When Schools Do Not Like Students: Direct and Spillover Effects of Cream-Skimming

Diego Escobar Salce* María-Adelaida Martínez

December 15, 2022

Link to most recent version.

Abstract

Cream skimming occurs when private or charter schools attract the best students from public schools, draining the latter of higher performing students. This paper studies the effect of school-side student selection on this pattern and its direct and spillover effects on students' outcomes. Using rich administrative data from Chile, we document strong evidence of school-side selection in publicly subsidized private schools, explaining up to 20 percent of the performance gap with respect to public schools. Leveraging centralized admission lotteries to simulate the distribution of potential classroom compositions, we estimate the impact of this screening on students' academic performance, college enrollment, and behavioral outcomes. While we document value-added benefits of attending selective schools, we find no difference between effects on low and high income students. These findings oppose school-student fit as the major driver for screening. In contrast, we find support for sizable peer effects in classrooms that received lottery-induced shocks to their class composition, which potentially explains schools' implementation of screening practices.

*Escobar Salce and Martínez: Harris School of Public Policy, University of Chicago. Emails: descobarsalce@uchicago.edu; mariaadelaidamc@uchicago.edu. All remaining errors are ours.

1 Introduction

There are several reasons why segregation arises in educational settings. On the demand side, families apply to different schools due to their income levels, residential segregation, preferences for educational quality, and even their preferences for peers (Abdulkadiroğlu et al., 2020, Idoux, 2022). On the supply side, schools often seek to enroll high-income students due to their higher willingness to pay tuition fees, their higher non-school inputs, and flat government subsidies that induce schools to enroll students who are less costly to educate (Epple and Romano, 2008). Moreover, the presence of social interactions and imperfect school quality signals also introduce incentives for schools to implement screening practices in order to enroll students that are attractive to other families (Epple and Romano, 1998; Allende, 2019).

In part due to limitations identifying the effects of these practices, the literature evaluating the effects of school choice on the allocation of students across schools has primarily focused on the aggregate effects of school choice on students' sorting (e.g., Hsieh and Urquiola, 2006) and on demand-side factors, such as the incentives for high achieving students to switch schools (e.g., Muralidharan and Sundararaman, 2015 and Altonji et al., 2015). However, disentangling these supply and demand mechanisms is crucial when designing educational policies. For example, while commonly employed informational campaigns to aid families in their school choice decisions can be highly effective against segregation arising from demand factors, they are futile against sorting emerging from supply-side selection mechanisms, called *screening*.

In this paper we aim to fill this gap by assessing the prevalence of supply-side cream-skimming induced by screening, its effects on the equilibrium allocation of students into schools, and its impact on benefited and displaced students. We estimate this by exploiting the staggered implementation of the School Admission System (*SAS*) in Chile starting in 2016, which forced all schools receiving public funding to join a centralized admission system, covering over 90 percent of the nationwide enrollment. This system implemented mandatory lotteries to allocate spots whenever a school receives more applications than their available spots, taking away schools' discretionary power to select students among their applicants.[1] Our analysis shows that the introduction of the SAS reduced the baseline achievement of students enrolling in publicly subsidized private schools (*voucher* schools) and high-performing schools. This finding confirms that a portion of the performance premium of these schools

[1] A fraction of these schools can still engage in screening through pricing policies, restricting access to low-income students. However, conditional on their applications, schools have no means to alter the selected subset of students.

comes directly from selection rather than improved school value-added. Moreover, the magnitude of the effect is sizable: the changes in incoming students' standardized scores are equivalent to 20 percent of the average test scores' gap between public and voucher schools.

The Chilean setting is particularly well suited to tackle these questions for several reasons. The drastic nationwide change from a largely unregulated admission system to a centralized one restricting schools' screening presents a vastly unusual scenario. In particular, Chile was one of the first to adopt a nationwide school choice system, where public schools (Public), publicly funded private schools (Voucher), and privately funded private schools (Private) must compete to fund themselves by charging tuition fees and by receiving government resources based on enrollment levels. This setting also allows us to use rich administrative data to track student and school-level outcomes over time, analyzing the applications to 31,032 classrooms in 6,123 schools across the country, sequentially reaching the entire population of students.

Sorting into schools has broadly been acknowledged as a primary mechanism for explaining the variance in student achievement (Nechyba, 2006). We propose a framework with two mechanisms to evaluate the impact of the changes in allocations induced by screening on students' outcomes: direct effects through school enrollment and spillover effects through classroom composition. First, we identify the direct value-added gains of attending selective schools by exploiting the randomness in offers induced by oversubscription lotteries to avert selection bias, as in Abdulkadiroğlu et al. (2017) and Angrist et al. (2017). Specifically, we simulate the empirical distribution of schools' admission offers under counterfactual random draws to find students with equal admission propensity to a given school but with different realized offers.[2] To test student-school mismatch theories as drivers for selection (e.g., Sander, 2004), our estimation allows for heterogeneous student-school value-added. Secondly, we expand this methodology in a novel way by exploiting the empirical distribution to identify classroom composition effects. In particular, we compare the performance of classrooms with an equivalent ex-ante empirical distribution of changes in classroom composition but with different realized shocks.

One fundamental challenge when assessing the impact of screening practices is that equilibrium allocations of students across schools depend on schools' capacity constraints: admitting a given student requires rejecting another in schools with excess demand. Consequently, evaluating changes to school admission systems requires assessing the impacts on students benefited and displaced by these policies. Defenders of these practices argue that they may

[2] Considering that students may reject their offers to apply in a second round, or enroll into pricier unsubsidized private school outside the system, we use the admission offers as an instrumental variable for actual enrollment, as detailed in Section 5.1.

improve allocative efficiency by improving the matching between schools and students, potentially benefiting all types of students. In contrast, detractors often claim that these policies introduce segregation, which can negatively impact students' development and increase inequality. Our results indicate that low-income students randomly obtaining access to selective schools benefit similarly to their high-income counterparts in terms of college enrollment and standardized test scores. Nevertheless, these changes in peers' allocations affect students through classroom composition: we find evidence that changes in peers' backgrounds affect their educational and behavioral outcomes, further enlarging the observed performance gap between selective and unselective schools.

This system implemented with the SAS is a nationwide centralized version of the Deferred Acceptance (DA) algorithm. Before the SAS, schools implemented several screening mechanisms ranging from academic selection, psychological assessments, tuition charges, religious factors, income level verification, and family background checks, among others. Instead, in the new system all publicly funded schools must report their slot availability before the application period and randomly assign spots among applicants in case of oversubscription. Unsubsidized private schools, which represented 7.8 percent of enrollment in 2015, can freely screen students even after the introduction of the SAS. Most public schools do not perform screening even before the reform, so their admission criteria are predominantly unaffected. Voucher schools could screen students at all levels prior to the reform, or at grades 7th or higher if they subscribe to a targeted subsidy program (52 percent of students enrolled in voucher schools in 2015 attended program-affiliated schools). In consequence, the SAS halted a large amount of subsidized private schools from employing screening practices.

When analyzing the impact of this screening prohibition, we find that the baseline standardized scores of students enrolled in selective voucher schools through the SAS decreased by 0.12 standard deviations. Conversely, the proportion of low-SES students in voucher schools increased by 8.2 percent. High-performing and highly demanded schools follow the same patterns as voucher schools in general, although with considerably larger magnitudes, given these schools' stricter admission policies before the reform. When interpreting these results, it is critical to consider that applicants' heterogeneity and schools' degree of oversubscription limit the changes to equilibrium allocations induced by the SAS. This constraint arises because the SAS affected the feasibility of supply-side cream skimming but not other factors inducing demand-side differentiation. In turn, demand-side heterogeneity depends on factors such as differences in family preferences, residential segregation, and willingness to pay tuition fees, which are not directly affected by the SAS. We see our results as complement to those by Kutscher et al. (2020), who directly measure the impact of the SAS implementation on school-level segregation, finding that its impacts depend critically on local school supply and residential segregation. However, they do not disentangle supply and demand factors and do not measure the impact on students' outcomes, which are part our main contributions.

Focusing on direct effects, we find that high and low-income students enrolling in high-performing schools improve their standardized test scores by up to 0.3 standard deviations. However, we find no significant differences between the gains for both groups. Our results regarding value-added gains in college enrollment and national admission exam scores are more muted, but there are no significant differences between low- and high-income students. On the other hand, consistently with more demanding standards at high-performing schools, we also observe a decrease in students' GPA and grade advancement rate. This drop is slightly larger among low-income students, partially explaining the decrease in self-reported levels of motivation observed among low-income students enrolling in high-performing schools. Our results are similar when we focus instead on other schools with a small proportion of low-income students or schools with more restrictive application processes, where low-income students were more likely to be rejected before the reform. We interpret the minor differences in the value-added gains obtained by high and low-income students as opposing commonly held mismatch theories of school selection as drivers behind screening.

When turning our attention to spillover effects, it is essential to consider that the SAS classroom allocations play a prominent role in the Chilean system because students share all of their subjects with the same classmates' group, usually remaining unchanged for several years. Our results indicate positive effects of high-achieving peers on college enrollment and grade advancement: an increase of one standard deviation in classmates' average standardized scores significantly increased their classmates' grade advancement by 8 percent, despite reducing their GPA ranking within the school. Similarly, improving classmates' average standardized scores increases college admission exam scores of their peers in Math and Reading college admission exams.

Besides impacts on students' academic performance, we find that an increase in classmates' average standardized scores significantly decreases students reported motivation and self-confidence, although it also reduces their school-behavior problems and increases attendance. These results are consistent with adverse effects on motivation for low-income students attending high-performing schools, reflecting that some students may feel discouraged when participating in classrooms where their peers perform better than them.

Regarding the effects on students attending selective schools, the literature has mainly focused on higher education. In this domain, several authors have identified significant benefits of attending more selective colleges (e.g., Black et al. (2020) in Texas and Otero et al. (2021) in Brazil), although the literature is divided regarding the conditions under which these policies are effective (see Arcidiacono and Lovenheim, 2016 for a literature survey). Indeed, the efficiency consequences of redistributing slots will depend on factors such as the complementarity between students' preparation and schools' value-added (Durlauf, 2008) and private information about student-school match quality (Arcidiacono and Lovenheim, 2016). The focus on higher education is partly due to the widespread application of screening practices in higher education and the comparatively more transparent college admission mechanisms in some countries.[3] However, evidence is scarce in the context of secondary education, particularly in the US, where residential and enrollment decisions are highly intertwined and identification often requires strong structural assumptions.

Our results contribute to the broad literature evaluating the impacts of school choice expansion. In particular, most of the literature has not been able to isolate the spillover effects of school choice on students remaining in public schools and those already in private schools (Muralidharan and Sundararaman, 2015). Their study provides an exception to this by exploiting the experimental expansion of a school choice program in India, finding null spillover effects. However, the majority of private schools in that context are low-cost and cater to non-affluent sections of the population, which is not the case in Chile or the US. Altonji et al. (2015) present similar non-experimental evidence in the US focusing on the cream-skimming effects of school choice, but they do not separate demand and supply side mechanisms behind the aggregate effects. Our results then contribute by isolating the effects of the screening channel and cream skimming on students' stratification and performance.

Our findings also contribute to the extensive literature on peer effects in education. In general, the self-selection of students and their parents into schools makes it difficult to disentangle the effects of peers from self-selection into schools. Three ways have been used in the literature to measure and identify peer effects models, experiments (Sacerdote, 2011, Zimmerman, 2003, Carrell et al., 2009, Duflo et al., 2011, Carrell et al., 2013, Feld and Zölitz, 2017, Garlick, 2018), quasi-experiments (Gould et al., 2009, Imberman et al., 2012, Jackson, 2013, Abdulkadiroğlu et al., 2014, Figlio and Özek, 2019), and social networks (Bramoullé et al., 2009, Calvó-Armengol et al., 2009). Although experimental peer effects studies in education have a clear identification strategy, most evidence focuses exclusively on post-secondary education in the US and often leverages the random assignment of roommates. This is problematic because peer effects seem to vary considerably depending on the social context (Sacerdote, 2014), presenting a threat to the external validity of these results. Our

[3] Examples of this are the systems in Brazil and Chile, where precise cutoffs based on national admission exams determine higher education assignments.

estimates provide valuable measures in the context of a middle-income education system, exploiting a much larger sample and a richer set of outcomes than most previous studies.

The remainder of the paper is structured as follows. Section 2 introduces the setting and provides details about the implementation of the SAS. Section 3 presents our data sources. Section 4 measures the impact of the implementation of the SAS and explains the empirical approach for these results. Section 5 presents our estimation method and results for the direct effects of school enrollment, and 6 shows our estimates of the spillover effects of classroom composition. Section 7 concludes.

2 Background: The Chilean System Before The SAS

There are three types of schools in Chile: public schools, private voucher schools, and non-subsidized private schools.[4] The first two types are publicly subsidized and represent over 90 percent of schools in Chile. The voucher system consists of monthly payments per student enrolled that vary depending on students' socioeconomic background and school attendance. Before 2015, the state fully funded all public schools, while 37.1 percent of private voucher schools had copayment systems. The copayment system allows subsidized private schools to charge fees to students on top of the public voucher. In 2019, just 18.1 percent of voucher schools charged a copayment in response to reforms to the educational system.

