# Go Your Own Way? Peer Effects, School Choice, and Educational Inequality*

Inbar Amit

Click Here for Latest Version

November 14, 2025

**Abstract**

High schools have lasting academic and economic consequences for their students. I use detailed Chilean administrative data to show that interactions among primary school peers shape how students sort into high schools. I exploit the random assignment of students into oversubscribed primary schools, together with repeated data on graduates' applications, to identify the causal effect of primary school peers on students' choice of high school. I find that studying with a particular peer significantly increases the probability that a student applies to and attends the same high school. Peer effects interact nonlinearly with the spatial distribution of high school quality, shaping educational inequality. Embedding peer effects into a model of school choice, I show that when high school quality is unevenly distributed across space, interactions among peers can reduce applications to high-quality schools from students in disadvantaged areas, amplifying inequality. Empirically, I find that peer effects account for one-sixth of the gap in the probability that low- and high-income students attend a high-quality high school. To address these gaps, I show that policies that expand students' access to high-quality high schools can benefit from increasing returns to scale through spillovers within peer networks, without generating negative externalities for untargeted students.

## 1 Introduction

The choice of high school is a key decision children and parents face. Students who attend higher-quality high schools tend to achieve better grades, continue to higher education more frequently,

*Inbar Amit: University of Oxford (inbar.amit@economics.ox.ac.uk)

I am very grateful for the guidance and helpful comments of Barbara Petrongolo, Alison Andrew and Romuald Méango. I also thank Áureo de Paula, Anaïs Fabre, Ana Gazmuri, Zach Bleemer, Jack Mountjoy, Rob Garlick, Julio Rodriguez, Marco Pariguana, Giulio Gottardo, Stephen Bond, Mathias Jensen, Howard Smith, Antonella Bancalari, participants and the 2nd TSE Economics of Education Workshop, the Stone Centre Conference on Education and Inequality, the 23rd Nordic Development Conference, the QMUL Economics and Finance Workshop, the Stockholm Uppsala Education Economics Workshop and seminar participants at the University of Oxford and the Institute for Fiscal Studies for their helpful comments. I also thank the Chilean Ministry of Education (Mineduc) for providing data access and support.

and experience better labor market outcomes (Angrist et al., 2016; Beuermann and Jackson, 2022). Households from different backgrounds, however, systematically choose schools of different quality. Across many countries, lower-income students disproportionately sort into worse-performing schools, contributing to large socioeconomic gaps in educational achievement (Chetty et al., 2016; Burgess et al., 2015). In Chile, children of high school dropouts score $0.62\sigma$ lower on standardized math tests and are twice as likely to drop out of high school compared to children of graduates. Conditioning on the high school that the child attends reduces these gaps by two-thirds, suggesting that they primarily reflect differences *between* the high schools that different children attend, rather than their underlying ability.[1]

Poorer students attend lower-quality schools for several reasons. On the supply side, higher-quality schools are usually located in wealthier neighborhoods, and often charge tuition fees or hold competitive admissions processes which tend to disadvantage poorer students. On the demand side, poorer students might be less sensitive to school quality, or might favor attending schools closer to home, which may be of lower quality (Agte et al., 2024; Allende et al., 2019; Nielson, 2021). Finally, given students' preferences and the distribution of schools, the mechanism that assigns students to schools can play a key role in reducing or increasing inequality (Burgess and Briggs, 2010; Calsamiglia et al., 2020).

In this paper, I focus on a novel demand-side factor: social interactions among primary school peers. School peers form an essential part of children's network, and influence among students on a large number of outcomes has been well documented in the literature (e.g., Sacerdote, 2011; Barrios-Fernandez, 2023). I focus on how primary school peers impact students' subsequent choices over high schools, and the resulting impact on socioeconomic inequality in education. Since students prefer to study close to home, peer networks can increase the tendency to enroll locally. Given the uneven distribution of school quality across neighborhoods, this can make poorer students more likely to attend lower quality high schools and wealthier students more likely to attend higher quality ones, increasing educational inequality.

Identifying social interactions in school choice is challenging for at least two reasons. First, peer groups do not form randomly; primary school peers (and their families) share similar backgrounds and preferences, which lead them to become peers in the first place. Moreover, by virtue of studying in the same environment, peers are exposed to shared advice from teachers or school counselors, which may generate correlations in choices absent any influence among peers. I overcome these by leveraging a key feature of Chile's school admission system – applicants to oversubscribed primary schools are admitted through randomized lotteries. This makes it possible to identify students' 'counterfactual' peers: students who are similarly selected as their actual peers, but randomly assigned to a different school. I combine this randomization with repeated data on students across primary school cohorts, which allows me to control for the effect of the school environment itself, to cleanly identify peer effects.

I find that peers significantly impact students' choice of high school. Studying with a peer

[1] The details of this decomposition can be found in Appendix A6.

increases the probability that a student ranks the same high school as their preferred option by 1.1 percentage points, and the probability they subsequently attend the same high school by 1.5 percentage points. These effects are qualitatively large; for reference, two randomly selected primary school students have only a 0.08 percentage point probability of attending the same high school.

This analysis alone does not explain how peer effects contribute to educational inequality in equilibrium. To better understand this relationship, I build a school choice model in which the utility of attending a high school depends, in part, on the (expected) share of a student's peers who will attend. I show that the relationship between the magnitude of peer effects and educational inequality depends on the spatial distribution of high school quality. When high-quality high schools are relatively accessible for both high-income and low-income students, peer effects pull all students *towards* high-quality schools and away from low-quality schools, reducing educational inequality. Alternatively, when inequality in access to high-quality high schools is high – i.e. high-income students have much better access to good high schools relative to low-income students – peer effects pull low-income students *away* from high-quality schools, and high-income students towards high-quality schools, increasing inequality.

To quantify the contribution of peer effects to educational inequality in equilibrium and to perform policy counterfactuals, I extend the model in two key ways. First, I allow peers to influence students' choices through two distinct channels: students may respond both to peers' *behavior* or to peers' *characteristics*. For instance, a student may apply to the same school as a peer because they prefer to study together, or because the peer lives next to the school and shares positive information about its quality, which induces both to apply. If students respond to peers' characteristics rather than behavior, policies targeting behavior alone are likely to be ineffective. A structural model allows me to disentangle these two channels, which are generally not separately identified in linear models – a challenge known as the *reflection problem* (Manski, 1993; Angrist, 2014). Second, I model the deferred acceptance algorithm that maps students' behavior into admission outcomes, which allows me to explicitly account for externalities arising from increased competition across high schools. As a result, the model is able to speak to the impact of policies that encourage lower-income students to apply to higher-quality schools on untreated students, who may be crowded-out.

I find that students are highly sensitive to their peers' choices: a student will be willing to travel to a high school that is over 900 meters farther from their home in order to study with 10% more of their primary school peers. In contrast, I find no evidence that students' behavior depends on their peers' characteristics. Using the estimated model, I show that peer effects significantly contribute to disparities in applications and admissions into high-quality high schools. Eliminating peer effects shifts low-income students' applications to higher quality schools located farther from their home. Overall, this reduces the gap in the probability that a low-income student attends a high-quality high school relative to high-income student by one-sixth.

The findings of this study have important implications for policy debates on educational inequality. First, to the extent that students' choices influence one another, the elasticity of an individual student's choice with respect to changes in a school's attributes is likely to be low. For instance,

students may be less inclined to attend a new school that opened nearby if it is not convenient for their peers. This mechanism helps rationalize why targeting supply-side barriers that make it harder for lower-income students to attend good schools, such as by centralizing admissions or reducing transportation costs, has not significantly reduced educational inequality across many contexts (Dustan and Ngo, 2018; Herskovic, 2020; Kutscher et al., 2023). Moreover, students' revealed preference to study with their peers suggests that while "moving to opportunity" initiatives may have large effects on students' academic and labor market outcomes, their true welfare effects may be lower than those implied by these outcomes alone (e.g., Chetty et al., 2016).

More optimistically, peer effects can be harnessed to increase the effectiveness of policy aimed at encouraging low-income students to enroll in better schools. A single student applying to a better school encourages their peers to apply, which suggests that an intervention aimed at re-allocating students across schools is likely to have a positive multiplier effect if it is sufficiently strong. To study this, I simulate a counterfactual policy which offers subsidies to cover the cost of transport to high-quality high schools to different shares of students in primary schools with predominantly low-income students.[2]

I find that this policy generates both increasing return to scale on treated students as well as spillovers on untreated students, which increase with the scale of the program. Treating 10% of students in a primary school cohort increases the probability that treated students attend a high-quality high school by 15% with no effect on untreated students. Treating 80% of students, in contrast, increases the probability that treated students attend a high-quality school by 18% and untreated students by 5%. At the same time, the policy does not reduce the probability that students in schools that predominantly enroll high-income students – who are never treated – are admitted into a high-quality high school. Since students prefer to study in the same high school as their peers, the policy increases applications to higher-quality, but relatively under-subscribed schools where students (and their peers) are *all* more likely to be admitted. As a result, it does not generate congestion externalities that crowd out untreated students. This suggests that effective policy can close socioeconomic gaps in high school quality without harming high-income students by explicitly taking spillovers within peer networks into account.[3]

This paper contributes to the large body of literature studying the sources of inequality in educational attainment. Studies have highlighted that inequality is the result of multiple interacting factors: lower-income parents may have weaker preferences for school quality (Nielson, 2021; Burgess et al., 2015), less accurate beliefs about school performance (Agte et al., 2024), or a preference for enrolling their children with peers of similar socioeconomic background (Allende, 2021); lower-

[2]Similar policies may include building schools in central areas within low-income neighborhoods. If these schools are built in a location that is sufficiently attractive to enough students within a single network, students will *pull in* their peers, generating positive spillovers. Moreover, the finding that primary school peers influence each others' educational trajectories has bearing for the optimal assignment of students into primary schools. In an environment with high inequality in access to primary schools, increasing the socioeconomic diversity of peer groups may help draw in more low-income students into high-quality high schools.

[3]See Bennett and Bergman (2021) and Avvisati et al. (2014) for examples of papers documenting multiplier effects of educational interventions.

income areas typically provide worse access to high-quality schools (Chetty et al., 2016; Fabre et al., 2024); and high tuition costs may put good schools out of reach for poorer students (Hoxby and Avery, 2012).[4] I focus on a novel factor: the role of students' peer network. I document that students prefer to attend high school with their primary school peers. Given the unequal distribution of school quality, peer effects result in lower-income students applying to closer but lower-quality high schools. I show that eliminating peer effects reduces inequality between high- and low-income students by one-sixth. Moreover, while spillovers through peer networks constrain the effectiveness of policy aimed at encouraging low-income students to attend better schools when rolled out at small scale, they can also be exploited to yield high returns at larger scales with minimal negative externalities on untreated students.

Second, I add to the literature studying the impact of peers on students' educational choices, including which schooling track to choose (Black et al., 2013), whether to enroll in high school (Bobonis and Finan, 2009), or whether to enroll in university (Carrell et al., 2018; Barrios-Fernández, 2022; Chetty et al., 2016).[5] I contribute to this literature in two ways. First, to the best of my knowledge, this is the first paper to study the effect of *cohort peers* on students' choice of school. Prior research has largely examined sequential spillovers, such as younger siblings following the educational paths of older siblings (Dustan, 2018; Gurantz et al., 2020; Altmejd et al., 2021) or neighbors responding to earlier enrollment decisions (Bobonis and Finan, 2009; Barrios-Fernández, 2022).[6] A separate strand analyzes how exposure to peers with particular characteristics – such as ability, gender composition, or disruptive behavior – affects outcomes (Hoxby, 2000; Black et al., 2013; Carrell et al., 2018). This distinction is important; I find that joint decision-making reinforces existing inequalities by keeping lower-income students in lower-quality high schools. To the extent that students also impact younger peers, these effects are likely to be amplified.

Methodologically, I develop a novel approach that exploits school admission lotteries to identify the causal effect of studying with a peer on high school choice.[7] This method builds on the use of admission lotteries to estimate the value-added of schools on students' outcomes (e.g., Deming, 2014; Angrist et al., 2017).[8] Identifying peer effects, however, requires separately identifying the

[4] In contexts where private, profit-seeking schools dominate, monopsonistic providers may also exploit households' preferences, beliefs, or students' ability in ways that advantages wealthier families (Nielson, 2019; Allende, 2021; Bau, 2022).

[5] A number of papers focus on social interactions in job search. Topa (2001) builds a model that explicitly incorporates local interactions and allows agents to exchange information about job openings within their social networks. More recently, Caria et al. (2023) combine a randomized job-search assistance treatment with a social network survey to study how unemployed individuals impact each others' job search, and how information propagates across networks.

[6] A related literature examines the role of peer pressure on simultaneous decisions, showing that students' choices can be shaped by the observability of their actions or by coordination within peer groups (Bursztyn and Jensen, 2015; Bursztyn et al., 2019).

[7] In settings exploiting sequential choices, instruments can typically be found for older peers' outcomes, such as grade cutoffs in university admission (Altmejd et al., 2021), or variation in eligibility rules for cash transfers or loan programs (Bobonis and Finan, 2009; Barrios-Fernández, 2022). Similarly, studies examining the impact of exposure to peers with particular characteristics have exploited admission cutoffs at oversubscribed schools (Abdulkadiroğlu et al., 2014), or random assignment of students to classrooms and dorms (Sacerdote, 2001; Booij et al., 2017).

[8] Relatedly, Kirkebøen et al. (2021) leverage admission cutoffs to identify the impact of college enrollment on the probability that college graduates marry other graduates from their institution and field of study.

impact of attending a primary school from the impact of studying with a particular peer. To address this, I follow a common approach in the peer effects literature and leverage data on the application decisions of students across successive primary school cohorts to control for features of primary schools that influence students' application behavior (e.g., Sacerdote, 2001; Michelman et al., 2021).[9] This framework can be applied in any setting with repeated data and randomization into schools, and can be used to estimate the impact of exposure to peers on outcomes such as grades, study effort, or smoking.

Third, I complement the literature studying families' preferences over schools. Studies have documented that households prefer to enroll students in schools that are closer to home, cheaper, and of higher quality (Chumacero et al., 2011; Hastings et al., 2009; Agte et al., 2024). Parents are also sensitive to the social composition of schools, preferring to enroll their children in schools with higher quality peers, and peers from a similar socioeconomic background (Burgess et al., 2015; Allende et al., 2019; Campos, 2024). I document that peers within a student's own network directly shape their choice of school; studying with a given peer significantly increases the probability that a student will choose the same high school. Using a structural model, I separately identify whether this impact is driven by exposure to peers' characteristics or behavior, and find that it is entirely driven by students' reaction to their peers (expected) behavior.

Finally, I contribute to the growing literature modeling school choice. School choice models incorporating rich preference heterogeneity have been used to study private schools in Pakistan (Bau, 2022), voucher schools in Chile (Nielson, 2021; Agte et al., 2024), and charter schools in the US (Walters, 2018). Closely related, Allende (2021) models parents' preference that their children study with peers of the same socioeconomic group in Peru, and shows how preferences interact with schools' profit-maximizing behavior to generate complex equilibrium dynamics. I complement this literature by embedding students' choice of high school within their immediate social networks: students do not choose schools independently, rather, the utility of attending a school depends on the share of their actual peers who (in expectation) will attend it. I show that repeated data on students' high school applications across primary schools can identify peer effects in this setting, even in the presence of group-level unobservables, such as school effects, and the joint influence of peers' behavior and characteristics.

The rest of this paper is structured as follows. Section 2 describes the education system in Chile and introduces the data and variables used in the analysis. Section 3 discusses the identification of peer effects on students' choice of high school. Section 4 introduces a conceptual framework that ties peer effects to educational inequality in equilibrium. Section 5 develops a structural model of school choice which incorporates peer effects. Section 6 presents the estimated parameters and evaluates the model's fit. Section 7 conducts counterfactual simulations to evaluate the contribution of peer effects to inequality in equilibrium, and the effectiveness of transport subsidies in reducing inequality. Section 8 concludes.

[9] This approach parallels the use of firm fixed effects in matched employer–employee data to separate worker quality from firm productivity (e.g., Abowd et al., 1999), and the use of classroom fixed effects to isolate teachers' causal effect on students' outcomes (e.g., Chetty et al., 2014).

# 2 Setting and Data

In this section, I describe the education system in Chile, introduce the data sources used in the analysis, and provide descriptive evidence on schooling inequality.

## 2.1 Education in Chile

In Chile, compulsory education is divided into primary school (ages 6-14) and high school (ages 14-18). Students must remain in school until they are 18, and graduation rates are high; in 2023, 95% of students successfully graduated from high school. Students transition through these stages in either the same institution or in different institutions. 54% of primary schools offer only primary education, and enroll 48% of students, while the remaining institutions offer both primary and secondary education. Schools are either publicly funded and managed by local municipalities (public), publicly funded but privately managed (private voucher), and privately managed and funded (private non-voucher). In 2023, 53% of students attended private voucher schools, 35% attended public schools, and 12% attended private non-voucher schools.

Chile's education system underwent a significant reform with the introduction of the Schooling Inclusion Law (Ley de Inclusión Escolar) in 2016. The law aimed to increase fairness in admissions and curb the negative effects of private schools' profit-maximizing behavior through two key reforms. First, it forbade private voucher schools from charging parents' fees on top of the voucher amount and required them to restructure as non-profits. Second, it introduced a centralized application system (Sistema de Admisión Escolar, SAE) for admission into all schools that receive funding from the state (public schools and private voucher schools). In 2023, the system accounted for admission into 89% of primary schools and 85% of high schools in the country, including almost all public schools and private schools that accept school vouchers (Agte et al., 2024).

To be admitted into a school through the SAE, students submit a ranked-ordered-list (ROL) indicating their preferences over schools, listing as many options as they wish.[10] Schools reserve a share of their seats for specific groups of students: (1) students with special educational needs, (2) high academic performers,[11] or (3) low-income applicants.[12] Within each quota group, students are ranked according to a nationally standardized priority order. The highest priority is given to students with siblings currently enrolled at the school, followed by children of school employees, and then students who have previously attended the school. All other students fall into a residual priority category. These priority rules apply both within and outside of quota groups and determine the order in which students are considered for available seats. If the number of applicants in a given quota–priority combination is less than or equal to the number of reserved seats, all students in

[10]The main application round starts in early August for admission in January, during which applications can be submitted and edited for one month. A secondary application round is held in November to allow students who remained unmatched or were dissatisfied with their main-round placement to reapply.

[11]Typically these correspond to students with grades in the top 30% of their cohort.

[12]Quota policies are partly regulated by the Ministry of Education. All schools are required to reserve at least 15 percent of seats for special needs students, while the reservation of seats for high-achieving or low-income students is optional and capped at 15 percent each.

that group are admitted. If the number of applicants exceeds the number of reserved seats, a lottery is held to allocate the available places.[13] Notably, there are no geographic barriers (e.g. catchment areas), or academic barriers (e.g. minimum grade requirement) for admission into any school.

### 2.2 Data and Variables

I link several registries maintained by the Chilean Ministry of Education (Mineduc) to create the dataset for this study. All of the datasets include unique and consistent identifiers for students, classes, schools, and teachers, allowing me to perform various linkages across them. I begin by obtaining the administrative data from the centralized application system, which includes information on students' ranking of schools as well as their location. I link this dataset to data on schools' characteristics (including size, teachers, location, etc.), and publicly available measures of school quality developed by the Chilean Agency for Quality in Education (ACE). I then link data on students' ranking to the registry of all students in the education system, which includes sociodemographic information, test scores, and attendance records. Using this dataset, I identify students' relevant peer groups. Finally, I link data on students who apply to high school before 2023 to data from the national standardized SIMCE exam. This includes standardized Spanish and Math test scores taken when students were in the fourth grade and, importantly, a survey of their parents, which contains detailed information about households, including income and parental education. Appendix Table A1 describes the variables available in each data source in more detail.

I observe each student's gender, date of birth, residential location, GPA, attendance record, and household income.[14] My primary measure for students' household income is whether they are categorized as *prioritario*, corresponding to students living in households in the bottom 6 deciles of the income distribution. For students in the 2019-2022 cohorts, I link the administrative registries to survey data collected from students' parents when students take the standardized SIMCE exams. For this subset of students, I am also able to observe their parents' level of education.

I consider that a set of students are *peers* if they attend the same primary school in the same cohort. A typical primary school cohort will have 45 students. In 74% of primary schools, peers study in a single classroom and take all their courses together. They are taught by the same teachers, follow the same schedule, and share break times. In the remaining schools, students are divided across two or more classrooms. Primary school peers tend to attend the same high schools. On average, an incoming high school student studies with 10.5 of their former primary school peers.

[13] The algorithm that assigns students to schools proceeds sequentially through each quota group and priority level. Students with the highest priority are considered first for the quotas for which they are eligible. Once a quota is filled, any remaining students are considered for regular admission seats, again in priority order. In practice, quota groups are also ordered by priority, with the highest priority given to students with special needs, followed by low-income students, and finally high achieving students. Applicants eligible for multiple quotas are first considered under the highest-priority quota. If not admitted, they subsequently compete for regular seats.

[14] For the majority of students, residential location is recorded on the basis of the IP address of the device from which they access the school application portal. For cases in which the location was not retrieved directly from the device, it is based on geolocating the reported addresses of families. To protect students' identity, the location that appears in the administrative files includes a small amount of noise that shifts the true location in a random direction by a median of 150 meters.

Correspondingly, 8.9 of their former primary school peers ranked the same high school as their most preferred option, suggesting that these patterns are the result of students' application behavior. Appendix Figure A2 plots the distribution of primary school peers who apply to and attend the same high school as each incoming student.

I observe the entire rank-ordered lists (ROL) submitted by each student on the SAE. Students list an average of 4.09 different high schools schools on their ROL for admission into grade 9. On average, 51% of students are admitted to their first choice school, 18% to their second choice, 13% to their third choice and 12% to their fourth choice.[15] A student is considered admitted to a school whenever they receive an offer from it. On average, 82% of students accept the offer they receive in the first application round.[16] A student is classified as enrolled when they accept an offer, and as attending when they appear in enrollment records in the following academic year.

I observe each school's number of enrolled students, number of open seats for new students, grades taught, specialization, religious orientation, and quality. My main measure of school quality corresponds to ratings from the Agency for Quality in Education (ACE). ACE serves as the primary body in charge of monitoring and rating the quality of schools in Chile. It classifies schools into four performance categories: high, medium, medium-low, and insufficient. These ratings are based on the performance of teachers, personal development indicators, and SIMCE exam results, adjusted for school-level student characteristics. Nationally, 15% of schools are rated as high-quality, 55% as medium-quality, 24% as medium low-quality, and 6% as having insufficient quality. These ratings are publicly available online, and are displayed to parents on the school application platform.[17] Quality ratings correlate strongly with value-added measures on test scores, and surveys indicate parents view them as key factors in their school choice decision (Agte et al., 2024).

### 2.3 Descriptive Evidence on Inequality in Education

The centralized admission system in Chile ensures that admission to public high schools is not constrained by a student's location or academic performance. Even so, lower-income students are significantly more likely to apply to and attend worse performing high schools. Figure 1 plots the difference in the probability that a student ranks, ranks as their preferred option, and is admitted into a school categorized as having *medium*, or *high* quality between low- and high-income students. Low-income students are 3 percentage points less likely to attend a higher-quality school, 6 per-

[15] Relative to applications for admission into primary schools, families rank more schools but face slightly worse placement chances. On average, 65% of applicants for admission into primary schools were admitted to their preferred school.

[16] Some applicants likely enroll in off-platform private schools or reapply in the secondary admission round.

[17] The SAE provides several tools to inform parents about the characteristics of schools. Since 2016, the SAE application platform has offered the "School Showcase" (Vitrina SAE) website, enabling parents to search for schools by name (Correa et al., 2022). The platform displays information on schools' available seats, tuition prices, and academic quality ratings. In 2021, the Ministry of Education introduced a new school explorer platform, Más Información Mejor Educación (MIME), developed by an EdTech NGO and hosted on the government website (Agte et al., 2024). This tool aggregates public data on all schools nationwide and allows parents to search for schools on a map rather than by name, enabling the discovery and comparison of nearby schools previously unknown to them. Additionally, MIME uses data on students' locations and characteristics to provide tailored information. For more details, see Agte et al. (2024).

centage points less likely to rank a higher-quality school as their preferred option, and 5 percentage points less likely to rank a higher-quality high school anywhere in their ROLs. These gap are even higher for students who live within the same commune, Chile's smallest administrative division.

Figure 1: Gap in applications and admissions to medium- and high-quality high schools




Notes: The figure present differences (low-income - high-income) in the probability that a student ranks (in any position) a school categorized as having *medium* or *high* quality, differences in the probability that a student ranks as their preferred option a school categorized as having *medium* or *high* quality, and differences in the probability that a student is admitted into a school categorized as having *medium* or *high* quality. The bars in orange correspond to raw gaps in the mean between high- and low- income students, and the bars in green correspond to estimates of the difference controlling for a students' commune. The number above each bar corresponds to the average probability among high-income students. The figure uses data from 2023 and all differences are significant at the 1% level.

Key to socioeconomic gaps in applications to higher-quality schools is the uneven spatial distribution of school quality. Chile's education system offers parents and students a high degree of choice. Within a 10 km radius, the average student can choose among 63 publicly funded high schools classified as having *medium* or *high* quality by the Agency for Quality in Education. Lower-income students, however, have markedly worse access to higher-quality schools. Table 1 shows that the nearest medium- or high-quality high school is, on average, 165 meters farther from low-income students than from high-income students. Low-income students also live next to 19 fewer medium- and high-quality high schools and have 1,000 fewer available seats in these high schools within a 10 km radius of their home. Controlling for the commune in which a student lives decreases but does not close these gaps, suggesting that even within the same town and neighborhood, lower-income

students tend to have worse schooling options.[18] In Section 4, I discuss how peer effects interact with students' uneven access to high-quality schools to amplify gaps in applications to high-quality high schools.

Table 1: Access to Higher Quality High Schools

| | Distance | | Schools | | Seats | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Variables* | | | | | | |
| Constant | 0.954 | | 72.4 | | 3,656.2 | |
| | (0.007) | | (0.383) | | (20.6) | |
| Low-Income | 0.165 | 0.069 | -19.0 | -1.31 | -980.8 | -79.0 |
| | (0.009) | (0.009) | (0.497) | (0.437) | (26.7) | (23.8) |
| *Fixed-effects* | | | | | | |
| Commune | No | Yes | No | Yes | No | Yes |
| *Fit statistics* | | | | | | |
| Observations | 107,832 | 107,034 | 110,391 | 109,585 | 110,391 | 109,585 |
| $R^2$ | 0.00293 | 0.21431 | 0.01299 | 0.65557 | 0.01205 | 0.63818 |

Notes: Columns (1) and (2) present the results of regressing the average distance to the nearest school categorized as having *medium* or *high* quality on whether a student in from a low-income household. Columns (3) and (4) present the results of regressing the number of schools categorized as having *medium* or *high* quality within a 10 km radius on whether a student in from a low-income household. Columns (5) and (6) present the results of regressing the number of available seats in schools categorized as having *medium* or *high* quality within a 10 km radius on whether a student in from a low-income household. The regressions use data from 2023 and all coefficients are significant at the 1% level.

