | 0.157*** | 0.121*** |
| | (0.008) | (0.006) | (0.005) |
| $F$-Statistic | 8,470 | 8,470 | 8,470 |
| N-Obs | 26,567 | 26,567 | 26,567 |
| *Panel D: Both are high achievers* | | | |
| Enrolled in 1st Choice | 0.024 | 0.009 | −0.006 |
| | (0.018) | (0.014) | (0.013) |
| Mean (Not enrolled) | 0.368*** | 0.160*** | 0.119*** |
| | (0.011) | (0.008) | (0.008) |
| $F$-Statistic | 4,154 | 4,154 | 4,154 |
| N-Obs | 11,891 | 11,891 | 11,891 |
| $p$-value for equal effects | 0.978 | 0.741 | 0.283 |

Notes: This table reports 2SLS estimates from equations (2) and (3) of the effects of applicants' nearest neighbor attending their most preferred school depending on applicants' and neighbors' eighth-grade GPA. High achiever is an indicator equals to one if the student is ranked in the top-performing 20% of the eighth-grade GPA distribution. Enrollment is instrumented with an indicator equals to one if the nearest neighbor received a seat offer. All models control for a saturated set of indicators for the assignment propensity score. Clustered standard errors at the neighbor level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 8: Heterogeneous Neighbor Spillovers: by Previous Math Scores

| | In $t$, Applicant: | | |
|---|---|---|---|
| | Ranks School Any (1) | Ranks School 1st (2) | Attends School (3) |
| *Panel A: Both below median score* | | | |
| Enrolled in 1st Choice | 0.015 | 0.005 | 0.009 |
| | (0.011) | (0.009) | (0.008) |
| Mean (Not enrolled) | 0.373*** | 0.161*** | 0.115*** |
| | (0.007) | (0.005) | (0.005) |
| $F$-Statistic | 8,897 | 8,897 | 8,897 |
| N-Obs | 32,047 | 32,047 | 32,047 |
| *Panel B: Only applicant below median score* | | | |
| Enrolled in 1st Choice | 0.028** | 0.021** | 0.021** |
| | (0.014) | (0.010) | (0.009) |
| Mean (Not enrolled) | 0.351*** | 0.139*** | 0.096*** |
| | (0.008) | (0.006) | (0.005) |
| $F$-Statistic | 6,590 | 6,590 | 6,590 |
| N-Obs | 22,018 | 22,018 | 22,018 |
| *Panel C: Only neighbor below median score* | | | |
| Enrolled in 1st Choice | 0.006 | 0.000 | 0.006 |
| | (0.014) | (0.011) | (0.010) |
| Mean (Not enrolled) | 0.389*** | 0.164*** | 0.124*** |
| | (0.009) | (0.007) | (0.006) |
| $F$-Statistic | 6,194 | 6,194 | 6,194 |
| N-Obs | 19,773 | 19,773 | 19,773 |
| *Panel D: Both above median score* | | | |
| Enrolled in 1st Choice | 0.005 | 0.008 | 0.010 |
| | (0.017) | (0.013) | (0.012) |
| Mean (Not enrolled) | 0.401*** | 0.162*** | 0.119*** |
| | (0.010) | (0.008) | (0.007) |
| $F$-Statistic | 4,955 | 4,955 | 4,955 |
| N-Obs | 14,835 | 14,835 | 14,835 |
| $p$-value for equal effects | 0.624 | 0.535 | 0.673 |

Notes: This table reports 2SLS estimates from equations (2) and (3) of the effects of applicants' nearest neighbor attending their most preferred school depending on applicants' and neighbors' previous math test scores. Enrollment is instrumented with an indicator equals to one if the closest neighbor received a seat offer. All models control for a saturated set of indicators for the assignment propensity score. Clustered standard errors at the neighbor level are reported in parenthesis. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 9: Heterogeneous Neighbor Spillovers: by School Characteristics

| | In $t$, Applicant: | | | | | |
|---|---|---|---|---|---|---|
| | Ranks School Any | | | Ranks School 1st | | |
| Neighbor's School Characteristics: | Tercile 1 (1) | Tercile 2 (2) | Tercile 3 (3) | Tercile 1 (4) | Tercile 2 (5) | Tercile 3 (6) |
| *Panel A: Distance to School* | | | | | | |
| Enrolled in 1st Choice | 0.092*** | 0.017* | −0.040*** | 0.056*** | 0.009 | −0.021*** |
| | (0.011) | (0.010) | (0.010) | (0.009) | (0.007) | (0.007) |
| $p$-value for equal effects | | 0.000 | | | 0.000 | |
| $F$-Statistic | 7,800 | 11,009 | 9,416 | 7,800 | 11,009 | 9,416 |
| N-Obs | 39,621 | 39,605 | 39,600 | 39,621 | 39,605 | 39,600 |
| *Panel B: Average 10th Grade Scores* | | | | | | |
| Enrolled in 1st Choice | 0.001 | 0.042*** | 0.062*** | 0.008 | 0.016** | 0.041*** |
| | (0.013) | (0.010) | (0.013) | (0.009) | (0.007) | (0.009) |
| $p$-value for equal effects | | 0.002 | | | 0.024 | |
| $F$-Statistic | 5,063 | 8,878 | 6,649 | 5,063 | 8,878 | 6,649 |
| N-Obs | 48,189 | 53,822 | 21,781 | 48,189 | 53,822 | 21,781 |
| *Panel C: School Value-Added* | | | | | | |
| Enrolled in 1st Choice | 0.018 | 0.062*** | 0.026** | 0.013 | 0.030*** | 0.026*** |
| | (0.012) | (0.010) | (0.012) | (0.009) | (0.007) | (0.009) |
| $p$-value for equal effects | | 0.009 | | | 0.293 | |
| $F$-Statistic | 5,963 | 9,075 | 6,869 | 5,963 | 9,075 | 6,869 |
| N-Obs | 54,726 | 47,457 | 21,545 | 54,726 | 47,457 | 21,545 |
| *Panel D: School Climate Index* | | | | | | |
| Enrolled in 1st Choice | −0.020* | 0.029** | 0.048*** | −0.020*** | 0.018** | 0.036*** |
| | (0.011) | (0.012) | (0.011) | (0.008) | (0.009) | (0.008) |
| $p$-value for equal effects | | 0.000 | | | 0.000 | |
| $F$-Statistic | 6,856 | 6,762 | 8,256 | 6,856 | 6,762 | 8,256 |
| N-Obs | 47,823 | 41,086 | 34,893 | 47,823 | 41,086 | 34,893 |

Notes: This table presents 2SLS estimates from equations (2) and (3) to investigate how spillovers depend on the school characteristics where neighbors enroll. Distance (measured in miles) corresponds to the euclidean distance between the applicant's residence and the neighbor's school. Average tenth-grade scores uses math and language scores in 2017 and 2018. School value-added on tenth-grade scores is constructed using data from the 2015-2018 cohorts of tenth-grade students (see the main text for details). The school climate index is created by the Ministry of Education and refers to students', teachers', and parents' perceptions about the school environment. Average tenth-grade scores and the school climate index are standardized to be mean zero and unit variance using all public and private schools with positive ninth-grade enrollment. All models include a saturated set of indicators for the assignment propensity score. Standard errors are clustered at the neighbor level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 10: Estimates of Effects on Applicants' Attended Schools

| | First Stage | Top-Ranked School | | | Attended School | | |
|---|---|---|---|---|---|---|---|
| | | Mean (Compliers) | ITT | 2SLS | Mean (Compliers) | ITT | 2SLS |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| 10th Grade Scores | 0.845*** | −0.145*** | 0.106*** | 0.125*** | −0.496*** | 0.114*** | 0.134*** |
| | (0.003) | (0.006) | (0.004) | (0.005) | (0.005) | (0.004) | (0.004) |
| $F$-Statistic | | | | 62,966 | | | 61,733 |
| N-Obs | 62,035 | | 110,367 | 107,879 | | 107,715 | 105,303 |
| Value-Added: College Enrollment | 0.883*** | −0.037*** | 0.317*** | 0.359*** | −0.061*** | 0.309*** | 0.350*** |
| | (0.002) | (0.001) | (0.004) | (0.005) | (0.001) | (0.004) | (0.004) |
| $F$-Statistic | | | | 99,923 | | | 99,008 |
| N-Obs | 62,684 | | 112,115 | 109,893 | | 109,319 | 107,178 |
| Value-Added: HS Graduation | 0.813*** | 0.003*** | 0.194*** | 0.238*** | −0.015*** | 0.223*** | 0.273*** |
| | (0.004) | (0.000) | (0.004) | (0.005) | (0.000) | (0.004) | (0.005) |
| $F$-Statistic | | | | 37,371 | | | 36,941 |
| N-Obs | 62,684 | | 112,115 | 109,893 | | 109,319 | 107,178 |
| School Climate Index | 0.832*** | 0.044*** | 0.212*** | 0.254*** | −0.140*** | 0.230*** | 0.276*** |
| | (0.003) | (0.007) | (0.004) | (0.005) | (0.007) | (0.004) | (0.005) |
| $F$-Statistic | | | | 51,580 | | | 51,443 |
| N-Obs | 61,695 | | 109,607 | 106,968 | | 106,889 | 104,338 |