An essential feature of the Chilean school system is that schools must compete to attract families. Given that families cannot easily distinguish the quality of the school from the skills of the students in it, schools have powerful incentives to select students by their academic performance and economic status. Unlike educational districts in the US, families do not face restrictions in choosing a school depending on their neighborhood or residence, enhancing competition across neighborhoods. However, distance is a relevant factor in families' decisions (e.g., Gallego and Hernando, 2009). On top of this, schools may also find it easier to educate students from more advantaged backgrounds who can pay higher tuition fees, placing additional incentives to select them. This difficulty was recognized by the Chilean authorities, leading them to implement a program with larger vouchers for students from lower socioeconomic status (SEP, by its acronym in Spanish). Despite this, the preferential subsidy only differentiates students into three broad income groups, leaving plenty of space for sorting within such groups. Moreover, the extra voucher payments do not necessarily offset the additional cost of educating these students, as reflected by schools' reticence to

[4] There also exist schools with delegated administration that are considered a separate category, representing 1.3 percent of enrollment in 2015. We merge this group into public schools for the analysis.

enroll these students.

Before the school system reform, school admission policies were highly unregulated for private voucher schools and unsubsidized private schools. In comparison, most public schools were not allowed to select students based on their characteristics. In turn, private schools in Chile have been practicing school-side selection since the system's original implementation. This sorting arises directly from copayments made by families, restricting access to students based on socioeconomic status and parents' valuation of education. However, tuition fees explain only part of the observed segregation. Other screening mechanisms, such as academic selection, psychological assessments, religious considerations, and family background checks, play a role even within schools with comparable prices and affect students' sorting more obscurely.

The Chilean school system has high levels of segregation by socioeconomic status. Using the Duncan Dissimilarity Index, Valenzuela et al. (2014) estimate that in order to have a homogeneous distribution of students in the lowest 30th percentile across schools, it would be necessary to transfer between 54 and 60 percent of these low-income students from schools with high to low concentrations of disadvantaged students. They also find that school SES segregation was comparatively more elevated than Chilean residential segregation. This fact indicates that segregation cannot be explained exclusively by location factors, where factors like cream skimming can play a role in exacerbating the differences.

Schools' side selection has been controversial in Chile for several years. Following significant reforms in 2009, Chile prohibited selection based on academic or socioeconomic factors for children up to 6th grade in all schools receiving public funding (i.e., all except fully private schools). However, selection remained admissible at 6th grade or below when it was allegedly based on other factors, such as religion or adherence to the institution's values. In turn, this vague definition opens a window for blurry screening mechanisms that can maintain schools' screening based on other hidden factors, including socioeconomic level.

Since its implementation in 2009, the Preferential Subsidy Law (SEP, for its acronym in Spanish) also forbids schools voluntarily adhering to a special subsidy for low-income students to select students based on academic or socioeconomic reasons. About half of students attending voucher schools in 2015 attended schools that opted into the program. This program places a double prohibition against screening in these schools up to 6th grade. In practice, however, the tolerance of screening by alternative motives and the difficulty monitoring have called the effectiveness of the verification mechanisms' into doubt (e.g., Carrasco et al., 2014). Despite this, we focus on 7th grade and above in our estimations to avoid comparing groups with highly imperfect compliance before the SAS implementation.

In 2015, the Chilean Ministry of Education promulgated the School Inclusion Law with a broad objective of equal access to education. This regulation changed the admission process for all publicly subsidized schools, representing over 90 percent of enrollment, implementing a series of changes in the education system. The major reforms were the gradual termination of schools' for-profit allowance, the gradual elimination of parents' copayments in voucher schools, and prohibiting selection based on social, religious, economic, or academic criteria through the implementation of the SAS. However, the deployment of the other programs followed different patterns than the SAS, allowing for more gradual adjustment periods for schools.

The School Admission System (SAS) is a nationwide system that adapts the Deferred Acceptance (DA) algorithm to Chilean law requirements. In particular, the SAS guarantees their current seats to students applying to switch schools and favors the assignment of siblings and children of parents who work in the same school. It also introduced reserved quotas by socioeconomic level, prioritizing 15 percent of vacancies for students from disadvantaged backgrounds.[5] For these adaptations, the SAS defines priority groups and runs independent lotteries on each group to break ties randomly.[6]

The Chilean Ministry of Education started implementing the SAS in 2016 with a staggering design across regions and grades. It started in the least populated region in the southern part of Chile, and in its first year only covered pre-k, kinder, first, seventh, and ninth grades. The school assignment mechanism expanded sequentially, adding four regions in 2017 and the remaining ten regions in 2018, as depicted in Figure 1. In 2019, the SAS was entirely in place for the whole country from Pre-Kinder to 12th grade.

The introduction of the SAS switched school applications from a completely decentralized system to a unique application platform where all students must submit their rank-ordered list, including as many entrances as desired. On the other hand, schools must declare their slots' availability to the Ministry of Education, and they are available on the platform at the time of application. The system then runs the DA algorithm to match schools and students, offering admission offers to students, placing them on a waitlist, or assigning them to the closest school with available spots. Once these vacancies are assigned, families can accept their allocation or participate in a second round. If families accept their admission offer, the offers become binding, and schools must enroll all those students. If families do not accept their offer, they must participate in the second stage or apply to unsubsidized

[5] The system also considers spots for students with special education needs and high achieving students. However, these only apply to a limited subset of schools.

[6] For more detailed information on the algorithm and the computational perspective, see (Correa et al., 2019) on the mechanism design of the school assignment algorithm.

schools outside the system. The process is then repeated, but families must compete for the vacancies that remain unassigned after the first round. Across the first two years of the SAS implementation, 91.1 percent of students got assigned to a school in their first-round applications. Overall, 69.3 percent of students enroll in the school assigned during the first stage. Hence, we focus on the assignments from the first application stage for the remainder of our analysis.

3 Data

Enrollment data. We use administrative panel data from the Ministry of Education, including student-level enrollment information from 2016 to 2022. The data includes records of all students in the country regardless of the type of institution. This source also contains GPA and school attendance data, SEP eligibility (targeted voucher for low SES students), and basic demographic information.

The GPA in the Chilean school system takes values between 1 and 7. Given that the relative position of a student's classroom achievement changes when the new students arrive, we use the raw and standardized GPA score in our estimates. Attendance is measured from 1 to 100 and is the percentage of school days a student attended during the academic year. We also complement this using national standardized test scores from the SIMCE exams from 4th, 6th, and 8th grades performed between 2015 and 2017. The exams include math and language sections and a third subject that varies across years and grades. We only use the former two exams since they are available for all grade levels in the sample. SIMCE exam takers also respond to a survey covering their household composition and socioeconomic status. Additionally, we also employ college enrollment data and college admission exam scores, which we connect to secondary students using anonymous identifiers.

Applicants' data. In addition to the data mentioned above, we also use data from the Ministry of Education containing individual-level applications' data. This data comprises the complete list of schools each student applies to, the order in which they list their preferences, and students' classification as vulnerable and high-performance. This source also includes detailed information about vacancies in each grade and school for each applicant type. Due to data availability, we focus on applications occurring in December of 2017, 2018, and 2019, covering students who start the following school and calendar year in March at their new institutions (2018, 2019, and 2020 academic years). Among them, we exclude students in sixth grade or below in our estimations since we focus on years where selection was previously allowed. Across all levels, there were 76,821 applicants in 2017, 274,990 applicants in 2018, 483,070 applicants in 2019. Among them, 30,317 applicants were above sixth grade in 2017, 107,165 in 2018, and 175,497 in 2019.

School's supply. Schools participating in the school admission system must inform the Ministry of Education of all the slots and vacancies available at each grade level. The sample's average slots per grade level are 56.5 (divided in some cases into several classrooms), and the average number of vacancies is 24.4. However, the median grade-level size and vacancies are smaller, at 40 and 12, respectively, indicating that most schools offer a single classroom per level. The number of vacancies also displays high variation across levels, mainly driven by the larger quantity offered in 9th grade by high schools. Overall, 68.8 percent of courses had at least half of their slots already occupied by current students. Moreover, 22.2 percent of the classrooms filled all their original vacancies (some new vacancies can open if current students switch to a new school).

Schools' Screening We use parents' survey data to identify schools using selectivity policies before the government implemented the centralized admission system. While schools do not directly report these policies on their own, parents of students taking the SIMCE standardized exams are surveyed about the requirements they had to fulfill when applying to their respective schools. We then focus on parents of new students entering the most recently available year before the implementation of the SAS to account for the most recent admission policies within these schools. The most common types of requirements were grade certificates and parent interviews. However, since most schools require these, it becomes less informative about their selectivity.

On the other hand, psychological assessments and evaluated games are most common among voucher schools. These assessments are associated with schools that filter their students more thoroughly. One caveat is that parents' ability to recall the application process can limit the reliability of this data. Carrasco et al. (2014) survey school principals and contrast the results with those reported in the SIMCE questionnaires, finding highly consistent results in both sources. They also find that schools differ in their preferred screening mechanisms, with schools taking students from higher SES backgrounds showing more selective policies.

Algorithm data. The Chilean implementation of the Deferred Acceptance (DA) algorithm allows no preferences from the schools' side, forcing them to select students based on a random number whenever there is oversubscription. As a result, these lotteries introduce randomness in the admitted applicants' set in oversubscribed schools. Specifically, 52.9 percent of students have at least some variation in their assigned schools when changing the corresponding seed in the random process.[7] Figure 3 shows the high degree of variation in the schools to which students are assigned. For example, this figure shows that around 2.5 percent of students are assigned to one of their choices just 40 percent of the time, meaning that the remaining 60 percent get assigned to a combination of different schools.

4 Effects of Screening On Students' Allocation

4.1 Empirical Approach: Exploiting Staggered Implementation

The new admission system induced a significant shift in students' enrollment patterns. Specifically, it allowed numerous students to enroll at schools that would have rejected them before the reform. Consequently, we start by measuring the effects of prohibiting schools' side selection on enrollment patterns. Specifically, we contrast enrollment changes generated in schools forced to halt their selection practices with those schools that did not implement exclusionary practices even before introducing the new system. In particular, the SAS was initially introduced only in a subset of regions and sequentially for specific grades within these regions, as illustrated in Figure 1. We then estimate the following model:

$$y_{i,j,t} = \alpha + \beta Selective_j \times SAS_{g,r,t} + \gamma Selective_j + \delta SAS_{g,r,t} + \theta_j + \eta_t + \phi_g + \varepsilon_{i,j,t} \quad (1)$$

Where $y_{i,j,t}$ represent the outcomes of incoming student i in school j at time t, $Selective_j$ indicates whether the school j implemented selection practices before the SAS, $SAS_{g,r,t}$ indicates whether the SAS was functioning on grade g in region r at time t, and θ_j, η_t, and ϕ_g are school, year, and grade fixed effects. This model thus captures the differential changes on incoming students induced by the SAS in selective and non-selective grades and schools, measured by the coefficient β. By comparing different grades within the same school we aim to isolate the effect of concurrent reforms, which did not follow a staggered implementation as the SAS.

We interpret the estimated effect as measuring the effect of supply-side responses on school enrollment. While selective schools were the most affected by the introduction of the SAS, one potential caveat to this interpretation is that the centralization of applications can also provoke changes in the demand side, leading families to submit different applications (e.g., Idoux, 2022). However, that would only represent a problem when estimating the

[7] Given that the algorithm works in a staggered fashion, it is necessary to replicate the entire allocation under a different seed to compare for randomness. This is because availability may cascade depending on whether students get assigned to their higher-ranked options.

supply side responses if the SAS implementation affected applications systematically different at those grades within a school that were exposed to the SAS at different years due to the staggered implementation. Moreover, outside options in the Chilean system are highly limited to high-income families who can pay for unsubsidized private schools since the system implementation occurred in entire regions. These schools outside the system are considerably more expensive, employ stricter screening processes, and only enroll less than 8 percent of students.

4.2 Estimation: Students' Sorting Across Schools

As detailed in Section 4.1, our specification exploits the staggered implementation of the SAS across grades and regions to measure the impact of school screening practices on students' allocations across schools. Specifically, we analyze the differential impact of the introduction of the SAS on schools and grades that previously performed screening and those that did not to estimate the effects of school selection, after accounting for school, grade, and year fixed effects.

We begin in Table 1 by showing the changes in the background characteristics of students enrolled in Voucher schools. Specifically, Panel A in this table start by showing that the introduction of the SAS significantly decreased the background achievement and income level of students enrolled in voucher schools compared to other school types. The baseline standardized scores and GPA of new students enrolled in voucher schools through the SAS decreased by 0.12 and 0.08 standard deviations. In comparison, the gap between the average standardized scores of public schools and voucher schools able to select students before the SAS was around 0.6 standard deviations, thus reducing this difference by 20 percent. Conversely, the proportion of low socioeconomic status students in voucher schools significantly increases by 2.1 percent, representing an 8 percent increase compared to their pre-SAS levels, while their income per capita is 9 percent lower. Overall, this confirms that the introduction of the SAS forced voucher schools to accept higher rates of disadvantaged students.