## 3 Exposure to peers and students' choice of high school

Identifying peer effects is challenging. First, peer groups do not form randomly. Students who study in the same primary school live close to each other, have similar socioeconomic backgrounds, and, likely, have similar (unobserved) preferences over school attributes. Observed similarities in the choices of a student and their peers may simply reflect these common factors. Second, even in cases where students are randomly assigned to peer groups, peers, by definition, study in similar environments which may also influence their choices. For instance, teachers or school counselors may provide advice on which high school to apply, generating a correlation in the choices of students in the same primary school.

In this section, I describe how admission lotteries for oversubscribed primary schools, together with data on students across different cohorts within primary schools, can identify peer effects. I begin by grounding my discussion in a potential outcomes framework that accounts for a causal

[18] Appendix Figure A1 shows that communes where higher shares of low-income students live have fewer medium- and high-quality high schools, fewer seats at medium- and high-quality schools, and fewer seats per high school applicant at these schools.

influence of exposure to a particular peer (peer effects), non-random selection of students into primary schools, and school effects. I then discuss how admission lotteries overcome selection into schools, and how repeated data on students in different cohorts within a primary school can isolate peer effects from common school-level influences. I conclude with estimates of the effect of exposure to a primary school peer on the probability that a student ranks and attends the same high school.

### 3.1 Identification

To motivate the analysis, let $y_{ij}$ denote whether student $i$ chooses the same high school as student $j$. Specifically, $y_{ij}$ may correspond to the probability that $i$ ranks as her most preferred option the same high school as $j$, the extent to which her rankings overlap with $j$'s, or whether she enrolls in the same high school as $j$. I decompose $y_{ij}$ into three components:

$$y_{ij} = \underbrace{\gamma D_{ij}}_{\substack{\text{Peer} \\ \text{effect}}} + \underbrace{\eta_{ij}}_{\substack{\text{Student} \\ \text{similarity}}} + \underbrace{\delta_{ij}}_{\substack{\text{School} \\ \text{effect}}} \tag{1}$$

where $D_{ij}$ denotes whether students $i$ and $j$ are peers, and $\gamma$ corresponds to the effect of studying with student $j$. $\eta_{ij}$ corresponds to the underlying similarity between $i$ and $j$'s preferences, reflecting shared background traits such as household aspirations, neighborhood influences, or residential location, which would lead $i$ to make a similar choice as $j$ regardless of the primary school she attends.[19] Finally, $\delta_{ij}$ corresponds to the effect of the primary school that $i$ and $j$ attend, reflecting school-level traits such as shared institutional practices or geographic proximity. School effects, $\delta_{ij}$, may take several forms. For example, students in the same school may receive similar recommendations from counselors or teachers, which will make them more likely to choose the same high school relative to students in different schools. Letting $A_i, A_j \in \mathcal{S}$ and $T_i, T_j \in \mathcal{T}$ denote the primary schools and cohorts of students $i$ and $j$, respectively, this implies that:

$$\mathbb{E}[\delta_{ij}|A_i = A_j,\, T_i = T_j] \geq \mathbb{E}[\delta_{ij}|A_i \neq A_j,\, T_i \neq T_j]$$

The raw difference in outcomes between peers and non-peers does not identify peer effects. To see this, note that we can decompose this difference as follows:

$$\begin{aligned} \mathbb{E}[y_{ij}|A_i = A_j, T_i = T_j] - \mathbb{E}[y_{ij}|A_i \neq A_j, T_i \neq T_j] = & \\ \underbrace{\gamma}_{\substack{\text{Peer} \\ \text{effects}}} + & \underbrace{\mathbb{E}[\eta_{ij}|A_i = A_j, T_i = T_j] - \mathbb{E}[\eta_{ij}|A_i \neq A_j, T_i \neq T_j]}_{\text{Selection bias}} \\ + & \underbrace{\mathbb{E}[\delta_{ij}|A_i = A_j, T_i = T_j] - \mathbb{E}[\delta_{ij}|A_i \neq A_j, T_i \neq T_j]}_{\text{Bias due to school effects}} \end{aligned} \tag{2}$$

[19] $\eta_{ij}$ also accommodates the possibility that some students have similar preferences over high schools due to common exposure to media, role models, or local norms.

The first term, $\gamma$, represents the causal effect of studying with student $j$ on the probability that student $i$ chooses the same high school. The second term captures selection bias: similar students, such as those living in the same neighborhood, are likely to attend the same primary school, and choose the same high schools. Finally, the third term reflects bias due to peers' exposure to the same schools. Students who attend the same primary school may be exposed to common institutional influences, such as recommendations from teachers or counselors, or school-wide norms which make them likelier to choose the same high school relative to students attending different primary schools.[20]

*Overcoming Selection with Admission Lotteries* Starting in 2016, Chile introduced the SAE platform which centralized admission into all schools. Under the new admission system, students who apply to an oversubscribed primary school – whom I refer to as *applicants* – are admitted by a random draw. As a result, applicants to the same oversubscribed primary school are, in expectation, just as similar to other students in the school – whom I refer to as *incumbents* – regardless of whether they are admitted.

Admission by lottery implies that any difference in outcomes between admitted and non-admitted applicants and incumbents are the result of admission into a primary school. Letting $S_i \in \mathcal{S}$ denote an applicant's preferred primary school, the expected similarity between applicant $i$ and student $j$ is independent of $i$'s admission outcome if $S_i$ is oversubscribed:

$$\mathbb{E}[\eta_{ij} \mid A_i, A_j = S_i] = \mathbb{E}[\eta_{ij} \mid A_j = S_i] \tag{3}$$

Formally, the difference in outcomes between admitted applicants and incumbent students at primary school $S_i$, and non-admitted applicants and incumbent students is given by:

$$\begin{aligned} &\mathbb{E}[y_{ij} \mid A_i = S_i, A_j = S_i] - \mathbb{E}[y_{ij} \mid A_i \neq S_i, A_j = S_i] \\ &\qquad = \gamma \; + \; \mathbb{E}[\delta_{ij} | A_i = S_i, A_j = S_i] - \mathbb{E}[\delta_{ij} | A_i \neq S_i, A_j = S_i] \end{aligned} \tag{4}$$

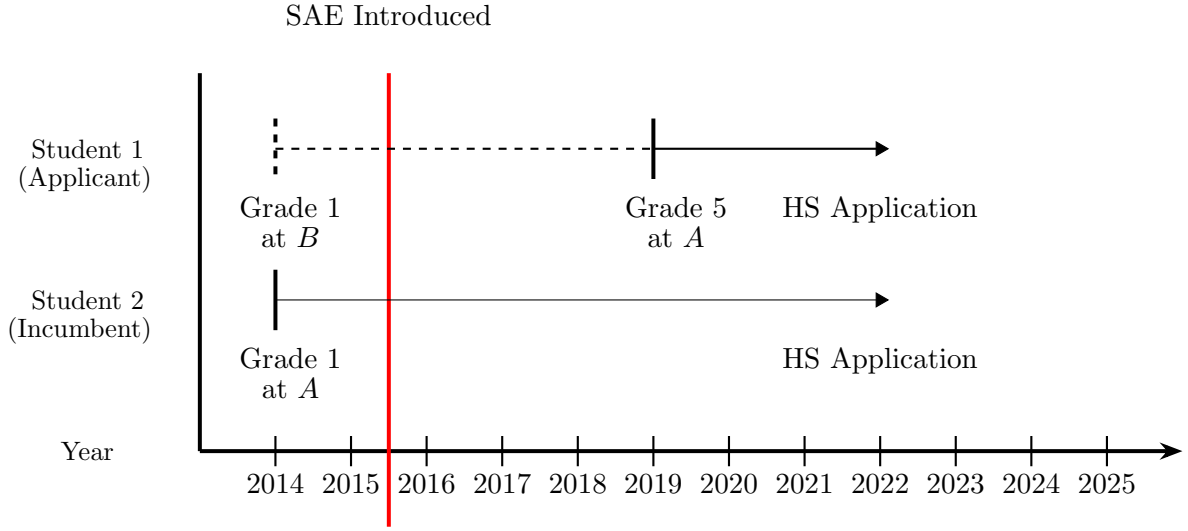
As before, $\gamma$, captures the causal effect of studying with student $j$. That is, the extent to which studying with student $j$ influences the similarity in $i$'s choice of high school. The second term reflects bias arising from school effects, corresponding to the additional similarity in choices that results from admitted applicants attending the same primary school as incumbents, independently of whether they influence one another. Notice, however, that relative to the raw comparison of means between peers and non-peers in Equation 2, the term corresponding to selection bias is gone.

Since students attend primary school for 8 years, cohorts who began primary school after the 2016 introduction of the SAE have not yet transitioned into high school. As a result, I identify applicants by focusing on students who *move* between primary schools after the introduction of the

[20] In the peer effects literature, this is often referred to as the issue posed by unobserved "correlated effects" (e.g. Barrios-Fernandez, 2023).

SAE system.[21] To fix ideas, Figure 2 presents an example of two students. Student 1, in the top row, began primary school in 2014 at School B, and transferred to School A in 2019, in the 5th grade. Since the transfer occurred after the introduction of the SAE, they were admitted through the centralized admission system and so I classify them as an *applicant*. Student 2, in the bottom row, was admitted to school A in 2014 and did not transfer schools. Since this preceded the introduction of the SAE, I classify the student as an *incumbent*.

Figure 2: Applicants and Incumbents at school A




*Disentangling Peer and School Effects* Separately identifying peer effects from school effects requires variation in exposure to different peers, while holding constant features of the school environment that influence students' application behavior. When school effects are constant from one year to the next, peer effects can be identified by using students in different cohorts as "counterfactual" incumbents – students who are exposed to the same school environment but do not overlap with admitted applicants.

There are several reasons to believe that school effects are plausibly constant within schools. First, primary school inputs are remarkably stable from year to year: the median school's level of funding, number of teachers, teaching hours, and enrollment change by less than 1% from one cohort to the next. Correspondingly, changes in these inputs across cohorts within a school are small compared to differences between schools. Within-school variation explains less than 1% of the total, while between-school variation explains about 98 to 99%. Teacher assignments also remain highly stable. Among students who completed primary school in 2023, 99.8% shared at least one teacher with students who graduated in 2022, and 80.8% shared at least half of their teachers. Overall, within a school, 64.3% of all teachers who taught students in 2023 also taught students in 2022. Appendix Table A2 presents estimates of within-school changes in more detail.

[21] In the case that a student moves between primary schools more than once, I keep data only about the last primary school to which they have applied.

Formally, imposing that school effects are constant over time within a school is equivalent to the following assumption:

> **Assumption FE (Fixed Effects):** The expectation of school effects, $\delta_{ij}$ is additively separable in schools and cohorts. That is, for any two students $i$ and $j$:
>
> $$\mathbb{E}[\delta_{ij}|A_i, A_j, T_i, T_j] = f_A(A_i, A_j) + f_T(T_i, T_j) \quad \forall A_i, A_j \in \mathcal{S},\ T_i, T_j \in \mathcal{T}$$

Assumption FE allows primary schools to differ in how they shape students' preferences but requires that these effects remain constant across cohorts within a given school. It is important to emphasize what Assumption FE does *not* impose. First, it does not restrict how students select into primary schools or how selection evolves over time. In particular, I do not assume that students' preferences over high schools are constant across cohorts within a primary school. For example, a change in a primary school's reputation from one year to the next could attract a different mix of applicants, making current students more similar to those in future cohorts than to those in earlier ones. Second, Assumption FE does not constrain how students in earlier cohorts influence the choices of later cohorts. For instance, if graduates from a primary school are admitted to a particularly prestigious high school, subsequent cohorts may adjust their applications in response. Assumption FE remains agnostic about the specific high schools to which students apply, imposing only that any cohort-to-cohort influences be stable over time. Finally, it does not require that school effects be homogeneous across pairs of students, only that their average-level does not vary with the interaction of their schools and cohorts.

To see how peer effects are identified under Assumption FE, consider the following four averages:

(1) $T_1$ : Average outcome of an applicant to primary school $S_i$, who is admitted, and an incumbent student in the same cohort and primary school.

$$T_1 := \mathbb{E}[y_{ij} \mid A_i = S_i, A_j = S_i, T_i = T_j]$$

(2) $T_0$ : Average outcome of an applicant to primary school, $S_i$ who is admitted, and an incumbent student in a *different cohort* within the same primary school.

$$T_0 := \mathbb{E}[y_{ij} \mid A_i = S_i, A_j = S_i, T_i \neq T_j]$$

(3) $C_1$ : Average outcome of an applicant to primary school $S_i$, who is *not admitted*, and an incumbent student in a the same cohort at primary school $S_i$.

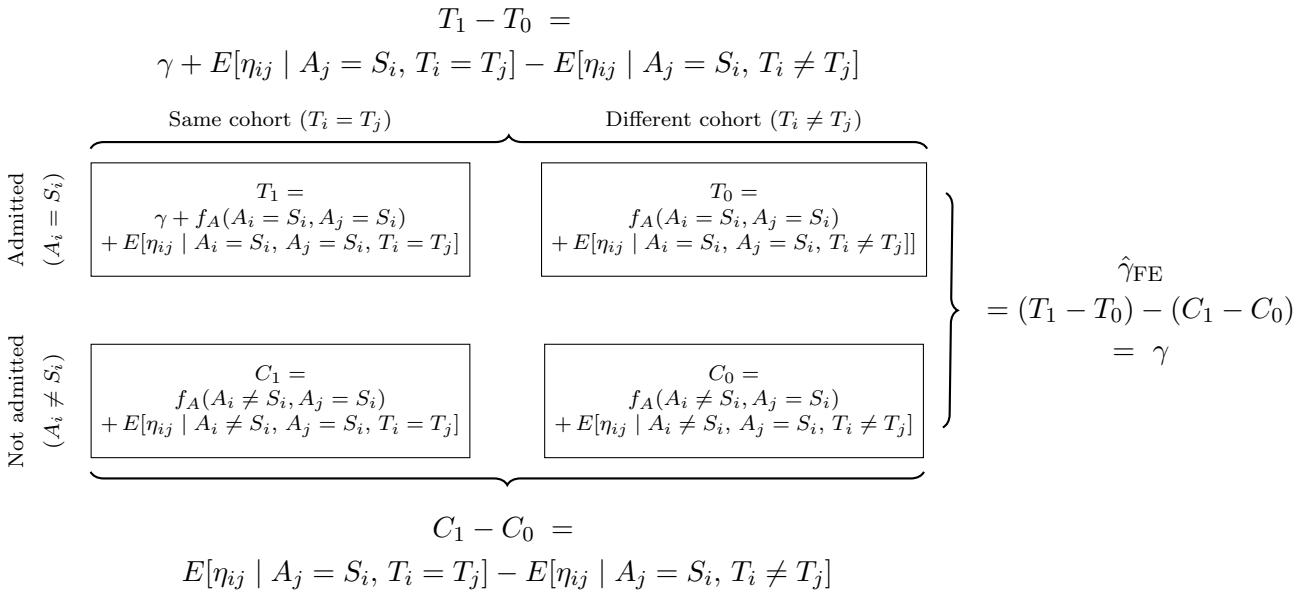
$$C_1 := \mathbb{E}[y_{ij} \mid A_i \neq S_i, A_j = S_i, T_i = T_j]$$

(4) $C_0$ : Average outcome of an applicant to primary school $S_i$ who is not admitted and an incumbent student in a different cohort at primary school $S_i$.

$$C_0 := \mathbb{E}[y_{ij} \mid A_i \neq S_i, A_j = S_i, T_i \neq T_j]$$

Under Assumption FE, the effect of attending school $S_i$ – independently of whether $i$ and $j$ are peers – on $y_{ij}$ is constant. As a result, comparing the outcomes of applicants admitted to primary school $S_i$ and incumbents in the same cohort, with the outcomes of admitted students and incumbents in a different cohort, $T_1 - T_0$, cancels the school effect. I illustrate this in the top row of Figure 3.[22]

Figure 3: Difference-in-Differences Illustration




Students in different cohorts, within a primary school, may nevertheless have systematically different unobserved preferences. However, since admission into oversubscribed schools is random among applicants, differences in the similarity of an applicant and an incumbent student's preferences in primary school $S_i$ across any cohort are the same whether or not the applicant is admitted. As a result, comparing the outcomes of an applicant who was not admitted to school $S_i$, and an incumbent from the same cohort versus an incumbent from another cohort, $C_1 - C_0$, identifies any differences in the similarity of applicants and incumbents across cohorts. I illustrate this in the bottom row of Figure 3. Putting this together, I show formally in Appendix A3.1 that the difference-in-differences estimator:

$$\hat{\gamma}_{\text{FE}} = (T_1 - T_0) - (C_1 - C_0)$$

identifies peer effects, $\gamma$.

## 3.2 Estimation sample

My estimation sample is constructed from the universe of students who are observed applying to *primary schools* through the SAE platform between 2016 and 2023 and who are later observed

[22] For simplicity, I set aside cohort effects, $f_T(T_i, T_j)$, noting that their presence does not change the substance of the analysis.

applying to *high schools* between 2022 and 2024. I restrict my sample of applicants to students who are randomized into primary schools. In practice, identifying these students is challenging. The probability that a student is admitted into a school depends on their eligibility for a quota seat and priority in admission. Overall, the system defines 20 different groups of students, based on eligibility for a quota seat and priority status, each of which may be subject to randomization in admission.

I construct my sample of applicants who are randomized into primary schools in two steps. First, I exclude all applicants with priority in admission, such as those with siblings already enrolled at the school. Since these students are first in line for admission, they are rarely subject to randomization.[23] I then partition the remaining students into four groups defined by eligibility for a particular quota, and identify groups subject to randomization. For each group $g$, I compare the number of students who rank a primary school as their preferred option to the number of reserved quota seats, after excluding applicants with priority.[24] If the number of quota seats exceed applicants in group $g$, all eligible applicants are admitted through the quota, and any surplus seats are reallocated. If eligible applicants exceed reserved seats, a random subset is admitted through the quota while the remainder compete for regular admission seats. I then pool eligible applicants who were not admitted through quotas, with regular applicants and compare their number to the number of available regular seats, net of those occupied by students with priority. If demand exceeds supply, admission to regular seats is randomized; otherwise, all applicants are admitted.

I illustrate this in Table 2. In school $A$, 3 low-income applicants fill all quota seats reserved for low-income students. Among 5 high-achieving applicants, 3 are admitted under the quota, while the remaining 2 compete with 4 regular applicants for 5 regular seats. Because the number of applicants competing for regular seats exceeds the number of available seats, both high-achieving and regular applicants are randomized in admission. While high-achieving students face a higher probability of admission than regular students, all applicants within the same group share the same ex-ante probability of admission. In school $B$ the situation is different. Only 1 low-income student applies and is admitted through a quota seat reserved for low-income students. This leaves 2 low-income seats unfilled, which are reallocated to regular students. As in school $A$, 5 high-achieving applicants apply, and 3 are admitted through quota seats, leaving 2 to compete for regular seats. However, in school $B$, 6 applicants (4 regular and 2 high-achieving) compete for 7 regular seats. Since available seats exceed the number of applicants, all are admitted, meaning that no group of students was randomized.

*Sample* Overall, my estimation sample is composed of 15,242 applicants across 1,370 primary schools. Of these, 13,216 compete for regular admission seats, 1,921 compete for quotas reserved

[23] I also exclude from my sample students who apply to a high school together with a sibling, as these students likely make a joint decision which is beyond the scope of my framework.

[24] Restricting attention to applications to students' preferred schools simplifies the identification of groups facing equal admission probabilities. Because the SAE employs a deferred acceptance algorithm, admission probabilities for second-choice schools depend on the first-choice ranking, which complicates the identification of groups with common probabilities.

Table 2: Illustration of Randomized Applicants

| | School $A$ | | | School $B$ | | |
|---|---|---|---|---|---|---|
| Group | Applicants | Reserved Seats | Randomized | Applicants | Reserved Seats | Randomized |
| Low-income | 3 | 3 | No | 1 | 3 | No |
| High achieving | 5 | 3 | Yes | 5 | 3 | No |
| Regular | 4 | 5 | Yes | 4 | 5 | No |

for low-income students, and a small residual group for quotas set for high-achieving students. Applicants apply to primary schools between 2016 and 2023, and again for high schools between 2022 to 2024. Appendix Table A3 provides a detailed breakdown of students by their quota type, the grade at which they are admitted, and the year at which they are admitted into primary school. Appendix Table A4 compares the estimation sample to the overall population of primary school students. Across most dimensions, applicants are remarkably similar to other students.[25] To verify that admission is random among students in my sample, Appendix Table A5 compares the characteristics of admitted and non-admitted applicants. Reassuringly, there are no statistically significant differences between the two groups.

### 3.3 Empirical specification

Assumption FE motivates the following difference-in-differences specification. For each applicant $i$ who ranked primary school $S_i$ as their preferred choice in cohort $T_i$, I consider all incumbent students $j$ who attended school $S_i$ in any cohort. I then regress $y_{ij}$ on an indicator for whether applicant $i$ was admitted to school $S_i$; an indicator for whether student $j$ is in the same cohort as student $i$; the interaction between the two, which corresponds for whether they are peers; and fixed effects for each primary school $S_i$, cohort $T_i$, and applicant group $g_i$ combination. Controlling for an applicant's school, cohort, group ensures that I am only performing comparisons among applicants who face the same probability of admission into a primary school.

$$y_{ij} = \beta_1 \cdot \mathbf{1}\{A_i = S_i\} + \beta_2 \cdot \mathbf{1}\{T_i = T_j\} + \gamma_{FE} \cdot \mathbf{1}\{A_i = S_i,\ T_i = T_j\} + \theta_{S_i,Ti,g_i} + \epsilon_{ij} \tag{5}$$

Here, $\gamma_{FE}$ captures the effect of studying with incumbent $j$ on the probability that applicant $i$ chooses the same high school. $\beta_1$ reflects the institutional effects of attending school $S_i$ that are common across time, while $\beta_2$ captures cohort-wide influences in year $T_i$. The term $\theta_{S_i,T_i,g_i}$ absorbs fixed differences across students in different admission groups who apply to different primary schools in different years.

Since only 82% of admitted applicants ultimately enroll in their most preferred primary school,

[25] The main differences between applicants and all students are that applicants, on average, live farther from the primary schools they attend and submit longer ROLs when applying to high school. The latter likely reflects their greater familiarity with the centralized application system.

a fixed effects regression only identifies intent-to-treat estimates. As a result, I complement my primary specification with an instrumental variables specification. In the first stage, I instrument enrollment at an applicants' preferred school with an indicator for whether the applicant was (randomly) admitted. I instrument for whether two students are peers with an indicator for whether a student was admitted to their preferred school in the same cohort as a given incumbent. In the second stage, I re-estimate the fixed effect regressions in Equation 5, replacing the dummies corresponding to whether applicant $i$ is admitted into incumbent $j$'s school and whether they are peers with the fitted values from the first stage. Accordingly, the instrumental variables regression identifies a local average treatment effect.

### 3.4 Results

I report estimates from Regression 5 in Table 3. Panel (A) corresponds to the effect on the probability that an applicant enrolls in the same high school as an incumbent student. Column (1) corresponds intent-to-treat estimates using admission to the same school as an incumbent as the key regressor. Column (2) corresponds to local average treatment effect, instrumenting for attendance in the same school as an incumbent with whether an applicant is admitted. I find that applicants who attend (are admitted) to their preferred primary school are 5.4 (3.2) percentage points more likely to be to enrolled in the same high school as an incumbent student.

I then decompose this total effect into three components in columns (3) and (4): (i) the effect of attending the same school as an incumbent, (ii) the effect of attending (any school) in the same cohort as an incumbent, indicating the presence of time varying selection, and (iii) the effect of attending the same school and cohort as an incumbent student. While attending the same primary school matters, studying in the same school and cohort – i.e., being the peer of an incumbent – produces a markedly larger increase in the probability of enrolling in the same high school. Attending the primary school in the same cohort as an incumbent student increases the probability that an applicant enrolls in the same high school by 1.2 percentage points. Interestingly, there is no effect from studying in the same cohort as an incumbent, suggesting that the underlying similarity in students' preferences is constant over time. Panel (B) re-estimates Regression 5, using whether an applicant ranks the same high school as their most preferred option as an outcome. I find that studying with an incumbent increases the probability that an applicant ranks the same high school as their preferred option by 1 percentage point.