Notes: This table reports 2SLS estimates from equations (7)-(8) of neighbors' school characteristics on applicants' top-ranked and enrolled school characteristics. Columns (2)-(4) display estimates of applicants' top-ranked school characteristics, and columns (5)-(7) show estimates of applicants' attended school characteristics. For each outcome, column (1) shows the first stage (coefficient $\rho$ in equation (8)). Columns (2) and (5) show mean outcomes for compliers computed following Abadie (2002). Columns (3) and (6) show coefficients from regressions of outcomes on the characteristics of the school where the neighbor received an offer. Columns (4) and (7) report 2SLS coefficients instrumenting the neighbor's attendance with the offer. All models include a saturated set of indicators for the assignment propensity score and cluster standard errors at the neighbor level. See the main text for details about the construction of each outcome. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

# References


A. Abadie. Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models. *Journal of the American Statistical Association*, 97:284–292, 2002.

A. Abdulkadiroğlu, J. Angrist, S. Dynarski, T. J. Kane, and P. Pathak. Accountability and Flexibility in Public Schools: Evidence from Boston's Charters And Pilots. *The Quarterly Journal of Economics*, 126(2):699–748, 2011.

A. Abdulkadiroğlu, J. D. Angrist, Y. Narita, and P. A. Pathak. Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation. *Econometrica*, 85:1373–1432, 2017.

A. Abdulkadiroğlu, P. A. Pathak, and C. Walters. Free to Choose: Can School Choice Reduce Student Achievement? *American Economic Journal: Applied Economics*, 10(1):175–206, 2018.

A. Abdulkadiroğlu, P. A. Pathak, J. Schellenberg, and C. Walters. Do Parents Value School Effectiveness? *American Economic Review*, 110(5):1502–39, 2020.

S. Agarwal, W. Qian, and X. Zou. Thy Neighbor's Misfortune: Peer Effect on Consumption. *American Economic Journal: Economic Policy*, 13(2):1–25, 2021.

J. Aguirre and J. Matta. Walking in Your Footsteps: Sibling Spillovers in Higher Education Choices. *Economics of Education Review*, 80:1020–62, 2021.

R. Ainsworth, R. Dehejia, C. Pop-Eleches, and M. Urquiola. Why Do Households Leave School Value Added on the Table? The Roles of Information and Preferences. *American Economic Review*, 113(4):1049–82, 2023.

A. Altmejd, A. Barrios-Fernández, M. Drlje, J. Goodman, M. Hurwitz, D. Kovac, C. Mulhern, C. Neilson, and J. Smith. O Brother, Where Start Thou? Sibling Spillovers on College and Major Choice in Four Countries. *The Quarterly Journal of Economics*, 136(3):1831–86, 2021.

T. Andrabi, J. Das, and A. I. Khwaja. Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets. *American Economic Review*, 107(6):1535–63, 2017.

M. Angelucci and G. De Giorgi. Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption? *American Economic Review*, 99(1):486–508, 2009.

J. Angrist, P. Hull, P. Pathak, and C. Walters. Credible School Value-Added with Undersubscribed School Lotteries. *The Review of Economics and Statistics*, 106(1):1–19, 2024.

F. Arteaga, A. J. Kapor, C. Neilson, and S. Zimmerman. Smart Matching Platforms and Heterogeneous Beliefs in Centralized School Choice. *The Quarterly Journal of Economics*, 137(3): 1791–1848, 2022.

A. Bacher-Hicks, S. B. Billings, and D. J. Deming. The School to Prison Pipeline: Long-Run Impacts of School Suspensions on Adult Crime. Working Paper 26257, National Bureau of Economic Research, 2019.

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

P. Bayer, S. L. Ross, and G. Topa. Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. *Journal of Political Economy*, 116(6):1150–1196, 2008.

D. W. Beuermann, C. K. Jackson, L. Navarro-Sola, and F. Pardo. What is a Good School, and Can Parents Tell? Evidence on the Multidimensionality of School Output. *The Review of Economic Studies*, 90(1):65–101, 2022.

S. B. Billings, D. J. Deming, and S. L. Ross. Partners in Crime. *American Economic Journal: Applied Economics*, 11(1):126–50, 2019.

G. J. Bobonis and F. Finan. Neighborhood Peer Effects in Secondary School Enrollment Decisions. *The Review of Economics and Statistics*, 91(4):695–716, 2009.

S. Burgess, E. Greaves, A. Vignoles, and D. Wilson. What Parents Want: School Preferences and School Choice. *Economic Journal*, 125(587):1262–1289, 2015.

K. Chay, P. McEwan, and M. Urquiola. The Central Role of Noise in Evaluating Interventions That Use Test Scores to Rank Schools. *American Economic Review*, 95(4):1237–1258, 2005.

R. Chetty, J. N. Friedman, and J. E. Rockoff. Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review*, 104(9):2593–2632, 2014.

S. Corcoran, J. L. Jennings, S. Cohodes, and C. Sattin-Bajaj. Leveling the Playing Field for High School Choice: Results from a Field Experiment of Informational Interventions. NBER Working Papers 24471, National Bureau of Economic Research, 2018.

J. Correa, R. Epstein, J. Escobar, I. Ríos, B. Bahamondes, C. Bonet, N. Epstein, N. Aramayo, M. Castillo, A. Cristi, B. Epstein, and F. Subiabre. School Choice in Chile. *Operations Research*, 70(2):1066–1087, 2022.

J. B. Cullen, B. A. Jacob, and S. Levitt. The Effect of School Choice on Participants: Evidence from Randomized Lotteries. *Econometrica*, 74(5):1191–1230, 2006.

G. B. Dahl, D.-O. Rooth, and A. Stenberg. Intergenerational and Sibling Spillovers in High School Majors. NBER Working Paper 27618, 2023.

D. J. Deming. Better Schools, Less Crime? *Quarterly Journal of Economics*, 126(4):2063–2115, 2011.

D. J. Deming, J. S. Hastings, T. J. Kane, and D. O. Staiger. School Choice, School Quality, and Postsecondary Attainment. *American Economic Review*, 104(3):991–1013, 2014.

W. Dobbie and J. Fryer, Roland G. Are High-Quality Schools Enough to Increase Achievement among the Poor? Evidence from the Harlem Children's Zone. *American Economic Journal: Applied Economics*, 3(3):158–87, 2011.

A. Dustan. Family Networks and School Choice. *Journal of Development Economics*, 134(C): 372–391, 2018.

A. Dustan, A. de Janvry, and E. Sadoulet. Flourish or Fail?: The Risky Reward of Elite High School Admission in Mexico City. *Journal of Human Resources*, 52(3):756–799, 2017.

B. Feigenberg, S. Rivkin, and R. Yan. Illusory Gains from Chile's Targeted School Voucher Experiment. *The Economic Journal*, 129(10):2805–2832, 2017.

D. Gale and L. S. Shapley. College Admissions and the Stability of Marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

J. Goodman, M. Hurwitz, J. Smith, and J. Fox. The Relationship Between Siblings' College Choices: Evidence From One Million SAT-Taking Families. *Economics of Education Review*, 48 (C):75–85, 2015.

G. Gray-Lobe, P. A. Pathak, and C. Walters. The Long-Term Effects of Universal Preschool in Boston. *The Quarterly Journal of Economics*, 138(1):363–411, 2023.

M. Grinblatt, M. Keloharju, and S. Ikäheimo. Social Influence and Consumption: Evidence from the Automobile Purchases of Neighbors. *The Review of Economics and Statistics*, 90(4):735–753, 2008.

O. Gurantz, M. Hurwitz, and J. Smith. Sibling Effects on High School Exam Taking and Performance. *Journal of Economic Behavior & Organization*, 178(C):534–549, 2020.

J. Hastings, T. J. Kane, and D. Staiger. Parental Preferences and School Competition: Evidence from a Public School Choice Program. Working Paper 11805, National Bureau of Economic Research, 2005.

J. S. Hastings and J. M. Weinstein. Information, School Choice, and Academic Achievement: Evidence from Two Experiments. *The Quarterly Journal of Economics*, 123(4):1373–1414, 2008.