Numerous practices associated with selective schools can act as a proxy for selectivity. Nonetheless, there is arguably no single variable that perfectly identifies schools that implemented exclusionary policies before preferences were collected, prior to the enactment of the SAS. Instead, we rely on a series of selective practices to construct a screening index that measures the number of screening practices that families had to undergo to enter their current schools. We construct this using parents' responses to a nationwide survey implemented alongside standardized exams. Figure 2 illustrates this index, reflecting the high asymmetry in screening practices by school dependence, tuition fees, and educational level. In particular, private and highly-priced schools are the most selective, while secondary schools are more selective than primary schools. This pattern directly reflects the prohibition for public schools and a share of voucher schools to screen their students before 7th grade. Henceforth, we focus on 7th grade and above in our analysis.

When interpreting these results, it is crucial to consider that the changes to equilibrium allocations induced by SAS are limited by the variability in the characteristics of the applicants to a given school and their degree of oversubscription. This variability depends on factors such as family preferences, residential segregation, tuition fees, and local schools' supply. This intuition is confirmed when focusing on high-performing and high-demand schools in Panels B and C from Table 1. High-performing schools correspond to those identified by the Ministry of Education as highly effective given their socioeconomic composition. High-demand schools correspond to those receiving more applicants than their available slots at any grade during the first three years of implementation of the SAS. This is a conservative definition, as some of these schools are only slightly oversubscribed and are hence randomize a small subset of their slots. Moreover, while most high-performing schools are oversubscribed, many oversubscribed schools are not high-performing.

Our results indicate that high-performing and high-demanded schools follow the same patterns as voucher schools, although of considerably larger magnitudes. Specifically, these schools experienced a substantial decrease in their admitted students' background achievement, reflected by a decrease in the average baseline standardized scores of 0.22 and 0.12 standard deviations when the SAS was introduced. Similarly, their lagged GPA decreased by 0.19 and 0.11 standard deviations and their students attended to class significantly less before they switched into their new schools. Finally, columns (5) through (7) indicate that the proportion of low-SES students in high-performing schools increased by 410 basis points, representing a 20.6 percent increase compared to the initial proportion of low-SES students at these schools. This change is equivalent to a 9.1 percent decrease in the average income of the students getting access to spots at publicly subsidized high-performing schools. This pattern is repeated among highly-demanded schools, where the proportion of admitted low-SES increased by 2.6 percentage points, representing a 10.4 percent increase compared to the baseline levels of low-SES enrollment.

Table 2 inquires further into the effects of the adoption of the SAS and the prohibition of school screening on the distribution of students across schools by splitting results by tuition charges. Consistently with the results above, Panel A shows that high-priced schools are the most affected by the implementation of the SAS, decreasing the average baseline scores of their newly admitted students by 0.15 standard deviations. Mid-priced schools also experienced a decrease in the baseline SIMCE scores of their admitted students, although the size was moderate, at 0.05 standard deviations. In contrast, schools that do not charge tuition, which do not normally implement high-screening practices, did no see significant differences in their enrolled students and potentially even experienced an increase in their students' average income.

We complement this analysis by comparing changes in enrollment of grades that forcefully adopted admission lotteries against those that did not within a given school and year. These results are displayed in Appendix A.1. While this comparison has the appeal of controlling by school-specific characteristics that vary over time, it is also more susceptible to within-school spillovers. For instance, the mandatory adoption of the SAS at specific grades may affect admission policies at other grades within the same school. Nevertheless, the results are robust to this alternative specification, confirming the significant changes in enrollment patterns at formerly selective schools introduced by the SAS.

5 Direct Effect: Heterogeneous Schools' Value-Added

5.1 Empirical Approach: Exploiting Admission Lotteries

Following Angrist et al. (2017) and Abdulkadiroğlu et al. (2020), we model the potential outcomes of student i in school j are defined as:

$$\begin{aligned} Y_{ij} &= \alpha_j + X_i'\beta_j + f(\mathbf{X}_j^{-i}) + \epsilon_{ij} && (2)\\ &= \bar{\alpha}_j + X_i'\bar{\beta}_j + \bar{\epsilon}_i + (\alpha_j - \bar{\alpha}_j) + X_i'(\beta_j - \bar{\beta}_j) + f(\mathbf{X}_j^{-i}) + (\epsilon_{ij} - \bar{\epsilon}_i) && (3)\\ &= A_i + (\alpha_j - \bar{\alpha}_j) + X_i'(\beta_j - \bar{\beta}_j) + f(\mathbf{X}_j^{-i}) + (\epsilon_{ij} - \bar{\epsilon}_i) && (4)\\ &= A_i + ATE_j + f(\mathbf{X}_j^{-i}) + M_{ij} && (5) \end{aligned}$$

Where $A_i = \frac{1}{J}\sum_j \alpha_j + X_i'\beta_j + \epsilon_{ij}$. This framework follows Abdulkadiroğlu et al. (2020), expanding on it by allowing peers to affect individual outcomes: $f(.)$ represents a general form for the peer effects of classmates of student i in school j. In a linear-in-means model, we would have that $f(\mathbf{X}_j^{-i}) = \kappa \bar{X}_j^{-i}$.

This decomposition allows us to express the average outcome at school j as

$$E[Y_i|S_i = j, \theta_i, \theta_{-i}] = Q_j + ATE_j + E[f(\mathbf{X}_j^{-i})|S_i = j] + E[M_{ij}|S_i = j, \theta_i, \theta_{-i}] \quad (6)$$

Where $Q_j = E[A_i|S_i = j]$ is the average ability of students enrolled at school j, $E[M_{ij}|S_i = j, \theta_i, \theta_{-i}]$ represents the average suitability of students in school j given their type $\theta_i = \theta$, $E[f(\mathbf{X}_j^{-i})|S_i = j]$ represents the expected peer effect on student i in school j.

Returning to the model above, we model expected individual outcomes as:

$$E[Y_{ij}|X_i, \mathbf{X}_j^{-i}, S_i] = \alpha_j + X_i^{'}\beta_j + E[f(\mathbf{X}_j^{-i})|X_i, \mathbf{X}_j^{-i}, S_i] + E[\epsilon_{ij}|X_i, \mathbf{X}_j^{-i}, S_i], \quad j = 1, ..., J \tag{7}$$

The direct estimation of this would model give biased estimates because $E[\epsilon_{ij}|X_i, \mathbf{X}_j^{-i}, S_i] \neq 0$. This occurs because students self-select into schools, and these schools further select among their applicants, potentially based on the match quality measured by ϵ_{ij}. Allowing the peer effects to vary linearly depending on peers' observable characteristics, we can model the peer effects function f as:

$$E[f(\mathbf{X}_j^{-i}|X_i, \mathbf{X}_j^{-i}, S_i] = \sum_{l \in j \setminus \{i\}} X_l^{'}\gamma_{il} \tag{8}$$

So the potential outcomes equation becomes

$$E[Y_{ij}|X_i, \mathbf{X}_j^{-i}, S_i] = \alpha_j + X_i^{'}\beta_j + \sum_{l \in j \setminus \{i\}} X_l^{'}\gamma_{il} + E[\epsilon_{ij}|X_i, \mathbf{X}_j^{-i}, S_i], \quad j = 1, ..., J \tag{9}$$

Following Angrist et al. (2017), we exploit the variation induced by the lotteries to obtain exogenous shifts on school assignments that are uncorrelated with potential outcomes once we account for students' preferences over schools, yielding an unbiased measure of value-added. Furthermore, we follow the method by Abdulkadiroğlu et al. (2017) to fully exploit the variation induced by lotteries in oversubscribed schools by using the entire distribution of admission offers instead of focusing on first-ranked offers. This allows us to exploit the entire assignments distribution rather than first-ranked option comparisons. In practice, the allocation probabilities have no closed-form solutions in the DA algorithm, so they must be approximated. We make this approximation by computing the assignment probability to a given school for all applicants over several runs of the algorithm with counterfactual lottery assignments. Once we calculate this, we then condition on propensity score to obtain conditionally exogenous variation on school admission, producing efficiency gains over alternative methods of exploiting lottery variations.

As in Abdulkadiroğlu et al. (2017), the individual-level stratified randomization introduced by equal treatment of equals (ETE) in the DA algorithm implies that:

$$P(D_i(S) = 1|X_i = x_i, \theta_i = \theta) = P(D_i(S) = 1|\theta_i = \theta) = p(\theta) \tag{10}$$

Where $D_i(S) = 1$ is the probability that student i will be offered a spot in school j. This indicates that allocation probabilities are independent of students' characteristics X_i once we condition on students' preferences θ. In other words, lottery offers are conditionally independent of student types. Essentially, the variation exploited by this method is parallel to that exploited by propensity score matching. However, the advantage of this method is that the equal treatment of equals (ETE) in the centralized randomization lotteries guarantees the validity of the conditional independence assumption.

The ETE property implies that admission offers are a valid instrument for school enrollment after controlling for lottery assignment strata, as in Angrist et al. (2017). Given that only oversubscribed schools implement lotteries, we can only use this method to compute the value-added measurements of a subset of schools. This implies that the external validity of the value-added effects in oversubscribed schools does not necessarily extend to undersubscribed schools. However, oversubscribed schools are the policy-relevant cases since they are required to understand counterfactual assignments given the observed students' preferences where slots are in dispute. The estimated model is the following:

$$Second \;\; Stage:$$

$$Y_{ij} = \alpha_j + AX_j + \epsilon_{ij} = \sum_j 1[S_i = j]1[j = Type](\alpha_j + AX_j) + \chi P_{ij} + \epsilon_{ij}$$

$$First \;\; Stage:$$

$$1[S_i = j] = \phi Admitted_{ij} \times 1[j = Type] + \chi P_{ij} + \eta_j + v_{ij}$$

In practice, we estimate this by splitting P_{ij} into bins because it is not continuous in our empirical setting, but results are mostly unchanged when controlling by P_{ij} continuously.

5.2 Estimation: Performance, College Enrollment, and Behavior

The quality and characteristics of schools are fundamental for students' future outcomes. Consequently, we continue our analysis in this section by measuring the impact of attending different types of schools. As explained in Section 5.1, we exploit the randomness in school admission offers to students with otherwise identical assignment probability that arise from the property of equal treatment of equals (ETE). We employ this to estimate and compare the effect of attending selective schools for different student types. Moreover, we explore the hypothesis of heterogeneous student-school value-added as a driver for screening by estimating and contrasting value-added at selective schools for benefited and displaced students.

We begin in Figure 4 focusing on the impacts of attending different school types on students' college enrollment and standardized test scores. We find that high-performing schools significantly increased Math scores in the national college admission exam for low-income students by 0.2 standard deviations. In contrast, the impact on Reading scores in this test is not significant on low or high-SES students, although the estimates are less precise. Our estimates of the effect on 8th-grade standardized scores show that all students enrolling in high-performing schools increased their test scores two years after school assignments. High-SES students significantly increased their math and reading scores by 0.29 and 0.31 standard deviations in Math and Reading. The increase in test scores is only significant in Reading for low-income students, reaching 0.26 standard deviations. These results confirm that these high-performing schools positively affect their students' performance on standardized tests. However, the differences in gains in college admission exams and test scores between high and low-SES students are not significant for any of these outcomes.

We further explore the effects of enrolling in high-performing schools in Table 4. Our estimates show that enrollment in high-performing schools had no significant impact on the rate at which students take the national college admission exam, their percentile on the exam, and whether they enrolled in college. Consistently with the value-added estimates of high-performing schools, Panel B of Table 3 shows similar effects on students enrolling in quota schools. These schools correspond to those with less than 15 percent of low-income students, for which the SAS mandated priority spots for low-income students. They are the prime candidates to screen out low-SES students: over 90 percent had at least two applicants per slot reserved for low-SES students, indicating that the previous scarcity of low-income students arises at least partly from school-side mechanisms. These schools do not seem as highly effective at raising their students' college admission outcomes or scores as those identified as high-performing, although part of the differences comes from more noisily estimated coefficients. On the other hand, high and low-income students increase their Reading scores when attending highly segregated quota schools by 0.13 and 0.23 standard deviations, respectively. These are large effects considering that these students attended their new schools for up to 2 years only. However, we still do not find differences in value-added gains between low and high-income students.

Finally, Panel C in Table 3 focuses on high-screening schools, defined according to parents' survey responses about the process they had to undergo to enroll at their respective schools.

Again, college admission outcomes follow similar patterns from previous panels with no significant effects. Perhaps strikingly, Math test scores decrease for high-income students enrolling at high-screening schools, and Reading test scores decrease for low-income students, but these estimates are only marginally significant.