I plot the estimates of the effect of studying together with an incumbent across multiple outcomes in Figure 4. Studying with a particular peer increases the probability that an applicant attends the same high school as the incumbent by 4 percentage point. This effect is *on top* of the effect of studying together in the same primary school, and is driven exclusively by exposure to particular incumbents. Examining applicants' ranking over high schools confirms that this pattern is driven by changes in applicants' behavior: exposure to an incumbent increases the probability to enroll in the same high school by 1.5 percentage points, the probability to be admitted into the same high school by 3.9 percentage points, the probability to rank the same high school as an

Table 3: Regression Results

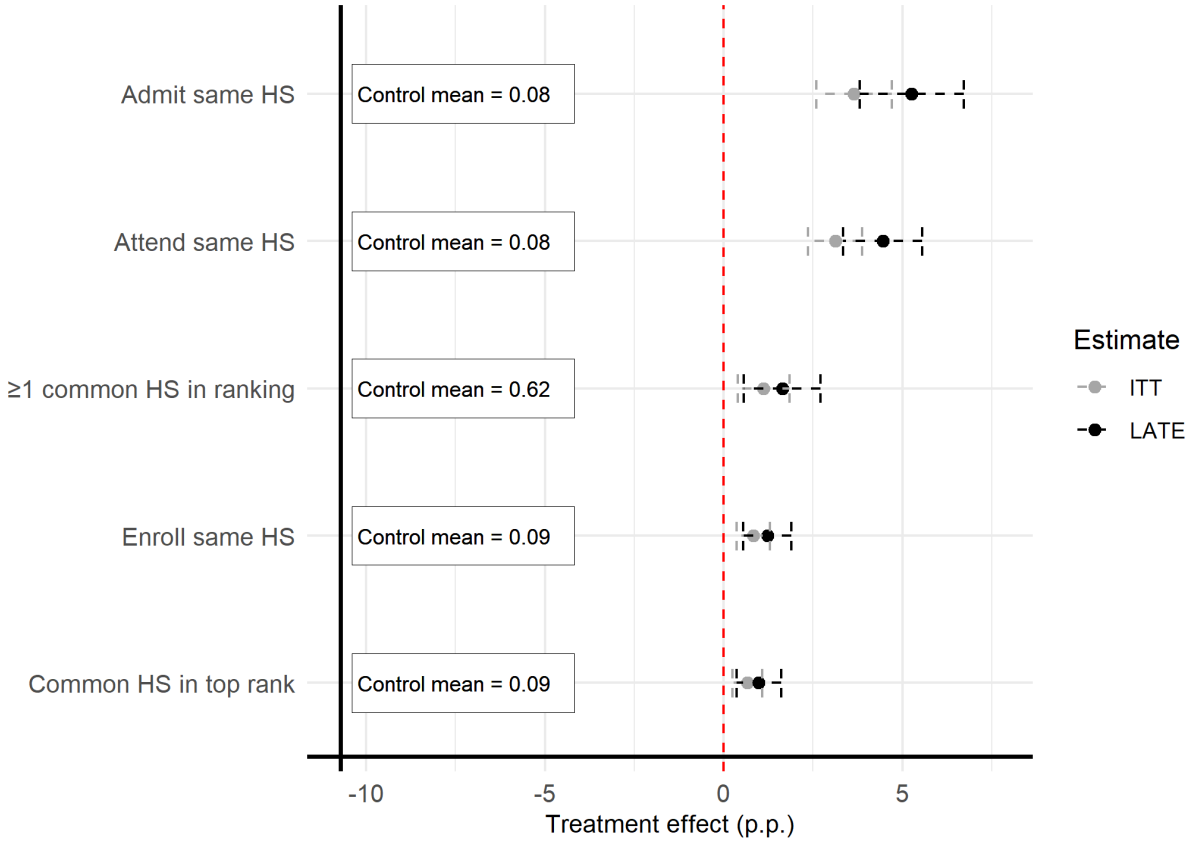
| | Panel A: Enrolled same HS | | | |
|---|---|---|---|---|
| | (1) ITT | (2) LATE | (3) ITT | (4) LATE |
| *Variables* | | | | |
| Admitted top PS | 0.0320*** | 0.0539*** | 0.0306*** | 0.0518*** |
| | [0.0204; 0.0436] | [0.0344; 0.0733] | [0.0191; 0.0422] | [0.0324; 0.0713] |
| Same year | | | 0.0008 | $-9.34 \times 10^{-5}$ |
| | | | [-0.0018; 0.0034] | [-0.0030; 0.0028] |
| Peers | | | 0.0083*** | 0.0122*** |
| | | | [0.0037; 0.0129] | [0.0055; 0.0189] |
| *Fixed-effects* | | | | |
| Top-ranked PS × Cohort × Group | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | |
| # Top-ranked PS × Cohort × Group | 3,427 | 3,427 | 3,427 | 3,427 |
| Observations | 1,909,872 | 1,909,872 | 1,909,872 | 1,909,872 |
| F-test (1st stage), Attended top PS | | 1,179,182.2 | | 393,069.7 |
| F-test (1st stage), Same year | | | | $5.2 \times 10^{36}$ |
| F-test (1st stage), peers | | | | 1,174,951.7 |
| $R^2$ | 0.17970 | 0.18260 | 0.17975 | 0.18269 |

| | Panel B: Top-rank same HS | | | |
|---|---|---|---|---|
| | (1) ITT | (2) LATE | (3) ITT | (4) LATE |
| *Variables* | | | | |
| Admitted top PS | 0.0196*** | 0.0329*** | 0.0185*** | 0.0313*** |
| | [0.0071; 0.0320] | [0.0117; 0.0542] | [0.0059; 0.0311] | [0.0098; 0.0528] |
| Same year | | | 0.0012 | 0.0004 |
| | | | [-0.0019; 0.0043] | [-0.0030; 0.0039] |
| Peers | | | 0.0067*** | 0.0099*** |
| | | | [0.0025; 0.0108] | [0.0036; 0.0161] |
| *Fixed-effects* | | | | |
| Top-ranked PS × Cohort × Group | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | |
| # Top-ranked PS × Cohort × Group | 3,427 | 3,427 | 3,427 | 3,427 |
| Observations | 1,944,540 | 1,944,540 | 1,944,540 | 1,944,540 |
| F-test (1st stage), Attended top PS | | 1,195,339.0 | | 398,452.3 |
| F-test (1st stage), Same year | | | | $6.29 \times 10^{36}$ |
| F-test (1st stage), peers | | | | 1,210,746.2 |
| $R^2$ | 0.21481 | 0.21636 | 0.21485 | 0.21642 |

Notes: This table presents estimates from running Regression 5 with different outcome variables. The outcome in panel (A) corresponds to whether an applicant enrolls in the same high school as an incumbent student in their preferred primary school. The outcome in Panel (B) corresponds to whether an applicant ranks the same high school as an incumbent as their most preferred option. Column (1) presents the estimated effect of being admitted into the applicant's preferred school. Column (2) presents the estimated effect of attending the applicant's preferred school, where attendance is instrumented for with whether the applicant is admitted. Column (3) presents the estimated effect of being admitted into an applicants' preferred school, studying in the same cohort, and the interaction of the two, representing the ITT effect of studying with an incumbent student. Column (4) presents the estimated effect of attending an applicant's preferred school, studying in the same cohort, and the interaction of the two, where these are instrumented by whether an applicant receives an offer of admission, whether they are in the same cohort as an incumbent, and the interaction of whether they receive an offer of admission and whether the applicant and incumbent are in the same cohort. All regressions include fixed effects for the interaction between an applicant's preferred primary school, cohort and application group. Clustered (Top-ranked PS × Cohort × Group) covariance matrix, 95% confidence intervals in brackets Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

applicant's most preferred option by 1.1 percentage points, and increases the fraction of schools appearing on both students' ranked lists by 1.3 percentage points.

Figure 4: Treatment effects of studying with an incumbent




Notes: Each row of this figure represents the estimated coefficient on the dummy variable corresponding to whether an applicant and an incumbent at the applicant's preferred primary school are peers in Regression 5. All regressions control for whether the applicant is admitted to their preferred primary school. Estimates in black correspond to the estimated coefficients of an instrumental variables regression, where attending an applicants' preferred primary school, and attending with a particular incumbent are instrumented by whether an applicant receives an offer of admission, and the interaction of whether they receive an offer of admission and whether the applicant and incumbent are in the same cohort, together with fixed effects for an applicants' preferred primary school, year and group. Estimates in gray correspond to the estimated coefficients of a fixed effects regression, where the primary variables are whether a student is offered admission into their preferred primary school and the interaction between this and a dummy variable for whether the applicant and incumbent are in the same cohort, together with fixed effects for an applicants preferred primary school, year and group. Error bars represent 95% confidence intervals.

To better understand the magnitude of these estimates I compare them to a population-wide benchmark: what is the probability that an applicant and incumbent pair, across any primary school, attend the same high school? Based on a random pair of 5000 applicants and incumbents, the likelihood that a randomly matched pair selects the same high school is less than 0.1 percentage points. In comparison, attending the same primary school increases this probability by

1.5 percentage points – over fifteen times larger. Moreover, assignment as a peer of an incumbent further increases the probability by 4 percentage points, a further forty-fold increase relative to the population-wide average.

Assumption FE may not hold in practice for a number of reasons. For instance, small changes in teacher composition may nonetheless make a meaningful impact on students' choices. Alternatively, some high schools may become more or less popular for students in different primary schools over time, inducing a correlation in the choices of students who in the same primary school cohort. In Appendix Section A5, I study the robustness of my results by explicitly controlling for variation in teacher composition across cohorts. I also employ an alternative identification strategy that exploits variation in peer exposure generated by the assignment of students into classes *within* a primary school cohort. In both analyses, I find that studying with a particular *peer* increases the probability that a student ranks and ultimately attends the same high school.

So far, I have shown that studying with a particular peer affects students' choice of high school. This effect, however, may be driven by students reacting to their peers' behavior or to their peers' characteristics. For example, a peer who lives near a particularly good high school may share information about the school, which would lead both the peer and student to apply. In the peer effects literature, these are referred to as *contextual peer effects*, whereas peer effects driven by a reaction to peers' behavior are referred to as *endogenous peer effects*. In linear models, contextual and endogenous peer effects are collinear and so cannot be disentangled from each other without instruments, a classical result referred to as the *reflection problem* (Manski, 1993; Angrist, 2014).[26] As a result, I revisit this question in Section 5, and use a structural model to separately identify contextual and endogenous peer effects, and assess the relative importance of each.

# 4 Peer Effects and Inequality in Education

In Section 3, I show that studying with a particular peer increases the likelihood that a student applies to and attends the same high school. It is not clear, however, whether or how this affects educational inequality.

In this section, I develop a simple conceptual framework that links peer effects, interpreted as a preference for studying with peers, to disparities in high school quality between low- and high-income students. The intuition is straightforward. If most of a student's peers live in a low-income neighborhood and have access primarily to low-quality high schools, they are likely to apply to these schools and, in turn, pull their peers with them. The same logic applies to students in high-income neighborhoods, leading these students to apply to high-quality high schools and amplifying inequality. Conversely, if students have good access to high-quality schools regardless of location, peer effects increase the likelihood that all students apply to better schools, raising overall attainment and reducing inequality. The qualitative impact of peer effects on aggregate attainment and educational segmentation in equilibrium is therefore not clear a priori, and depends on the

[26] See for instance, De Giorgi et al. (2010) for approach that exploits partially overlapping peer groups to isolate variation in peers' choices and characteristics at the individual level.

underlying spatial distribution of school quality, students' preference for quality, and distaste for distance.

Consider the choice of student $i$, currently studying in primary school $s$, over which high school to attend. For simplicity, assume that there are no capacity constraints so that every student gets into their chosen school. Let the utility of high school $j$ be a linear function of its characteristics (e.g. distance) and the share of student $i$'s peers who, in expectation, attend the school:[27]

$$U_{ij} = \alpha q^j - \beta d_i^j + \gamma m_s^j + \nu_i^j \tag{6}$$

where $m_s^j$ corresponds to the expected share of students from primary school $s$ that will attend high school $j$; $d_i^j$ to the distance from $i$'s home to school $j$; $q^j$ to the quality of school $j$; and $\nu_i^j$ to an idiosyncratic error term capturing unobserved individual heterogeneity. I assume that $\nu_i^j$ follows a type-I extreme value distribution.

To illustrate how peer effects impact inequality, consider a simple example with two high schools and two primary schools. Assume that one high school is of low-quality, with quality normalized to 0, and another of high-quality, with quality 1. There are two groups of students, attending two distinct primary schools, $A$ and $B$. The main difference between the two groups is that students in primary school $B$ must travel farther to attend the high-quality high school relative to students in primary school $A$, that is, $d_B^H > d_A^H$.

Students apply to the school that maximizes their utility. I assume that students are aware of $q_j$ and $d_i^j$, but not the idiosyncratic utility of the school for other students, $\nu_i^j$. As such, students' expectation about the share of their peers who will attend school $j$ is equal to the average probability that a student from primary school $s$ applies to school $j$ given $q_j, d_i^j$ and $m_s^j$. This will constitute a Bayesian Nash equilibrium, in which each student applies optimally given the expected application behavior of their peers. Since I assume that $\nu_i^j$ follows an extreme-type I distribution, the share of students from primary school $s$ that will attend the high-quality high school satisfies:

$$m_s^H = \frac{\exp(\alpha - \beta d_s^H + \gamma m_s^H)}{\exp(\alpha - \beta d_s^H + \gamma m_s^H) + \exp(\beta d_s^L + \gamma(1 - m_s^H))} \tag{7}$$

In this case, the effect of increasing peer effects, $\gamma$, on the share of students who attend high school H will depend on the distance of students to the school. Specifically,

> **Result 1:** When $m_s^H$ is unique, there exists a unique *cut-off distance* to high school H, $\bar{d}$, such that peer effects will increase the share of students attending if they live closer than $\bar{d}$, and will decrease the share of students attending if they live farther away. That is:

$$\begin{cases} \frac{\partial m_s^H}{\partial \gamma} > 0 & \text{if } d_s^H < \bar{d} \\ \frac{\partial m_s^H}{\partial \gamma} < 0 & \text{o.w.} \end{cases} \tag{8}$$

[27] I relax this assumption in the structural model presented in Section 5.

The value of $\bar{d}$ is the distance that makes $m_s^H \geq 0.5$ for a given value of $\gamma$.

The proof can be found in Appendix A3.2.

Result 1 suggests that the relationship between peer effects and inequality depends on how far away the high-quality high school is from students. To see this, assume that students from primary school $A$ live sufficiently close to high school $H$ such that, absent any peer effects ($\gamma = 0$), the majority of students attend it ($d_A^H < \bar{d}$). Now suppose that students attending primary school $B$ also live sufficiently close to high school $H$ such that, without any peer effects, the majority of students in primary school $B$ attend it, ($d_B^H < \bar{d}$). In this case, peer effects increase the utility of high school $H$ for students in both primary schools $A$ and $B$, since more than half of their peers are expected to attend. If students choose high schools while taking into account their peers' choices ($\gamma > 0$), the previously marginal student in both primary schools would apply to high school $H$, increasing the average quality of the high schools that *all* students attend, and reducing educational inequality. I illustrate this in panel (a) of Figure 5.

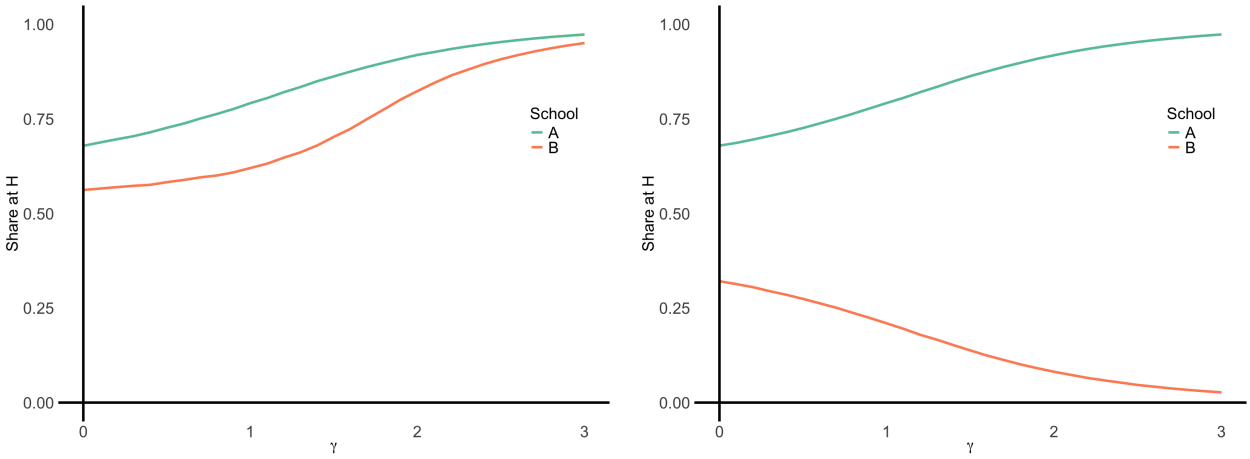
Now consider the opposite case in which, at baseline, the majority of students in primary school $B$ attend the low-quality high school, ($d_b^H > \bar{d}$), while the majority of students in primary school $A$ attend the high-quality high school ($d_a^H < \bar{d}$). In this case, peer effects increase the utility of the low-quality high school, $L$, for students in primary school $B$, since more than half of their peers are expected to attend it, while decreasing it for students in primary school $A$. If students choose high schools while taking into account their peers' choices ($\gamma > 0$), the previously marginal student in primary school $B$ would apply to high school $L$. Students in primary school, $A$, in contrast would become more likely to apply to high school $H$, increasing educational inequality. I illustrate this in panel (b) of Figure 5. Generalizing, it follows that:

> **Corollary 1:** If school quality is sufficiently unequally distributed across space, peer effects increase schooling inequality. That is:

$$\begin{cases} \frac{E(q|A)-E(q|B)}{\partial\gamma} \geq 0 & \text{if} \quad d_a^H \leq \bar{d}, \quad d_b^H \geq \bar{d} \\ \frac{E(q|A)-E(q|B)}{\partial\gamma} \leq 0 & \text{o.w.} \end{cases} \tag{9}$$

Although this framework has been agnostic about the characteristics of students, Section 2.3 documents that high-income students have, on average, substantially better access to high-quality schools near their homes. If spatial inequality in access to high-quality schools between low- and high-income students is sufficiently severe – relative to students' preference for quality and distaste for distance – peer effects will increase the likelihood that low-income students attend low-quality schools and that high-income students attend high-quality ones, thereby increasing the gap in high school quality between students from rich and poor households.

Figure 5: Simulated share of students attending high-quality school




(a) Low inequality

(b) High-inequality

Note: The figure plots the estimated share of students from two primary schools that will attend the high-quality high school as a function of the magnitude of peer effects, $\gamma$. The orange line corresponds to the share of students from primary school a that will attend the high-quality high school, while the teal line corresponds to the share of students from primary school b that will attend the high-quality high school. Plot (a) corresponds to the case where inequality is low, and a majority of students from primary school b apply to the high-quality high school without peer effects. Plot (b) corresponds to the case where inequality is high, and a majority of students from primary school b apply to the low-quality high school without peer effects.

# 5 A Model of School Choice with Peer Effects

In this section, I develop an empirical model of demand for high schools which incorporates social interactions among primary school peers. My goal is to characterize how students trade off attributes of schools, such as distance, with a preference to study together with their peers.

## 5.1 Set-up

A student $i$ is characterized by the primary school which they currently attend, $s \in \mathcal{S}$, and the cohort at which they apply to high school, $t \in \mathcal{T}$. The utility of attending high school $j \in \mathcal{J}$ is a linear function of the school's observable and unobservable characteristics. A school's observable characteristics, $X_{ij}$, capture attributes such as its quality, the distance between each student and a school, and whether a student has a parent or a sibling at the school. To capture additional unobserved reasons students may prefer school $j$ over other schools in their choice set, I model an unobserved component of utility, $u_{ij}$. This term includes school attributes that are unobserved to me but relevant to a student's choice, such as its reputation, sport facilities or physical infrastructure.

I make three departures from a standard school choice model. First, I allow students to have a preference to study together with their peers. I model the utility of a high school as depending on students' belief about the share of peers who will be admitted to school $j$, $\hat{m}_{ij}$. This type of peer

effects is often referred to as endogenous peer effects. Second, I allow students' preferences over high schools to depend on their primary school peers' characteristics, $\bar{X}_{stj}$, referred to as contextual peer effects. I incorporate two types of contextual effects into my model: (1) effects driven by exposure to particular characteristics of peers, for instance, peers who live near a high school and may share information about the school; and (2) effects driven by exposure to a particular group of peers, for example, exposure to high-income students which may encourage a student to apply to high schools with more high-income students.

Finally, I allow students' unobserved preferences over high schools, $u_{ij}$, to vary by the primary school they attend. This captures two important features of my environment. First, since I focus on all primary school students, not just randomly selected applicants, students in the same primary school likely share similar (unobserved) background characteristics, and may therefore have similar (unobserved) preferences over high schools. Second, the choice of students in the same primary school is likely to be influenced by the primary school environment itself. For instance, students may interact with the same teachers or counselors who would impact their preferences over high schools. This suggests that students' unobserved preferences are likely to be correlated within a primary school. To capture this, I assume that unobserved preferences for a particular high school are independently and identically drawn from a distribution with a primary school-specific mean:

$$u_{ij} \sim Gumbel(\delta_{stj}, 1)$$

Putting this together, I write the utility of attending high school $j$ as:

$$U_{ij} = \underbrace{\beta X_{ij}}_{\substack{\text{Preference} \\ \text{for observable} \\ \text{school characteristics}}} + \underbrace{\gamma_0 \hat{m}_{ij}}_{\substack{\text{Endogenous} \\ \text{peer} \\ \text{effects}}} + \underbrace{\gamma_1 \bar{X}_{stj}}_{\substack{\text{Contextual} \\ \text{peer} \\ \text{effects}}} + \underbrace{u_{ij}}_{\substack{\text{Unobserved} \\ \text{preferences}}} \tag{10}$$

*Beliefs* I model the interaction among students as a game of incomplete information. I assume that each student has complete information about their own characteristics, $X_{ij}$ and $u_{ij}$, as well as the average unobserved preference for a high school within their primary school, $\delta_{stj}$. I also assume that students know the probability of admission into each high school.[28] Unlike a coordination game, I do not assume that students know the exact choices of their peers, their characteristics, $X_{kj}$, or their idiosyncratic preferences for a particular school $u_{kj}$. Instead, students know the distribution of their peers' characteristics $H(X_{kj})$, and unobserved preferences.[29]

[28] While the assumption that students have correct information about model primitives is a common assumption in many school choice models (e.g., Bau, 2022; Allende et al., 2019; Nielson, 2021; Walters, 2018). More recent work focusing on primary school choice and university choice highlights the importance of beliefs and consideration sets (Agte et al., 2024; Fabre et al., 2024).

[29] I assume that individuals respond to the expected behavior of others rather than their realized behavior. This assumption is grounded in the idea that each individual observes their own random shock but not the shocks of others. An alternative framework proposed by Soetevent and Kooreman (2007) assumes that individuals observe the shocks of others, with shocks unobservable only to the econometrician. Under their framework, each individual's expectation of others' average behavior aligns with the realized average behavior in equilibrium. However, this approach becomes computationally infeasible in my setting due to the large number of students and schools, which leads to a large

I assume that, in equilibrium, students' belief about the share of their peers who will be admitted into each school is consistent with the actual admission outcome. That is, students' belief about the share of their peers who will be admitted into high school $j$ is given by:

$$\hat{m}_{ij} = \frac{1}{|G_i|} \sum_{k \in G_i} \mathbb{E}[\mathbb{1}\{A_k = j\} | X_{ij},\ \delta_{stj}] \quad \forall i, j \tag{11}$$

where $G_i$ corresponds to the set of students who are i's peers and $A_k$ corresponds to the school to which student $k$ is eventually admitted.

Since students' idiosyncratic preferences, $u_{ij}$ are independent of one another, and all students know the distribution of their peers' characteristics and preferences, students have no private information that would lead them to hold any other belief than the common belief about the share of students who will be admitted to a school. This belief simply corresponds to the average probability of admission into school $j$ for students within primary school $s$ at time $t$:

$$\hat{m}_{stj} = \int \mathbb{P}(A_i = j) dH(X_{ij}) \quad \forall s, t, j \tag{12}$$

*Ranking Probability* I assume that students' observed rank-ordered lists (ROLs) accurately reflect their true preference over schools. This follows from the strategy-proof nature of the student-proposing deferred acceptance (DA) mechanism used in the assignment process (Correa et al., 2022). Under DA, students are always assigned to the highest-ranked school in their list in which there is a seat available. Crucially, misrepresenting preferences by ranking a less preferred school in a higher rank does not increase the chance of being admitted to a preferred school. Given this, students have no incentive to strategically distort their rankings based on admission probabilities – allowing observed ROLs to be interpreted as direct measures of preference.

In my framework, the presence of peer effects does not alter this property. Students form a belief about the share of their peers who will be admitted into each high school that is independent of students' actual choices. As such, students' belief about the share of their peers who will be admitted into a high school is no different from any other characteristic of a high school. Since students cannot influence their peers' behavior by their *actual* choices, there is no incentive to strategically manipulate rankings.[30]

The school that student $i$ ranks as their most preferred option, denoted by $R_i(1)$, is the school that maximizes their utility within their choice set $J_i$:

$$R_i(1) = \arg\max_{j \in Ji} U_{ij}$$

Students may rank as many schools as they like, but they are not required to rank all available

number of equilibria.

[30] This stands in contrast to examples such as siblings coordinating their applications. In that case, since students primarily care about studying together and are aware of each others' choices, strategic ranking that accounts for admission probabilities is beneficial.

schools. To model the decision to only rank school up to position $r \geq 2$, I introduce an outside option, denoted by $\emptyset$, which represents the choice not to rank any schools. For any rank $r \geq 2$, students may select any school that has not been previously ranked or the outside option. The choice set of student $i$ at rank $r$, denoted by $S_r$, is therefore:

$$S_r = J_i / \{R_i(1), ..., R_i(r-1)\} \cup \emptyset\}$$

The school ranked in position $r > 1$ is the maximizer of students' utility within the restricted choice set:

$$R_i(r) = \arg\max_{j \in S_r} U_{ij} \quad \forall i$$

The probability that student $i$ ranks school $j$ as their most preferred option is given by:

$$\begin{aligned} \mathbb{P}(R_i(1) = j) &= F(\beta X_{ij} + \gamma_0 \hat{m}_{ij} + \gamma_1 \bar{X}_{stj} + \delta_{stj}) \\ &= \frac{\exp(\beta X_{ij} + \gamma_0 \hat{m}_{ij} + \gamma_1 \bar{X}_{stj} + \delta_{stj})}{\sum_{k \in J_i} \exp(\beta X_{ik} + \gamma_0 \hat{m}_{ik} + \gamma_1 \bar{X}_{stj} + \eta_{stk})} \end{aligned} \tag{13}$$

Correspondingly, the probability to observe a ROL, $R_i = \{R_i(1), ..., R_i(J_i - 1)\}$, is given by:

$$\mathbb{P}(R_i) = \prod_{r=1}^{J_i - 1} \mathbb{P}(R_i(r) = j_r | S_r), \tag{14}$$

where $j_r$ corresponds to the school ranked in position $r$.

*Admission Probability* A student's admission to a given high school depends not only on their own ranking but also on how other students rank schools, their priority at each school, and eligibility for a quota. Because students submit ROLs and school capacities are limited, the assignment process is shaped by the interactions of all applicants' preferences and priority-quota group. Letting $R_i$ denote the ROL of a student, let all possible rankings of schools by student $i$ be denoted by $\mathcal{R}_i$. Correspondingly, let the rankings submitted by all other students be denoted by $\tilde{R}_{-i} = \{R_1, ..., R_N\}$ and the set of all possible rankings submitted by other students by $\tilde{\mathcal{R}}_{-i}$ .

Let $\phi_{ij}(R_i, \tilde{R}_{-i})$ denote the probability that student $i$ is admitted into school $j$ if they submit ranking $R_i$ when other students submit rankings $\tilde{R}_{-i}$. The centralized admission system differentiates between students based on their priority-quota group at a high school. As a result, two students who submit the same rankings, have the same priority, and eligibility for a quota, will therefore have the same probability of being admitted into a school. I therefore write $\phi_{ij}(R_i, \tilde{R}_{-i}) = \phi_{pj}(R_i, \tilde{R}_{-i})$, where p denotes the priority-quota group of a student at school j.

Using the law of total probability, the probability that a student who submits ranking $R_i$ is admitted into a school can be re-written by taking the expectation over all possible rankings

submitted by other students:

$$\mathbb{P}(A_i = j|R_i) = \sum_{\tilde{R}_{-i} \in \tilde{\mathcal{R}}_{-i}} \phi_{pj}(R_i, \tilde{R}_{-i}) \mathbb{P}(\tilde{R}_{-i}|R_i) \tag{15}$$

where the sum is taken across all possible permutations of rankings submitted by all students other than $i$. Note that since students react to their peers' *expected* behavior, rather than realized rankings, it follows that, $\mathbb{P}(\tilde{R}_{-i}|R_i) = \mathbb{P}(\tilde{R}_{-i})$. For the same reason, this probability is equivalent to the product of individual ranking probabilities:

$$\mathbb{P}(\tilde{R}_{-i}) = \prod_{k \neq i} \mathbb{P}(R_k) \tag{16}$$

Putting this together, I write the unconditional probability that a student is admitted into a high school as follows:

$$\begin{aligned} \mathbb{P}(A_i = j) &= \sum_{R_i \in \mathcal{R}} \mathbb{P}(A_i = j|R_i) \mathbb{P}(R_i) \\ &= \sum_{R_i \in \mathcal{R}} \sum_{\tilde{R}_{-i} \in \tilde{\mathcal{R}}_{-i}} \phi_{pj}(R_i, \tilde{R}_{-i}) \left( \prod_{k \neq i} \mathbb{P}(R_k) \right) \mathbb{P}(R_i) \end{aligned} \tag{17}$$

where the sum is taken across all possible rankings that (1) a student, and (2) all other students can submit.