J. S. Hastings, T. J. Kane, and D. O. Staiger. Gender and Performance: Evidence from School Assignment by Randomized Lottery. *American Economic Review*, 96(2):232–236, 2006.

J. K. Hellerstein, M. McInerney, and D. Neumark. Neighbors and Coworkers: The Importance of Residential Labor Market Networks. *Journal of Labor Economics*, 29(4):659–695, 2011.

N. Honey and A. Carrasco. A New Admission System in Chile and Its Foreseen Moderate Impact on Access for Low-Income Students. *Educational Evaluation and Policy Analysis*, 45(1):108–133, 2023.

G. Imbens and J. Angrist. Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62(2):467–75, 1994.

R. Jensen. The (Perceived) Returns to Education and the Demand for Schooling. *The Quarterly Journal of Economics*, 125(2):515–548, 2010.

J. S. Joensen and H. S. Nielsen. Spillovers in Education Choice. *Journal of Public Economics*, 157 (C):158–183, 2018.

T. J. Kane and D. O. Staiger. Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation. Working Paper 14607, National Bureau of Economic Research, 2008.

A. J. Kapor, C. Neilson, and S. Zimmerman. Heterogeneous Beliefs and School Choice Mechanisms. *American Economic Review*, 110(5):1274–1315, 2020.

P. Kuhn, P. Kooreman, A. Soetevent, and A. Kapteyn. The Effects of Lottery Prizes on Winners and Their Neighbors: Evidence from the Dutch Postcode Lottery. *American Economic Review*, 101(5):2226–47, 2011.

M. Kutscher, S. Nath, and S. Urzúa. Centralized Admission Systems and School Segregation: Evidence from a National Reform. *Journal of Public Economics*, 221:104863, 2023.

R. Lalive and M. A. Cattaneo. Social Interactions and Schooling Decisions. *The Review of Economics and Statistics*, 91(3):457–477, 2009.

D. S. Lee, J. McCrary, M. J. Moreira, and J. Porter. Valid $t$-Ratio Inference for IV. *American Economic Review*, 112(10):3260–90, 2022.

J. A. Lincove, J. Valant, and J. M. Cowen. You Can't Always Get What You Want: Capacity Constraints in a Choice-Based School System. *Economics of Education Review*, 67(C):94–109, 2018.

E. F. P. Luttmer. Neighbors as Negatives: Relative Earnings and Well-Being. *The Quarterly Journal of Economics*, 120(3):963–1002, 2005.

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

A. Mizala and F. Torche. Means-Tested School Vouchers and Educational Achievement: Evidence from Chile's Universal Voucher System. *The ANNALS of the American Academy of Political and Social Science*, 674(1):163–183, 2017.

A. Mizala and M. Urquiola. School Markets: The Impact of Information Approximating Schools' Effectiveness. *Journal of Development Economics*, 103(C):313–335, 2013.

C. Neilson. Targeted Vouchers, Competition Among Schools, and the Academic Achievement of Poor Students. Working Paper, 2023.

C. Neilson, C. Allende, and F. Gallego. Approximating the Equilibrium Effects of Informed School Choice. Working Paper, Princeton University, Department of Economics, Industrial Relations Section, 2019.

C. Nicoletti and B. Rabe. Sibling Spillover Effects in School Achievement. *Journal of Applied Econometrics*, 34(4):482–501, 2019.

J. A. Qureshi. Additional Returns to Investing in Girls' Education: Impact on Younger Sibling Human Capital. *Economic Journal*, 128(616):3285–3319, 2018.

I. M. Schmutte. Job Referral Networks and the Determination of Earnings in Local Labor Markets. *Journal of Labor Economics*, 33(1):1–32, 2015.