We continue in Table 4 by evaluating the impact of school enrollment on students' school performance. These outcomes are available for the entire sample, allowing for a more comprehensive comparison. Our estimates show that low and high-SES students enrolling in high-performing schools decreased their GPA, GPA-Rank, and grade advancement rate. The raw GPA of high and low SES students enrolling in high-performing schools decreased by 0.16 and 0.22 points on the 1-7 scale used in Chile compared to other students from similar backgrounds who did not enroll at these schools. Given that different schools potentially have different grading standards, we complement this with a class-standardized GPA measurement, revealing a more considerable decrease of 0.41 and 0.58 standard deviations. Similarly, our estimates also show that students enrolling in high-performing schools decreased their grade advancement rates by 2.3 percent for high-income students and 5.8 percent for low-income students, significantly affecting low-income students more than their high-income counterparts. Despite the adverse effects on GPA, low-income students increase class attendance by 1.2 percent, or 0.12 standard deviations. These performance gaps are unsurprising because the more demanding environment in high-performing schools affects low-income students more intensively than high-income students, mainly driven by the lower academic standards of their alternative schools. However, they may explain the belief that low-income students underperform at these selective schools compared to their higher-income peers despite scarce evidence of this pattern in terms of more comparable measurements such as standardized tests and college admission exams.

Further inquiring into the effect of enrollment in selective schools, Panels B and C of Table 4 present an analogous comparison for segregated schools that forcefully reserved spots for low-income students and schools implementing high-screening practices. Our results show a similar negative effect on students' GPA when enrolling at these more demanding schools but null effects on grade advancement and school attendance.

To understand the consequences of school change on students' non-academic outcomes, we present in Table 5 the impact of school enrollment on students' motivation, self-confidence, school satisfaction, discrimination, and behavioral problems at school. Each of these indices comprises a set of 8th-grade students' survey responses. While we do not find any statistically significant impact on high-SES students, Panel A shows that low-income students decrease their reported motivation and behavior problems by around 0.25 standard deviations. This decrease is consistent with observed decreases in GPA and GPA rank within their classrooms, impacting students' motivation. On the other hand, decreases in behavioral problems at school are also concordant with increases in attendance by these low-SES students. When focusing on segregated schools forced to implement the affirmative action quota and high screening schools in Panels B and C, we do not observe any statistically significant impact on students' non-academic outcomes, except for a reduction in behavioral problems. These results suggest that part of the improvements in outcomes experienced by low-income students are likely to emerge from their better behavior due to changes in their school environment, even when the decrease in performance relative to their classmates may negatively affect their motivation.

These results generally confirm that enrollment in selective schools positively affects their students' performance on standardized tests and college admission exams. Although more limited in power, we find no evidence of higher performance by high-SES students, who traditionally enrolled at these schools, and low-SES students, who mostly gained access through the application of the SAS.

6 Spillover Effects: Changes in Classroom Composition

6.1 Empirical Approach: Lottery Induced Shocks to Classrooms' Composition

Returning to the model above, we can have that:

$$Y_{ij} = \alpha_j + \beta_j X_i + \gamma_{i,-i} X^j_{-i} + \epsilon_{ij} \tag{11}$$

Where $\gamma_{i,-i} = [\gamma_{i1}, \ldots, \gamma_{i,i-1}, 0, \gamma_{i,i+1}, ..., \gamma_{iN}]$ represents the usual modeling of linear effects of peers on individuals' outcomes (Blume et al. (2015)). When estimating this model, a problem arises because computing the effect of counterfactual allocations requires computing potential outcomes that depend on school effectiveness (α_j, β_j), exogenous peer effects (Γ), and self selection parameters (ϵ_j). Unfortunately, individual variation induced by school assignment does not allow us to identify separately the effect β_j and Γ. This is because changes to school assignment also modify the entire set of classmates X^j_{-i}. Instead, $\gamma_{i,-i}$ can be identified by exploiting variation in peers characteristics X^j_{-i} induced by the lotteries, while maintaining school assignment j unaltered.

Evidently, the set of applicants to any given school j is not randomly assigned. Instead, families apply to schools depending on their characteristics and their own preferences. Given that the DA algorithm has several rules giving special preferences to individual such as siblings or alumni, among others, no closed solution exist for the allocation probabilities. However, given the set of applications and capacity constraints, we can use the assignment algorithm under alternative random draws to obtain estimates of the empirical distribution of students across schools. This follow the same logic as Angrist et al. (2017) but expands on it but allowing to estimate the effect of peers' background on their classmates' performance.

Similarly to our value-added estimates, we use instrumental variables to estimate the impact of the characteristics of newly enrolled peers' on their classmates in the following second-stage model:

$$\begin{aligned} Y_{i,c,t} = \alpha + \beta \bar{X}_{-i,c,t} + \gamma X_i + \rho_1 \bar{X}^{Appl}_{sg,t} + \rho_2 SAS\_prop_{c,t} \\ + \rho_3 SAS\_prop_{c,t} \times \bar{X}^{Appl}_{c,t} + \theta_{r,g,t} + \varepsilon_{i,c,t} \end{aligned}$$

Where $Y_{i,c,t}$ is the outcome of student i in classroom c and year t, $\bar{X}_{-i,c,t}$ is the average background in of classmates of student i in class c and year t, $\bar{X}^{Appl}_{j,t}$ represents applicants' average background, school j and year t, $SAS\_prop_{c,t}$ is the fraction of new students assigned by SAS on classroom c and year t, X_i are lagged standardized test-score of student i, and θ_j and τ_t are school and year fixed effects

We instrument $\bar{X}_{-i,c,t}$ with the following first stage:

$$\begin{aligned} \bar{X}_{-i,c,t} = \phi_1 + \phi_2 \bar{X}^{SAS}_{c,t} + \phi_2 SAS\_prop_{c,t} + \phi_3 SAS\_prop_{c,t} \times \bar{x}^{SAS}_{c,t} \\ + \phi_4 \bar{X}^{Appl}_{sg,t} + \phi_5 \bar{X}^{Appl}_{sg,t} \times SAS\_prop_{c,t} + \phi_6 X_i + \pi_{rgt} + u_{i,c,t} \end{aligned}$$

Where $\bar{x}^{SAS}_{c,t}$ is the mean of previous standardized test-scores of student i's new classmates (randomly) assigned by SAS at classroom c in year t.

Acknowledging the non-linear nature of peer effects (Hoxby and Weingarth, 2005; Imberman et al., 2012), we also estimate a more flexible model. First, we classify incumbent students by their previous achievement terciles using the most recent pre-SAS standardized SIMCE test scores. Then we estimate the peer effects for each tercile using the following specification that allows for a heterogeneous effect by incumbent student i achievement level:

$$y_{i,c,t} = \sum_{k=1}^{3}(\beta_k \; tercile\_k_i \times \bar{x}_{-i,c,t}) + \gamma x_{i,t-1} + \rho_1 \bar{x}_{sg,t}^{Appl} + \rho_2 SAS\_proption_{c,t}$$
$$+ \rho_3 SAS\_proption_{c,t} \times \bar{x}_{sg,t}^{Appl} + \theta_{rgt} + \varepsilon_{i,c,t}$$

Where $tercile\_k_i$ is a dummy variable that takes the value of 1 if the student i standardized test score is on tercile k on the school-year standardized test score distribution.

Section A.3 in the appendix presents an analysis of the validity of this instrument.

6.2 Estimation: Exogenous Peer Effects

Arguably the most affected by policies modifying the school admission system are those students whose assignment changes with the introduction of the SAS. However, the allocation of students across schools also impacts their classmates through the presence of social interactions. This is particularly relevant in the case of large-scale shifts to admission mechanisms, such as the one introduced by the SAS. Moreover, group-level interaction provides an alternative explanation for the high degrees of selectivity in schools, particularly in light of the scarce value-added differentials between usually selected and rejected students. Our findings from Section 4.2 suggest that incoming students' characteristics changed classrooms' composition in several schools. We study how these changes affected students remaining in non-selective schools and in selective receiving schools by measuring the impact of changes in classroom composition on their classmates' college enrollment and other academic and behavioral outcomes. We answer this by exploiting the random allocation of students to schools in oversubscribed schools to estimate a linear-in-means model of peer background effects, also called exogenous peer effects, following the specification from Section 6.1. In our main estimation, we focus on changes in peers' standardized test scores, GPA-Rank within students' previous schools, and the proportion of low-income classmates.

We start by focusing on the impact of classmates' shifts on students' school performance and standardized test scores in Table 6. The results in Panel A indicate that an increase of one standard deviation in classmates' average standardized scores statistically significantly reduced their classmates' GPA rank by up to 0.7 standard deviations, a direct consequence of the admission of highly competitive peers. On the other hand, they also increased their classmates' grade advancement by 8 percent. We observe similar effects in Panel B when analyzing the impact of a one standard deviations increase in classmates' average GPA-

Rank in their previous schools: GPA-rank decreases by around 0.6 standard deviations in response to shifts toward students with better rankings positions in their previous schools; Grade advancement is positively affected, increasing by 0.6 percent in response to higher performing peers.

Despite confirming the relevance of high-achieving peers towards the academic achievement of their classmates, Table 6 also reveals that varying classmates' income levels notoriously affect their classmates' outcomes. In particular, Panel C reveals that an increase of 10 percent in average class income per capita induces an increase in GPA of 0.14 points, or 0.13 standard deviations of class-standardized GPA ranking. Similarly, classmates' attendance and grade advancement rates increase by 0.8 and 0.01 percentage points in response to the same average income shift. These results then confirm the high relevance of social interactions in the educative process, partly explaining schools' screening decisions.

Besides the impacts on academic performance reported in the previous table, we assess the effect of exogenous changes in classmates' backgrounds on self-declared behavioral outcomes of their peers in Table 7. We observe that an increase in their classmates' average standardized scores (SIMCE) by one standard deviation produces a decrease in their reported motivation by 0.08 index points and self-confidence by 0.098, equivalent to 0.12 and 0.13 standard deviations, respectively. Panel B shows that the impact of classmates' past GPA rank has similar impacts on students' behavioral outcomes, although the effects are minor and not as statistically robust. On the other hand, the incorporation of high-performing classmates generally improved classroom behavior, as measured by the number of disciplinary faults committed by students. Finally, Panel C presents strong evidence showing that increasing classmates' income per capita by 10 percent decreases their peers' motivation, self-confidence, and school satisfaction by 0.017, 0.016, and 0.017 index points, corresponding to 0.07, 0.13, and 0.17 standard deviations, respectively. Combined, this evidence suggests that, despite their positive effects on academic outcomes., the presence of more highly prepared peers can potentially undermine their classmates' non-academic inputs.

In terms of college enrollment, Table 9 shows that improving classmates' average standardized scores and GPA rank by one standard deviation produces an increase in college admission exams of their peers of a similar magnitude in both Math and Reading. This is equivalent to a 34 percent improvement in these students' college admission exam percentile. Moreover, point estimations of the effects of peers' GPA-rank report similarly strong responses, although the estimates are highly noisy and not significant. On the other hand, Panel C reports that enrollment of classmates with 10 percent higher household income per capita is associated with an increase in the college exam percentile of 4.7 percent points, corresponding to 0.13 and 0.09 percent in the Math and Reading exams, respectively.

In light of the mixed impacts caused by peers, we measure in Table 8 the effects of changes in peers' backgrounds on school switching patterns. Specifically, we ask whether students shift their enrollment patterns in response to classmates' scores, GPA, and household income, as in the previous analysis. These results indicate that increasing classmates' average standardized scores by one standard deviation decreases school switching by 10 percent, while increasing peers' GPA rank by one standard deviation decreases school switching by 14.7 percent. Moreover, this effect arises from changes in schools switching to public and voucher schools. Finally, Panel C presents similar patterns when analyzing responses to changes in peers' per capita income: increasing average income by 10 percent reduced school switching by 7.3 percent, particularly affecting those switching to voucher schools. These results indicate that families display preferences for high-performing and high-income peers despite their adverse short-term effects on measures such as GPA rank or students' motivation.

To understand the heterogeneous effects of changes in peers, Figure 5 decomposes the effects by estimating the model differentiating by students' tercile in standardized test scores. These estimations show a relatively flat effect of peers' standardized scores on college admission exam scores, higher education enrollment, and GPA. However, students from the higher-performing tercile seem to be more affected by their peers' standardized test scores.

Finally, we perform a similar exercise comparing the effects of changes in the proportion of classmates from each tercile on students' outcomes. These results are displayed in Figure 6, reflecting that most of the changes are driven by changes in the first and third terciles of the distribution, while mid-performing students appear to be less impactful on their classmates. Interestingly, negative effects on college enrollment, college admission exams, standardized test scores, and school switching appear to be driven mostly by students in the lowest part of the distribution. In contrast, negative effects on peers' motivation are driven by students from the top tercile, who possibly outperform their peers' achievement.

7 Conclusion

Segregation in schools extensively impacts students' academic and labor outcomes later in life, particularly among minorities. There are several reasons why this segregation may arise in educational settings. On the demand side, families apply to different schools due to their willingness to pay tuition fees, distance to schools due to residential segregation, and preferences for educational quality. On the other hand, market incentives may lead schools to implement screening practices. Specifically, these practices allow selective schools to capture students from more advantaged backgrounds who are less costly to educate, also helping them to attract other high-performing peers. Although some literature has addressed the general existence of cream skimming (e.g., Altonji et al., 2015), our empirical setting has the advantage of isolating the effects of supply-side cream skimming, presenting novel evidence of this mechanism in a competitive market.