*Equilibrium* An equilibrium belief will always exist in my model. To see this, note that students' belief about the share of their peers admitted into each high school can be re-written as satisfying the following "belief-mapping function":

$$\Phi(\hat{m}_{stj}) = \int \mathbb{P}(A_i = j) dH(X_{ij}^{st}) \quad \forall s, t, j \tag{18}$$

This belief-mapping function is continuous and maps the interval [0,1] onto itself. As a result, Brouwer's fixed-point theorem implies that it must have at least one solution. Moreover, for all values of $\gamma_0$ that make the belief-mapping function a contraction, this solution will be unique. This allows me to prove the following result about equilibrium beliefs:

> **Result 2 (Equilibrium Existence and Uniqueness):** An equilibrium belief always exists. Furthermore, the equilibrium within a primary school cohort is unique if $\gamma_0 \leq \left(\bar{R} \cdot |G_{st}|\right)^{-1}$, where $\bar{R}$ corresponds to the longest possible ROL and $|G_{st}|$ to the size of the cohort.

The full proof can be found in Appendix A3.3. Note that this is a sufficient and not a necessary condition. In fact, for many parameterizations, a unique equilibrium will exist for values of $\gamma_0$ greater than $\left(\bar{R} \cdot |G_{st}|\right)^{-1}$. However, the result provides a universal bound that always guarantees

uniqueness.

*Supply of Schools* I treat the supply of schools as exogenous and allow it to vary over time. Since I focus exclusively on school that are organized as not-for-profit entities and are publicly funded, I do not consider strategic behavior on the part of schools.[31]

## 5.2 Identification

*Beliefs* Result 2 establishes that an equilibrium belief must exist. Since students will be correct about the share of their peers who are admitted to each school, it is possible to consistently estimate this belief with the sample share of students who are admitted into a school, $\bar{m}_{ij}$, defined as follows:

$$\bar{m}_{ij} = \frac{1}{|G_i|} \sum_{i \in G_i} \mathbf{1}\{A_i = j\} \tag{19}$$

For large values of $\gamma_0$ there may be multiple different beliefs about the share of students admitted into a high school that constitute an equilibrium within a given primary school. Note that my model is not equipped with any equilibrium selection mechanism. However, given $\hat{m}_{stj}$, $X_{ij}$ and $\delta_{stj}$, there is no need to solve for an equilibrium to predict students' ranking probabilities. This is because these variables uniquely pin down the (deterministic) utility of each school for every student and so it is not necessary to solve the equilibrium system expressed in Equation 12 to predict students' behavior. This is possible since only $\hat{m}_{stj}$ is endogenously determined in equilibrium, and can be consistently estimated using $\bar{m}_{ij}$. A similar approach is adopted by Bhattacharya et al. (2024) to estimate peer effects in the take up of mosquito nets in Kenya.[32]

*Endogenous Peer Effects* Endogenous peer effects, $\gamma_0$, are not identified from students' rankings without imposing additional structure. To see this, recall that the utility of a high school includes a term capturing the mean unobserved utility of the school for all students in a given primary school, $\delta_{stj}$. Because the share of peers admitted to each high school depends on $\delta_{stj}$, it is not possible to distinguish whether students rank a high school because their peers do so, or because of its mean unobserved utility. This is precisely the identification challenge discussed in Section 3.

To address this in the structural model, I impose an analogous assumption to Assumption FE, which identified peer effects in the reduced form model:

**Assumption S-FE:** The unobserved mean utility of each high school $j$ is constant

[31] A large literature has focused on the incentives underlying private schools' decision to provide a level of quality and charge a particular price. For instance, Nielson (2021) focuses on the strategic incentives of voucher schools in Chile to provide a given price and quality. Allende et al. (2019) builds on this framework to study on how private schools target students of different socioeconomic groups in Peru. The key difference relative to my context is that I focus on publicly funded schools that, since the introduction of the 2015 Schooling Inclusion Law, are forbidden from operating on a for-profit basis and therefore likely face less of an incentive to compete.

[32] An alternative approach relies on fixed-point iteration to endogenously determine beliefs and estimates structural parameters using the simulated beliefs (see, for e.g., Seim, 2006).

within a primary school $s$ across time. That is,

$$\delta_{stj} = \delta_{sj} \quad \forall s \in \mathcal{S},\ t \in \mathcal{T},\ j \in \mathcal{J}$$

Assumption S-FE allows me to identify $\gamma_0$ by exploiting variation within primary schools in the expected share of students who will be admitted into each high school across cohorts. By construction, it imposes that variation in the expected share of peers admitted into high schools across cohorts within a primary school must be driven by changes in the distribution of peers' observable characteristics, such as residential location, from one cohort to the next, which students take into account when forming a belief about the share of their peers who will be admitted into each high school.

In the context of this model, the mean unobserved utility of a high school reflects both school effects and unobserved similarity in the preferences of students who select into primary schools. In other words, I do not distinguish whether a high school is popular because teachers encourage students to apply, or because students share latent preferences for it. Assumption S-FE therefore imposes that both the causal effect of primary schools on high school choice is constant over time, and that students across cohorts within the same primary school share the same distribution of preferences over high schools. As such, it is stronger than Assumption FE, which identified peer effects in the reduced-form model presented in Section 3 by imposing that school effects are constant across cohorts within a primary school, without restricting the variation in students' unobserved preferences.

The validity of the second part of Assumption S-FE can be assessed using the difference-in-differences design described in Section 3. Under Assumption FE, differences across cohorts in the probability that an applicant and an incumbent choose the same high school – between (randomly) non-admitted applicants and incumbents at applicants' preferred primary school – reflect changes in students' high school preferences that are independent of the effect of primary schools. As such, by comparing whether non-admitted applicants are more likely to choose the same high school as an incumbent in their preferred primary school in the *same* cohort relative to incumbents in a *different* cohort, I can test for the existence of time-varying selection on high school preferences into primary schools. The corresponding estimate appears in row 2 of Table 3. The results indicate that the probability that a non-admitted applicant ranks the same high school as an incumbent is statistically indistinguishable across cohorts, providing strong evidence *against* the presence of dynamic selection into primary schools.

In Appendix A3.4.1, I show that under two additional technical assumptions: (i) that the distribution of students' observable characteristics is non-degenerate, and (ii) that it varies across cohorts within a school, $\beta$, $\gamma_0$, $\gamma_1$, and $\delta_{sj}$ are identified from the probability that each high school is ranked as a student's preferred option.[33]

[33]The proof is related to the proof of Brock and Durlauf (2007) that peer effects *cannot* be separately identified from unobserved group-level variables in discrete choice models. However, whereas the authors focus on identification with cross-sectional data, I show that longitudinal data allows peer effects to be identified if unobserved group-level

*Contextual Peer Effects* In my model, the (expected) share of peers that are admitted into a particular school is pinned down by Equation 12, a non-linear function of the distribution of peers' characteristics and of individual admission probabilities. As a result, across primary schools and cohorts, the expected share of peers admitted into different high schools does not vary *co-linearly* with peers' average characteristics. Instead, variation in the expected share of peers admitted into each high school is the result of variation in peers' characteristics – through a non-linear mapping – or of changes in admission probabilities (e.g., changes in school capacity) that are orthogonal to peers' characteristics. In Appendix A3.4.1, I show that variation from either source is sufficient to separately identify contextual peer effects, $\gamma_1$, from endogenous peer effects, $\gamma_0$.[34]

*Outside Option* I identify the deterministic utility of the outside option, which I denote by $V_0$, directly from the number of schools that a student ranks. Intuitively, conditional on knowing the characteristics of each student-school pair, the structural parameters and the fixed effects, the expected number of schools that a student ranks will depend only on the utility of the outside option. I prove this formally in Appendix A3.5.

### 5.3 Estimation

*Estimation Strategy* I estimate $\beta, \gamma_0, \gamma_1$ and $\delta$ using a Generalized Method of Moments approach, leveraging moments based on the difference between students' predicted and observed ranking over high schools. A key challenge to this approach, however, is the presence of high-dimensional fixed effects, which pose a significant computational challenge. For example, with only 100 primary schools and 50 high schools, estimating separate fixed effects for each primary school-high school combination requires estimating 5,000 parameters.

To address this issue, I treat fixed effects as nuisance parameters and estimate them in a two-step iterative procedure. In a first step, I treat the structural parameters, $\theta = \{\beta, \gamma_0, \gamma_1\}$, as given and estimate the value of the fixed effects, $\delta(\theta)$, which set the observed probability a student ranks each high school as their preferred option equal to the observed probability across primary schools.[35] Letting $J_s$ correspond to the number of high schools in the choice set of students in primary school $s$, this yields $\sum_{s \in \mathcal{S}}(J_s - 1)$ moment conditions of the form:

$$m_{sj}(\delta) = \sum_{i \in \mathcal{I}_s} \left[ \mathbb{1}\{R_i(1) = j\} - F(X_{ij}, \bar{m}_{stj}, \bar{X}_{stj}) | \delta, \theta) \right] \tag{20}$$

variables are constant across time.

[34] This proof is closely related to Brock and Durlauf (2007) who show that both contextual effects that drive preferences for a specific school and endogenous peer effects can be identified in discrete choice models. I extend their result to show the presence of an admission system that maps from choices to the expected outcomes of peers does not change the result. Moreover, it is also possible to identify contextual effects that drive preferences for a class of schools, such as the impact of exposure to high-income peers on students' preference for high-quality schools.

[35] Rather than leveraging students' entire ROLs, I focus exclusively on the probability that a student ranks a high school as their most preferred option. Conditional on the model parameters, this probability has a closed-form solution which simplified computation. Focusing on the entire ranking can increase power at the cost of additional computation time, as the probability to rank a school ranks 2 and above does not have a closed-form solution.

where $\mathcal{I}_s$ denotes the set of students in primary school s; $\mathbb{1}\{R_i(1) = j\}$ is a dummy equal to 1 if student $i$ selects school $j$ as their preferred school; and $F(X_{ij}, \bar{m}_{ij}, \bar{X}_{stj}|\delta, \theta)$ is the probability that a student will select school $j$ conditional on the school's observable characteristics, $X_{ij}$, the expected share of peers who will be admitted to the school, $\bar{m}_{ij}$, and peers' characteristics, $\bar{X}_{stj}$. In order to identify all of the fixed effects, I normalize the value of the fixed effect for one high school in each primary school to zero. Since fixed effects are specific to each primary school-high school pair, I estimate them in parallel, which greatly reduces computation time.

In a second step, given $\hat{\delta}(\theta)$, I find the value of $\theta$ that minimizes a criterion function that exploits a set of orthogonality conditions associated with each structural parameter. These moments exploit that residuals, defined as the difference between the observed and predicted probability that a student ranks a school as their preferred option, should be independent of each variable. Formally, this yields the following K+1 moment conditions:

$$\begin{aligned} m_1(\theta) &= \sum_i \sum_j \left[\mathbb{1}\{R_i(1) = j\} - F(X_{ij}, \bar{m}_{ij}, \bar{X}_{stj}|\hat{\delta}(\theta), \theta)\right] x^1_{ij} \\ &\vdots \\ m_{K+1}(\theta) &= \sum_i \sum_j \left[\mathbb{1}\{R_i(1) = j\} - F(X_{ij}, \bar{m}_{ij}, \bar{X}_{stj}|\hat{\delta}(\theta), \theta)\right] \bar{m}_{ij} \end{aligned} \tag{21}$$

where K corresponds to the total number of school and peer characteristics in the vectors $X_{ij}$ and $\bar{X}_{stj}$.[36]

Since I estimate the structural parameters of the model using an estimate of the fixed effects, $\hat{\delta}(\theta)$, rather than the true value, whenever the size of observations per group is small, the estimated structural parameters will be inconsistent, a common challenge referred to as the *incidental parameters problem* (Neyman and Scott, 1948). To account for this bias, I implement a split-sample jackknife correction. Following Dhaene and Jochmans (2015), I partition my sample into two disjoint subsamples, the first containing observations for 2019, 2020 and the second for 2021, 2022, 2023. I then estimate the model separately within each subsample, and construct an estimate of the bias by removing from the full-panel estimate a weighted average of the estimates from the subsamples.[37]

Finally, I estimate the value of the outside option $V_0$, by matching the expected number of options ranked with the actual number of options ranked:

$$m_1(\theta) = \sum_i \sum_j \sum_r \left[\mathbb{1}\{R_i(r) = j\} - \mathbb{P}(V_{ij} - V_0 > 0\,|\hat{\delta}(\hat{\theta}), \hat{\theta})\right] \tag{22}$$

[36] To ensure $\theta$ is estimated efficiently, I weigh these moment conditions using a Continuous Updating Estimator, which minimizes the GMM criterion function directly over both $\theta$ and the weighting matrix simultaneously (Hansen et al., 1996).

[37] A key advantage of the split-sample jackknife in my context is that it does not require an explicit analytical form for the bias.

*Parameterization* I define a student's choice set according to two criteria. First, a high school must have available places in the year a student applies and be located within 10 km of the student's geo-coded residence.[38] Second, the high school must have had at least five students from the student's primary school ever apply. This restriction is motivated by a desire to reduce incidental parameter bias generated by the large number of primary school-high school fixed effects, $\delta_{sj}$.

I include in $X_{ij}$ three variables that capture the student-school match: (1) the distance between a student and a school, (2) a dummy indicator for whether a student has a sibling studying at the school, (3) a dummy indicator for whether a student has a parent working at the school. Note that any characteristics of schools that do not vary over time, such as quality, are absorbed by the fixed effects, $\delta_{sj}$.[39] I consider two sources of contextual peer effects: (1) the share of peers who live near a high school, and (2) the interaction between the share of high-income peers and whether a high school is classified as having high quality.

Due to the computational burden of the estimation, I estimate the model on students in 100 randomly sampled primary schools in Santiago, the capital and largest city of Chile. To increase the model's power to estimate fixed effects, I use data on high school applications submitted between the 2019 and 2024 application rounds.

# 6 Estimation Results and Model Fit

In this section, I present the estimated model parameters and evaluate the model's fit by comparing predicted and observed school choice patterns.

## 6.1 Parameter Estimates

Table 4 presents the estimated parameters of the structural model, which capture how students trade off different school attributes when applying to high school. Panel (A) confirms that students systematically prefer schools that are closer to home, where a sibling attends or where a parent works. Since the scale of these parameters is not identified, to understand their relative importance, I normalize them by the estimate for students' distaste for distance. This re-parameterization expresses the importance of each variable in terms of students' "willingness to commute". I find that students are willing to commute almost 22 km to study with their sibling, and over 15 km to study where their parent works.

Panel (B) corresponds to the estimates of peer effects. The results suggest that students are highly sensitive to their peers' choices: a student will be willing to travel to a school over 900 meters farther from their home in order to study with 10% more of their primary school peers. Under Assumption S-FE, these estimates capture the *causal* effect of peers' expected admission outcomes on the utility of a high school. As such, they do not simply reflect the fact that students

[38] In 2023, 92% of students' only ranked schools within 10 km of their residence.

[39] While, in principle, school quality does vary over time, in the sample used to estimate model parameters there are no changes in the quality of high schools over the study period.

Table 4: Parameter Estimates

| Variable | Estimate | Willingness to Commute |
|---|---|---|
| **Panel A: School characteristics** | | |
| Distance (km) | -0.20 | – |
| | [-0.197, -0.196] | |
| Sibling | 4.30 | 21.88 |
| | [4.275, 4.318] | [21.76, 21.99] |
| Parent | 3.12 | 15.89 |
| | [1.622, 4.618] | [8.26, 23.51] |
| **Panel B: Peer effects** | | |
| Peer share | 1.78 | 9.06 |
| | [1.770, 1.788] | [9.01, 9.11] |
| Share of peers near school | -0.07 | -0.34 |
| | [-0.933, 0.801] | [-4.75, 4.08] |
| Share of high-income peers × high-quality school | 0.04 | 0.19 |
| | [-3.849, 3.925] | [-19.60, 19.99] |
| **Panel C: Outside option** | | |
| $V_0$ | 1.20 | – |
| | [1.190, 1.211] | |

*Note:* The table presents estimates of the structural parameters of the model described in Section 5. The first column corresponds to estimates of the parameters, and the second column corresponds to estimates of the willingness to commute. Panel (A) corresponds to estimates relating to characteristics of schools: (1) distance to a high school, (2) having a sibling at the high school, and (3) having a parent working at the high school. Panel (B) corresponds to peer effects: (1) the share of peers expected to attend the school, (2) the share of peers who live close to the school, (3) the interaction of the share of high-income peers, and a school being of high-quality. Panel (C) corresponds to the estimated deterministic utility of the outside option. Confidence intervals correspond to 95% confidence intervals.

apply to the same schools as their peers because they share similar backgrounds or live in the same neighborhoods. Rather, the expectation that applying to a school will lead students to study together with their peers affects students' choice. The second and third rows of Panel (B) correspond to estimates of contextual peer effects. I am not able to reject the null hypothesis that exposure to peers living near a high school influences students' choices. Likewise, studying with more high-income peers does not change students' preferences over high-quality schools. This suggests that peer effects in this context are primarily endogenous, that is, peers react to each others' choices, rather than characteristics.

## 6.2 Model Fit

To simulate rankings over high schools, I compute each school's deterministic utility according to $\hat{\beta}$ and $\hat{\gamma}_0$ and draw students' unobserved preferences for each high school from a Gumbel distribution centered around $\hat{\delta}_{sj}$. Since I am not able to reject that $\gamma_1 = 0$ and $\gamma_2 = 0$, I set these to zero in the

simulation.[40]

To map from ranking into admission outcomes, I simulate a student-proposing deferred acceptance algorithm that assigns students to schools based on their ranked preferences, priority and eligibility for a quota, mirroring the Chilean school admission system (SAE). A student's ranking contains all schools whose utility exceeds the utility of the outside option. Each student is initially unmatched and iteratively applies to their most preferred high school. Schools tentatively accept students according to their priority and eligibility for a quota, and reject lower-priority applicants when capacity constraints bind. If a school is oversubscribed within a quota-priority group, students are assigned random lottery numbers that act as tie breakers. Rejected students then propose to their next preferred school.[41] I repeat this process 100 times, and average across simulations to obtain the average probability that each student ranks and is admitted into each high school. I discuss the algorithm in more detail in Appendix A4.2.

As a first test of the validity of the model, I compare its predictions to the actual ranking patterns observed in the data. Panel (a) of Figure 6 examines how the probability of ranking a school as students' preferred choice varies with distance. Schools are grouped into twenty bins based on their distance from students' homes. The observed data and the model predictions align closely, showing a strong negative relationship – students are significantly less likely to apply to schools that are farther away. Panel (b) examines how ranking probabilities vary with the share of a student's peers who are admitted into a school. Again, schools are grouped into twenty bins based on the share of peers who are admitted. As expected, since students prefer to attend high school with their primary school peers, the probability to rank a school increases monotonically with the fraction of peers who attend. Again, the predicted probability tracks the observed probability closely. The close alignment between the model and the observed data in both panels indicates that the model provides a strong approximation of students' school choice behavior.

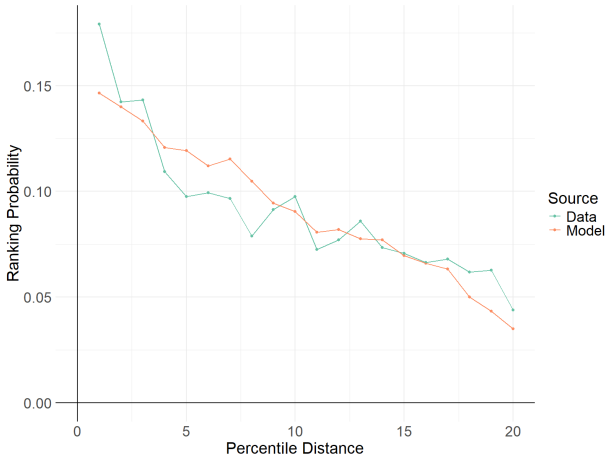
I also evaluate the fit of the model using actual admission probabilities, which were not targeted in estimation, rather than ranking probabilities. Panel (c) of Figure 6 examines how the probability of being admitted into a school varies with distance. As before, I group schools into twenty bins based on their distance from students' homes. Panel (d) examines how admission probabilities vary with the share of peers who are admitted into each school. As expected, the probability of being admitted into a school decreases with distance and increases with the fraction of peers who are admitted. In both cases the predicted probability tracks the observed probability closely.

For the model to be useful in discussing socioeconomic gaps in school quality, it should be able to replicate the distribution of quality among the schools to which low- and high-income students are admitted. To assess this, I plot in Appendix Figure A3 the average simulated share of high- and low-income students who are admitted to high schools of each level of quality and compare it to the true share. The model predictions closely match observed shares: high-income students are

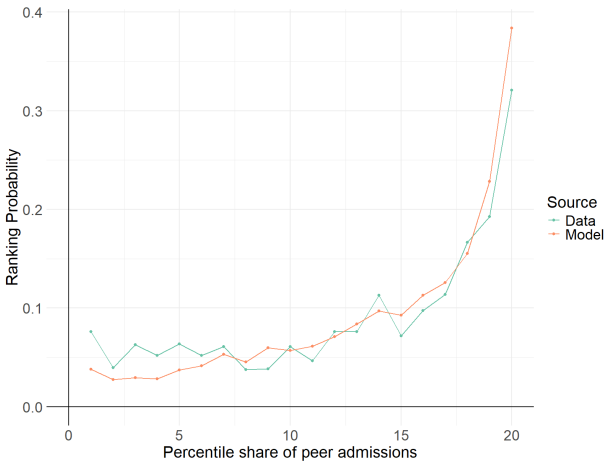
[40] To match the equilibrium observed in the data, I assume that the observed share of students admitted into each high school is the equilibrium value. I relax this assumption in the counterfactuals I present in Section 7.

[41] Rejected students that are eligible for a quota then compete for regular admission seats, and if rejected again, proposed to their next preferred school

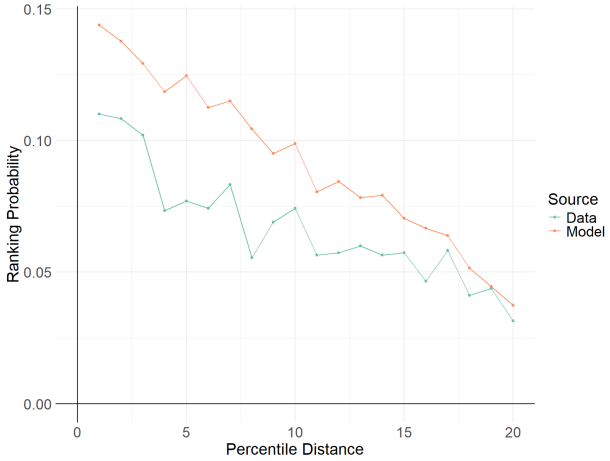
Figure 6: Model fit



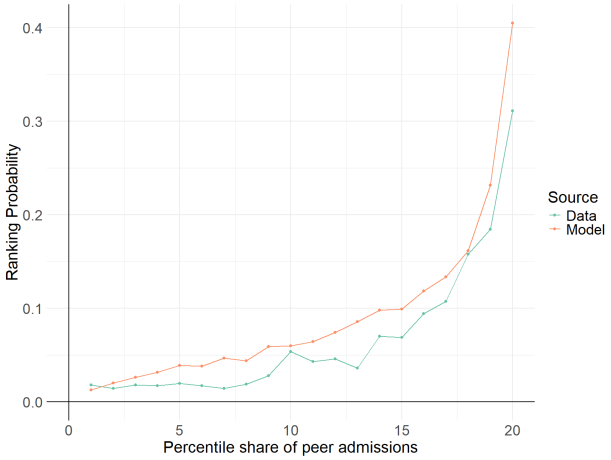

(a) Ranking probability and distance




(b) Ranking probability and share of peer admissions




(c) Admission probability and distance




(d) Admission probability and share of peer admissions

Note: The top two figures correspond to two moments that are implicitly targeted in the estimation. Plot (a) corresponds to the probability that a student ranks a school as their most preferred option at every percentile of distance. Plot (b) corresponds to the probability that a student ranks a school as their most preferred option at every percentile of the share of peers admitted to the school. Plot (c) corresponds to the probability that a student is admitted to a school at every percentile of distance. Plot (d) corresponds to the probability that a student is admitted to a school every percentile of the share of peers admitted to the school. The orange line corresponds to the true data, whereas the teal line corresponds to the simulated probability, based on the model with the estimated parameters.

predicted to be 2.5 percentage points more likely to attend a school classified as having *high* quality and 1 percentage point less likely to attend a school classified as having *low* quality.

# 7 Counterfactuals

In this section, I first use the model to quantify how peer effects interact with the uneven spatial distribution of school quality to generate educational inequality. To do so, I simulate a counterfactual in which I remove peer effects, holding all other factors constant. Second, I use the model to assess potential spillovers from expanding access to high-quality schools by simulating a counterfactual policy in which different shares of students receive transport vouchers to attend high-quality schools. I show that in the presence of peer effects, this policy has increasing returns to scale and does not impose negative externalities on untargeted students.

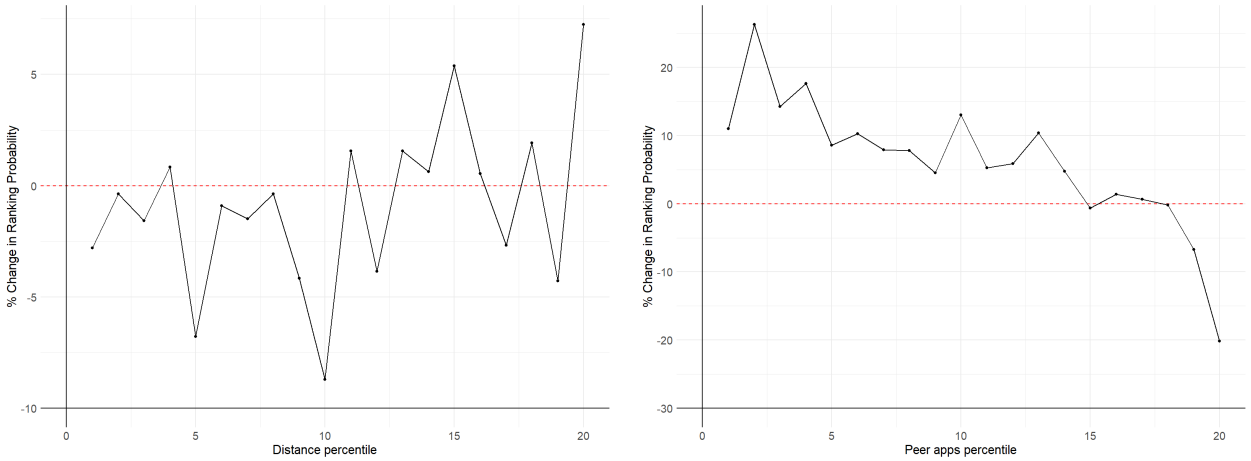
## 7.1 Peer Effects and Inequality in Education

I first focus on how peer effects impact the types of schools to which students apply. I re-estimate the model setting the endogenous peer effects parameter, $\gamma_0 = 0$, holding all other parameters constant.[42] I first examine the effect on students' rankings. I compare the percentage change in the probability that a student ranks a high school as their preferred option between the simulated model with peer effects and the simulated model without peer effects. Panel (a) of Figure 7 summarizes how students' rankings change based on the distance (in percentiles) of the high school from their home. A positive value indicates that students are more likely to rank a school as their most preferred option when peer effects are removed. Removing peer effects encourages students to rank high schools located farther away from their homes, and decreases the probability that they rank nearby high schools. Panel (b) of Figure 7 summarizes the change in ranking based on the share of a student's peers (in percentiles) that are eventually admitted to each high school. As expected, removing peer effects leads students to rank high schools where fewer of their peers eventually attend. Appendix Figure A4 plots the corresponding change in admission probabilities at different percentiles of distance and shares of admitted peers. The results indicate that peer effects reinforce clustering, leading students to apply closer to home and to the same high schools as their peers.