C. Walters. The Demand for Effective Charter Schools. *Journal of Political Economy*, 126(6):2179 – 2223, 2018.

# A Online Appendix

Neighbors' Spillovers on High School Choice

Juan Matta Alexis Orellana

## A.1 Additional Figures

Figure A.1: Number of Applicants by Grade

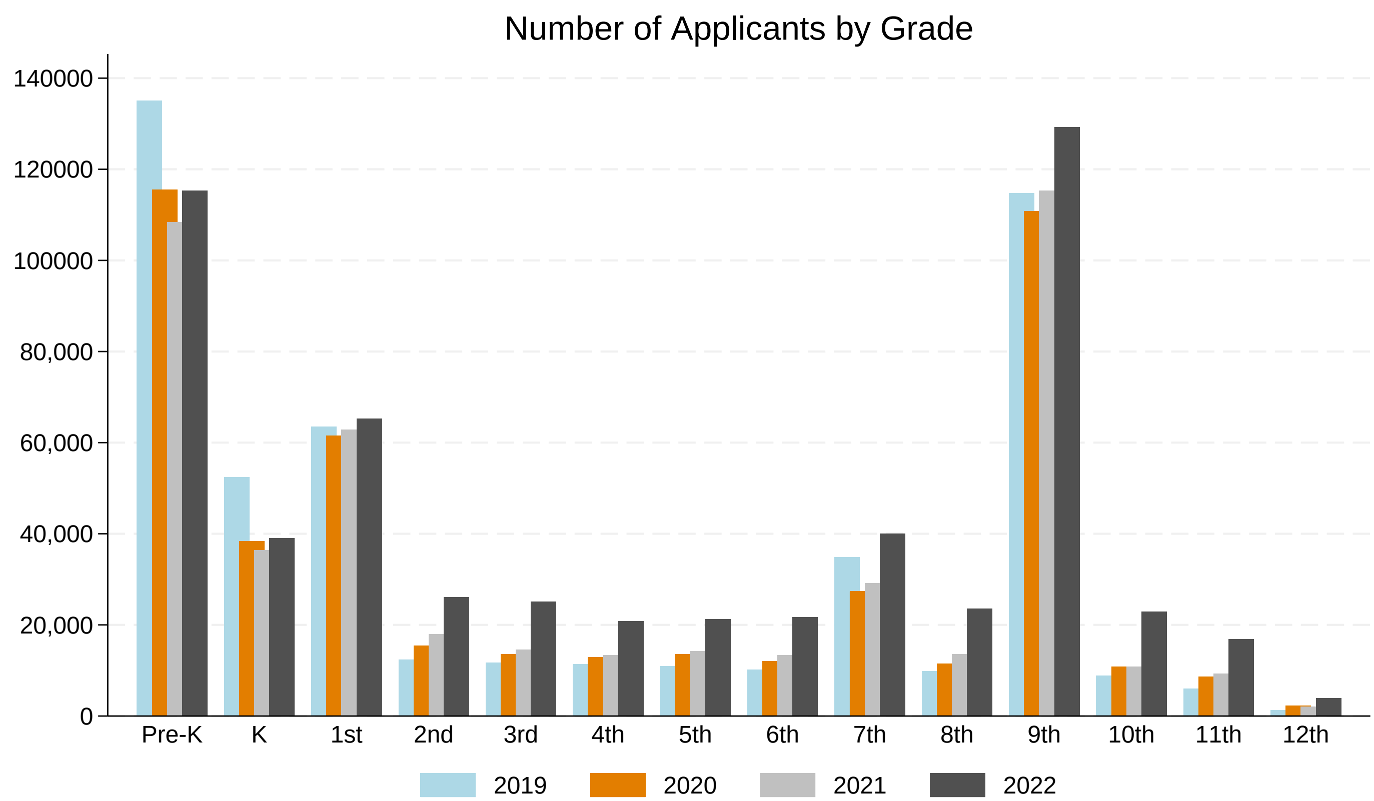


Figure A.2: Number of Participating Schools by Grade

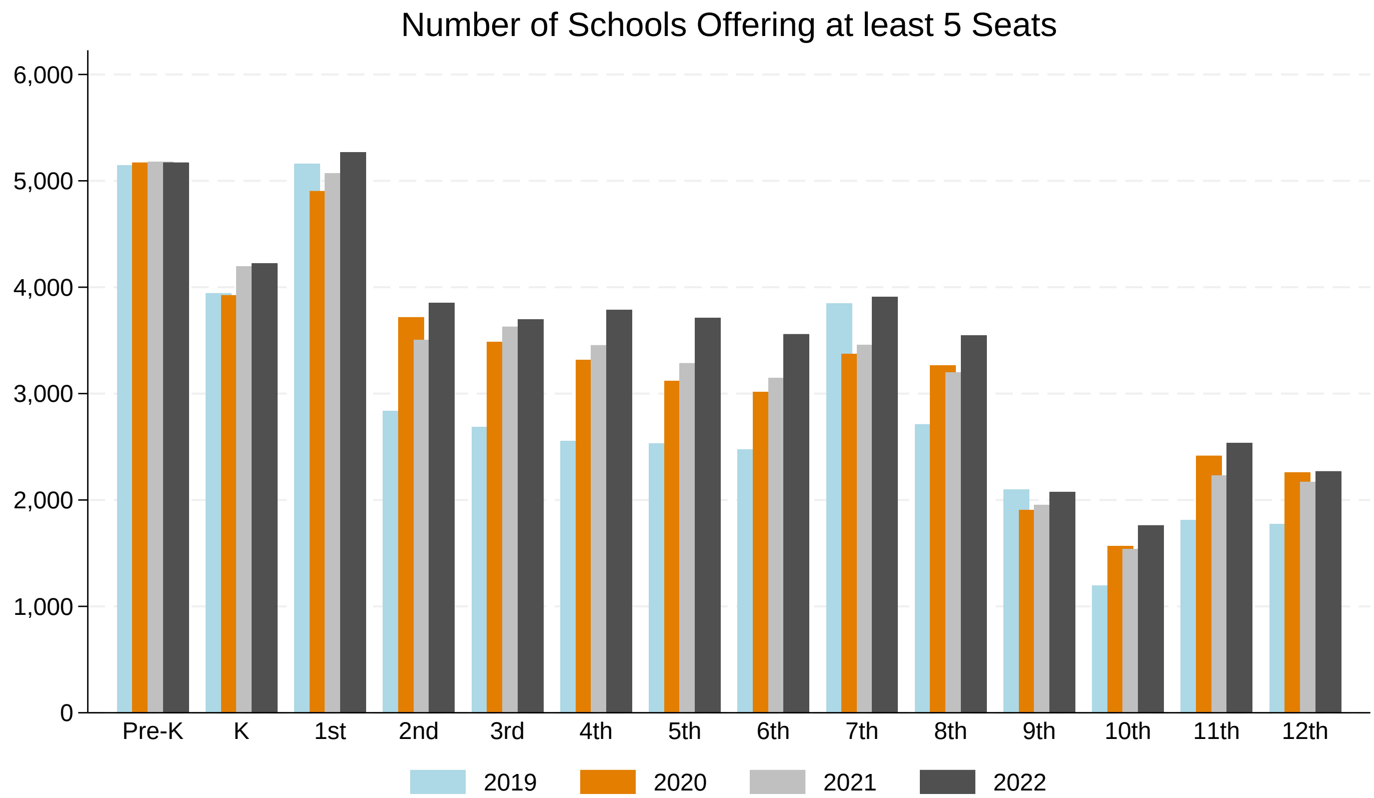
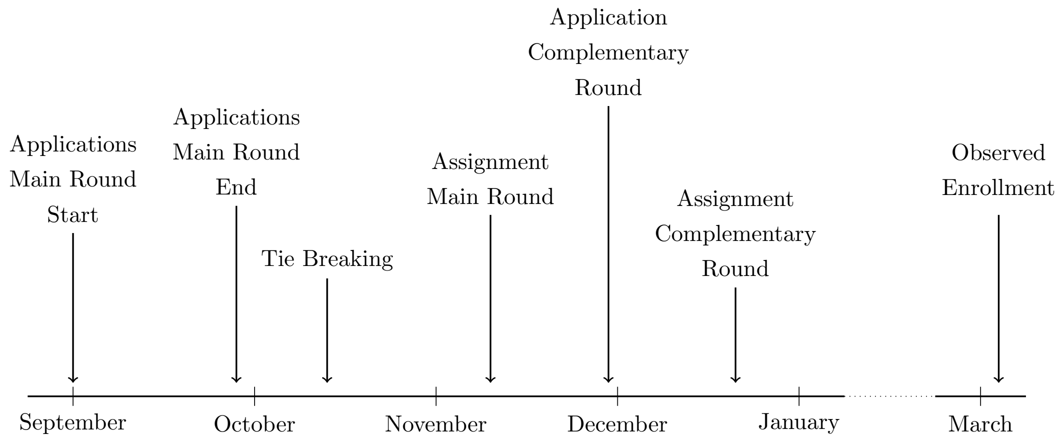


Figure A.3: Timeline of the Application Process




Source: Correa et al. (2022)

Figure A.4: Priority Groups

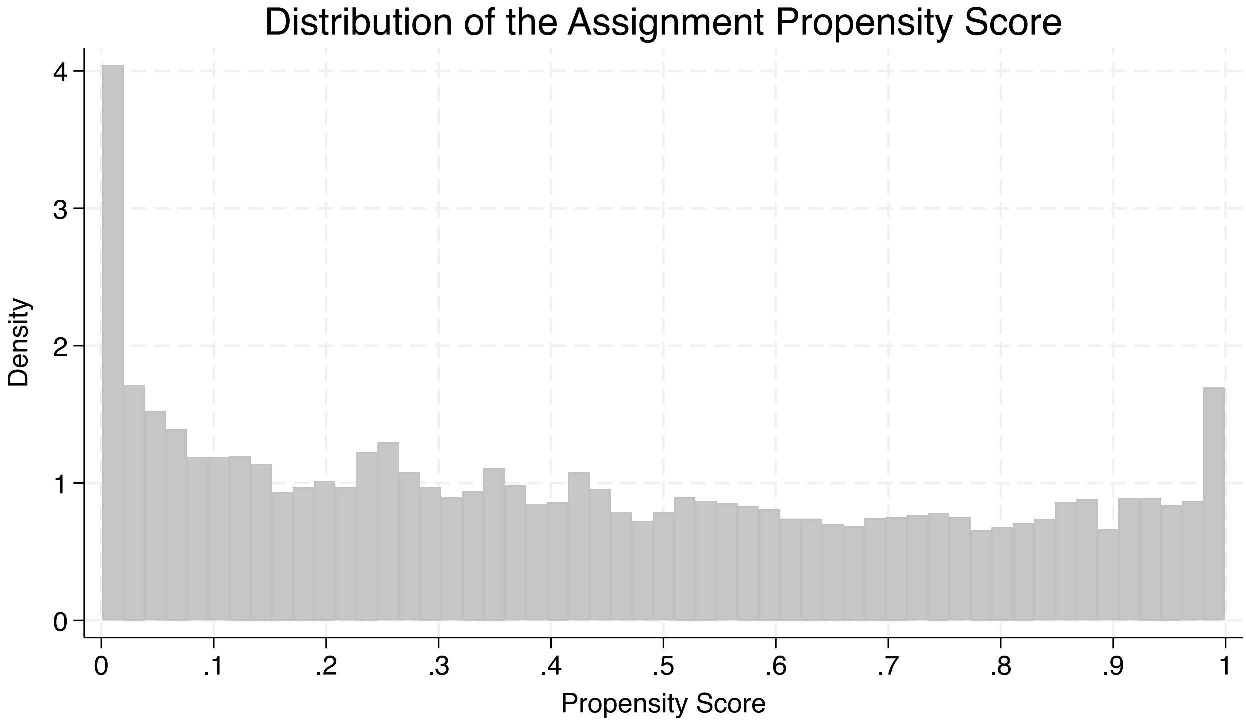
**Table 1.** Weak Priorities by Type-Specific Seats

| Priority | Special needs | Academic excellence | Disadvantaged | No trait |
|---|---|---|---|---|
| 1 | Current school | Current school | Current school | Current school |
| 2 | Special needs | Academic excellence | Siblings | Siblings |
| 3 | Siblings | Siblings | Disadvantaged | Working parent |
| 4 | Working parent | Working parent | Working parent | Returning students |
| 5 | Returning students | Returning students | Returning students | No priority |
| 6 | No priority | No priority | No priority | |

*Note.* Lower numbers indicate higher priority.