Disentangling these demand and supply factors behind cream skimming is crucial because it leads to different policy recommendations. In particular, school districts often attempt to alter segregation patterns through policies such as reserved spots for minority students, diminishing costs of attending selective schools through busing or scholarships, and even admission lotteries, as in Chile. However, the efficacy of these policies is limited depending on whether segregation arises from demand (families and students) or supply factors (schools), stressing the importance of distinguishing between these sources.

Our results imply that supply-side cream skimming contributes to the country's large socioeconomic status (SES) segregation, expanding the achievement gap between vulnerable and wealthier students. The induced segregation is problematic because research has shown that students attending more segregated schools have lower graduation, achievement, and college attendance rates (Billings et al., 2014, Johnson, 2011). Furthermore, extensive literature documents the benefits of less segregated schools for minority, low-income, and low-achieving students (e.g., Hoxby, 2000, Hanushek et al., 2009). Our results support these patterns, showing that voucher and private schools' ability to engage in cream skimming is one of the sources of this increase in segregation and affects their classmates through classroom composition.

The efficiency consequences of redistributing slots will depend on factors such as the complementarity between students' preparation and schools' value-added (Durlauf, 2008) and private information about student-school match quality (Arcidiacono and Lovenheim, 2016). However, evidence is scarce in the context of secondary education, partly due to the difficulty of untwisting the equilibrium effects of such policies. Our results present novel evidence of the effects of redistributing students across schools through a centrally designed public program, separately estimating the value-added benefits for different student types and the indirect impact of this redistribution on students through spillover effects. In particular, our estimations confirm the benefits of attending selective secondary schools in test scores and college enrollment for low- and high-income students in secondary education. However, selective practices appear to transfer slots from lower to higher-income students without increasing overall educational achievement, based on the small differences in value-added between students.

In our context, the lack of value-added differences between higher and lower socioeconomic students admitted into selective schools suggests that the fit of students and schools does not drive school screening practices on average. Instead, the presence of social interactions supports the idea of families' preferences and group-level dynamics as a motive behind supply-side cream skimming. While it may become difficult to justify denying schooling options to high-achieving students simply because their departure from public schools affects those left behind (Ladd, 2002), it may be similarly challenging to justify rejecting high-performing students based on their low-income backgrounds simply because their characteristics are not as beneficial or attractive to their peers.

References

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

Abdulkadiroğlu, A., Angrist, J. D., Narita, Y., and Pathak, P. A. (2017). Research design meets market design: Using centralized assignment for impact evaluation. *Econometrica*, 85(5):1373–1432.

Abdulkadiroğlu, A., Pathak, P. A., Schellenberg, J., and Walters, C. R. (2020). Do parents value school effectiveness? *American Economic Review*, 110(5):1502–39.

Allende, C. (2019). Competition under social interactions and the design of education policies. *Job Market Paper*.

Altonji, J. G., Huang, C.-I., and Taber, C. R. (2015). Estimating the cream skimming effect of school choice. *Journal of Political Economy*, 123(2):266–324.

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

Arcidiacono, P. and Lovenheim, M. (2016). Affirmative action and the quality-fit trade-off. *Journal of Economic Literature*, 54(1):3–51.

Billings, S. B., Deming, D. J., and Rockoff, J. (2014). School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg*. *The Quarterly Journal of Economics*, 129(1):435–476.

Black, S. E., Denning, J. T., and Rothstein, J. (2020). Winners and losers? the effect of gaining and losing access to selective colleges on education and labor market outcomes. Technical report, National Bureau of Economic Research.

Blume, L. E., Brock, W. A., Durlauf, S. N., and Jayaraman, R. (2015). Linear social interactions models. *Journal of Political Economy*, 123(2):444–496.

Bramoullé, Y., Djebbari, H., and Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, 150(1):41–55.

Calvó-Armengol, A., Patacchini, E., and Zenou, Y. (2009). Peer Effects and Social Networks in Education. *Review of Economic Studies*, 76(4):1239–1267.

Carrasco, A., Bogolasky, F., Flores, C., Gutiérrez, G., and San Martín, E. (2014). Selección de estudiantes y desigualdad educacional en chile:¿ qué tan coactiva es la regulación que la prohíbe. *Informe resumen proyecto FONIDE*, 711286.

Carrell, S. E., Fullerton, R. L., and West, J. E. (2009). Does your cohort matter? measuring peer effects in college achievement. *Journal of Labor Economics*, 27(3):439–464.

Carrell, S. E., Sacerdote, B. I., and West, J. E. (2013). From natural variation to optimal policy? the importance of endogenous peer group formation. *Econometrica*, 81(3):855–882.

Correa, J., Epstein, R., Escobar, J., Rios, I., Bahamondes, B., Bonet, C., Epstein, N., Aramayo, N., Castillo, M., Cristi, A., et al. (2019). School choice in chile. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 325–343.

Duflo, E., Dupas, P., and Kremer, M. (2011). Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *American Economic Review*, 101(5):1739–1774.

Durlauf, S. N. (2008). Affirmative action, meritocracy, and efficiency. *Politics, Philosophy & Economics*, 7(2):131–158.

Epple, D. and Romano, R. (2008). Educational vouchers and cream skimming. *International Economic Review*, 49(4):1395–1435.

Epple, D. and Romano, R. E. (1998). Competition between private and public schools, vouchers, and peer-group effects. *American Economic Review*, pages 33–62.

Feld, J. and Zölitz, U. (2017). Understanding peer effects: On the nature, estimation, and channels of peer effects. *Journal of Labor Economics*, 35(2):387–428.

Figlio, D. and Özek, U. (2019). Unwelcome guests? the effects of refugees on the educational outcomes of incumbent students. *Journal of Labor Economics*, 37(4):1061–1096.

Gallego, F. A. and Hernando, A. (2009). School Choice in Chile: Looking at the Demand Side. *SSRN Electronic Journal*.

Garlick, R. (2018). Academic peer effects with different group assignment policies: Residential tracking versus random assignment. *American Economic Journal: Applied Economics*, 10(3):345–69.

Gould, E. D., Lavy, V., and Daniele Paserman, M. (2009). Does immigration affect the long-term educational outcomes of natives? quasi-experimental evidence. *The Economic Journal*, 119(540):1243–1269.

Hanushek, E. A., Kain, J. F., and Rivkin, S. G. (2009). New evidence about brown v. board of education: The complex effects of school racial composition on achievement. *Journal of labor economics*, 27(3):349–383.

Hoxby, C. M. (2000). Peer effects in the classroom: Learning from gender and race variation.

Hoxby, C. M. and Weingarth, G. (2005). *Taking race out of the equation: School reassignment and the structure of peer effects*.

Hsieh, C.-T. and Urquiola, M. (2006). The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program. *Journal of Public Economics*, 90(8-9):1477–1503.

Idoux, C. (2022). *Integrating New York City schools: The role of admission criteria and* *family preferences*. PhD thesis, Massachusetts Institute of Technology.

Imberman, S. A., Kugler, A. D., and Sacerdote, B. I. (2012). Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees. *American Economic Review*, 102(5):2048–2082.

Jackson, C. K. (2013). Can higher-achieving peers explain the benefits to attending selective schools? evidence from trinidad and tobago. *Journal of Public Economics*, 108:63–77.

Johnson, R. C. (2011). Long-run Impacts of School Desegregation & School Quality on Adult Attainments. Technical Report w16664, National Bureau of Economic Research.

Kutscher, M., Nath, S., and Urzua, S. (2020). Centralized admission systems and school segregation: Evidence from a national reform.

Ladd, H. F. (2002). School vouchers: A critical view. *Journal of economic perspectives*, 16(4):3–24.

Muralidharan, K. and Sundararaman, V. (2015). The aggregate effect of school choice: Evidence from a two-stage experiment in india. *The Quarterly Journal of Economics*, 130(3):1011–1066.

Nechyba, T. J. (2006). Income and peer quality sorting in public and private schools. *Handbook of the Economics of Education*, 2:1327–1368.

Otero, S., Barahona, N., and Dobbin, C. (2021). Affirmative action in centralized college admission systems: Evidence from brazil. Technical report, Working paper.

Sacerdote, B. (2011). Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far? In *Handbook of the Economics of Education*, volume 3, pages 249–277. Elsevier.

Sacerdote, B. (2014). Experimental and quasi-experimental analysis of peer effects: two steps forward? *Annu. Rev. Econ.*, 6(1):253–272.

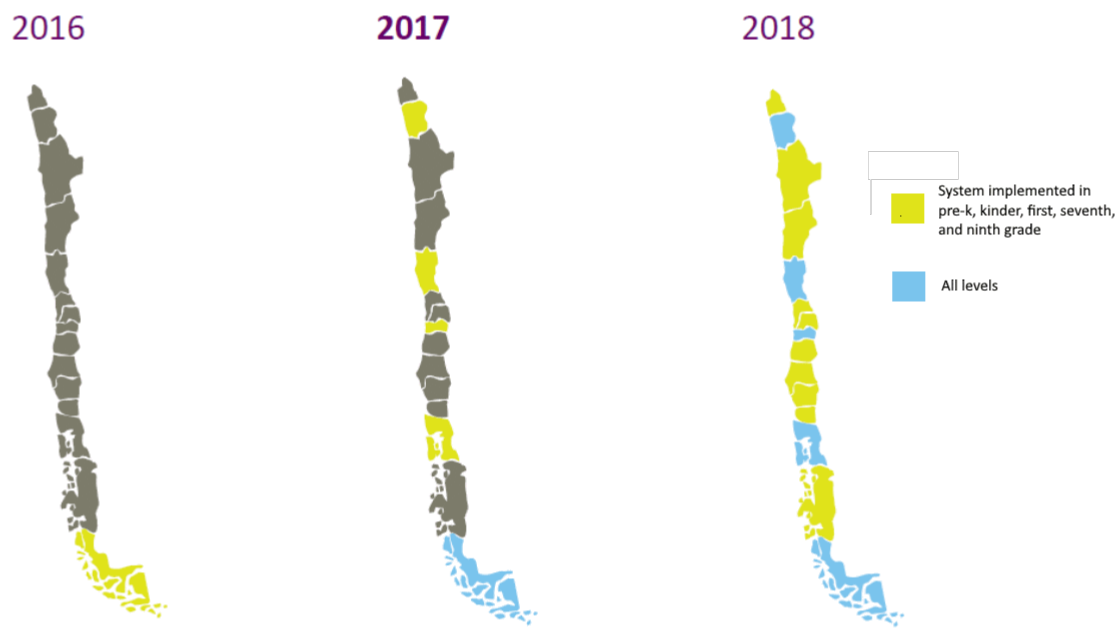
Sander, R. H. (2004). A systemic analysis of affirmative action in american law schools. *Stan. L. Rev.*, 57:367.

Valenzuela, J., Bellei, C., and Ríos, D. (2014). Socioeconomic school segregation in a market-oriented educational system. The case of Chile. *Journal of Education Policy*, 29(2):217–241. Publisher: Routledge _eprint: https://doi.org/10.1080/02680939.2013.806995.

Zimmerman, D. J. (2003). Peer Effects in Academic Outcomes: Evidence from a Natural Experiment. *Review of Economics and Statistics*, 85(1):9–23.