To assess how peer effects contribute to inequality, I examine how the quality of the high schools that students rank changes when peer effects are removed. First, I convert the Agency for Quality in Education's measure of school quality into a numeric (1-4) scale and study the effect of removing peer effects on the distribution of quality in students' rankings. Panel (A) of Table 5 summarizes the effect on the variance of the distribution. A higher variance indicates more dispersion between students in the quality of the high schools which they rank. Removing peer effects results in a small, 4% decrease in the overall variance of school quality in students' applications. However, this small effect masks the fact that removing peer effects increases the variance in quality *within* primary schools by 8% while decreasing it *between* primary school by 8%. Panel (B) presents the corresponding effect on the variance of quality among the schools to which students are admitted. Overall, peer effects increase the likelihood that students within the same primary schools apply and are admitted into high schools of the same quality, while decreasing this probability across

[42] For all counterfactuals, I also set $\gamma_1 = 0$, since I am not able to reject the null of no contextual peer effects.

Figure 7: Impact of Peer Effects on Applications




(a) Ranking probability and distance

(b) Ranking probability and share of peer admissions

Note: The plots correspond to the percentage difference in the probability to rank a school as a student's preferred option between the simulated model without peer effects and the simulated model with peer effects. A positive percentage difference indicates that the probability is higher in the model without peer effects. Plot (a) corresponds to the percentage difference in ranking probability at each percentile of distance. Plot (b) corresponds to the percentage difference in ranking probability at each percentile of the share of peers admitted into a school.

primary schools.

Second, I study what peer effects imply for socioeconomic gaps in the probability that students apply to and attend high-quality schools.[43] At baseline, high-income students (top 4 deciles of the income distribution) are 4.4 percentage points more likely to list a high-quality school anywhere in their ROL relative to low-income students (bottom 6 deciles). Removing peer effects decreases this gap by 11%. Likewise, students whose mother completed high school – a common proxy for socioeconomic status – are 3.9 percentage points more likely to rank a high-quality high school at baseline relative to students whose mother dropped out. Removing peer effects decreases this gap by 20%. The effect on the probability that a student ranks a high-quality high school as their preferred option is more pronounced. Removing peer effects decreases the gap in the probability that high- and low-income students, and high- and low-SES students rank a high-quality school as their top choice by 16% and 23% respectively. Panel (B) maps these differences into actual admission outcomes. The effects are similar. Removing peer effects decreases the gap in the probability that high- and low-income students are admitted into a high-quality school by 14%. Correspondingly, it reduces the gap in the probability that high- and low-SES students are admitted into a high-quality school by 19%.

[43] I define a high-quality school as a school rated as having *medium* or *high* quality by the Agency for Quality in Education.

Table 5: Impact of Peer Effects on Inequality

| Statistic | Baseline | Counterfactual | % difference |
|---|---|---|---|
| **Panel A: Applications** | | | |
| Variance | | | |
| Overall | 7.149 | 6.842 | -4.296 |
| Within | 1.680 | 1.809 | 7.679 |
| Between | 5.470 | 5.034 | -7.973 |
| Prob. Apply Good School | | | |
| High Income - Low Income | 4.414 | 3.914 | -11.325 |
| High SES - Low SES | 3.876 | 3.111 | -19.728 |
| High Income - Low Income (Top rank) | 4.449 | 3.755 | -15.599 |
| High SES - Low SES (Top rank) | 4.339 | 3.339 | -23.041 |
| **Panel B: Admissions** | | | |
| Variance | | | |
| Overall | 5.365 | 5.119 | -4.584 |
| Within | 0.459 | 0.584 | 27.269 |
| Between | 4.906 | 4.534 | -7.564 |
| Prob. Admit Good School | | | |
| High Income - Low Income | 4.605 | 3.962 | -13.962 |
| High SES - Low SES | 4.619 | 3.738 | -19.074 |

*Note:* The table compares the distribution of school quality under a counterfactual in which peer effects are removed. Panel (A) corresponds to the effects of removing peer effects on students' applications. The first row corresponds to the overall variance of the distribution of quality of schools that students rank. The second row corresponds to the average within - primary school variance, whereas the third row corresponds to the average between - primary school variance. The fourth row corresponds to the average gap in the probability to attend a high-quality school between students in the top-6 deciles, and the bottom 4 deciles. The fifth row corresponds to gaps between students whose mothers completed high school and students whose mothers dropped out of high school. The sixth- and seventh- rows repeat this analysis considering only the top school that students rank, rather than all schools in their ranked-ordered list. Panel (B) repeats this analysis focusing on admission outcomes rather than applications.

### 7.2 Multiplying Policy Through Peer Networks

Equalizing educational opportunities for high- and low-income students is often cited as a key policy objective. A common lever to achieve this goal is to directly target the uneven spatial distribution of school quality, either by building high-quality schools in low-income areas, or by making it easier for lower-income students to access high-quality schools, such as through transport subsidies. The presence of peer effects, however, implies that these interventions can have complicated equilibrium consequences and may have nonlinear returns to scale. For instance, if enrollment decisions are interdependent, students may not enroll in a new school that is accessible to them but not their peers. Conversely, a policy may generate positive spillover effects if it targets enough students within a network, as one student's decision to enroll in a better school can induce similar choices by peers.

To study this trade-off, I simulate a policy where students are given transport subsidies that

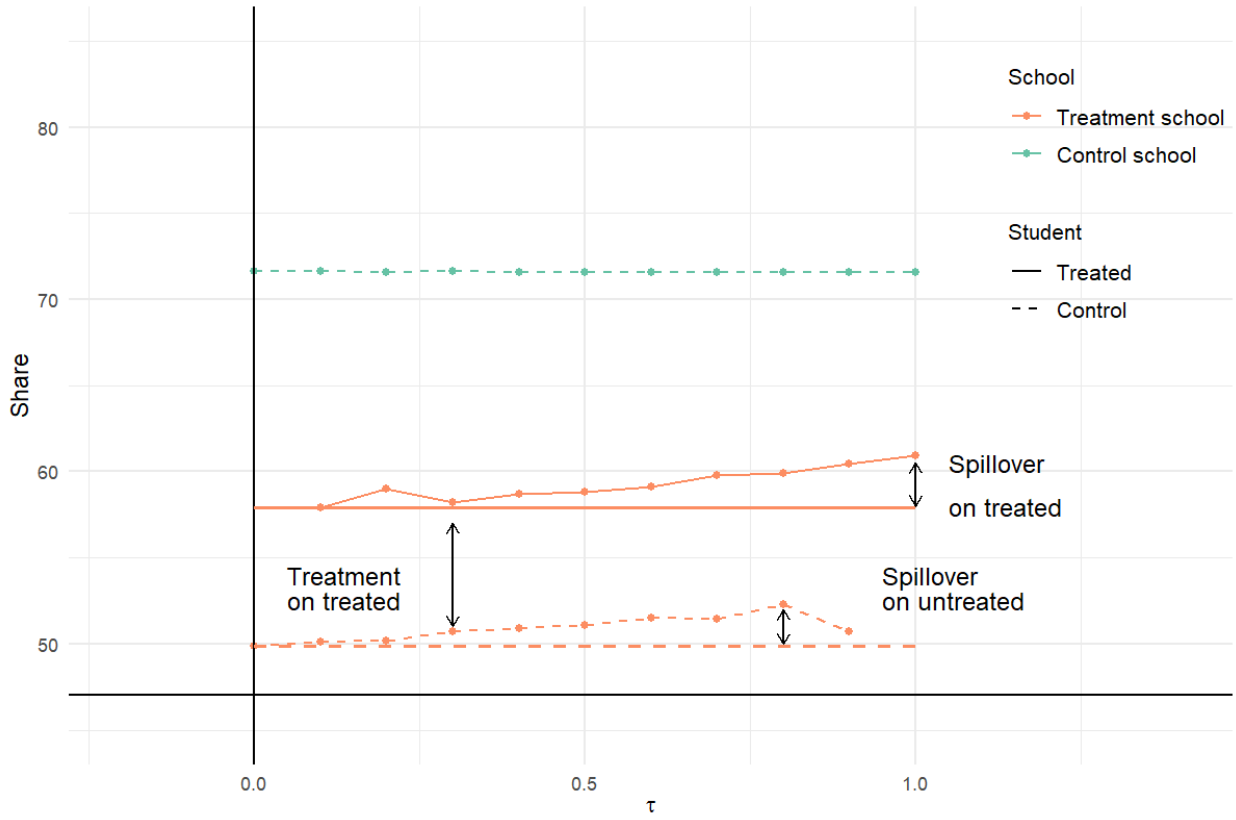
compensate them for the cost associated with attending a high-quality school farther from home. To do this, I offset the utility-cost associated with travel to a high-quality school. To understand the equilibrium dynamics of this policy, I assign the subsidy to students studying in deprived primary schools, which I identify as primary schools where the share of low-income students is above the median, and vary the fraction, $\tau$, of students who are offered the subsidy. Since my parameter estimates suggest that multiple equilibria are possible, I simulate choices starting from 100 randomly chosen starting values for students' beliefs about the share of their peers who will attend each high school, and average outcomes across the simulations.

Figure 8 presents the effect of the policy on the probability that students are admitted into a high-quality school for $\tau$ between 10% and 90%. The policy has a strong effect on students who receive the subsidy. When 10% of students are treated, treated students become 15% more likely to be admitted into high-quality school relative to untreated students. This nearly halves the gap in the likelihood students in deprived and non-deprived primary schools are admitted into a high-quality high school. As the share of treated students increases so does the effect of treatment. When 80% of students are treated, treated students become 18% more likely to be admitted into a high-quality school relative to the baseline with no students targeted. Since most students already attend high-quality schools, treated peers induce each other to apply, increasing the effectiveness of the policy. Similarly, the policy generates strong spillovers on non-treated students. When 10% of students are treated, the effect on untreated students is negligible. On the other hand, when 80% of students are treated, untreated students become 5% more likely to be admitted into a high-quality school, as they are pulled in by their treated peers. Appendix Figure A5 confirms that these patterns are driven by changes in students' behavior. Increasing $\tau$ increases the probability that both treated and untreated students in treatment schools rank a high-quality high school as their most preferred option, with no impact on the behavior of students in control schools.

Notice, however, that treating more students does *not* reduce the probability that untreated students – either in treatment or control schools – are admitted into a high-quality school. At first glance, this is surprising; since school capacity is fixed, an increase in applications to high-quality schools should reduce the individual probability of admission, generating a negative externality for untreated students who face greater competition. To better understand this pattern, Figure 9 plots the change in the number of applications per school when $\tau = 0.8$, relative to when $\tau = 0$. I order schools by the percentile of their pre-policy competitiveness, measured by the difference between the number of applicants (students ranking the school as their preferred option) and available seats. As expected, the policy reduces applications to low-quality schools at all percentiles of competition. On the other hand, the policy increases applications to the least competitive high-quality schools, but has almost no impact on applications to the most competitive schools.

This pattern is consistent with peers' preference to study together: the policy leads students to apply to higher-quality but relatively under-subscribed schools, where the probability of being admitted *with their peers* is high. Consequently, the policy generates minimal externalities for untreated students, whose chances of admission to high-quality schools remain largely unchanged.

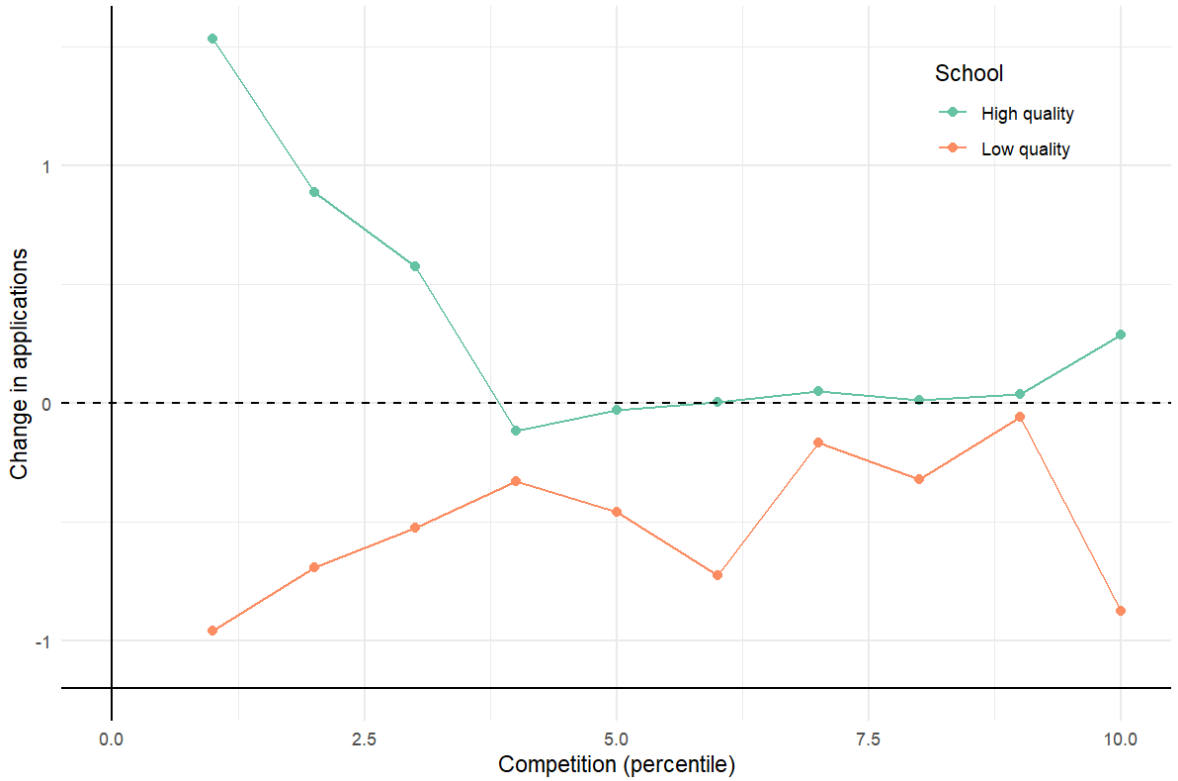
Figure 8: Impact of Transport Subsidies




*Note:* This figure corresponds to a counterfactual simulation in which different shares of students, $\tau$, in schools with an above-median share of low-income students are given transport subsidies to attend high-quality schools. The orange line corresponds to the probability that a student in a treated school is admitted into a high-quality high school. The solid line corresponds to the probability for students who are treated, whereas the dashed line corresponds to untreated students. The teal line corresponds to the probability that a student in a control school (with a below-median share of low-income students) is admitted into a high-quality high school.

To confirm that this pattern is not driven by a correlation between the competitiveness of a school and other attributes, such as being closer to treated students, I regress the change in admission probability for a student - high school pair on the characteristics of the school and its pre-policy level of competition. Appendix Table A7 presents the results. I find that the policy significantly increases the probability that students apply to schools of higher quality, which are located farther from home, and with a lower pre-policy level of competition.

Figure 9: Impact of Transport Subsidies on Competition




*Note:* This figure corresponds to the change in applications to different schools when $\tau = 0.8$. The x-axis corresponds to the pre-policy competitiveness, measured by the difference between applicants (students ranking the school as their preferred option) and available seats. The y-axis corresponds to the change in the number of applications. The teal line corresponds to high-quality high schools, defined as those having *medium* or *high* quality by ACE. The orange line corresponds to low-quality high schools, defined as those having *insufficient* or *low* quality by ACE.

# 8 Conclusion

This paper shows that primary school peers play a crucial role in shaping students' choice of high school. Using administrative data from Chile's centralized school admission system, I provide causal evidence that studying with a particular peer increases the probability that a student applies to and attends the same high school. To quantify the equilibrium impact of these social interactions, I incorporate peer effects into a structural model of school choice. I find that social interactions have a significant impact on the equilibrium allocation of students to schools. If students were to make decisions independently of their peers, the gap in high school quality between low- and high-income students would shrink by 14%.

A key limitation of my study is the lack of specialized network data. I assume that students care about all of their peers equally, whereas in reality, the choices of some peers are almost surely more important than others. Data that explicitly maps social connections would allow for more precise identification of peer effects and a better understanding of which peers matter most in shaping

school choices, expanding the scope of feasible policy instruments, such as targeting the behavior of *key players* (Ballester et al., 2006).

My results suggest that interventions which take advantage of peer networks can have meaningful impacts on inequality in education. Increasing access to high-quality schools in under-served areas – whether by expanding transportation options or building new schools – may benefit from significant spillover effects within peer networks, with minimal negative externalities on untreated students. Recognizing that school choice is not an isolated decision but one embedded in a social context is crucial for designing policies that effectively address educational inequality.

## References


Abdulkadiroğlu, A., J. D. Angrist, and P. A. Pathak (2014). The elite illusion: Achievement effects at boston and new york exam schools. *Econometrica 82*(1), 137–196.

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Agte, P., C. Allende, A. Kapor, C. Neilson, and F. Ochoa (2024, July). Search and biased beliefs in education markets. Working Paper 32670, National Bureau of Economic Research.

Allende, C. (2021). Competition under social interactions and the design of education policies.

Allende, C., F. Gallego, and C. Nielson (2019). Approximating the equilibrium effects of informed school choice.

Altmejd, A., A. Barrios-Fernández, M. Drlje, J. Goodman, M. Hurwitz, D. Kovac, C. Mulhern, C. Neilson, and J. Smith (2021). O brother, where start thou? sibling spillovers on college and major choice in four countries. *Quarterly Journal of Economics 136*(3), 1831–1886.

Angrist, J. D. (2014). The perils of peer effects. *Labour Economics 30*, 98–108.

Angrist, J. D., S. R. Cohodes, S. M. Dynarski, P. A. Pathak, and C. R. Walters (2016). Stand and deliver: Effects of boston's charter high schools on college preparation, entry, and choice. *Journal of Labor Economics 34*, 275–318.

Angrist, J. D., P. D. Hull, P. A. Pathak, and C. R. Walters (2017). Leveraging lotteries for school value-added: Testing and estimation. *Quarterly Journal of Economics 132*(2), 871–919.

Avvisati, F., M. Gurgand, N. Guyon, and E. Maurin (2014). Getting parents involved: A field experiment in deprived schools. *Review of Economic Studies 81*(1), 57–83.

Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006). Who's who in networks. wanted: The key player. *Econometrica 74*(5), 1403–1417.

Barrios-Fernandez, A. (2023, 10). Peer effects in education.

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

Bau, N. (2022). Estimating an equilibrium model of horizontal competition in education. *Journal of Political Economy 130*(7).

Bennett, M. and P. Bergman (2021). Better together? social networks in truancy and the targeting of treatment. *Journal of Labor Economics 39*(1), 1–36.

Beuermann, D. and C. K. Jackson (2022). The short and long-run effects of attending the schools that parents prefer. *Journal of Human Resources 57*(3), 725–746.

Bhattacharya, D., P. Dupas, and S. Kanaya (2024). Demand and welfare analysis in discrete choice models with social interactions. *Review of Economic Studies 91*(2), 748–784.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2013). Under pressure? the effect of peers on outcomes of young adults. *Journal of Labor Economics 31*(1), 119–153.

Bobonis, G. J. and F. Finan (2009). Neighborhood peer effects in secondary school enrollment decisions. *The Review of Economics and Statistics 91*(4), 695–716.

Booij, A. S., E. Leuven, and H. Oosterbeek (2017). Ability peer effects in university: Evidence from a randomized experiment. *Review of Economic Studies 84*(2), 547–578.

Brock, W. A. and S. N. Durlauf (2007). Identification of binary choice models with social interactions. *Journal of Econometrics 140*, 52–75.

Burgess, S. and A. Briggs (2010). School assignment, school choice and social mobility. *Economics of Education Review 29*(4), 639–649.

Burgess, S., E. Greaves, A. Vignoles, and D. Wilson (2015). What parents want: School preferences and school choice. *The Economic Journal 125*(587), 1262–1289.

Bursztyn, L., G. Egorov, and R. Jensen (2019). Cool to be smart or smart to be cool? understanding peer pressure in education. *Review of Economic Studies 86*(4), 1487–1526.

Bursztyn, L. and R. Jensen (2015). How does peer pressure affect educational investments? *Quarterly Journal of Economics 130*(3), 1329–1367.

Calsamiglia, C., C. Fu, and M. Güell (2020). Structural estimation of a model of school choices: The boston mechanism versus its alternatives. *Journal of Political Economy 128*(2), 642–680.

Campos, C. (2024). Social interactions, information, and preferences for schools: Experimental evidence from los angeles.

Caria, S., S. Franklin, and M. J. Witte (2023). Searching with friends. *Journal of Labor Economics 41*(4), 887–922.

Carrell, S. E., M. Hoekstra, and E. Kuka (2018). The long-run effects of disruptive peers. *American Economic Review 108*(11), 3377–3415.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014, September). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American Economic Review 104*(9), 2593–2632.

Chetty, R., N. Hendren, and L. Katz (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review 106*(4), 855–902.

Chumacero, R. A., D. Gómez, and R. D. Paredes (2011). I would walk 500 miles (if it paid): Vouchers and school choice in chile. *Economics of Education Review 30*(5), 1103–1114. Special Issue on Education and Health.

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

De Giorgi, G., M. Pellizzari, and S. Redaelli (2010, April). Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics 2*(2), 241–75.

Deming, D. J. (2014). Using school choice lotteries to test measures of school effectiveness. *American Economic Review 104*(5), 406–411.

Dhaene, G. and K. Jochmans (2015). Split-panel jackknife estimation of fixed-effect models. *The Review of Economic Studies 82*(3), 991–1030.

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

Dustan, A. and D. K. Ngo (2018). Commuting to educational opportunity? school choice effects of mass transit expansion in mexico city. *Economics of Education Review 63*, 116–133.

Fabre, A., T. Larroucau, C. A. Neilson, and I. Rios (2024). College application mistakes and the design of information policies at scale.

Gurantz, O., G. T. Henry, M. F. Larsen, R. Romano, and J. Smith (2020). Sibling spillovers in education: Evidence from advanced placement exams. *American Economic Journal: Applied Economics 12*(3), 142–168.

Hansen, L. P., J. Heaton, and A. Yaron (1996). Finite-sample properties of some alternative gmm estimators. *Journal of Business & Economic Statistics 14*(3), 262–280.

Hastings, J., T. J. Kane, and D. O. Staiger (2009). Heterogeneous preferences and the efficacy of public school choice. *NBER working paper 2145*, 1–46.

Herskovic, L. (2020). The effect of subway access on school choice. *Economics of Education Review 78*, 102021.

Hoxby, C. M. (2000). Peer effects in the classroom: Learning from gender and race variation. *National Bureau of Economic Research Working Paper* (7867).

Hoxby, C. M. and C. Avery (2012). The missing" one-offs": The hidden supply of high-achieving, low income students. Technical report, National Bureau of Economic Research.

Kirkebøen, L., E. Leuven, M. Mogstad, and J. Mountjoy (2021). College as a marriage market. NBER Working Paper 28688, National Bureau of Economic Research.

Kutscher, M., S. Nath, and S. Urzúa (2023). Centralized admission systems and school segregation: Evidence from a national reform. *Journal of Public Economics 221*.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies 60*(3), 531–542.

Michelman, V., J. Price, and S. D. Zimmerman (2021, 12). Old boys' clubs and upward mobility among the educational elite*. *The Quarterly Journal of Economics 137*(2), 845–909.

Neyman, J. and E. L. Scott (1948). Consistent estimates based on partially consistent observations. *Econometrica 16*(1), 1–32.

Nielson, C. (2019). Targeted vouchers, competition among schools and the academic achievement of poor students.

Nielson, C. (2021). Targeted vouchers, competition among schools, and the academic achievement of poor students. Working paper.

Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates. *Quarterly Journal of Economics 116*(2), 681–704.

Sacerdote, B. (2011). Chapter 4 - peer effects in education: How might they work, how big are they and how much do we know thus far? Volume 3 of *Handbook of the Economics of Education*, pp. 249–277. Elsevier.

Seim, K. (2006). An empirical model of firm entry with endogenous product-type choices. *The RAND Journal of Economics 37*(3), 619–640.

Soetevent, A. R. and P. Kooreman (2007). A discrete-choice model with social interactions: with an application to high school teen behavior. *Journal of Applied Econometrics 22*(4), 599–624.

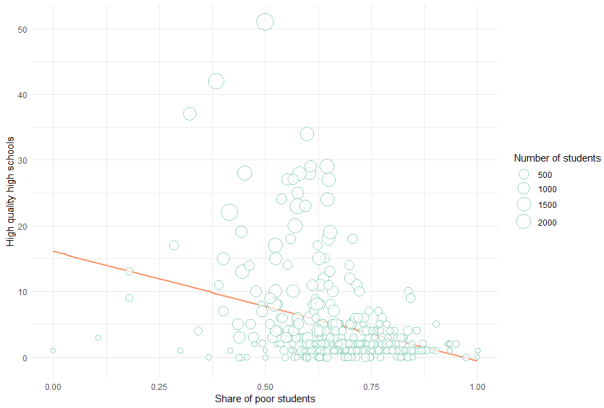
Topa, G. (2001). Social interactions, local spillovers and unemployment. *The Review of Economic Studies 68*(2), 261–295.

Walters, C. R. (2018). The demand for effective charter schools. *Journal of Political Economy 126*(6), 2179–2223.