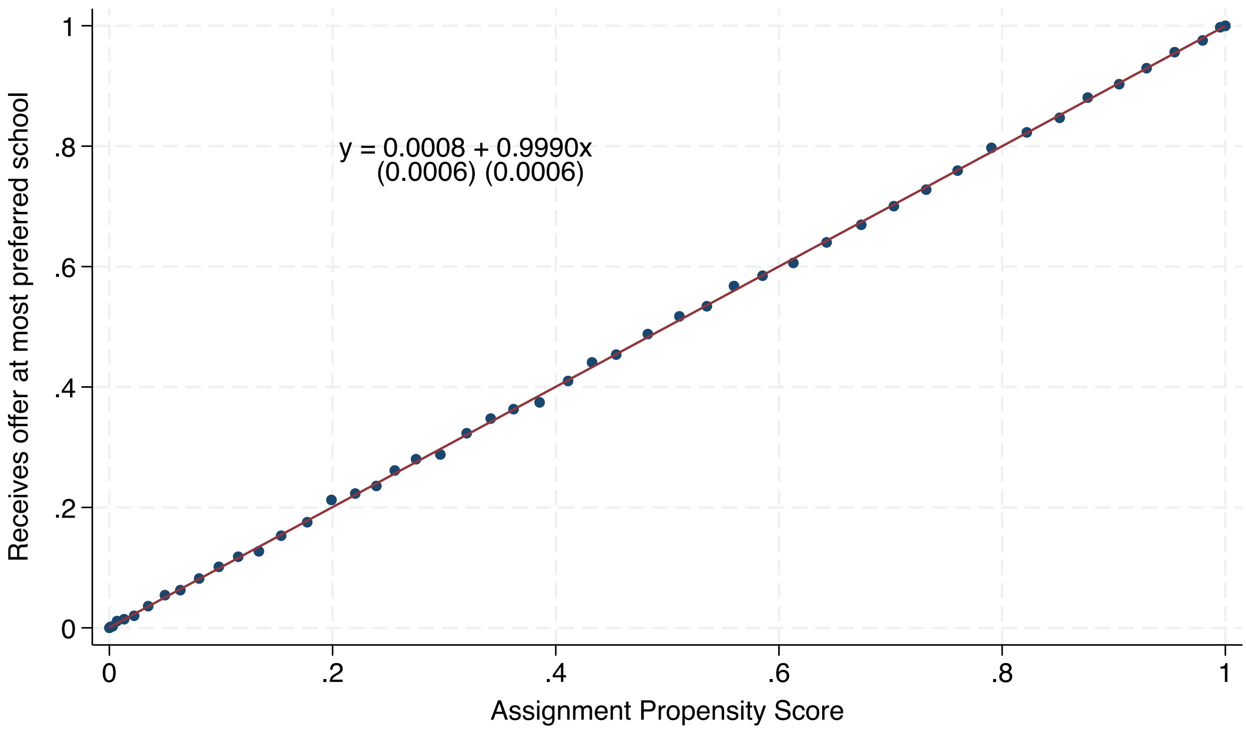
Source: Correa et al. (2022)

Figure A.5: Propensity Score Values for Ninth-Grade Applicants



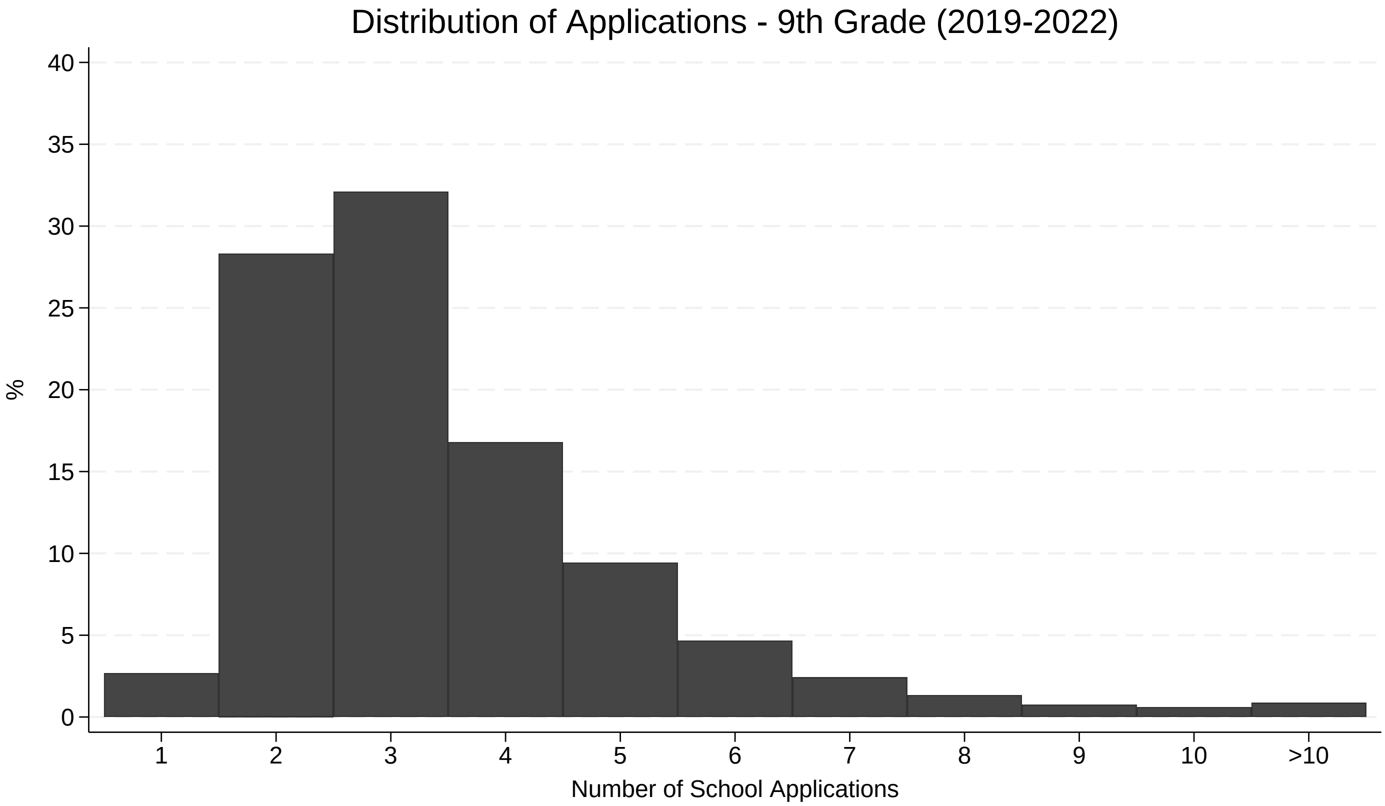

Notes: This plot shows the distribution of assignment propensity scores for all randomized (excluding strict zero or one values) ninth-grade applicants between 2019 and 2022. We simulate the assignment system 1,000 times for each applicant changing the random tie-breaking number in each iteration and keeping fixed preferences and priorities. We record the school allocation in each iteration and compute the propensity score as the fraction of times an applicant receives an offer in their highest ranked school.