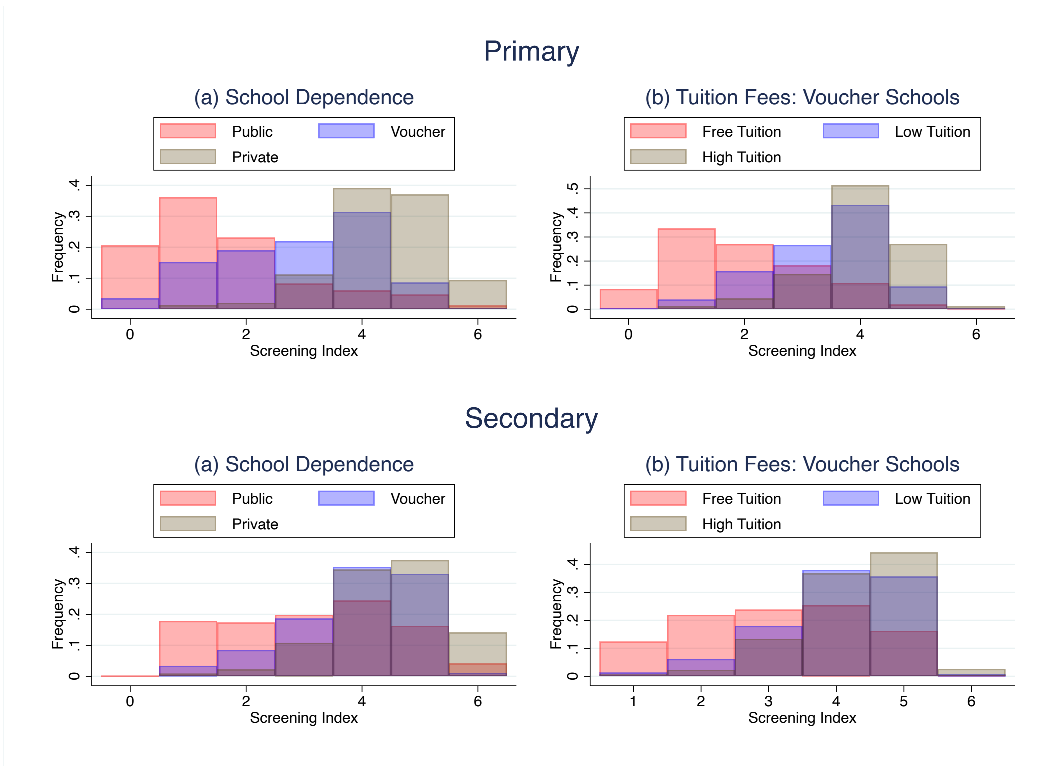
8 Figures

Figure 1: Staggered Implementation of the School Admission System (SAS)



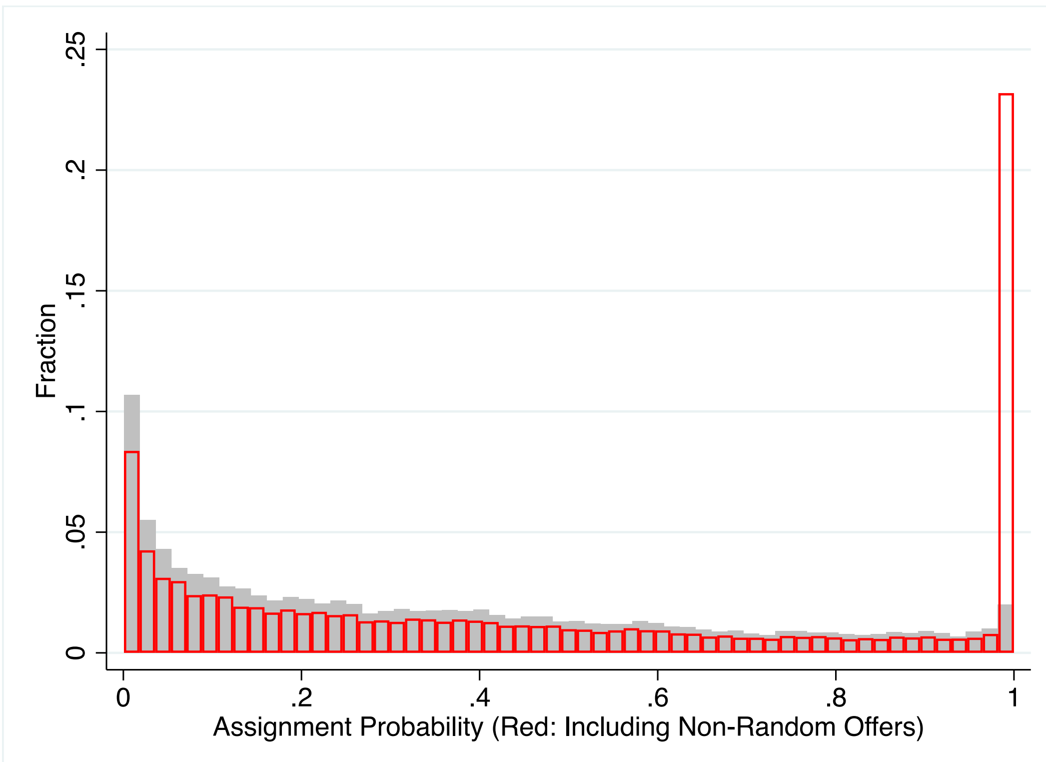
Source: Ministry of Education of Chile. Divisions in the map represent the administrative geographical divisions of regions in Chile.

Figure 2: School Screening Index by Level, School Dependence, and Tuition Fees



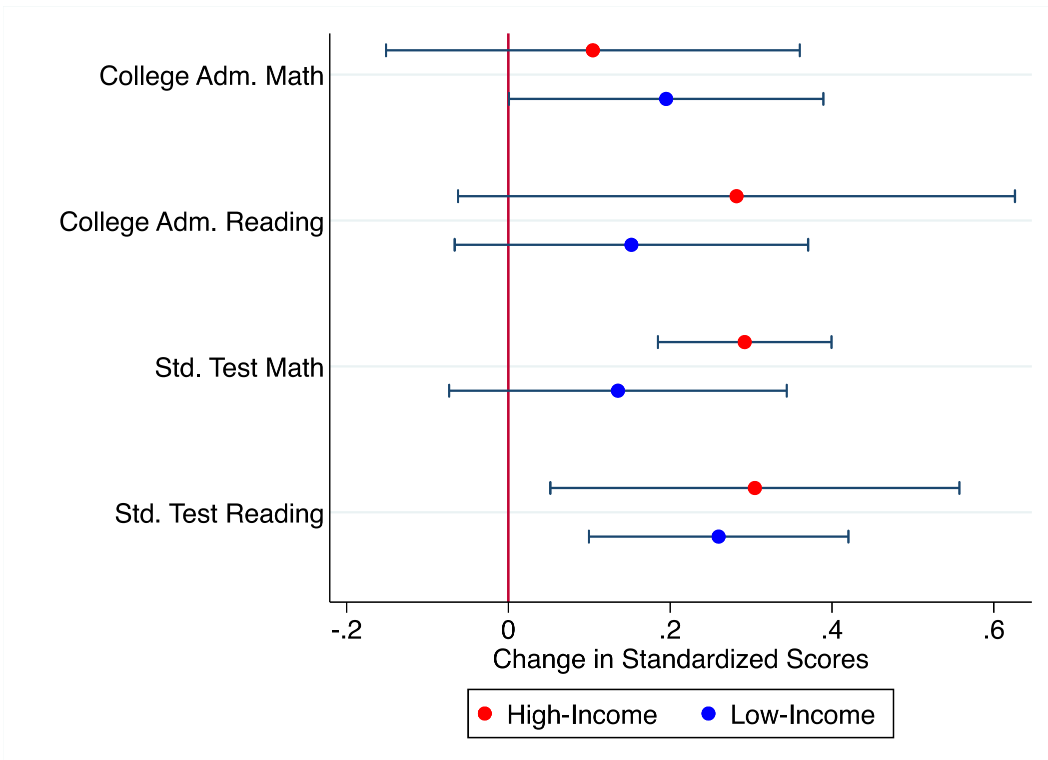
Note: Selectivity index based on parents response about the screening process partaken when joining the school before the implementation of the centralized admission system. The index measures whether parents report having going through a screening process including each of the following: grades certificate, personal interview, preschool certificate, admission exam, psychological assessment, and game dynamics. Panel (a) pools all school dependences and panel (b) includes exclusively voucher schools, since public schools do not charge tuition fees and private schools' admission system was not affected by the SAS.

Figure 3: Assignments' Variability: Proportion of Assignments to a Given School



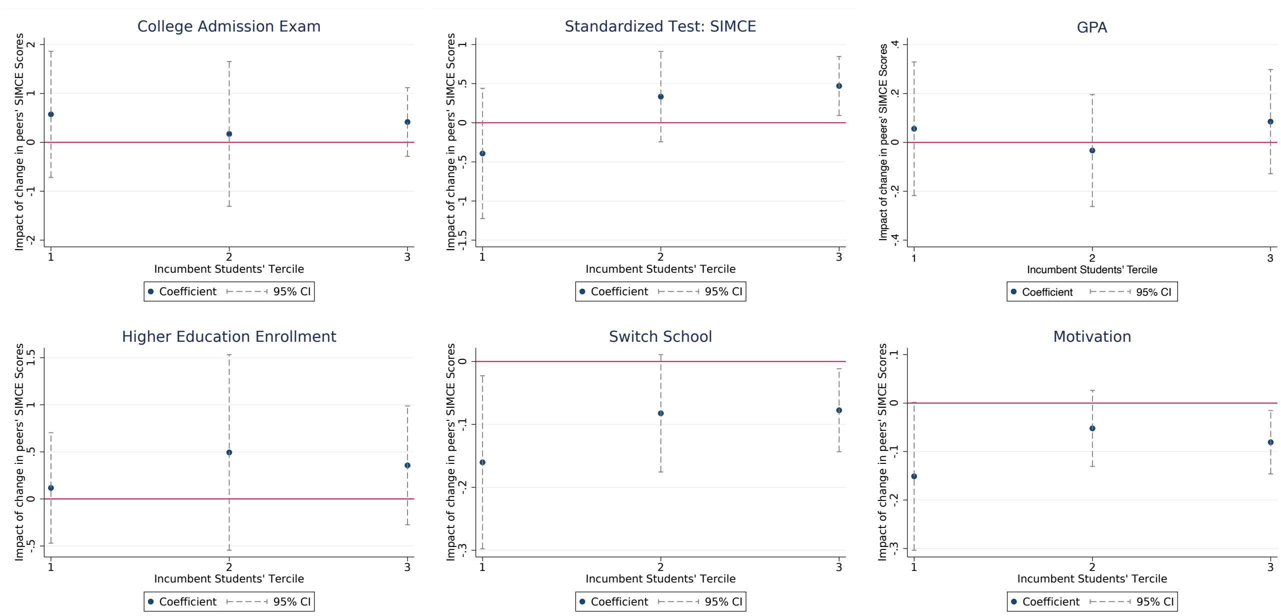
Note: the horizontal axis represents the proportion of the simulations that a student gets assigned to a given school and the vertical axis is the fraction of occurrences of such case. For example, students' that get assigned to a given school 80 percent of the simulations represent about 1 percent of the sample.

Figure 4: Impact of Enrolling in High-Performing Schools by Student Income Level



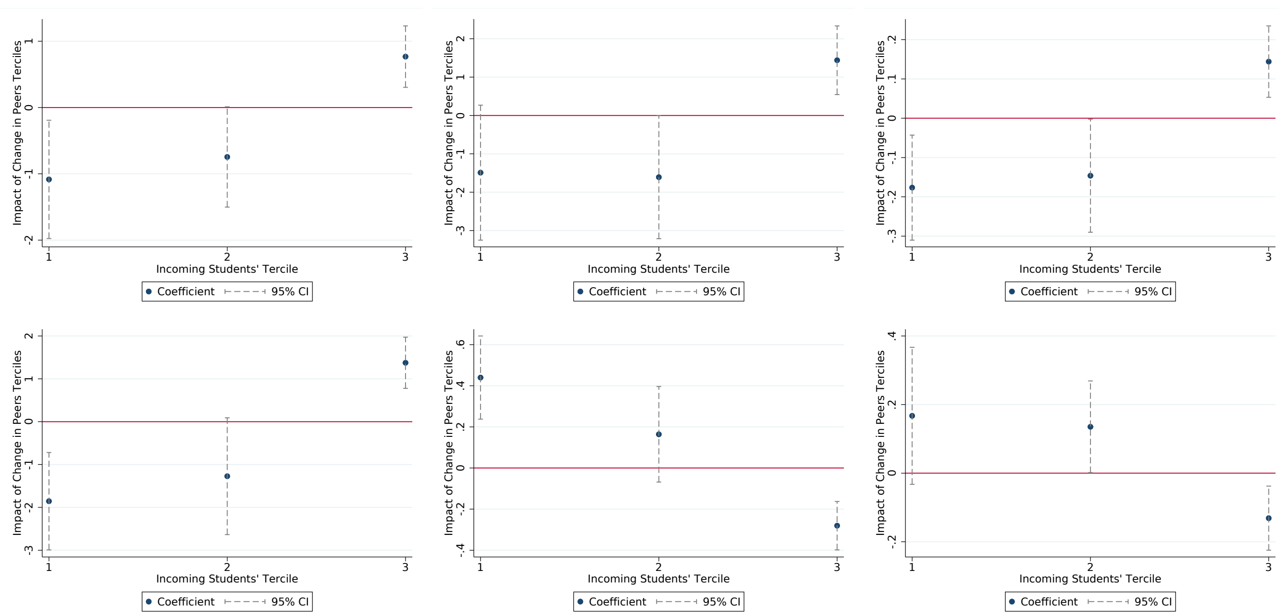
Note: this corresponds to the coefficient from our instrumental variables regression as specified in Section 5.1

Figure 5: Impact of Peers' Standardized Scores - Heterogeneity by Incumbents' Test-Scores Tercile



Each observation corresponds to student-year outcomes. The X-axis corresponds to incumbent student achievement tercile Y-axis corresponds to the impact of changes on peers' average scores on standardized tests. Regressions control for fixed effects by grade-year, decile of applicants' heterogeneity, region, school dependence, and SEP program adherence. Actual incoming students characteristics are instrumented using deviation in each school-grade admitted set of students relative to its applicants. Robust standard errors clustered at the school-grade level in parenthesis. *** $p<0.01$ ** $p<0.05$ * $p<0.1$.