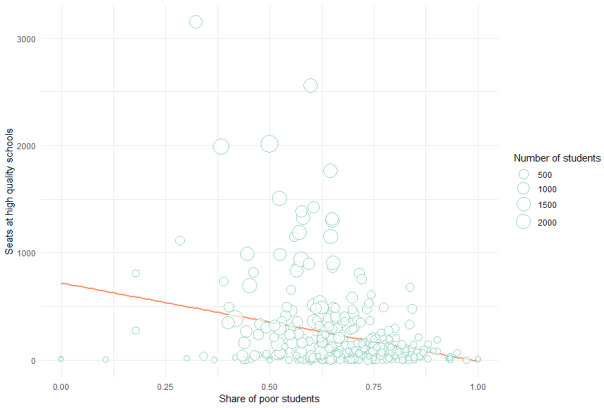
# Appendix

# A1 Appendix figures

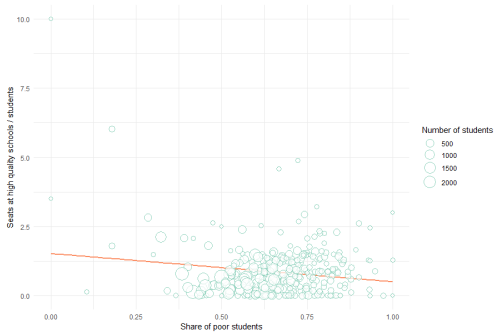
Figure A1: Access to high-quality schools across communes




(a) MH Schools



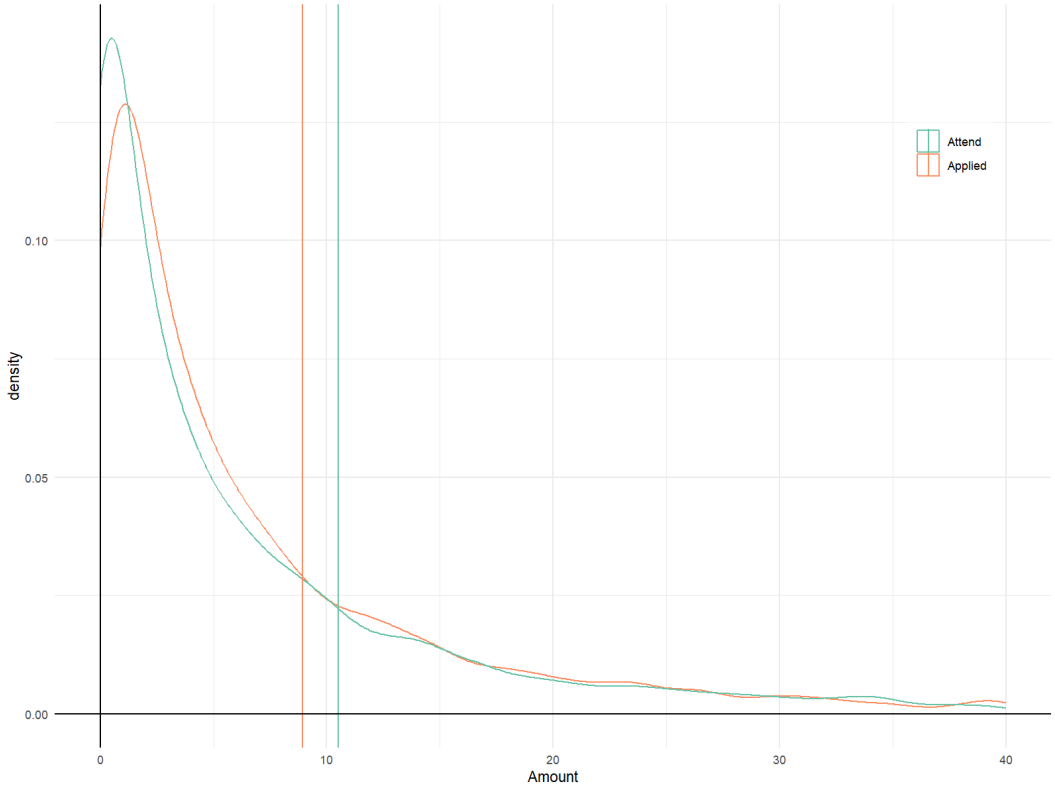

(b) Seats at MH Schools




(c) Seats per student at MH Schools

*Note:* The solid lines represent the average number of schools within each distance bin of students' homes. The dashed lines represent the average number of schools ranked as *medium- or high-quality*. The orange lines correspond to averages for high-income students, while the teal lines represent averages for low-income students.

Figure A2: Distribution of primary school peers in high school



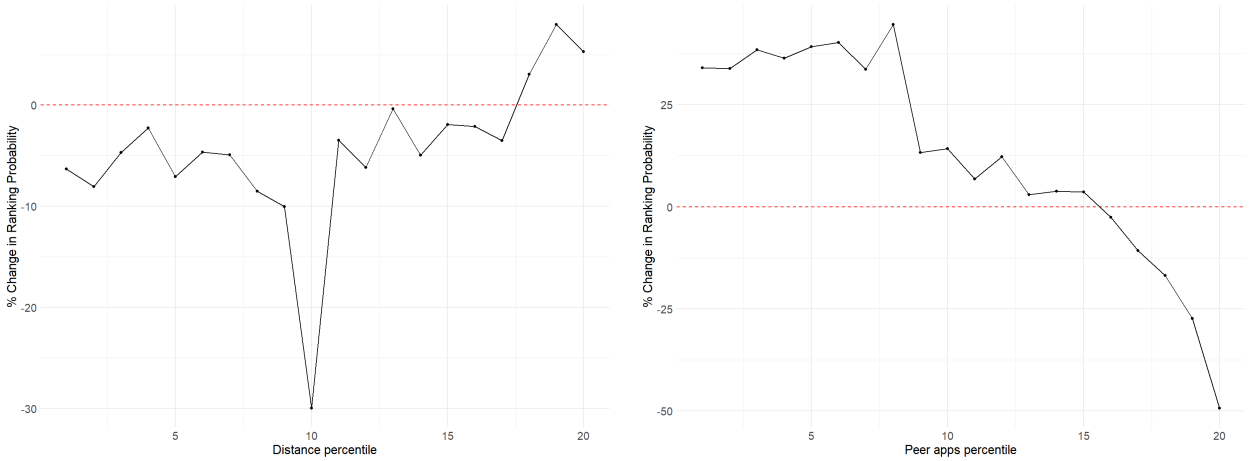

Note: The teal line represents the number of primary school peers with which every incoming high school student studies. The orange line represents the number of primary school peers who have ranked the incoming student's high school as their most preferred option. The vertical lines represent averages.

Figure A3: Difference in admission probability by quality and income




Note: Each bar corresponds to the difference in the probability to be admitted into a school between high- and low-income students. The quality bins correspond to the definition provided by the Agency for Quality in Education (ACE). The orange bars corresponds to the true data, whereas the teal bars corresponds to the simulated share based on the model simulated with the estimated parameters. I restrict the sample to students in treatment schools.

Figure A4: Impact of Peer Effects on Admissions

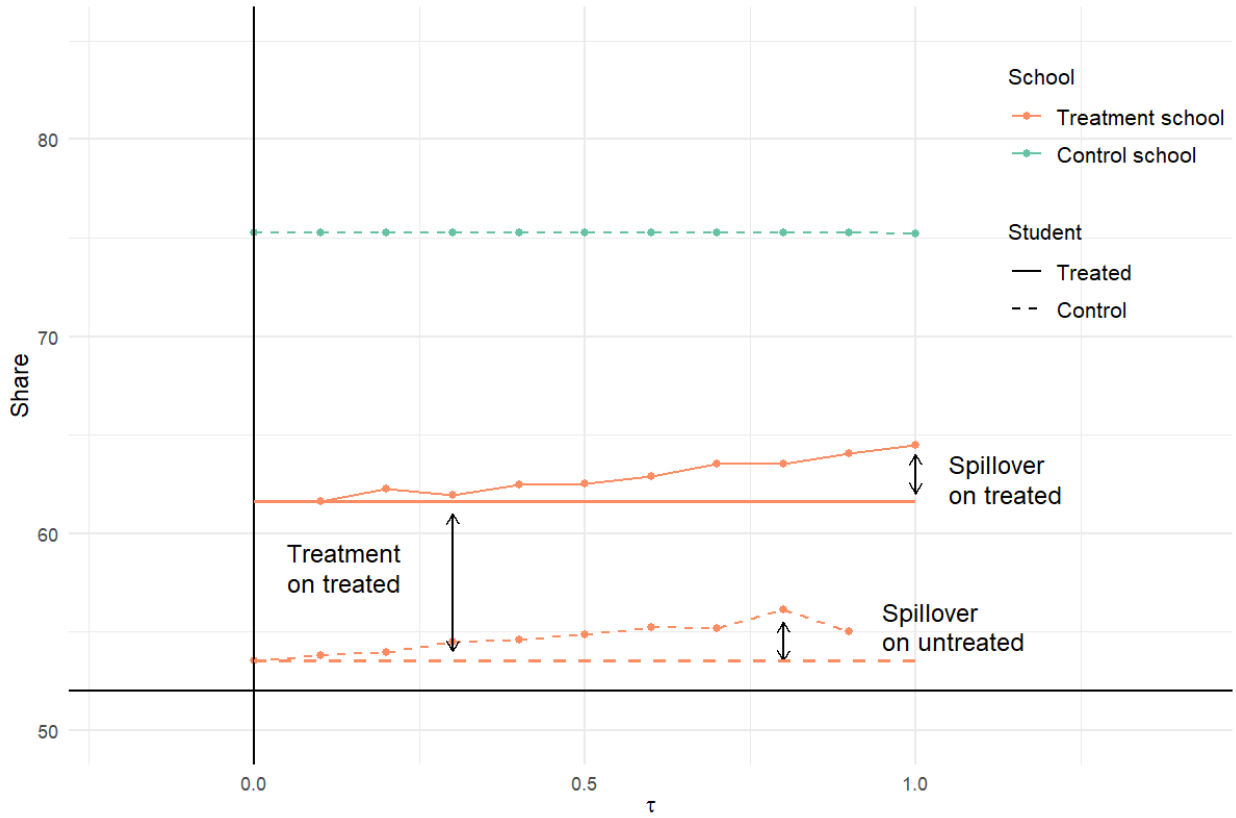



(a) Admission probability and distance

(b) Admission probability and share of peer admissions

Note: The plots correspond to the percentage difference in the probability to be admitted into a high school between the simulated model without peer effects and the simulated model with peer effects. A positive percentage difference indicates that the probability is higher in the model without peer effects. Plot (a) corresponds to the percentage difference in ranking probability at each percentile of distance. Plot (b) corresponds to the percentage difference in ranking probability at each percentile of the share of peers who are admitted into a school.

Figure A5: Impact of Transport Subsidies on Ranking




*Note:* This figure corresponds to a counterfactual simulation in which different shares of students, $\tau$, in schools with an above-median share of low-income students are given transport subsidies to attend high-quality schools. The orange line corresponds to the probability that a student in a treated school ranks a high-quality school as their preferred option. The solid line corresponds to the probability for students who are treated, whereas the dashed line corresponds to untreated students. The teal line corresponds to the probability that a student in a control school (with a below-median share of low-income students) rank a high-quality high school as their preferred option.

## A2 Appendix tables

Table A1: Data sources

| Dataset | Relevant Variables |
|---|---|
| Centralized school admission records | Students' choice of high school, students' geolocation |
| SIMCE Grades | Standardised scores for tests taken at 4th, 8th, and 10th grade. Rich sociodemographic information including parents' education |
| Administrative student records | Sex, age, commune of residence, school & class identifiers, low-income flag (from 2009), non-standardized grades, monthly attendance, graduation |
| Administrative school records | aggregates of students, teachers, teaching assistants, measures of quality from ministry of education, averages of students socioeconomic status, subsidies, fees charged |

Table A2: Changes in Primary School Inputs Across Cohorts

| | % Change (Within-school) | | Share of Variance | | |
|---|---|---|---|---|---|
| Variable | Median | IQR | Schools | Years | Residual |
| Funding | -0.681 | 6.911 | 0.985 | 0.005 | 0.005 |
| Teachers | -1.079 | 9.216 | 0.983 | 0.000 | 0.009 |
| Teaching hours | -0.638 | 8.848 | 0.984 | 0.001 | 0.009 |
| Enrollment | -0.112 | 3.853 | 0.990 | 0.000 | 0.004 |
| Funding (per student) | -0.352 | 6.071 | 0.964 | 0.022 | 0.012 |
| Teachers (per student) | -0.482 | 10.108 | 0.965 | 0.003 | 0.041 |
| Teaching Hours (per student) | -0.507 | 9.708 | 0.958 | 0.004 | 0.041 |

*Note:* Each row corresponds to a different school input. The first two columns present the median and inter-quartile range of within-school percentage changes in a variable between consecutive years, after removing common year effects. The remaining columns report the share of total variance in each input explained by persistent differences between primary schools, common year trends, and residual within-school fluctuations. Values near one in the column corresponding to the share of variance explained by schools indicate that nearly all variation in inputs arises from time-invariant differences across schools rather than changes within schools over time.

Table A3: Student Breakdown by Type and Grade

| Variable | Level | Students | Share |
|---|---|---|---|
| Type | | | |
| | High achieving | 105 | 0.7% |
| | Low income | 1921 | 12.6% |
| | Regular | 13216 | 86.7% |
| Grade | | | |
| | 2 | 9 | 0.1% |
| | 3 | 108 | 0.7% |
| | 4 | 618 | 4.1% |
| | 5 | 992 | 6.5% |
| | 6 | 1187 | 7.8% |
| | 7 | 7524 | 49.4% |
| | 8 | 4804 | 31.5% |
| High school admission year | | | |
| | 2023 | 2985 | 19.6% |
| | 2024 | 4632 | 30.4% |
| | 2025 | 7625 | 50.0% |

Table A4: Summary Statistics

| | Applicants | | All students | |
|---|---|---|---|---|
| Variable | Mean | Median | Mean | Median |
| Age | 13.28 | 13.00 | 13.39 | 13.00 |
| Female | 0.52 | 1.00 | 0.49 | 0.00 |
| In Santiago | 0.41 | 0.00 | 0.38 | 0.00 |
| Low-income | 0.61 | 1.00 | 0.67 | 1.00 |
| Distance from primary school (km) | 2.07 | 1.11 | 1.79 | 0.95 |
| GPA | 50.54 | 58.00 | 49.64 | 56.00 |
| Attendance | 82.47 | 90.00 | 85.71 | 93.00 |
| High-Achieving | 0.27 | 0.00 | 0.26 | 0.00 |
| Number of HS Ranked | 4.61 | 4.00 | 4.10 | 3.00 |
| Quality of HS Ranked | 2.85 | 3.00 | 2.85 | 3.00 |

*Note:* The first two columns correspond to applicants, while the last two columns correspond to the full sample. Age refers to a student's age at the time of high school choice. Female indicates the student's gender. In Santiago identifies whether a student lives in Santiago, Chile's largest metropolitan area. Low-income indicates whether a student belongs to the bottom 6 deciles of the income distribution. Distance from primary school measures the average distance between a student's home and their primary school at the time of application. GPA refers to the student's average grade (out of 70) in the year before applying to high school. Attendance represents the percentage of school days attended (out of 100). High-achieving indicates whether a student ranks in the top 3 deciles of their primary school cohort's grade distribution. Number of high schools ranked corresponds to the total number of schools a student lists on their application. Finally, Quality of high schools ranked refers to the average quality of the schools a student applies to, where quality is defined on a 1–4 scale based on the ratings of Chile's Agency for Quality in Education.

Table A5: Characteristics of admitted applicants

| Variable | Difference | Std. Error | p-value |
|---|---|---|---|
| Female | 0.012 | 0.062 | 0.849 |
| Poor | -0.006 | 0.062 | 0.918 |
| Attendance | -2.678 | 1.303 | 0.041 |
| GPA | 1.517 | 1.708 | 0.375 |
| Distance to PS (KM) | 1.189 | 1.122 | 0.290 |
| In PS commune | -0.009 | 0.065 | 0.896 |
| # of schools nearby | -0.061 | 0.286 | 0.830 |
| # of good schools nearby | 0.001 | 0.066 | 0.983 |

Table A6: Balance of Student Characteristics Across Classes

| Variable | Difference | Std. Error | p-value |
|---|---|---|---|
| Same Gender | -0.002 | 0.003 | 0.507 |
| Same socioeconomic status | 0.002 | 0.003 | 0.467 |
| Same age | 0.000 | 0.003 | 0.989 |
| Live in same comune | 0.004 | 0.003 | 0.147 |
| Distance | 0.303 | 0.281 | 0.283 |

Table A7: Change in admission probability

| Dependent Variable: | Change in Admission Probability |
|---|---|
| Model: | (1) |
| *Variables* | |
| Constant | -0.031*** |
| | [-0.041; -0.021] |
| Competition (pre-policy) | $-7.56 \times 10^{-5}$*** |
| | [-0.0001; $-3.61 \times 10^{-5}$] |
| Distance (km) | 0.003*** |
| | [0.002; 0.005] |
| Share of admitted peers (pre-policy) | -0.045** |
| | [-0.080; -0.010] |
| Sibling | 0.012* |
| | [-0.001; 0.025] |
| Low quality school | 0.001 |
| | [-0.007; 0.010] |
| Medium quality school | 0.034*** |
| | [0.025; 0.042] |
| High quality school | 0.030*** |
| | [0.020; 0.039] |
| *Fit statistics* | |
| Observations | 11,440 |
| $R^2$ | 0.24841 |

*Clustered (primary_school_id) co-variance matrix, 95% confidence intervals in brackets*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

*Note:* The table shows how the probability that a student is admitted into a different school when $\tau = 0.8$ relative to when $\tau = 0$, changes as a function of their characteristics. Competition is defined as the difference between the number of applicants and number of available seats when $\tau = 0$. The definition of school quality follows the ACE measures. The missing category corresponds to schools categorized as having *insufficient* quality.

# A3 Proofs

## A3.1 Identification of Peer Effects under Assumption FE

Consider the following four groups of students.

(1) $T_1$ : Average outcome of an applicant to primary school $S_i$, who is admitted, and an incumbent student in the same cohort and primary school.

$$T_1 := \mathbb{E}[y_{ij} \mid A_i = S_i, A_j = S_i, T_i = T_j]$$

Under Assumption FE, we can re-write $T_1$ follows:

$$T_1 = \gamma + E[\eta_{ij}|A_i = S_i, A_j = S_i, T_i = T_j] + E[\delta_{ij}|A_i = S_i, A_j = S_i, T_i = T_j]$$

Since admission is randomized, we can re-write:

$$E[\eta_{ij}|A_i = S_i, A_j = S_i, T_i = T_j] = E[\eta_{ij}|A_j = S_i, T_i = T_j]$$

Moreover, under assumption FE, we can re-write:

$$E[\delta_{ij}|A_i = S_i, A_j = S_i, T_i = T_j] = f_A(A_i = S_i, A_j = S_i) + f_T(T_i, T_j = T_i)$$

Putting this together, we can re-write $T_1$ as:

$$T_1 = \gamma + E[\eta_{ij}|A_j = S_i, T_i = T_j] + f_A(A_i = S_i, A_j = S_i) + f_T(T_i, T_j = T_i)$$

(2) $T_0$ : Average outcome of an applicant to primary school, $S_i$ who is admitted, and an incumbent student in a *different cohort* within the same primary school.

$$T_0 := \mathbb{E}[y_{ij} \mid A_i = S_i, A_j = S_i, T_i \neq T_j]$$

Under my assumptions, we can re-write $T_0$ as follows:

$$T_0 = E[\eta_{ij}|A_j = S_i, T_i \neq T_j] + f_A(A_i = S_i, A_j = S_i) + f_T(T_i, T_j \neq T_i)$$

(3) $C_1$ : Average outcome of an applicant to primary school $S_i$, who is *not admitted*, and an incumbent student in a the same cohort at primary school $S_i$.

$$C_1 := \mathbb{E}[y_{ij} \mid A_i \neq S_i, A_j = S_i, T_i = T_j]$$

Under my assumptions, we can re-write $C_1$ as follows:

$$C_1 = E[\eta_{ij}|A_j = S_i, T_i = T_j] + f_A(A_i \neq S_i, A_j = S_i) + f_T(T_i, T_j = T_i)$$

(4) $C_0$ : Average outcome of an applicant to primary school $S_i$ who is not admitted and an incumbent student in a different cohort at primary school $S_i$.

$$C_0 := \mathbb{E}[y_{ij} \mid A_i \neq S_i, A_j = S_i, T_i \neq T_j]$$

Under my assumptions, we can re-write $C_0$ as follows:

$$C_0 = E[\eta_{ij}|A_j = S_i, T_i \neq T_j] + f_A(A_i \neq S_i, A_j = S_i) + f_T(T_i, T_j \neq T_i)$$

Consider the estimator $\gamma_{FE}$, defined as follows:

$$\hat{\gamma}_{\text{FE}} = (T_1 - T_0) - (C_1 - C_0)$$

Plugging in the values for $T_1, T_0, C_1, C_0$, this can be expressed as follows:

$$\begin{aligned}
\hat{\gamma}_{\text{FE}} &= (T_1 - T_0) - (C_1 - C_0) \\
&= \gamma + \mathbb{E}[\eta_{ij} \mid \ S_i = A_j, \ T_i = T_j] + f_A(A_i = S_i, A_j = S_i) + f_T(T_i, T_j = T_i) \\
&\quad - E[\eta_{ij}|A_j = S_i, T_i \neq T_j] + f_A(A_i = S_i, A_j = S_i) + f_T(T_i, T_j \neq T_i) \\
&\quad - E[\eta_{ij}|A_j = S_i, T_i = T_j] + f_A(A_i \neq S_i, A_j = S_i) + f_T(T_i, T_j = T_i) \\
&\quad + E[\eta_{ij}|A_j = S_i, T_i \neq T_j] + f_A(A_i \neq S_i, A_j = S_i) + f_T(T_i, T_j \neq T_i) \\
&= \gamma + f_T(T_i, T_j = T_i) - f_T(T_i, T_j \neq T_i) - (f_T(T_i, T_j \neq T_i) - f_T(T_i, T_j \neq T_i)) \\
&= \gamma
\end{aligned}$$

### A3.2 Proof of Result 1

The share of students from school $s$ that pick school H is given by:

$$m_s^H = \frac{\exp(\alpha - \beta d_s + \gamma m_s^H)}{\exp(\alpha - \beta d_s^H + \gamma m_s^H) + \exp(\beta d_s^L + \gamma(1 - m_s^H))} \tag{A1}$$

Taking the derivative w.r.t. to $\gamma$ yields:

$$\begin{aligned}
\frac{\partial m_s^H}{\partial \gamma} = &\frac{\exp\left(\alpha - \beta d_s + \gamma m_s^H\right)\left(m_s^H + \gamma \frac{\partial m_s^H}{\partial \gamma}\right)\left(\exp\left(\alpha - \beta d_s^H + \gamma m_s^H\right) + \exp\left(\beta d_s^L + \gamma\left(1 - m_s^H\right)\right)\right)}{\left(\exp\left(\alpha - \beta d_s^H + \gamma m_s^H\right) + \exp\left(\beta d_s^L + \gamma\left(1 - m_s^H\right)\right)\right)^2} \\
&- \frac{\exp\left(\alpha - \beta d_s + \gamma m_s^H\right)\left[\exp\left(\alpha - \beta d_s^H + \gamma m_s^H\right)\left(m_s^H + \gamma \frac{\partial m_s^H}{\partial \gamma}\right) + \exp\left(\beta d_s^L + \gamma\left(1 - m_s^H\right)\right)\left(\left(1 - m_s^H\right) - \gamma \frac{\partial m_s^H}{\partial \gamma}\right)\right]}{\left(\exp\left(\alpha - \beta d_s^H + \gamma m_s^H\right) + \exp\left(\beta d_s^L + \gamma\left(1 - m_s^H\right)\right)\right)^2}
\end{aligned} \tag{A2}$$

Re-arranging and substituting the definition of $m_s^H$ and $m_s^L$, we can write this as:

$$\frac{\partial m_s^H}{\partial \gamma} = m_s^H(1-m_s^H)(m_s^H + \gamma\frac{\partial m_s^H}{\partial \gamma}) - m_s^H m_s^L(m_s^L + \frac{\partial m_s^L}{\partial \gamma})$$

Substituting $m_s^L = 1 - m_s^H$ we can write this as:

$$\frac{\partial m_s^H}{\partial \gamma} = m_s^H(1-m_s^H)(2m_s^H - 1 + 2\gamma\frac{\partial m_s^H}{\partial \gamma})$$

Re-arranging:

$$\frac{\partial m_s^H}{\partial \gamma} = \frac{m_s^H(1-m_s^H)(2m_s^H-1)}{1-2\gamma m_s^H(1-m_s^H)}$$

For $\gamma \in c(0,4)$, it follows that if $m_A^H > 0.5$, then $\frac{\partial m_s^H}{\partial \gamma} > 0$. Since $m_s^H$ is strictly monotonic, for every value of $\{d_A^H, d_B^H\}$ and parameters $\{\alpha, \beta, \gamma\}$, there exists a unique value of $m_s^H$. Therefore, there is a unique cut-off $\bar{d}(\gamma)$ s.t. if $d_s > \bar{d}(\gamma)$, $m_s^H < 0.5$.

□

### A3.3 Proof of Result 2

By the Contraction Mapping Theorem, if the function $\Phi(\hat{m}_{jst})$ is a contraction then it has a unique fixed-point satisfying:

$$\hat{m}_{jst} = \Phi(\hat{m}_{jst})$$

For $\Phi(\hat{m}_{jst})$ to be a contraction mapping, we require:

$$|\Phi(\hat{m}_{jst}) - \Phi(\hat{m}'_{jst}))| \leq \rho|\hat{m}_{jst} - m'_{st}| \quad \text{for some } \rho \in (0,1) \tag{A3}$$

This will be true whenever the derivative of $\Phi(\hat{m}_{jst})$ is less than 1:

$$\left|\frac{d\Phi(\hat{m}_{stj})}{d\hat{m}_{stj}}\right| \leq 1 \tag{A4}$$

In the remainder of this proof, I show the condition on $\gamma_0$ that will make $\Phi(\hat{m}_{jst})$ a contraction.