Figure A.6: Simulated Probabilities and Observed School Offers



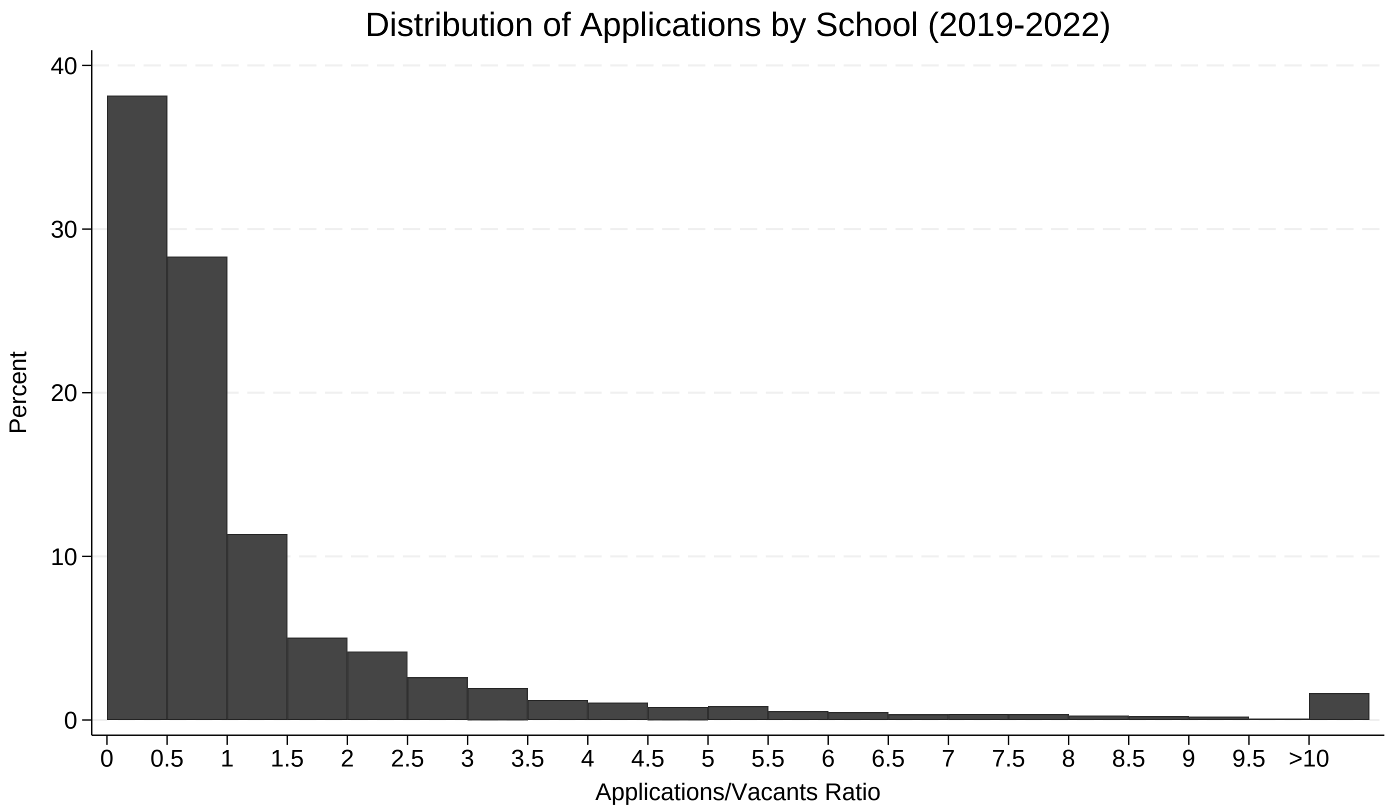

Notes: This plot shows the relationship between the simulated probability of receiving an offer in the most preferred school and the observed offer receipt. We simulate the assignment system 1,000 times using application data from the 2019-22 rounds and compute the average probability of receiving an offer for each ninth-grade applicant (assignment propensity score). The estimates displayed show the fit of a bivariate regression of an offer indicator onto the simulated probability.

Figure A.7: Distribution of School Applications - 9th Grade




Notes: This plot shows the distribution of the number of schools submitted by ninth-grade applicants. This plot pools the 2019-22 application rounds.

Figure A.8: Applicants/Seats Ratio Across Schools




Notes: This plot shows the distribution of the applicants/seats ratio for schools offering ninth grade, restricted to schools offering at least five vacant seats. The number of applicants considers only first-rank preferences.

## A.2 School Value-Added

We use information from tenth-grade cohorts between 2015 and 2018 to construct a proxy of school effectiveness for high school graduation and college attendance. For the 2015, 2016, and 2017 cohorts we observe test scores in eight grade, while for the 2018 cohort we observe the same variables in fourth grade. Based on this information, we estimate school value-added models of the form:

$$y_{ist} = X_{ist}'\beta + \theta_s + \theta_t + \xi_{ist} \tag{9}$$

Our outcomes $y_{ist}$ are tenth-grade standardized test scores, indicators equal to one when student $i$ in cohort $t$ graduated on time from high school $s$ and attended college the next year, respectively. The vector $X_{ist}$ includes a third-order polynomial in math and language lagged test scores, GPA in eighth and seventh grades, and indicators for gender, and repetition in eighth and seventh grades. $\theta_t$ corresponds to cohort fixed effects. We estimate equation (9) and recover the raw school fixed effects $\hat{\theta}_s$.

As it is common practice in the teacher and school value-added literature (Kane and Staiger, 2008; Chetty et al., 2014; Bacher-Hicks et al., 2019), we generate empirical Bayes (EB) shrunken estimates of $\hat{\theta}_s$ to account for sampling error and minimize mean square prediction errors. Following Abdulkadiroğlu et al. (2020), we assume that the distribution of the true school-specific parameters $\theta_s$ is given by the following hierarchical Bayesian model:

$$\hat{\theta}_s|\theta_s \sim N(\theta_s, \Omega_s) \tag{10}$$

$$\theta_s \sim N(\mu_\theta, \Sigma_\theta) \tag{11}$$

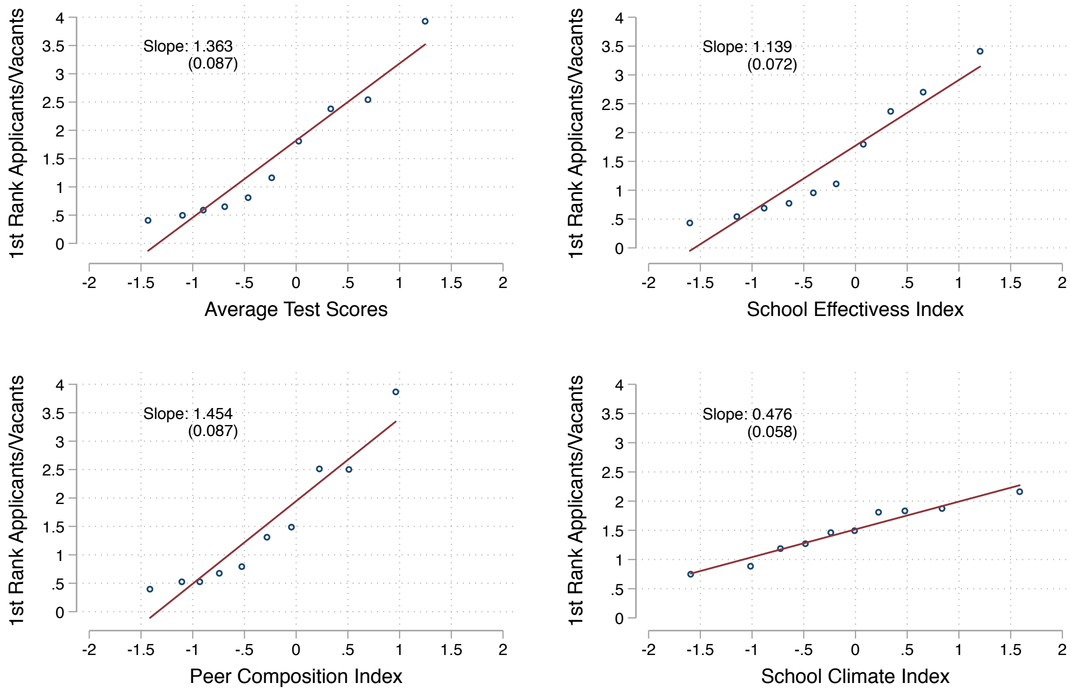
Where $\Omega_s$ is the sampling variance of the estimator $\hat{\theta}_s$, while $\mu_\theta$ and $\Sigma_\theta$ are the mean and variance of the distribution of the underlying parameters $\theta_s$. We compute the posterior mean for each school, $\hat{\theta}_s^{EB}$ as the weighted average of the OLS estimate and the prior mean, where the weight corresponds to the signal-to-noise ratio:

$$\hat{\theta}_s^{EB} = \frac{\Omega_s^{-1}}{\Omega_s^{-1} + \Sigma_\theta^{-1}}\hat{\theta}_s + \frac{\Sigma_\theta^{-1}}{\Omega_s^{-1} + \Sigma_\theta^{-1}}\mu_\theta \tag{12}$$

In practice, we construct the sample estimates of the hyperparameters $\mu_\theta$ and $\Sigma_\theta$ using the distribution of estimated fixed effects $\{\hat{\theta}_s\}_{s=1}^{S}$, while we employ the standard error of $\hat{\theta}_s$ to estimate $\Omega_s$. We plug $\hat{\Omega}_s, \hat{\mu}_\theta, \hat{\Sigma}_\theta$ into (12) and use the EB posterior means as regressors in our analysis of heterogeneity effects by school characteristics in section 4.3.3.

## A.3 School Attributes and Demand

Figure A.9: Correlation between School Attributes and Demand




Notes: This plot shows the correlation between each school index described in section 4.3.3 and the ratio of first-rank applicants to the number of seats available in ninth grade in each school. We exclude schools with less than five vacant seats. Average tenth grade scores correspond to the school-level math and language scores observed in each school in 2018 (last available year). The school effectiveness index uses the average math and language school-level scores in tenth grade, school value-added on high-school graduation and college attendance. The peer composition index uses cohort-level average math and language scores, the proportion of students with college-educated mothers, and the proportion of students whose parents expect them to attend college. The school climate index is reported by the Ministry of Education and refers to students', teachers', and parents' perceptions about the school environment. The effectiveness and peer composition indexes are constructed using a principal component model and the Bartlett method. We standardize each outcome to be mean zero and unit variance using all public and private schools with positive ninth-grade enrollment.