Figure 6: Peers' Standardized Test Scores Effect: Changes in the Proportion of Students From Each Test-Score Tercile



Each observation corresponds to student-year outcomes. The X-axis corresponds to the tercile of peers' achievement and the Y-axis corresponds to the impact of changes in the proportion of students from that tercile. Regressions control for fixed effects by grade-year, decile of applicants' heterogeneity, region, school dependence, and SEP program adherence. Actual incoming students characteristics are instrumented using deviation in each school-grade admitted set of students relative to its applicants. Robust standard errors clustered at the school-grade level in parenthesis. *** p<0.01 ** p<0.05 * p<0.1.

9 Tables

Table 1: SAS Adoption: Changes in New Students Enrollment by School Type

	(1) Std. Scores SIMCE	(2) Raw GPA [1-7]	(3) GPA Rank	(4) Attend. [1-100]	(5) Low SES	(6) Income per capita	(7) Mother w/ High School
Panel A: Voucher Schools							
SAS Level	0.003	0.051***	0.063***	0.546***	0.056***	-0.029***	-0.010***
	(0.007)	(0.006)	(0.008)	(0.065)	(0.003)	(0.005)	(0.003)
Voucher × SAS Level	-0.122***	-0.095***	-0.083***	-1.075***	0.021***	-0.091***	-0.027***
	(0.024)	(0.013)	(0.018)	(0.090)	(0.005)	(0.009)	(0.004)
N	583,730	1,015,469	1,012,693	1,015,471	1,401,676	708,213	738,928
Panel B: High-Performing Schools							
SAS Level	0.027***	0.048***	0.061***	0.492***	0.035***	-0.054***	-0.017***
	(0.007)	(0.005)	(0.008)	(0.067)	(0.003)	(0.005)	(0.003)
High-Performing × SAS Level	-0.222***	-0.174***	-0.189***	-1.496***	0.041***	-0.103***	-0.023***
	(0.046)	(0.022)	(0.033)	(0.171)	(0.013)	(0.023)	(0.008)
N	406,142	758,618	756,610	758,616	836,523	524,171	545,475
Panel C: High Demand Schools							
SAS Level	0.007	0.056***	0.066***	0.583***	0.055***	-0.028***	-0.011***
	(0.007)	(0.006)	(0.009)	(0.069)	(0.003)	(0.005)	(0.003)
High-Demand × SAS Level	-0.118***	-0.107***	-0.103***	-1.136***	0.026***	-0.082***	-0.022***
	(0.022)	(0.012)	(0.017)	(0.092)	(0.005)	(0.009)	(0.004)
N	553,560	973,504	970,768	973,506	1,330,157	684,546	706,235

Each observation corresponds to a student who switched schools at a given year to compare trend changes in students enrolled at each school. Outcomes correspond to the background characteristics of the incoming students enrolled at each school. Voucher schools are publicly subsidized private schools. High-performing schools are those identified by the Ministry of Education as having good test scores given the socioeconomic composition. High-demand schools are those experiencing oversubscription at any of their classrooms during the first three years of the SAS implementation. Robust standard errors clustered at the school level in parenthesis. *** p<0.01 ** p<0.05 * p<0.1.

Table 2: SAS Adoption: Changes in New Students Enrollment by Tuition Fees

	(1) Std. Scores SIMCE	(2) Raw GPA [1-7]	(3) GPA Rank	(4) Attend. [1-100]	(5) Low SES	(6) Income per capita	(7) Mother w/ High School
				Panel A: High-Priced Schools			
SAS Level	0.000	0.048***	0.058***	0.506***	0.057***	-0.029***	-0.011***
	(0.007)	(0.006)	(0.008)	(0.065)	(0.003)	(0.005)	(0.003)
High-Priced	-0.148***	-0.098***	-0.057*	-1.284***	0.011	-0.137***	-0.030***
× SAS Level	(0.033)	(0.019)	(0.030)	(0.116)	(0.008)	(0.013)	(0.005)
N	578,996	1,007,044	1,004,296	1,007,046	1,387,544	703,103	733,536
				Panel B: Mid-Priced Schools			
SAS Level	-0.008	0.044***	0.057***	0.452***	0.057***	-0.037***	-0.013***
	(0.007)	(0.006)	(0.008)	(0.064)	(0.003)	(0.005)	(0.003)
Mid-Priced	-0.051**	-0.071***	-0.061***	-0.818***	0.020***	-0.042***	-0.011*
× SAS Level	(0.025)	(0.015)	(0.020)	(0.123)	(0.007)	(0.013)	(0.006)
N	578,996	1,007,044	1,004,296	1,007,046	1,387,544	703,103	733,536
				Panel C: Free-Tuition Schools			
SAS Level	-0.008	0.042***	0.057***	0.438***	0.058***	-0.042***	-0.014***
	(0.007)	(0.006)	(0.008)	(0.066)	(0.003)	(0.005)	(0.003)
Free Tuition	-0.017	-0.011	-0.027	-0.187	-0.006	0.033***	0.005
× SAS Level	(0.022)	(0.013)	(0.018)	(0.162)	(0.006)	(0.009)	(0.006)
N	578,996	1,007,044	1,004,296	1,007,046	1,387,544	703,103	733,536

Each observation corresponds to a student who switched schools at a given year to compare trend changes in students enrolled at each school. Outcomes correspond to the background characteristics of the incoming students enrolled at each school. High-priced schools are those charging between over 50,000 CLP monthly tuition fees and mid-priced schools charge any tuition fees up to 50,000 CLP. Regressions control by school fixed effects, year fixed effects, and grade fixed effects. Robust standard errors clustered at the school level in parenthesis. *** p<0.01 ** p<0.05 * p<0.1.

Table 3: School Enrollment Effect: College Enrollment

	(1) Takes Coll. Adm. Exam	(2) Enrolled in College	(3) Exam Percentile	(4) Coll. Adm. Exam Math	(5) Coll. Adm. Exam Reading	(6) 8th Grade Std. Test Math	(7) 8th Grade Std. Test Reading
Panel A: High-performing Schools							
High-performing * High-SES	-0.047	0.064	5.968	0.104	0.282	0.292***	0.305**
	(0.071)	(0.095)	(5.312)	(0.130)	(0.176)	(0.055)	(0.129)
High-performing * Low-SES	-0.006	-0.021	3.975	0.195**	0.152	0.135	0.260***
	(0.099)	(0.038)	(4.864)	(0.099)	(0.112)	(0.106)	(0.082)
Difference High-Low	-0.041	0.085	1.993	-0.091	0.130	0.157	0.045
P-Value	0.640	0.402	0.783	0.575	0.512	0.141	0.782
N	42,339	42,339	20,834	20,834	20,834	5,962	5,912
Panel B: Segregated Schools - Affirmative Action Quota							
Has Quota * High-SES	-0.035	0.055	10.032	0.137	0.338	0.077	0.129**
	(0.091)	(0.084)	(9.923)	(0.235)	(0.269)	(0.094)	(0.064)
Has Quota * Low-SES	-0.003	0.002	6.686	0.116	0.137	0.148	0.234**
	(0.075)	(0.057)	(10.659)	(0.237)	(0.260)	(0.110)	(0.092)
Difference High-Low	-0.032	0.053	3.346	0.021	0.201	-0.070	-0.105
P-Value	0.670	0.389	0.590	0.912	0.339	0.623	0.303
N	42,339	42,339	20,834	20,834	20,834	5,962	5,912
Panel C: High-Screening Schools							
High-Screening * High-SES	0.009	0.034	8.963	0.180	0.261	-0.242*	0.187
	(0.077)	(0.074)	(8.080)	(0.193)	(0.213)	(0.143)	(0.188)
High-Screening * Low-SES	-0.000	-0.002	5.570	0.178	0.131	-0.100	-0.600*
	(0.074)	(0.057)	(9.395)	(0.199)	(0.231)	(0.318)	(0.334)
Difference High-Low	0.009	0.036	3.394	0.001	0.130	-0.142	0.788**
P-Value	0.894	0.519	0.517	0.993	0.454	0.673	0.037
N	42,339	42,339	20,834	20,834	20,834	5,962	5,912
Outcome Mean	0.605	0.170	40.229	-0.146	-0.118	0.096	0.160
Outcome SD	0.489	0.375	26.713	0.867	0.855	0.884	0.938

Each observation corresponds to a student who applied through the SAS on a given year. Regressions contains fixed effects grouping students of equivalent SEP status (low-income indicator) and similar assignment propensity to a given classroom. Actual enrollment in the school is instrument using random admission offers. High-performing schools are those identified by the Ministry of Education as having good test scores given the socioeconomic composition. High-screening schools are defined according to parents' survey responses about the process they had to undergo to enroll at their respective schools. Robust standard errors clustered at the school-grade level in parenthesis. *** p<0.01 ** p<0.05 * p<0.1.

Table 4: School Enrollment Effect: Students' Performance

	(1) Raw GPA	(2) Std. GPA	(3) Pass Year	(4) Attendance
Panel A: Selective Schools - High Performing Schools				
High-performing * High-SES	-0.166***	-0.403***	-0.030***	0.762
	(0.050)	(0.036)	(0.011)	(0.501)
High-performing * Low-SES	-0.232***	-0.582***	-0.072***	1.213***
	(0.063)	(0.057)	(0.019)	(0.449)
Difference High-Low	0.065*	0.179***	0.042***	-0.451
P-Value	0.085	0.000	0.002	0.229
N	336,912	336,591	348,263	336,913
Panel B: Segregated Schools - Affirmative Action Quota				
Quota-School * High-SES	-0.054	-0.158***	-0.007	0.468
	(0.037)	(0.034)	(0.009)	(0.428)
Quota-School * Low-SES	-0.057	-0.145***	-0.013	0.108
	(0.048)	(0.046)	(0.015)	(0.549)
Difference High-Low	0.003	-0.013	0.006	0.360
P-Value	0.922	0.724	0.553	0.402
N	336,912	336,591	348,263	336,913
Panel C: High-screening Schools				
High-Screening * High-SES	-0.096***	-0.201***	-0.011	0.299
	(0.036)	(0.033)	(0.009)	(0.459)
High-Screening * Low-SES	-0.130***	-0.259***	-0.022	-0.026
	(0.046)	(0.044)	(0.015)	(0.624)
Difference High-Low	0.034	0.058	0.012	0.325
P-Value	0.248	0.117	0.281	0.468
N	336,912	336,591	348,263	336,913
Outcome Mean	5.578	0.014	0.934	91.893
Outcome SD	0.801	0.990	0.249	10.306

Each observation corresponds to a student who applied through the SAS on a given year. Regressions contains fixed effects grouping students of equivalent SEP status (low-income indicator) and similar assignment propensity to a given classroom. Actual enrollment in the school is instrument using random admission offers. High-performing schools are those identified by the Ministry of Education as having good test scores given the socioeconomic composition. High-screening schools are defined according to parents' survey responses about the process they had to undergo to enroll at their respective schools. Robust standard errors clustered at the school-grade level in parenthesis. *** p<0.01 ** p<0.05 * p<0.1.

Table 5: School Enrollment Effect: Students' Behavior

	(1) Motivation	(2) Self-Confid.	(3) School Satisf.	(4) Discrim.	(5) Behavior Prob.
	Panel A: High-Performing Schools				
High-Perf. * High-SES	-0.017	-0.008	0.014	0.004	-0.012
	(0.027)	(0.023)	(0.014)	(0.009)	(0.010)
High-Perf. * Low-SES	-0.034**	-0.011	0.007	0.005	-0.021**
	(0.015)	(0.022)	(0.022)	(0.010)	(0.008)
Difference High-Low	0.017	0.003	0.007	-0.001	0.009
P-Value	0.335	0.765	0.558	0.931	0.367
N	6,657	6,636	6,671	6,620	6,599
	Panel B: Affirmative Action Schools				
Quota-School * High-SES	-0.012	-0.007	0.008	0.005	-0.014*
	(0.022)	(0.020)	(0.011)	(0.009)	(0.009)
Quota-School * Low-SES	-0.013	-0.005	0.004	-0.008	-0.018***
	(0.018)	(0.020)	(0.019)	(0.011)	(0.007)
Difference High-Low	0.001	-0.002	0.004	0.013	0.004
P-Value	0.925	0.844	0.751	0.328	0.657
N	6,657	6,636	6,671	6,620	6,599
	Panel C: High-Screening Schools				
High-Screening* High-SES	0.033	0.023	0.001	0.038	0.025
	(0.038)	(0.025)	(0.024)	(0.034)	(0.025)
High-Screening* Low-SES	-0.036	-0.043	-0.045	-0.004	0.064*
	(0.037)	(0.037)	(0.034)	(0.052)	(0.035)
Difference High-Low	0.069	0.067	0.046	0.042	-0.038
P-Value	0.212	0.150	0.265	0.536	0.260
N	6,657	6,636	6,671	6,620	6,599
ControlMean	0.690	0.746	0.511	0.086	0.249
ControlSD	0.138	0.123	0.104	0.122	0.086

Each observation corresponds to a student who applied through the SAS on a given year. Regressions control for fixed effects by grade-year, decile of applicants' heterogeneity, region, school dependence, and SEP program adherence. Actual incoming students characteristics are instrumented using deviation in each school-grade admitted set of students relative to its applicants. Robust standard errors clustered at the school-grade level in parenthesis. *** p<0.01 ** p<0.05 * p<0.1.