The equilibrium belief function is given by:

$$\Phi(\hat{m}_{stj}) = \int \mathbb{P}(A_i = j)dH(X_{ij}^{st}) \quad \forall s,t,j$$

Using definition of $\mathbb{P}(A_i = j)$, we can re-write this as:

$$\Phi(\hat{m}_{stj}) = \int \sum_{R_i \in \mathcal{R}_i} \sum_{\tilde{R}_{-i} \in \tilde{\mathcal{R}}_{-i}} \phi_{pj}^R \mathbb{P}(\tilde{R}_{-i})\mathbb{P}(R_i)dH(X_{ij}^{st}) \quad \forall s,t,j$$

where, in order to simplify notation, I write $\phi_{pj}^R = \phi_{pj}(R_i, \tilde{R}_{-i})$. Using the product-rule, we can take the derivative of this and write:

$$\frac{\partial \Phi(\hat{m}_{stj})}{\partial \hat{m}_{stj}} = \int \sum_{R_i \in \mathcal{R}_i} \sum_{\tilde{R}_{-i} \in \tilde{\mathcal{R}}_{-i}} \frac{\partial \phi^R_{pj}}{\partial \hat{m}_{stj}} \mathbb{P}(\tilde{R}_{-i}) \mathbb{P}(R_i) dH(X^{st}_{ij}) \int \sum_{R_i \in \mathcal{R}_i} \sum_{\tilde{R}_{-i} \in \tilde{\mathcal{R}}_{-i}} \phi^R_{pj} \frac{\partial \mathbb{P}(\tilde{R}_{-i})}{\partial \hat{m}_{stj}} \mathbb{P}(R_i) dH(X^{st}_{ij}) \int \sum_{R_i \in \mathcal{R}_i} \sum_{\tilde{R}_{-i} \in \tilde{\mathcal{R}}_{-i}} \phi^R_{pj} \mathbb{P}(\tilde{R}_{-i}) \frac{\partial \mathbb{P}(R_i)}{\partial \hat{m}_{stj}} dH(X^{st}_{ij}) \tag{A5}$$

Notice first that conditional on students' ranking, the probability of admission into schools is constant. Therefore,

$$\frac{\partial \phi^R_{pj}}{\partial \hat{m}_{stj}} = 0$$

We can also prove the following two lemmas:

**Lemma 1**: The derivative of the probability to observe a ranking $R_i$, with respect to $\hat{m}_{stj}$ is bounded by $\gamma_0 \bar{R} \mathbb{P}(R_i)$, where $\bar{R}$ is the maximum length of a ranking

**Lemma 2**: The derivative of the probability to observe a ranking $\tilde{R}_{-i}$ across students with respect to $\hat{m}_{stj}$ is bounded by $\gamma_0 \bar{R} |G_i| \mathbb{P}(\tilde{R}_{-1})$, where $\bar{R}$ is the maximum length of a ranking and $|G_i|$ corresponds to the size of student i's peer group

Using the Lemmas, we can write this derivative as:

$$\left| \frac{\partial \Phi(\hat{m}_{stj})}{\partial \hat{m}_{stj}} \right| \leq \int \sum_{R_i \in \mathcal{R}_i} \sum_{\tilde{R}_{-i} \in \tilde{\mathcal{R}}_{-i}} \phi^R_{pj} \left( \gamma_0 \bar{R} (|G_i| \cdot R \mathbb{P}(\tilde{R}_{-i}) \right) \mathbb{P}(R_i) dH(X^{st}_{ij}) \int \sum_{R_i \in \mathcal{R}_i} \sum_{\tilde{R}_{-i} \in \tilde{\mathcal{R}}_{-i}} \phi^R_{pj} \mathbb{P}(\tilde{R}_{-i}) \gamma_0 \bar{R} dH(X^{st}_{ij}) \tag{A6}$$

$$\left| \frac{\partial \Phi(\hat{m}_{stj})}{\partial \hat{m}_{stj}} \right| \leq \gamma_0 \bar{R} |G_i| \cdot R \int \sum_{R_i \in \mathcal{R}_i} \sum_{\tilde{R}_{-i} \in \tilde{\mathcal{R}}_{-i}} \phi^R_{pj} \mathbb{P}(\tilde{R}_{-i}) \mathbb{P}(R_i) dH(X^{st}_{ij}) \gamma_0 \bar{R} \int \sum_{R_i \in \mathcal{R}_i} \sum_{\tilde{R}_{-i} \in \tilde{\mathcal{R}}_{-i}} \phi^R_{pj} \mathbb{P}(\tilde{R}_{-i}) \mathbb{P}(R_i) dH(X^{st}_{ij}) \tag{A7}$$

Then,

$$\left| \frac{\partial \Phi(\hat{m}_{stj})}{\partial \hat{m}_{stj}} \right| \leq \gamma_0 \bar{R} (|G_i| + 1) \int \sum_{R_i \in \mathcal{R}_i} \sum_{\tilde{R}_{-i} \in \tilde{\mathcal{R}}_{-i}} \phi^R_{pj} \mathbb{P}(\tilde{R}_{-i}) \mathbb{P}(R_i) dH(X^{st}_{ij}) \tag{A8}$$

Notice that the expression within the integral is simply the probability that a student is admitted into school $j$, which we can re-write as:

$$\left|\frac{\partial\Phi(\hat{m}_{stj})}{\partial\hat{m}_{stj}}\right| \leq \gamma_0\bar{R}(|G_i|+1)\int \mathbb{P}(A_i=j)dH(X_{ij}^{st}) \tag{A9}$$

Since the integral is just a convex combination of probabilities, it is bounded by 1. Therefore,

$$\left|\frac{\partial\Phi(\hat{m}_{stj})}{\partial\hat{m}_{stj}}\right| \leq \gamma_0\bar{R}(|G_i|+1) \tag{A10}$$

since $|G_i|+1$ is simply the size of the primary school cohort attended by student i, we can re-write this as:

$$\left|\frac{\partial\Phi(\hat{m}_{stj})}{\partial\hat{m}_{stj}}\right| \leq \gamma_0\bar{R}|G_{st}| \tag{A11}$$

It follows then that for a given school $s$ and cohort $t$, if $\gamma_0 \leq (\bar{R}|G_{st}|)^{-1}$, then $\left|\frac{\partial\Phi(\hat{m}_{stj})}{\partial\hat{m}_{stj}}\right| < 1$. As a result, $\Phi(\hat{m}_{stj})$ is a contraction, and so has a unique fixed-point. □

*Proof of Lemma 1:* The probability of a full ranking $R_i$ occurring follows the ranked-ordered logit model:

$$\mathbb{P}(R_i) = \prod_{r=1}^{\bar{R}} \mathbb{P}(R_i(r)|S_r) \tag{A12}$$

where $R_i(r)$ is the school chosen in position $r$ in ranking $R_i$, and $S_r$ is the set of available schools at step $r$. Each conditional choice probability follows a logit form:

$$\mathbb{P}(R_i(r)|S_r) = \frac{\exp(V_{iR_i(r)})}{\sum_{\ell\in S_r}\exp(V_{i\ell})}. \tag{A13}$$

Differentiating $\mathbb{P}(R_i)$ with respect to $V_{ij}$ using the product rule:

$$\frac{\partial\mathbb{P}(R_i)}{\partial V_{ij}} = \sum_{r=1}^{J_i}\frac{\partial\mathbb{P}(R_i(r)=r_j|S_r)}{\partial V_{ij}}\prod_{l\neq r}\mathbb{P}(R_i(l)|S_l) \tag{A14}$$

By the definition of logit choice probabilities we can write:

$$\frac{\partial\mathbb{P}(R_i(r)=j_r|S_r)}{\partial V_{ij}} = \mathbb{P}(R_i(r)=j_r|S_r)(\mathbf{1}\{j_r=j\}-\mathbb{P}(R_i(r)=j|S_r))$$

where $\mathbf{1}\{j_r=j\}$ is an indicator function equal to 1 if $j_r$ is equal to $j$ and 0 otherwise. We can therefore write the derivative as:

$$\frac{\partial\mathbb{P}(R_i)}{\partial V_{ij}} = \sum_{r=1}^{J_i}\mathbb{P}(R_i(r)=j_r|S_r)(\mathbf{1}\{j_r=j\}-\mathbb{P}(R_i(r)=j|S_r))\prod_{l\neq r}\mathbb{P}(R_i(l)|S_l) \tag{A15}$$

$$\frac{\partial\mathbb{P}(R_i)}{\partial V_{ij}} = P(R_i)\sum_{r=1}^{\bar{R}}(\mathbf{1}\{j_r=j\}-\mathbb{P}(R_i(r)=j|S_r)) \tag{A16}$$

Notice that since $j_r$ can be equal to $j$ for at most 1 rank, it follows that the maximum value of this expression is $P(R_i)$. Alternatively, the smallest value is $-P(R_i)\bar{R}$. Therefore, we can say that this expression is bounded

by:

$$\left|\frac{\partial \mathbb{P}(R_i)}{\partial V_{ij}}\right| = \bar{R}P(R_i) \tag{A17}$$

From the definition of utility in Equation 10, we have that:

$$\frac{\partial V_{ij}}{\partial \hat{m}_{stj}} = \gamma_0 \tag{A18}$$

It follows that:

$$\left|\frac{\partial \mathbb{P}(R_i)}{\partial m_{stj}}\right| = \gamma_0 \bar{R}P(R_i) \tag{A19}$$

□

*Proof of Lemma 2:* The probability of a full ranking across students, $\tilde{R}_{-i}$ is given by:

$$\mathbb{P}(\tilde{R}_{-i}) = \prod_{\substack{k=1 \\ k\neq i}}^{N} \mathbb{P}(R_k) \tag{A20}$$

Taking the derivative this yields:

$$\frac{\partial \mathbb{P}(\tilde{R}_{-i})}{\partial \hat{m}_{stj}} = \sum_{k=1}^{N} \frac{\partial P(R_i)}{\partial \hat{m}_{stj}} \prod_{l\neq k} \mathbb{P}(R_l) \tag{A21}$$

However, note that $\hat{m}_{stj}$ does not impact the probability of students who are not in primary school $s$ at cohort $t$, therefore we can write this as:

$$\frac{\partial \mathbb{P}(\tilde{R}_{-i})}{\partial \hat{m}_{stj}} = \sum_{k=1}^{|G_i|} \frac{\partial P(R_k)}{\partial \hat{m}_{stj}} \prod_{l\neq k} \mathbb{P}(R_l) \tag{A22}$$

where $|G_i|$ is the number of students in student i's cohort. Following Lemma 1, we can re-write this as:

$$\left|\frac{\partial \mathbb{P}(\tilde{R}_{-i})}{\partial \hat{m}_{stj}}\right| \leq \sum_{k=1}^{|G_i|} \gamma_0 \bar{R} \cdot \mathbb{P}(R_k) \prod_{l\neq k} \mathbb{P}(R_l) \tag{A23}$$

From the definition of $\mathbb{P}(\tilde{R}_{-1})$ in Equation A20.

$$\left|\frac{\partial \mathbb{P}(\tilde{R}_{-i})}{\partial \hat{m}_{stj}}\right| \leq \gamma_0 \bar{R}|G_i|\mathbb{P}(\tilde{R}_{-1}) \tag{A24}$$

□

## A3.4 Identification of Model parameters

### A3.4.1 Identification of Structural Parameters and Fixed Effects

Consider the following utility function:

$$U_{ij} = \beta X_{ij} + \gamma_0 \hat{m}_{stj} + \gamma_1 \bar{X}_{stj} + \gamma_2 \bar{X}_{st} \cdot Y_j + \delta_{sj} + u_{ij}$$

As before, $X_{ij}$ corresponds to characteristics of school $j$ that are specific to student $i$. $\hat{m}_{stj}$ corresponds to the expected share of peers that will apply to a school. $\bar{X}stj$ corresponds to some average characteristic of students in primary school $s$ at time $t$ that is specific to school $j$ (for instance, the number of peers that live close to the school or have siblings at the school). $\bar{X}st$ corresponds to some average characteristic of students in class $s$ at time $t$ that is not specific to school $j$ (for instance, the share of high-income peers), and $Y_j$ is some characteristic of school $j$ (e.g. share of high-income students). In this specification, $\gamma_0$ corresponds to endogenous peer effects, $\gamma_1$ to contextual peer effects that drive preferences for a specific school, and $\gamma_2$ to contextual peer effects that drive preferences for a type of school.

I claim that the probability that a student ranks each high school as their most preferred option point identifies the vector of parameters $\theta = \{\beta, \gamma_0, \gamma_1, \gamma_2, \delta_{sj}\}$. In other words if:

$$\mathbb{P}(R_i(1) = j|\beta, \gamma_0, \gamma_1, \gamma_2, \delta_{sj}) = \mathbb{P}(R_i(1) = j|\tilde{\beta}, \tilde{\gamma_0}, \tilde{\gamma_1}, \tilde{\gamma_2}, \tilde{\delta}_{sj})$$

then:

$$\beta = \tilde{\beta}, \gamma = \tilde{\gamma_0}, \gamma_1 = \tilde{\gamma_1}, \gamma_2 = \tilde{\gamma_2}, \delta_{sj} = \tilde{\delta}_{sj}$$

I show that this will be the case if Assumption S-FE is satisfied, together with two technical assumptions on the distribution of students' observable characteristics within a school.

> **Non-degenerate distribution of student-school characteristics (NSD):** The distribution of students' observable characteristics, $H(X_{ij}^{st})$, is not degenerate.
>
> **Variation in the distribution of student-school characteristics (VSD):** The distribution of students' observable characteristics, $H(X_{ij}^{st})$, varies across time within a school:
>
> $$H(X_{ij}^{st}) \neq H(X_{ij}^{st'})) \quad \forall s, t' \neq t$$
>
> This is a sufficient but not necessary condition, as the expected share of peers admitted from one year to the next may also vary because of expected changes in admission probabilities.

*Identification of $\beta$* First consider the identification of $\beta$. Without loss of generality, consider the set of students in primary school $s_0$ and time $t_0$ considering whether to apply to some high school $j_0$. Define $k_0 = \gamma_0 m_{s_0 t_0 j_0} + \gamma_1 \bar{X}_{s_0, j_0} + \gamma_2 \bar{X}_{s_0 t_0} Y_{j_0} + \delta_{s_0 j_0}$. This is just a constant value as the expected share of peers who are admitted into a given high school does not vary within a primary school cohort. Consider some alternative $\tilde{k}_0$ and $\tilde{\beta}$. Assume that within this set of students:

$$F(k_0 + \beta X_{ij}) = F(\tilde{k}_0 + \tilde{\beta} X_{ij})$$

where $F(\cdot)$ corresponds to the probability to rank a school as a student's preferred option. As defined in Equation 13 this is simply given by the multinomial choice probabilities. Since $F$ is strictly monotonic, this implies that:

$$k_0 + \beta X_{ij} = \tilde{k}_0 + \tilde{\beta} X_{ij}$$

Because e we observe the choice probability for at least two different students in every class, it must then be the case that:

$$k_0 + \beta X_{1j_0} = \tilde{k}_0 + \tilde{\beta} X_{1j_0}$$

$$k_0 + \beta X_{2j_0} = \tilde{k}_0 + \tilde{\beta} X_{2j_0}$$

Subtracting one equation from the other:

$$\beta(X_{1j_0} - X_{2j_0}) = \tilde{\beta}(X_{1j_0} - X_{2j_0})$$

Under Assumption NSD we can always find two students in a class such that $X_{1j_0} \neq X_{2j_0}$. As such, this equation can only hold if $\beta = \tilde{\beta}$.

*Identification of $\gamma_0$ and $\gamma_1$* I now consider the identification of $\gamma_0$ and $\gamma_1$. Note first that if two sets of parameters yield the same probability to rank a school as a student's most preferred option, it must also yield the same probability to observe students' full ranking. That is:

$$P(R_i(1) = j|\tilde{\theta})) = P(R_i(1) = j|\theta) \quad \forall j \in S \quad \implies \quad P(R_i|\tilde{\theta}) = P(R_i|\theta) \quad \forall R_i \in \mathcal{R}_i$$

This follows directly from the definition of choice probabilities in the rank ordered logit model:

$$P(R_i(1) = j|\tilde{\theta})) = P(R_i(1) = j|\theta) \quad \iff \quad \frac{\exp(V_{ij}(\tilde{\theta}))}{\sum_j \exp(V_{ij}(\tilde{\theta}))} = \frac{\exp(V_{ij}(\theta))}{\sum_j \exp(V_{ij}(\theta))}$$

Re-arranging, this implies that the "weight" of each school, $\exp(V_{ij})$, under parameters $\tilde{\theta}$ is proportional to the weight under parameters $\theta$ :

$$\exp(V_{ij}(\tilde{\theta})) = \underbrace{\frac{\sum_j \exp(V_{ij}(\tilde{\theta}))}{\sum_j \exp(V_{ij}(\theta))}}_{:=c} \exp(V_{ij}(\theta))$$

Recall that the probability of observing a full ranking, $R_i$, is given by:

$$P(R_i|\tilde{\theta}) = \prod_{r=1}^{|J_i|-1} P(R_i(r) = j|\tilde{\theta})$$

Where the probability to observe school $j$ ranked in position $r$ is simply the probability that school $j$ is preferred among the remaining alternatives:

$$P(R_i|\tilde{\theta}) = \prod_{r=1}^{|J_i|-1} P(R_i(1) = j|S_r, \tilde{\theta})$$

Where

$$S_r = J_i / \{R_i(1), ..., R_i(r-1)\} \cup \emptyset\}$$

Substituting the expression for the probability to rank a school as a student's most preferred option, this yields:

$$\begin{aligned} P(R_i|\tilde{\theta}) &= \prod_{r=1}^{|J_i|-1} \frac{\exp(V_{ij}(\tilde{\theta}))}{\sum_{j \in S_r} \exp(V_{ij}(\tilde{\theta}))} \\ &= \prod_{r=1}^{|J_i|-1} \frac{c \cdot \exp(V_{ij}(\theta))}{\sum_{j \in S_r} c \cdot \exp(V_{ij}(\theta))} \end{aligned}$$

Since the coefficient $c$ cancels from both the numerator and denominator, this is equivalent to:

$$\begin{aligned} P(R_i|\tilde{\theta}) &= \prod_{r=1}^{|J_i|-1} \frac{\exp(V_{ij}(\theta))}{\sum_{j \in S_r} \exp(V_{ij}(\theta))} \\ &= P(R_i|\theta) \end{aligned} \tag{A25}$$

If the probability to observe a ranking is the same under both sets of parameters, then it must also be the case that the expected share of admitted peers is the same. To see this, recall from Equations 17 that the share of peers admitted into a school can be written as:

$$\begin{aligned} \tilde{m}_{stj}(\tilde{\theta}) &= \int \sum_{R \in \mathcal{R}} \phi_{pj}(R_i) \mathbb{P}(R_i|\tilde{\theta}) dH(X_{ij}^{st}) \\ &= \int \sum_{R \in \mathcal{R}} \phi_{pj}(R_i) \mathbb{P}(R_i|\theta) dH(X_{ij}^{st}) \\ &= \hat{m}_{stj}(\theta) \end{aligned} \tag{A26}$$

where the equality between the first line and the second is given by the fact that the probability to observe each ranking remains unchanged under both sets of parameters.

Consider first w.l.o.g. that $j$ is a school such that $Y_j = 0$. In this case, since the two sets of parameters imply the same probability a school is selected as a student's top option, the following equality holds:

$$F(\gamma_0 \hat{m}_{st_0j} + \gamma_1 \bar{X}_{st_0j} + \delta_{sj}) = F(\tilde{\gamma} \hat{m}_{st_0j} + \tilde{\gamma}_1 \bar{X}_{st_0j} + \tilde{\delta}_{sj})$$

Since $F$ corresponds to the multinomial logit choice probability, it is monotonically increasing. As such, the following must hold:

$$\gamma_0 \hat{m}_{st_0j} + \gamma_1 \bar{X}_{st_0j} + \delta_{sj} = \tilde{\gamma} \hat{m}_{st_0j} + \tilde{\gamma}_1 \bar{X}_{st_0j} + \tilde{\delta}_{sj}$$

Re-arranging:

$$(\gamma_0 - \tilde{\gamma_0}) \hat{m}_{st_0j} = -(\gamma_1 - \tilde{\gamma_1}) \bar{X}_{st_0j} + \tilde{\delta}_{sj} - \delta_{sj} \quad \forall j, s, t \tag{A27}$$

Since this must be the case for all students, it will also be true for another student in the same school in a different cohort:

$$(\gamma_0 - \tilde{\gamma_0}) \hat{m}_{st_1j} = -(\gamma_1 - \tilde{\gamma_1}) \bar{X}_{st_1j} + \tilde{\delta}_{sj} - \delta_{sj}$$

Subtracting the two equations from each other, we get that:

$$(\gamma_0 - \tilde{\gamma_0})(\hat{m}_{st_1j} - \hat{m}_{st_0j}) = (\gamma_1 - \tilde{\gamma_1})(\bar{X}_{st_0j} - \bar{X}_{st_1j}) \tag{A28}$$

Assumption VSD guarantees that there is some variation in $\bar{X}_{stj}$ across time, which will also generate variation in $\hat{m}_{stj}$ across time within a primary school.Then we can always find values of $\hat{m}_{stj}$ s.t. $\hat{m}_{st_1j} \neq \hat{m}_{st_0j}$ and $\bar{X}_{st_1j} \neq \bar{X}_{st_0j}$.

In this case, equation A28 can hold under two conditions. First, it may be the case that $\hat{m}_{stj}$ is a linear function of $\bar{X}_{stj}$. Second, it may be the case that the parameters are equal to each other. However from the definition of $\hat{m}_{stj}$ in Equation 18, it is clear that $\hat{m}_{stj}$ is a nonlinear function and so, Equation A28 can only hold if,

$$\begin{aligned}\gamma_0 &= \tilde{\gamma_0} \\ \gamma_1 &= \tilde{\gamma_1}\end{aligned} \tag{A29}$$

*Identification of* $\gamma_2$ By the same argument as before exploiting the strict monotonicity of F, the following condition must hold for a school where $Y_j = 1$:

$$\beta X_{ij} + \gamma_0 \hat{m}_{stj} + \gamma_1 \bar{X}_{stj} + \gamma_2 \bar{X}_{st} \cdot Y_j + \delta_{sj} = \beta X_{ij} + \gamma_0 \hat{m}_{stj} + \gamma_1 \bar{X}_{stj} + \tilde{\gamma_2} \bar{X}_{st} + \tilde{\delta}_{sj}$$

$$\gamma_2 \bar{X}_{st} + \delta_{sj} = \tilde{\gamma_2} \bar{X}_{st} + \tilde{\delta}_{sj} \quad \forall j, s, t \tag{A30}$$

Since this must hold for all students, schools and cohorts, we can find two students within the same primary school across two separate cohorts, such that:

$$\gamma_2 \bar{X}_{st_0} + \delta_{sj} = \tilde{\gamma_2} \bar{X}_{st_0} + \tilde{\delta}_{sj}$$

and

$$\gamma_2 \bar{X}_{st_1} + \delta_{sj} = \tilde{\gamma_2} \bar{X}_{st_1} + \tilde{\delta}_{sj}$$

Taking the difference of the two we get,

$$(\gamma_2 - \tilde{\gamma_2})(\bar{X}_{st_1} - \bar{X}_{st_0}) = 0$$

By Assumption VSD we can always find two cohorts within a school where $\bar{X}_{st_1} \neq \bar{X}_{st_0}$. Therefore Equation A30 can only hold if $\gamma_2 = \tilde{\gamma_2}$

*Identification of* $\delta_{sj}$ By the same argument as before, strict monotonicity implies that if the choice probability for two sets of fixed effects are the same, then:

$$\beta X_{ij} + \gamma_0 \hat{m}_{stj} + \gamma_1 \bar{X}_{stj} + \gamma_2 \bar{X}_{st} \cdot Y_j + \delta_{sj} = \beta X_{ij} + \gamma_0 \hat{m}_{stj} + \gamma_1 \bar{X}_{stj} + \gamma_2 \bar{X}_{st} + \tilde{\delta}_{sj}$$

which can only be satisfied if:

$$\delta_{sj} = \tilde{\delta}_{sj}$$

□

### A3.5 Identification of Outside Option $V_o$

I show that $V_o$ is uniquely identified by the number of schools that a student ranks. The probability that an individual ranks inside option $j$ before selecting the outside option is given by

$$p_{ij}(V_o) = \mathbb{P}(U_{ij} > U_{i0}) = \frac{\exp(V_{ij} - V_o)}{1 + \exp(V_{ij} - V_o)} \tag{A31}$$

Since $p_{ij}(V_o)$ is a function of $V_o$, its derivative with respect to $V_o$ is

$$\begin{aligned}\frac{\partial p_{ij}(V_o)}{\partial V_o} &= -\frac{\exp(V_{ij} - V_o)}{(1 + \exp(V_{ij} - V_o))^2} \\ &< 0\end{aligned} \tag{A32}$$

Thus, $p_{ij}(V_o)$ is strictly decreasing in $V_o$.

The expected number of ranked inside options for individual $i$ is given by

$$\mathbb{E}[K_i|X_i, \hat{m}_{stj}, \delta_{stj}] = \sum_{j=1}^{J_i} p_{ij}(V_o) \tag{A33}$$

Since each $p_{ij}(V_o)$ is strictly decreasing in $V_o$, it follows that

$$\frac{\partial \mathbb{E}[K_i|X_i, \hat{m}_{stj}, \delta_{stj}]}{\partial V_o} = \sum_{j=1}^{J_i} \frac{\partial p_{ij}(V_o)}{\partial V_o} < 0 \tag{A34}$$

Thus, $\mathbb{E}[K_i|X_{ij}, \hat{m}_{stj}, \delta_{stj}]$ is a strictly decreasing function of $c$ and is as such monotonic. This suggests that if:

$$\mathbb{E}[K_i|X_{ij}, \hat{m}_{stj}, \delta_{stj}, \tilde{V}_o] = \mathbb{E}[K_i|X_{ij}, \hat{m}_{stj}, \delta_{stj}, V_o]$$

then it must be the case that $V_o = \tilde{V}_o$

□

# A4 Simulation

## A4.1 Simulating equilibrium choice probabilities

In my conceptual framework, the probability that a student chooses a high school depends on the share of their peers who choose the high school. To estimate this probability, I implement a fixed-point algorithm to determine the equilibrium peer distribution, where the probability of each student applying to a school is updated iteratively. The core of the algorithm lies in solving for the stable set of probabilities $P_{ij}$. These probabilities represent the likelihood that student $i$ attending school $s$ applies to school $j$ and are influenced by both the school's characteristics and the decisions of other students in the same peer group.

The algorithm begins by initializing $P_{ij}^{(0)}$ for each student and school using uniform probabilities. I then compute $m_{sj}^1$, the average *predicted* number of peers who rank high school $j$ as their most preferred option, for every primary school:

$$m_{ij}^{(1)} = \frac{1}{N_s} \sum_{k \in \mathcal{G}_i} P_{kj}^{(0)}$$

where the summation is over all students $k$ that are in the peer group of student i, $\mathcal{G}_i$. Using these updated values of $m_{ij}^{(l)}$, I compute the next iteration of $P_{ij}^{(l+1)}$ as:

$$P_{ij}^{(1)} = F(X_{ij}, m_{ij}^{(1)}, \theta_s, \delta_s)$$

The algorithm iterates between recalculating $m_{sj}^{(l)}$ and updating $P_{ij}^{(l+1)}$ until convergence. Convergence is achieved when the difference between successive iterations of $P_{ij}^{(l)}$ falls below a predefined tolerance level:

$$\max_{i,j} \left| P_{ij}^{(l+1)} - P_{ij}^{(l)} \right| < \varepsilon$$

where $\varepsilon$ is a small positive number. Since each iteration of $P_{ij}^{(l)}$ implies a particular $m_{sj}^{(l)}$, the algorithm inherently captures the feedback loop between individual probabilities and the peer distribution. Since the probability function is continuous and monotonic in $m_{sj}$ , the existence of a fixed point is guaranteed under standard conditions. It is easy to show that when $\gamma \leq 4$ there is a unique equilibrium. This proof can be found in Bhattacharya et al. (2024).

## A4.2 Simulating admission probabilities

In this section, I provide a detailed description of the student-proposing Deferred Acceptance algorithm used to assign students to schools based on their ranked preferences and school priorities. This implementation closely follows the structure of Chile's centralized school admission system (SAE).