## A.4 Additional Tables

Table A.1: Summary of Acceptances by School Grade

| | 2019 | | 2020 | | 2021 | | 2022 | |
|---|---|---|---|---|---|---|---|---|
| | Accepted in 1st-3rd options (1) | Accepted in 1st option (2) | Accepted in 1st-3rd options (3) | Accepted in 1st option (4) | Accepted in 1st-3rd options (5) | Accepted in 1st option (6) | Accepted in 1st-3rd options (7) | Accepted in 1st option (8) |
| *School Level* | | | | | | | | |
| Pre-K and K | 85% | 59% | 91% | 68% | 92% | 70% | 92% | 68% |
| Elementary | 76% | 38% | 78% | 39% | 79% | 40% | 77% | 35% |
| Middle School | 81% | 42% | 83% | 44% | 81% | 42% | 76% | 32% |
| High School | 87% | 60% | 87% | 59% | 86% | 57% | 82% | 49% |

Notes: This table summarizes the assignment process for different school levels. Columns (1), (3), (5), and (7) show the proportion of applicants who were allocated and accepted a seat in one of their top three choices. Columns (2), (4), (6), and (8) show the proportion of applicants who accepted a seat in their top choice.

Table A.2: Application Cohorts and Data Availability

| | Calendar Year | | | | | | |
|---|---|---|---|---|---|---|---|
| Application Cohort | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 |
| 2017 | 7th | **8th** | 9th | 10th | 11th | 12th | post-HS |
| 2018 | **6th** | 7th | 8th | 9th | 10th | 11th | 12th |
| 2019 | 5th | 6th | 7th | **8th** | 9th | 10th | 11th |
| 2020 | 4th | 5th | **6th** | 7th | 8th | 9th | 10th |
| 2021 | 3th | **4th** | 5th | 6th | 7th | 8th | 9th |
| 2022 | 2nd | 3rd | **4th** | 5th | 6th | 7th | 8th |

Notes: This table presents data availability for different cohorts of eighth-graders. Grey cells represent the eighth-grade cohorts participating in the school assignment under the Deferred Acceptance mechanism. For each row, grades in bold indicate when we observe previous test scores and background information for the respective cohort.

Table A.3: Application Patterns by Socioeconomic Status

| | | All Schools Ranked | | Top-Ranked School | |
|---|---|---|---|---|---|
| | Number of Applications (1) | 10th Grade Scores (Language) (2) | 10th Grade Scores (Math) (3) | 10th Grade Scores (Language) (4) | 10th Grade Scores (Math) (5) |
| Low-SES | $-0.434^{***}$ | $-0.206^{***}$ | $-0.211^{***}$ | $-0.278^{***}$ | $-0.271^{***}$ |
| | (0.009) | (0.004) | (0.004) | (0.005) | (0.005) |
| Girl | $0.050^{***}$ | $0.133^{***}$ | $0.055^{***}$ | $0.154^{***}$ | $0.070^{***}$ |
| | (0.011) | (0.003) | (0.003) | (0.005) | (0.005) |
| High Achiever | $0.037^{***}$ | $0.186^{***}$ | $0.189^{***}$ | $0.236^{***}$ | $0.236^{***}$ |
| | (0.011) | (0.004) | (0.004) | (0.005) | (0.005) |
| Distance to school | | $0.003^{***}$ | $0.011^{***}$ | 0.000 | $0.009^{***}$ |
| | | (0.001) | (0.001) | (0.001) | (0.001) |
| Low-SES × High Achiever | $0.044^{***}$ | -0.005 | $-0.010^{**}$ | -0.009 | $-0.017^{***}$ |
| | (0.013) | (0.005) | (0.005) | (0.007) | (0.006) |
| Low-SES × Girl | $0.033^{***}$ | -0.004 | -0.003 | $-0.021^{***}$ | $-0.032^{***}$ |
| | (0.012) | (0.004) | (0.004) | (0.006) | (0.006) |
| Low-SES × Distance to school | | $-0.006^{***}$ | $-0.012^{***}$ | -0.002 | $-0.008^{***}$ |
| | | (0.001) | (0.001) | (0.001) | (0.001) |
| Mean Outcome | 3.37 | -0.06 | -0.04 | -0.14 | -0.10 |
| Priority Controls | Yes | Yes | Yes | Yes | Yes |
| Number of applications | No | Yes | Yes | No | No |
| N-Obs | 394,189 | 1,307,240 | 1,307,151 | 389,574 | 389,565 |
| R-Squared | 0.04 | 0.04 | 0.04 | 0.06 | 0.06 |

Notes: This table reports OLS estimates of differences in the number and quality of schools ranked by ninth-grade applicants based on socioeconomic status and other background characteristics. The outcome in column (1) is the number of schools submitted to the centralized system. Outcomes in columns (2)-(3) measure the average school-level tenth-grade test scores across all schools included in the application. Columns (4) and (5) restrict the sample to consider only top-ranked school for each applicant. Priority controls include indicators for having a sibling enrolled in the school, a parent working in the school, and applying to a school previously attended. Students with special needs are excluded from the sample. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.4: Effect of Winning a Seat on School Characteristics

| | Attended School | | |
|---|---|---|---|
| | Mean (Compliers) | ITT | 2SLS |
| | (1) | (2) | (3) |
| 10th Grade Scores (s.d.) | $-0.509^{***}$ | $0.395^{***}$ | $0.587^{***}$ |
| | (0.003) | (0.004) | (0.006) |
| $F$-Statistic | | | 96,716 |
| N-Obs | | 202,634 | 202,634 |
| School Value-Added: College Enrollment (s.d.) | $-0.082^{***}$ | $0.018^{***}$ | $0.027^{***}$ |
| | (0.000) | (0.001) | (0.001) |
| $F$-Statistic | | | 96,626 |
| N-Obs | | 202,629 | 202,629 |
| School Value-Added: HS Graduation (s.d.) | $-0.004^{***}$ | $0.020^{***}$ | $0.029^{***}$ |
| | (0.000) | (0.000) | (0.000) |
| $F$-Statistic | | | 96,626 |
| N-Obs | | 202,629 | 202,629 |
| School Climate Index (s.d.) | $-0.298^{***}$ | $0.380^{***}$ | $0.565^{***}$ |
| | (0.004) | (0.005) | (0.007) |
| $F$-Statistic | | | 96,716 |
| N-Obs | | 202,634 | 202,634 |

Notes: This table reports ITT and 2SLS estimates of the effect of receiving an offer in a student's top choice on some characteristics of the school they attend in ninth grade. We estimate equations (5)-(6) using students' offer receipt as an instrument for attendance. The sample includes all randomized applicants without restricting by the quality of the residential location variable. Robust standard errors in parenthesis. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.5: OLS Estimates of Neighbor Spillovers

| | In $t$, Applicant: | | |
|---|---|---|---|
| | Ranks School Any | Ranks School 1st | Attends School |
| In $t-1$, Neighbor: | (1) | (2) | (3) |
| Enrolled in 1st Choice | 0.069*** | 0.060*** | 0.094*** |
| | (0.004) | (0.002) | (0.002) |
| Mean (Not enrolled) | 0.347*** | 0.124*** | 0.076*** |
| | (0.008) | (0.004) | (0.003) |
| N-Obs | 128,194 | 128,194 | 128,194 |
| N-Clusters | 73,255 | 73,255 | 73,255 |

Notes: This table shows OLS estimates of neighbors' spillovers on applicants' decisions observed the following year, excluding the full set of propensity score indicators. Enrolled is an indicator equal to one if the closest neighbor enrolled at their most preferred school. Clustered standard errors at the neighbor level are reported in parenthesis. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.6: Estimates of Neighbor Spillovers: Robustness to Different Samples

| | In $t$, Applicant: | | |
|---|---|---|---|
| | Ranks School Any (1) | Ranks School 1st (2) | Attends School (3) |
| *Panel A: Using K-12 schools* | | | |
| Enrolled to 1st Choice | 0.019*** | 0.014*** | 0.012*** |
| | (0.006) | (0.004) | (0.004) |
| Mean (Not Enrolled) | 0.352*** | 0.145*** | 0.106*** |
| | (0.004) | (0.003) | (0.002) |
| $F$-Statistic | 25,144 | 25,144 | 25,144 |
| N-Obs | 153,211 | 153,211 | 153,211 |
| N-Clusters | 81,414 | 81,414 | 81,414 |
| *Panel B: Using attendance data* | | | |
| Enrolled to 1st Choice | 0.022*** | 0.014*** | 0.012*** |
| | (0.007) | (0.005) | (0.004) |
| Mean (Not Enrolled) | 0.368*** | 0.149*** | 0.108*** |
| | (0.004) | (0.003) | (0.003) |
| $F$-Statistic | 19,407 | 19,407 | 19,407 |
| N-Obs | 128,085 | 128,085 | 128,085 |
| N-Clusters | 73,146 | 73,146 | 73,146 |
| *Panel C: Excluding siblings priorities* | | | |
| Enrolled to 1st Choice | 0.022*** | 0.014*** | 0.013*** |
| | (0.006) | (0.005) | (0.004) |
| Mean (Not Enrolled) | 0.369*** | 0.150*** | 0.105*** |
| | (0.004) | (0.003) | (0.003) |
| $F$-Statistic | 22,715 | 22,715 | 22,715 |
| N-Obs | 115,449 | 115,449 | 115,449 |
| N-Clusters | 68,800 | 68,800 | 68,800 |

Notes: This table presents 2SLS estimates from our main specification (2) and (3) using alternative sample selections. Panel A for all students enrolled in K-12 non-private schools. All models control for a saturated set of indicators for the assignment propensity score. Standard errors clustered at the neighbor level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.7: ITT and 2SLS Estimates of Neighbor Spillovers: Using Linear Propensity Score Controls

| | First Stage | ITT | | | 2SLS | | |
|---|---|---|---|---|---|---|---|
| | In $t-1$, Neighbor: | In $t$, Applicant: | | | In $t$, Applicant: | | |
| In $t-1$, Neighbor: | Enrolled in 1st Choice (1) | Ranks School Any (2) | Ranks School 1st (3) | Attends School (4) | Ranks School Any (5) | Ranks School 1st (6) | Attends School (7) |
| Admitted to 1st Choice | 0.673*** | 0.016*** | 0.010*** | 0.010*** | | | |
| | (0.004) | (0.005) | (0.003) | (0.003) | | | |
| Enrolled in 1st Choice | | | | | 0.023*** | 0.014*** | 0.014*** |
| | | | | | (0.008) | (0.005) | (0.004) |
| Mean (Not Enrolled) | 0.075*** | 0.300*** | 0.096*** | 0.019*** | 0.366*** | 0.150*** | 0.109*** |
| | (0.002) | (0.003) | (0.002) | (0.001) | (0.005) | (0.003) | (0.003) |
| $F$-Statistic | | | | | 12,836 | 12,836 | 12,836 |
| N-Obs | 73,255 | 128,194 | 128,194 | 128,194 | 128,194 | 128,194 | 128,194 |
| N-Clusters | | 73,255 | 73,255 | 73,255 | 73,255 | 73,255 | 73,255 |

Notes: This table reports intent-to-treat (ITT) and two-stage least squares (2SLS) estimates of neighbors' spillovers on applicants' decisions. Column (1) presents the OLS estimate from a regression where the dependent variable is an indicator equal to one if the neighbor enrolled in ninth grade in the same school where they received an offer and the explanatory variable is the offer receipt. Columns (2)-(4) display OLS estimates of regressions where the variable of interest is an indicator equal to one if the closest neighbor received an offer in their most preferred school. Columns (5)-(7) report 2SLS coefficients instrumenting neighbors' enrollment with the offer. Standard errors are clustered at the neighbor level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.8: Heterogeneous Neighbor Spillovers: by Previous Language Scores

| | In $t$, Applicant: | | |
|---|---|---|---|
| | Ranks School Any | Ranks School 1st | Attends School |
| | (1) | (2) | (3) |
| *Panel A: Both below median score* | | | |
| Enrolled in 1st Choice | 0.037*** | 0.010 | 0.011 |
| | (0.012) | (0.009) | (0.008) |
| Mean (Not enrolled) | 0.368*** | 0.165*** | 0.118*** |
| | (0.008) | (0.006) | (0.005) |
| $F$-Statistic | 7,526 | 7,526 | 7,526 |
| N-Obs | 27,238 | 27,238 | 27,238 |
| *Panel B: Only applicant below median score* | | | |
| Enrolled in 1st Choice | −0.008 | 0.012 | 0.014* |
| | (0.014) | (0.010) | (0.009) |
| Mean (Not enrolled) | 0.369*** | 0.138*** | 0.090*** |
| | (0.008) | (0.006) | (0.005) |
| $F$-Statistic | 7,033 | 7,033 | 7,033 |
| N-Obs | 22,482 | 22,482 | 22,482 |
| *Panel C: Only neighbor below median score* | | | |
| Enrolled in 1st Choice | 0.012 | 0.010 | 0.011 |
| | (0.014) | (0.011) | (0.010) |
| Mean (Not enrolled) | 0.388*** | 0.160*** | 0.128*** |
| | (0.009) | (0.007) | (0.006) |
| $F$-Statistic | 6,274 | 6,274 | 6,274 |
| N-Obs | 20,500 | 20,500 | 20,500 |
| *Panel D: Both above median score* | | | |
| Enrolled in 1st Choice | 0.009 | 0.011 | 0.005 |
| | (0.015) | (0.011) | (0.010) |
| Mean (Not enrolled) | 0.384*** | 0.157*** | 0.118*** |
| | (0.009) | (0.007) | (0.006) |
| $F$-Statistic | 6,040 | 6,040 | 6,040 |
| N-Obs | 18,088 | 18,088 | 18,088 |
| $p$-value for equal effects | 0.103 | 0.998 | 0.918 |

Notes: This table reports 2SLS estimates from equations (2) and (3) of the effects of each applicant's closest neighbor attending her most preferred school depending on applicant's and neighbor's previous language test scores. Each column reports the main estimate and an interaction between the main effect and an indicator variable of whether applicant and neighbor scored above the median in the corresponding test score distribution. Enrollment is instrumented with an indicator equals to one if the closest neighbor received a seat offer. All models control for a saturated set of indicators for the assignment propensity score. Clustered standard errors at the neighbor level are reported in parenthesis. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.9: Estimates of Effects on Applicants' Submitted Schools

| | Top-Ranked School | | | Second-Ranked School | | | Third-Ranked School | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean<br>(1) | ITT<br>(2) | 2SLS<br>(3) | Mean<br>(4) | ITT<br>(5) | 2SLS<br>(6) | Mean<br>(7) | ITT<br>(8) | 2SLS<br>(9) |
| 10th Grade Scores | −0.145*** | 0.106*** | 0.125*** | −0.104*** | 0.087*** | 0.102*** | −0.094*** | 0.081*** | 0.095*** |
| | (0.006) | (0.004) | (0.005) | (0.006) | (0.004) | (0.005) | (0.007) | (0.005) | (0.006) |
| $F$-Statistic | | | 62,966 | | | 62,191 | | | 55,051 |
| N-Obs | | 110,367 | 107,879 | | 108,249 | 105,793 | | 79,407 | 77,544 |
| Value-Added: College Enrollment | −0.037*** | 0.317*** | 0.359*** | −0.029*** | 0.297*** | 0.336*** | −0.027*** | 0.298*** | 0.336*** |
| | (0.001) | (0.004) | (0.005) | (0.001) | (0.004) | (0.005) | (0.001) | (0.005) | (0.005) |
| $F$-Statistic | | | 99,923 | | | 99,778 | | | 90,194 |
| N-Obs | | 112,115 | 109,893 | | 110,228 | 108,028 | | 80,993 | 79,314 |
| Value-Added: HS Graduation | 0.003*** | 0.194*** | 0.238*** | 0.003*** | 0.167*** | 0.206*** | 0.005*** | 0.136*** | 0.169*** |
| | (0.000) | (0.004) | (0.005) | (0.000) | (0.004) | (0.005) | (0.000) | (0.004) | (0.005) |
| $F$-Statistic | | | 37,371 | | | 36,431 | | | 29,855 |
| N-Obs | | 112,115 | 109,893 | | 110,228 | 108,028 | | 80,993 | 79,314 |
| School Climate Index | 0.044*** | 0.212*** | 0.254*** | 0.021*** | 0.181*** | 0.218*** | −0.009 | 0.174*** | 0.210*** |
| | (0.007) | (0.004) | (0.005) | (0.007) | (0.004) | (0.004) | (0.008) | (0.004) | (0.005) |
| $F$-Statistic | | | 51,580 | | | 51,363 | | | 42,458 |
| N-Obs | | 109,607 | 106,968 | | 107,467 | 104,865 | | 78,842 | 76,862 |

Notes: This table reports 2SLS estimates from equations (2) and (3) of neighbors' school characteristics on the characteristics of schools submitted by applicants. Columns (2)-(3) display estimates on the applicant's top-ranked school characteristics, and columns (5)-(6) and (8)-(9) show estimates on the applicant's second- and third-ranked schools, respectively. Columns (1), (4), and (7) show mean outcomes for compliers computed following Abadie (2002). Columns (2), (5), and (8) show coefficients from regressions of outcomes on the characteristics of the school where the neighbor received an offer. Columns (3), (6), and (9) report 2SLS coefficients instrumenting neighbors' attendance with offer receipt. All models include a saturated set of indicators for the assignment propensity score and cluster standard errors at the neighbor level. See the main text for details about the construction of each outcome. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.