To begin, I first generate rank-ordered lists (ROLs), denoted by $\hat{R}_i$, which represent students preferences over schools. In each simulation, I draw a vector of idiosyncratic preferences, $\epsilon$, for each student over each school as well as the outside option. I then compute the utility for each option based on the estimated structural parameters $\{\hat{\beta}, \hat{\gamma}\}$ and primary school – high school fixed effects, $\hat{\delta}$, as follows:

$$\hat{U}_{ij} = \begin{cases} X_{ij}\hat{\beta} + \bar{m}_{ij}\hat{\gamma} + \hat{\delta}_{sj} + \epsilon_{ij} & \text{if } j \in J_i \\ \hat{V}_0 + \epsilon_{ij} & \text{if } j = \emptyset \end{cases} \tag{A35}$$

where $X_{ij}$ represents observable school characteristics, $\bar{m}_{ij}$ instruments for students' beliefs, and $\epsilon_{ij}$ follows a Type-I extreme value distribution. I then construct a ranking, $\hat{R}_i$ by sorting schools based on their computed utility values $\hat{U}_{ij}$. Formally, the ranking $\hat{R}_i$ is constructed by ordering the set of schools $J_i$ and the outside option $\emptyset$ in descending order of utility:

$$\hat{R}_i = \arg\operatorname{sort}_{j \in J_i \cup \{\emptyset\}} \hat{U}_{ij}, \quad \text{such that} \quad \hat{U}_{ij_1} \geq \hat{U}_{ij_2} \geq \cdots \geq \hat{U}_{ij_{|J_i|}} \tag{A36}$$

Given $\hat{R}_i$, I simulate a student-proposing Deferred Acceptance algorithm, which assigns students to schools based on their ranked choices and school-specific priority. Students first apply to their highest ranked school, $\hat{R}_i(1)$, and schools tentatively accept students according to their priority rankings $\pi_j(i)$. Priorities are defined to match those in the Chilean school admission system, so top priority is given to students whose siblings study at the school, followed by students whose parents work at the school. If the number of applicants exceeds a school's capacity $q_j$, the lowest-ranked students are rejected and must apply to their next preferred school. If a school is oversubscribed within a priority group, students are assigned random lottery numbers $\lambda_i$ and those with lottery numbers lower than the number of remaining seats are accepted. Rejected students remove the rejected school from their preference list and apply to their next highest ranked school. The following provides an overview of the algorithm:

1. **Initialization:** All students start unmatched:

$$A_i = \emptyset, \quad \forall i \in \mathcal{I} \tag{A37}$$

where $A_i$ denotes the current assignment of student $i$ (or $\emptyset$ if unmatched).

2. **Proposals:** Each unmatched student applies to their highest-ranked available school:

$$\mathcal{P}_j = \{i \in \mathcal{I} \mid A_i = \emptyset, \hat{R}_i(1) = j\} \tag{A38}$$

where $\mathcal{P}_j$ represents the set of students applying to school $j$ in the current round.

3. **Evaluation:** Each school ranks all applicants based on their priority order $\pi_j(i)$, using random lottery numbers $\lambda_i$ to break ties when necessary. Formally let $\mathcal{G}_{pj}$ correspond to all applicants within priority group p at school j:

$$\mathcal{G}_{pj} = \{i \in \mathcal{P}_j \mid \pi_j(i) = p\}$$

Within each priority group, schools ranks all applicants based on a random lottery draws $\lambda_i$ for tie-breaking. The lottery numbers are assigned as a random permutation over the applicants:

$$\lambda : \mathcal{G}_{pj} \to \{1, 2, \ldots, |\mathcal{G}_{pj}|\} \tag{A39}$$

Formally, the position of student $i$ in schools' preference ranking is given by:

$$\mathcal{A}_j = \text{sort}_{i \in \mathcal{P}_j} (\pi_j(i), \lambda_i) \tag{A40}$$

Let $\mathcal{A}_j(i)$ corresponds to the rank of student $i$ within school's preference list. That is:

$$i \succ_j \tilde{i} \quad \Longleftrightarrow \quad \mathcal{A}_j(i) < \mathcal{A}_j(\tilde{i}) \tag{A41}$$

4. **Admission & Rejection:** Schools accept the top $q_j$ students in their ranked list, $\mathcal{A}_j$. Formally,

$$A_i = \begin{cases} j, & \text{if } \mathcal{A}_j(i) \leq q_j \\ \emptyset, & \text{otherwise} \end{cases} \tag{A42}$$

Students in the first priority group are automatically admitted if the size of the priority group is below the size of the school. If a school is oversubscribed, there will be a marginal priority group, consisting of students whose priority level places them at the boundary of the school's remaining available seats. Within the marginal priority group, students will be admitted if their lottery number is lower than the remaining number of seats.

5. **Reapplications:** Rejected students remove the school that rejected them from their preference list and apply to their next most preferred school:

$$\hat{R}_i \leftarrow \hat{R}_i \setminus \{j\}, \quad \text{where } j = \hat{R}_i(1), \quad \forall i \text{ rejected} \tag{A43}$$

6. **Termination Condition:** The process repeats until no students are rejected in a round, ensuring that all unmatched students have either been assigned to a school or have exhausted their preference lists.

I repeat this process S = 100 times and then compute the average probability that a student ranks a school in positions 1-3 as follows:

$$\bar{\mathbb{P}}(R_i(r) = j) = \frac{1}{S} \sum_s \mathbf{1}\{R_i(r) = j\}$$

and the average probability that a student is admitted to a school as follows:

$$\bar{\mathbb{P}}(A_i = j) = \frac{1}{S} \sum_s \mathbf{1}\{A_i = j\}$$

where $s$ corresponds to a simulation draw and S to the total number of simulations.

# A5 Robustness

## A5.1 Teacher Effects

To address the concern that similarity in peers' preferences may be driven by exposure to the same teachers, I incorporate direct controls for the set of teachers that teach each pair of students. I link teacher records to school cohorts and construct a measure of overlap between the sets of teachers associated with each applicant–incumbent pair. While all students in a given cohort are taught by the same pool of teachers, there is variation across cohorts in the extent to which different students are taught by the same subset of teachers. As before, I estimate two specifications: (1) a fixed effects model where the outcome is the expected share of teacher overlap between students, if an applicant were to take up their admission offers; and (2) an instrumental variables specification in which the actual overlap is instrumented with the interaction of expected overlap and admission to students' preferred school. Table A8 presents the results. Across all outcomes, the LATE estimate corresponding to studying with a peer remains positive and statistically significant. Similarly, for most outcomes, the ITT estimate remains significant, but for a subset the estimation lacks the power to disentangle peer and teacher effects. Taken together, these results indicate that the estimated peer effects do not simply reflect shared exposure to the same teachers.

Table A8: Baseline regression controlling for teacher overlap

| Dependent Variables: | Attended same HS | | Enrolled same HS | | Admitted same HS | | Top-rank same HS | | Overlap in HS Ranking | |
|---|---|---|---|---|---|---|---|---|---|---|
| | ITT | LATE | ITT | LATE | ITT | LATE | ITT | LATE | ITT | LATE |
| Model: | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| *Variables* | | | | | | | | | | |
| Admitted top ps | 0.0055 | 0.0579*** | 0.0280*** | 0.0651*** | -0.0216*** | 0.0654*** | 0.0167* | 0.0436*** | 0.0351*** | 0.0884*** |
| | [-0.0094; 0.0205] | [0.0443; 0.0716] | [0.0154; 0.0406] | [0.0514; 0.0788] | [-0.0348; -0.0083] | [0.0520; 0.0788] | [-0.0004; 0.0338] | [0.0227; 0.0644] | [0.0208; 0.0493] | [0.0742; 0.1026] |
| Same year | 0.0246*** | 0.0216*** | 0.0006 | -0.0008 | 0.0268*** | 0.0223*** | 0.0005 | -0.0006 | 0.0067*** | 0.0042*** |
| | [0.0196; 0.0296] | [0.0167; 0.0265] | [-0.0015; 0.0026] | [-0.0033; 0.0017] | [0.0226; 0.0309] | [0.0185; 0.0261] | [-0.0023; 0.0033] | [-0.0041; 0.0028] | [0.0043; 0.0091] | [0.0014; 0.0071] |
| Peers | 0.0141** | 0.0572*** | 0.0053 | 0.0232*** | 0.0030 | 0.0809*** | 0.0038 | 0.0185* | 0.0094* | 0.0383*** |
| | [0.0018; 0.0265] | [0.0427; 0.0717] | [-0.0028; 0.0135] | [0.0070; 0.0394] | [-0.0082; 0.0143] | [0.0665; 0.0953] | [-0.0058; 0.0134] | [-0.0016; 0.0385] | [-0.0004; 0.0192] | [0.0218; 0.0548] |
| Teacher overlap | 0.0310** | -0.0263** | 0.0106 | -0.0114 | 0.0765*** | -0.0403*** | 0.0077 | -0.0109 | -0.0010 | -0.0363*** |
| | [0.0074; 0.0547] | [-0.0467; -0.0060] | [-0.0046; 0.0258] | [-0.0315; 0.0087] | [0.0543; 0.0986] | [-0.0580; -0.0225] | [-0.0095; 0.0249] | [-0.0373; 0.0156] | [-0.0195; 0.0175] | [-0.0588; -0.0138] |
| *Fixed-effects* | | | | | | | | | | |
| Top-ranked PS X Cohort X Group | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | | | | | | | |
| Observations | 2,142,333 | 2,138,236 | 2,287,913 | 2,283,038 | 1,907,870 | 1,904,332 | 2,326,928 | 2,322,049 | 2,326,928 | 2,322,049 |
| F-test (1st stage), attended$_i$*ps*$_i$ | | 2,149,133.8 | | 2,320,642.5 | | 2,012,614.5 | | 2,360,310.5 | | 2,360,310.5 |
| F-test (1st stage), Same year | | $3.66 \times 10^{36}$ | | $3.56 \times 10^{37}$ | | $1.12 \times 10^{37}$ | | $2.94 \times 10^{36}$ | | $2.94 \times 10^{36}$ |
| F-test (1st stage), peers | | 1,622,751.8 | | 1,904,163.8 | | 1,458,736.8 | | 1,957,946.0 | | 1,957,946.0 |
| F-test (1st stage), Teacher overlap | | 3,131,142.0 | | 3,933,842.8 | | 2,965,008.2 | | 4,015,033.5 | | 4,015,033.5 |
| $R^2$ | 0.18292 | 0.18583 | 0.16380 | 0.16733 | 0.14524 | 0.14611 | 0.19654 | 0.19811 | 0.27094 | 0.27973 |
| Adjusted $R^2$ | 0.18142 | 0.18434 | 0.16236 | 0.16590 | 0.14349 | 0.14435 | 0.19519 | 0.19676 | 0.26971 | 0.27852 |
| Wu-Hausman, p-value | | | | | | | | $2.67 \times 10^{-17}$ | | $4.12 \times 10^{-91}$ |

*Clustered (Top-ranked PS X Cohort X Group) co-variance matrix, 95% confidence intervals in brackets*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

## A5.2 Time Varying School Effects

In Chile, students admitted to a given primary school are allocated into class groups, known as *grupos-cursos*, at the start of primary school. Once formed, these class groups are highly stable: over 80% of students remain in the same class group throughout the entire primary cycle. By law, schools have full discretion over how to allocate students into class groups, and there are no explicit national guidelines governing the process. Moreover, publicly available internal regulations from schools rarely mention any formal criteria for class assignment.

To better understand how class groups are formed in practice, I submitted freedom of information requests to two of Chile's 346 municipalities inquiring about the internal regulations guiding the assignment. The responses reveal that schools prioritize assigning incoming students to the class with the most available space, based on capacity figures determined in advance by the school, or to the class that will preserve gender balance across classrooms. Importantly, this allocation takes place before students arrive at the school – usually in February – and is based solely on the information available in the School Admission System (SAE). As a result, schools are unable to condition class assignments on unobserved student characteristics. Furthermore, reassignments between classes are rare, and most students remain with the same peer group over time. This institutional arrangement generates variation in classroom peer composition within school – cohort cells that, while not random, is plausibly orthogonal to student characteristics. Indeed, Appendix Table A6 shows that applicants admitted to their preferred primary school are no more similar to classmates in their assigned class than to students in other classes.

If applicants are effectively randomized into classes at the point of admission, then comparisons across classes within a primary school cohort can identify the causal effect of studying with a particular peer. A second assumption we require in order to identify peer effects from inter-class comparisons is that any school effects are constant at the cohort level and do not vary by class. Provided that class assignment is orthogonal to unobserved student, it is plausible that any school-time varying shocks would influence all students the same way. Moreover, students across classes within a cohort are taught by highly overlapping sets of teachers. In multi-class schools, 99.2% of students share at least one teacher with students in another class and 93.9% share at least half of their teachers. On average, students across classes share 72.8% of teachers. Formally, we can state the identifying assumptions as follows:

> **Assumption OCA (Orthogonal Class Assignment):** Within each school–cohort cell, students' allocation to classes is orthogonal to their unobserved similarity to other students.
>
> $$\mathbb{E}[\eta_{ij} \mid C_i,\ A_i = A_j,\ T_i = T_j] = \mathbb{E}[\eta_{ij} \mid A_i = A_j,\ T_i = T_j]$$
>
> **Assumption CSA (Constant school effects):** School effects are constant across classes within a primary school cohort:
>
> $$\mathbb{E}[\delta_{ij}|C_i, C_j, A_i, A_j, T_i, T_j] = f_A(A_i, A_j) + f_T(T_i, T_j) \quad \forall A_i, A_j \in \mathcal{S},\ T_i, T_j \in \mathcal{T}$$

Under these assumptions, any systematic difference in ranking similarity between students in the same class and in different classes – within the same school and cohort – can be attributed to peer effects rather than institutional or compositional differences. As a result, we can identify $\gamma$ by comparing the similarity in high school rankings between students $i$ admitted to school $s$ in class $c$ and cohort $t$, and $j$, where $j$ is an incumbent student in the same class, to the similarity between $i$ and $j'$, where $j'$ is an incumbent student in a different class, in the same school and cohort.

Let $i$ index an applicant student $j$ is fixed as an incumbent student in school $A_i$ at cohort $T_i$. Because class assignment is assumed to be random, and school-level influences $\delta(A_i, T_i, A_i, T_i)$ are constant across classes, this expression isolates $\gamma$, the causal effect of peer exposure.

$$\begin{aligned}
&\mathbb{E}[y_{ij} \mid A_i = A_j,\ T_i = T_j,\ C_i = C_j] - \mathbb{E}[y_{ij} \mid A_i = A_j,\ T_i = T_j,\ C_i \neq C_j] \\
&= \gamma + f_A(A_i, A_j = A_i) + f_T(T_i, T_j = T_i) + \mathbb{E}[\eta_{ij} \mid A_i = A_j,\ T_i = T_j,\ C_i = C_j] \\
&\quad - (f_A(A_i, A_j = A_i) + f_T(T_i, T_j = T_i) + \mathbb{E}[\eta_{ij} \mid A_i = s, A_i = A_j,\ T_i = T_j,\ C_i \neq C_j]) \\
&= \gamma + \mathbb{E}[\eta_{ij} \mid A_i = A_j,\ T_i = T_j] - \mathbb{E}[\eta_{ij} \mid A_i = A_j,\ T_i = T_j] \\
&= \gamma
\end{aligned} \tag{A44}$$

Assumptions OCA and CA motivate the following triple-difference regression specification. For each applicant $i$ who ranked school $S_i$ as their preferred choice in year $T_i$, I regress whether they rank the same high schools as an incumbent student $j$ in school $S_i$ on: (1) an an indicator for whether they were admitted to the school; (2) an indicator for whether they studied in the same cohort; (3) a double interaction term for whether they studies together in the same class and cohort; (4) an indicator for whether they studies in the same class; and (5) a triple-difference indicator for whether they studies in the same class, in the same cohort and in the same school; and (6) fixed effects for every primary school, cohort and applicant group combination.

$$\begin{aligned}
y_{ij} = \beta_1 \cdot \mathbf{1}\{A_i = S_i\} + \beta_2 \cdot \mathbf{1}\{C_i = C_j\} + \beta_3 \cdot \mathbf{1}\{T_i = T_j\} + \beta_4\{A_i = S_i, T_i = T_j\} + \\
+ \gamma_{OCA} \cdot \mathbf{1}\{A_i = S_i,\ C_i = C_j,\ T_i = T_j\} + \theta_{S_i, T_i, g_i} + \epsilon_{ij}
\end{aligned} \tag{A45}$$

Here, the coefficient $\beta_1$ reflects institutional effects of attending school $s$ that are constant across classes, while $\beta_2$ captures any overall within-school class-level correlation. Fixed effects $\theta_{R_i, t_i}$ control for unobserved differences across students who apply to different schools in different cohorts.

To exploit these assumptions, I focus on applicants who applied to primary schools with multiple classes and were admitted. I then construct applicant-incumbent pairs among students within the same cohort, and create a dummy for whether applicant $i$ and incumbent $j$ are class peers. Finally, I regress each outcome on the class peer along with primary school fixed effects. I present the results in Table A9. Reassuringly, admitted applicants are significantly more likely to and rank, be admitted to, enroll in, and attend the same high schools as their class peers, relative to other students in their primary school cohort.

Table A9: Main Regression Exploiting Variation Across classes

| Dependent Variables:<br>Model: | Attended same HS<br>(1) | Enrolled same HS<br>(2) | Admitted same HS<br>(3) | Top-rank same HS<br>(4) | Overlap in HS Ranking<br>(5) |
|---|---|---|---|---|---|
| *Variables* | | | | | |
| Class Peers | 0.0082** | 0.0067** | 0.0087** | 0.0123*** | 0.0056** |
| | [0.0003; 0.0161] | [0.0013; 0.0120] | [0.0015; 0.0159] | [0.0041; 0.0205] | [0.0009; 0.0103] |
| *Fixed-effects* | | | | | |
| Top-ranked PS X Cohort X Group | Yes | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | | |
| Observations | 87,274 | 138,816 | 83,998 | 143,523 | 143,523 |
| $R^2$ | 0.22582 | 0.22354 | 0.21461 | 0.27997 | 0.34196 |
| Adjusted $R^2$ | 0.21592 | 0.21323 | 0.20416 | 0.27073 | 0.33351 |

*Clustered (Top-ranked PS X Cohort X Group) co-variance matrix, 95% confidence intervals in brackets*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

## A6 Schools and educational inequality

This section reviews my framework for decomposing differences in the raw means of educational outcomes between groups into differences in the average 'quality' of schools students attend and differences in the average match-effect and 'advantage' of students. Let $Y_{ij}$ denote the potential outcomes of student $i$ were they to attend school $j \in \mathcal{J}$. We can decompose $Y_{ij}$ into three components:

$$Y_{ij} = \underbrace{\alpha_i}_{\text{Student effect}} + \underbrace{v_j}_{\text{School effect}} + \underbrace{\epsilon_{ij}}_{\text{Match effect}} \tag{A46}$$

where $\alpha_i$ is student i's average potential outcome across all schools; it vertically differentiates students, reflecting traits such as ability and ambition, but also parent's ability to support their children. $v_j$ correspond to school effects, which capture each school j's average causal value-addded on students' outcomes. Finally, $\epsilon_{ij}$ captures the quality of the match between student $i$ and school $j$, it corresponds to the causal effect of attending school $j$ for student $i$. This specification is similar to the specification commonly adopted in the value-added literature, but is more flexible as it does not restrict the value-added of school to be homogeneous across individuals (see, for example, Abowd et al., 1999; Chetty et al., 2014a; Angrist et al., 2016, 2017). [44]. Chetty et al. (2014b)

Differences in outcomes between different groups of students can represent differences in each of these dimensions. Poorer students are likely to attend worse schools and, at the same time, have lower potential outcomes on average. Letting $S_i = j$ indicate that student $i$ attends school $j$, and $g_i$ the group that student $i$ belongs to, we can decompose the raw mean outcomes between students in groups $g_1$ and $g_0$ as follows:

$$\underbrace{\mathbb{E}[Y_{ij}|g_1] - \mathbb{E}[Y_{ij}|g_0]}_{\substack{\text{Difference in raw means} \\ :=\Delta Y}} = \underbrace{\mathbb{E}[\alpha_i|g_1] - \mathbb{E}[\alpha_i|g_0]}_{\substack{\text{Difference in average} \\ \text{student potential outcomes} \\ :=\Delta\alpha}} + \underbrace{\mathbb{E}[v_j|g_1] - \mathbb{E}[v_j|g_0]}_{\substack{\text{Difference in school} \\ \text{vertical quality} \\ :=\Delta v}} + \underbrace{\mathbb{E}[\epsilon_{ij}|g_1] - \mathbb{E}[\epsilon_{ij}|g_0]}_{\substack{\text{Difference in school} \\ \text{match quality} \\ :=\Delta\epsilon}} \tag{A47}$$

[44] See Kirkebøen et al. (2021), Bonhomme et al. (2019), Lamadon et al. (2019), Hull (2020), Abdulkadiroglu et al. (2020), Goldsmith-Pinkham et al. (2021) for recent approaches relaxing value-added homogeneity.

Without imposing any restrictions on sorting, we cannot identify $\Delta\alpha+\Delta\epsilon$ using a fixed effects estimator. However, one way to overcome selection issues is to take the mean of school-level comparisons between different groups of students. The idea is that while any comparison may be biased due to selection, by the law of iterated expectation, the global average should not be influenced by selection (even if different groups select differently). To see this note that,

$$\begin{aligned}
E_j\{E_i[Y_{ij}|g_1,j]-E_i[Y_{ij}|g_0,j]\} = & \\
&= E_j\{E_i[\alpha_i|g_1,j]+E_i[\epsilon_{ij}|g_1,j]+v_j-(E_i[\alpha_i|g_0,j]+E_i[\epsilon_{ij}|g_0,j]+v_j)\} \\
&= E_j\{E_i[\alpha_i|g_1,j]-E_i[\alpha_i|g_1,j]+E_i[\epsilon_{ij}|g_0,j]-E_i[\epsilon_{ij}|g_0,j]\} \\
&= E_j\{E_i[\alpha_i|g_1,j]\}-E_j\{E_i[\alpha_i|g_1,j]\}+E_j\{E_i[\epsilon_{ij}|g_0,j]\}-E_j\{E_i[\epsilon_{ij}|g_0,j]\} \\
&= E_i[\alpha_i|g_1]-E_i[\alpha_i|g_0]+E_i[\epsilon_{ij}|g_1]-E_i[\epsilon_{ij}|g_0] \\
&= \Delta a+\Delta\epsilon
\end{aligned} \tag{A48}$$

where the Law of Iterated Expectations is invoked to move from the fourth line to the fifth line.

**Estimation:** We can estimate $E_j\{E_i[Y_{ij}|g_1,j]-E_i[Y_{ij}|g_0,j]\}$, the average school-level difference in outcomes, directly using the following estimator:

$$\hat{\mu}=\frac{1}{J}\sum_j\left[\frac{1}{N_j^1}\sum_{\substack{G_i=g_1\\S_i=j}}Y_{ij}-\frac{1}{N_j^0}\sum_{\substack{G_i=g_0\\S_i=j}}Y_{ij}\right] \tag{A49}$$

where J corresponds to the total number of schools and $N_j^{G_i}$ corresponds to the number of students of type $G_i$ in school j. The variance of this estimator is given by:

$$Var(\hat{\mu})=\frac{1}{J^2}\sum_j\left[\frac{\sigma^2_{1,j}}{N_j^1}+\frac{\sigma^2_{0,j}}{N_j^0}\right] \tag{A50}$$

where $\sigma^2_{G_i,j}$ corresponds to the variance of outcomes of students of type $G_i$ in school j. This estimator, however, requires full treatment support, which in our context means there should be students of each group in every school. Without further assumptions on heterogeneity, we require at least 2 students of each type in every school. Under the assumption of homogeneous variance within each group, we only require 1 student of each type in every school.

I apply the method described above to decompose the role of schools in generating gaps in standardized test scores. The SIMCE is a standardized exam that students take during the second year of high school (when students are typically 15-16 years old). It is graded on a scale running from 0 to 500, where the mean and median are, respectively, 264 and 258 points. In 2018, students whose mother had a high school education, constituting 70.7% of test-takers, scored, on average, 277 points, whereas students whose mother had less than a high school education scored only 239 points, corresponding to a difference of $0.6\sigma$. Table A10 presents the breakdown of standardized test scores to identify schools' influence on the achievement gap. Column (1) compares raw means between high- and low-SES students, defined on the basis of maternal education, while Column (2) provides the estimate of the difference in a regression that includes school fixed effects. Using this method, schools appear to account for around 76.5% of the test score difference between the groups. Columns (3) to (5) restrict the sample to the subsample with full support, that is, schools where

there is at least one student from each group (86% of schools, 92% of students). Column (3) again looks at the difference in raw means, which is smaller than gap in raw means for the whole sample, indicating that this sample is positively selected. Column (4) corresponds to the gap accounting for school fixed effects, and suggests that schools explain 72% of the achievement gap between students in the restricted subsample. Column (5) corresponds to the mean within-school average gap, using this approach, schools appear to explain only 66% of the achievement gap. Overall, the analysis indicates that schools significantly contribute to explaining the achievement gap between low- and high-SES students (over 60% across all specifications) and suggests non-symmetric selection patterns into schools across socioeconomic groups; either higher-SES students with higher average potential scores are more likely to select into better schools or higher-SES students enjoy greater match effects at schools.

| | Mathematics Test scores | | |
|---|---|---|---|
| | Base | FE | Within-school comparison |
| | (1) | (2) | (3) |
| $\Delta\hat{Y}$ | 32.46 | | |
| | [31.70, 33.22] | | |
| $\Delta\hat{\alpha} + \Delta\hat{\epsilon}$ | | 9.08 | 11.14 |
| | | [8.44, 9.69] | [10.03, 12.24] |
| $(1 - (\Delta\hat{\alpha} + \Delta\hat{\epsilon})/\Delta\hat{Y})$ | | 72.03% | 65.58% |
| *100 % | | [65.50,78.56] | [60.88 70.28] |
| N | 149,016 | 149,016 | 149,016 |
| J | 2,498 | 2,498 | 2,498 |

Table A10: Decomposition of standardized exam scores

*Note:* Note: The table corresponds to the decomposition of Maths test scores into a between-school and within-school component. The first row corresponds to the raw gap in scores between high- and low- SES students – defined based on whether a student's mother completed high school. The second row corresponds to estimates of the average within-school difference in test scores estimated using the approach outlined in Section A6. The third row corresponds to estimates of the average between-school component, obtained by subtracting the within-school estimate from the raw gap and normalizing by the size of the gap. Confidence intervals correspond to 95% confidence level.

# References

Abdulkadiroglu, A., P. A. Pathak, J. Schellenberg, and C. R. Walters (2020). Do parents value school effectiveness? *American Economic Review 110*, 1502–1539.

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Angrist, J. D., S. R. Cohodes, S. M. Dynarski, P. A. Pathak, and C. R. Walters (2016). Stand and deliver: Effects of boston's charter high schools on college preparation, entry, and choice. *Journal of Labor Economics 34*, 275–318.

Angrist, J. D., P. D. Hull, P. A. Pathak, and C. R. Walters (2017). Leveraging lotteries for school value-added: Testing and estimation. *Quarterly Journal of Economics 132*(2), 871–919.

Bhattacharya, D., P. Dupas, and S. Kanaya (2024). Demand and welfare analysis in discrete choice models with social interactions. *Review of Economic Studies 91*(2), 748–784.

Bonhomme, S., T. Lamadon, and E. Manresa (2019). A distributional framework for matched employer employee data. *Econometrica 87*, 699–739.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014a). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American Economic Review 104*(9), 2593–2632.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014b, September). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American Economic Review 104*(9), 2593–2632.

Goldsmith-Pinkham, P., P. Hull, and M. Kolesar (2021). On estimating multiple treatment effects with regression. Technical report, Working paper.

Hull, P. (2020). Estimating hospital quality with quasi-experimental data. Technical report, Working paper.

Kirkebøen, L., E. Leuven, M. Mogstad, and J. Mountjoy (2021). College as a marriage market. NBER Working Paper 28688, National Bureau of Economic Research.

Lamadon, T., M. Mogstad, and B. Setzler (2019). Imperfect competition, compensating differentials, and rent sharing in the u.s. labor market. Technical report, Working paper.