Table 6: Peers' Background Effect: Students' Performance

	(1)	(2)	(3)	(4)	(5)	(6)
	School Performance				Standardized Tests	
	Raw GPA	GPA Rank	Attend.	Pass Year	Math	Reading
Panel A: Classmates' Standardized Tests Score (SIMCE)						
Classmates' Scores	0.086	-0.668***	-0.252	0.078***	0.174	0.082
	(0.094)	(0.073)	(3.235)	(0.028)	(0.248)	(0.251)
N	282,672	282,519	282,672	287,388	39,473	39,264
Panel B: Classmates' GPA-Rank (previous school)						
Classmates' GPA-Rank	0.541***	-0.584***	4.844	0.062**	0.226	0.221
	(0.149)	(0.028)	(3.323)	(0.029)	(0.198)	(0.176)
N	916,489	916,485	916,485	922,390	41,178	40,943
Panel C: Classmates's Households Income per Capita						
Household Income	1.380***	1.280***	8.127***	0.095***	0.346	0.085
	(0.194)	(0.346)	(3.050)	(0.033)	(0.381)	(0.340)
N	745,117	744,887	745,114	753,856	37,813	37,490
OutcomeMean	5.751	0.002	92.616	0.964	-0.119	-0.037
SD	0.785	0.984	9.878	0.186	0.940	0.959
IncludesLags	Yes	Yes	Yes	No	Yes	Yes

Each observation corresponds to a student-year. Regressions control for fixed effects by grade-year, decile of applicants' heterogeneity, region, school dependence, and SEP program adherence. Actual incoming students characteristics are instrumented using deviation in each school-grade admitted set of students relative to its applicants. Robust standard errors clustered at the school-grade level in parenthesis. *** p<0.01 ** p<0.05 * p<0.1.

Table 7: Peers' Background Effect: Behavior Outcomes

	(1) Motivation	(2) Self-Confid.	(3) School Satisf.	(4) Discrim.	(5) Behavior Prob.
Panel A: Classmates' SIMCE Score					
Classmates SIMCE Score	-0.080**	-0.098***	-0.060	0.030	-0.042**
	(0.033)	(0.035)	(0.037)	(0.023)	(0.020)
N	39,165	38,994	39,219	38,961	38,856
Panel B: Classmates' GPA-Rank					
Classmates' GPA-Rank	-0.038	-0.047*	-0.022	0.004	-0.030*
	(0.026)	(0.026)	(0.032)	(0.019)	(0.016)
N	45,860	45,665	45,928	45,622	45,492
Panel C: Classmates' Households Income per Capita					
Household Income	-0.174***	-0.156**	-0.174**	0.018	0.027
	(0.065)	(0.066)	(0.074)	(0.042)	(0.048)
N	39,995	39,816	40,037	39,776	39,684
OutcomeMean	0.692	0.745	0.516	0.083	0.255
SD	0.136	0.122	0.104	0.118	0.089
IncludesLags	No	No	No	No	No

Each observation corresponds to a student-year. Regressions control for fixed effects by grade-year, decile of applicants' heterogeneity, region, school dependence, and SEP program adherence. Actual incoming students characteristics are instrumented using deviation in each school-grade admitted set of students relative to its applicants. Robust standard errors clustered at the school-grade level in parenthesis. *** $p<0.01$ ** $p<0.05$ * $p<0.1$.

Table 8: Peers' Background Effect: School Switching

	(1) Any School	(2) Public Sch.	(3) Voucher Sch.	(4) Priv. Sch.
Panel A: Classmates' SIMCE Score				
Classmates SIMCE Score	-0.099***	-0.061***	-0.036*	-0.002
	(0.033)	(0.019)	(0.021)	(0.004)
N	274,773	274,773	274,773	274,773
Panel B: Classmates' GPA-Rank				
Classmates' GPA-Rank	-0.147***	-0.061**	-0.054**	0.002
	(0.044)	(0.026)	(0.026)	(0.004)
N	866,980	866,980	866,980	866,980
Panel C: Classmates' Households Income per Capita				
Household Income	-0.731***	-0.288***	-0.393***	-0.003
	(0.094)	(0.047)	(0.055)	(0.006)
N	715,370	715,370	715,370	715,370
OutcomeMean	0.135	0.060	0.057	0.003
SD	0.342	0.237	0.233	0.053
IncludesLags	No	No	No	No

Each observation corresponds to a student-year. Regressions control for fixed effects by grade-year, decile of applicants' heterogeneity, region, school dependence, and SEP program adherence. Actual incoming students characteristics are instrumented using deviation in each school-grade admitted set of students relative to its applicants. Robust standard errors clustered at the school-grade level in parenthesis. *** $p<0.01$ ** $p<0.05$ * $p<0.1$.

Table 9: Peers' Background Effect: College Enrollment

	(1) Take College Adm. Exam	(2) Enroll in College	(3) Coll. Exam Percentile	(4) College Admission Exam Reading	(5) College Admission Exam Math
Panel A: Classmates SIMCE Scores					
Classmates' SIMCE Scores	0.260	0.413	34.191**	126.466**	110.551**
	(0.289)	(0.269)	(13.288)	(55.417)	(50.935)
N	71,622	71,622	44,913	44,913	44,913
Panel B: Classmates GPA-Rank					
Classmates' GPA-Rank	0.080	0.632	59.395	341.828	122.425
	(0.645)	(0.814)	(122.028)	(526.585)	(369.880)
N	199,202	199,202	137,522	137,522	137,522
Panel C: Classmates' Households Income per Capita					
Household Income per Capita	0.149	0.320	46.757***	156.560***	170.941***
	(0.223)	(0.196)	(15.057)	(59.723)	(62.381)
N	154,821	154,821	110,770	110,770	110,770
OutcomeMean	0.749	0.298	45.744	476.786	474.114
SD	0.434	0.458	28.093	122.191	127.771
IncludesLags	No	No	No	No	No

Each observation corresponds to a student-year. Regressions control for fixed effects by grade-year, decile of applicants' heterogeneity, region, school dependence, and SEP program adherence. Actual incoming students characteristics are instrumented using deviation in each school-grade admitted set of students relative to its applicants. Robust standard errors clustered at the school-grade level in parenthesis. *** p<0.01 ** p<0.05 * p<0.1.

A Appendix

A.1 SAS Adoption: Within School Comparison

Table 10: SAS Adoption: School Characteristics

	(1) Std. Scores SIMCE	(2) Raw GPA [1-7]	(3) GPA Rank	(4) Attend. [1-100]	(5) Low SES	(6) Income per capita	(7) Mother w/ High School
				Panel A: Voucher Schools			
SAS Level	0.037***	0.125***	0.117***	0.981***	0.035***	-0.029***	-0.018***
	(0.008)	(0.006)	(0.009)	(0.096)	(0.004)	(0.007)	(0.004)
Voucher	-0.114***	-0.128***	-0.003	-1.526***	0.025**	-0.076***	0.024***
× SAS Level	(0.023)	(0.015)	(0.026)	(0.200)	(0.010)	(0.019)	(0.009)
N	580,595	1,012,627	1,009,876	1,012,629	1,398,771	703,414	734,346
				Panel B: High-Performing Schools			
SAS Level	0.053***	0.105***	0.108***	0.771***	0.029***	-0.050***	-0.015***
	(0.008)	(0.006)	(0.009)	(0.092)	(0.004)	(0.007)	(0.004)
High-Performing	-0.170***	-0.153***	-0.043	-1.757***	0.007	-0.111***	0.006
× SAS Level	(0.047)	(0.031)	(0.043)	(0.269)	(0.020)	(0.040)	(0.015)
N	404,562	756,539	754,533	756,537	834,717	521,734	543,121
				Panel C: High Demand Schools			
SAS Level	0.036***	0.122***	0.115***	0.960***	0.035***	-0.025***	-0.017***
	(0.008)	(0.006)	(0.009)	(0.098)	(0.004)	(0.007)	(0.004)
High-Demand	-0.095***	-0.123***	-0.015	-1.401***	0.018*	-0.073***	0.020**
× SAS Level	(0.023)	(0.015)	(0.025)	(0.198)	(0.010)	(0.019)	(0.009)
N	550,696	970,917	968,203	970,919	1,327,468	680,210	701,983

Standard errors clustered at classroom level in parentheses. *** p<0.01 ** p<0.05 * p<0.1.

Each observation corresponds to a student who switched schools at a given year to compare trend changes in students enrolled at each school. Regressions include year times school fixed effects to compare outcome within a school-year application period. Outcomes correspond to the background characteristics of the incoming students enrolled at each school. Robust standard errors clustered at the school level in parenthesis. *** p<0.01 ** p<0.05 * p<0.1.

A.2 Peer Background Effect: Income Level

Table 12: Peers' Background Effect: Income

	(1)	(2)	(3)	(4)	(5)	(6)
	School Performance				Standardized Tests	
	GPA	GPA Rank	Attend.	Pass Year	Math	Reading
Panel A: Proportion of Low-SES Classmates						
Prop. of Low-SES Classmates	0.370	1.395***	-12.983	0.020	-1.201	0.583
	(0.439)	(0.299)	(8.028)	(0.092)	(1.735)	(1.616)
N	929,660	928,771	929,657	954,290	41,870	41,624
Panel B: Classmates's Mother High School Degree						
Mother High School Ed.	1.629***	0.177	10.282	0.214***	0.703	0.783
	(0.411)	(0.255)	(7.221)	(0.074)	(0.738)	(0.601)
N	770,535	770,291	770,531	779,651	38,852	38,536
Panel C: Classmates's Mother College Degree						
Mother College Ed.	0.817*	-0.601*	18.650**	0.043	2.289	4.226
	(0.457)	(0.312)	(8.872)	(0.076)	(2.964)	(3.804)
N	770,535	770,291	770,531	779,651	38,852	38,536
OutcomeMean	5.751	0.002	92.616	0.964	-0.119	-0.037
SD	0.785	0.984	9.878	0.186	0.940	0.959
IncludesLags	Yes	Yes	Yes	No	Yes	Yes

Each observation corresponds to a student-year. Regressions control for fixed effects by grade-year, decile of applicants' heterogeneity, region, school dependence, and SEP program adherence. Actual incoming students characteristics are instrumented using deviation in each school-grade admitted set of students relative to its applicants. Robust standard errors clustered at the school-grade level in parenthesis. *** p<0.01 ** p<0.05 * p<0.1.

Table 13: Peers' Background Effect

	(1) Motivation	(2) Self-Confid.	(3) School Satisf.	(4) Discrim.	(5) Behavior Prob.
Panel A: Proportion of Low-SES Classmates					
Low-SES Classmates	0.021	0.042	0.094	0.061	-0.053
	(0.155)	(0.148)	(0.182)	(0.135)	(0.144)
N	47,310	47,104	47,380	47,056	46,920
Panel B: Classmates' Mother High School Degree					
Mother High School Ed.	-0.133*	-0.062	-0.048	0.071	-0.081
	(0.077)	(0.069)	(0.094)	(0.052)	(0.056)
N	41,134	40,951	41,183	40,912	40,814
Panel C: Classmates' Mother College Degree					
Mother College Ed.	0.247	-0.251	-0.274	0.235	-0.094
	(0.375)	(0.329)	(0.449)	(0.268)	(0.229)
N	41,134	40,951	41,183	40,912	40,814
OutcomeMean	0.692	0.745	0.516	0.083	0.255
SD	0.136	0.122	0.104	0.118	0.089
IncludesLags	No	No	No	No	No

Each observation corresponds to a student-year. Regressions control for fixed effects by grade-year, decile of applicants' heterogeneity, region, school dependence, and SEP program adherence. Actual incoming students characteristics are instrumented using deviation in each school-grade admitted set of students relative to its applicants. Robust standard errors clustered at the school-grade level in parenthesis. *** p<0.01 ** p<0.05 * p<0.1.

Table 11: SAS Adoption: School Characteristics

	(1) Std. Scores SIMCE	(2) Raw GPA [1-7]	(3) GPA Rank	(4) Attend. [1-100]	(5) Low SES	(6) Income per capita	(7) Mother w/ High School
				Panel A: High-Priced Schools			
SAS Level	0.035***	0.121***	0.114***	0.928***	0.036***	-0.027***	-0.017***
	(0.008)	(0.006)	(0.009)	(0.093)	(0.004)	(0.007)	(0.004)
High-Priced	-0.150***	-0.136***	0.039	-1.741***	0.018	-0.134***	0.022**
× SAS Level	(0.025)	(0.018)	(0.031)	(0.212)	(0.011)	(0.024)	(0.009)
N	575,892	1,004,228	1,001,505	1,004,230	1,384,665	698,350	728,999
				Panel B: Mid-Priced Schools			
SAS Level	0.025***	0.114***	0.116***	0.850***	0.036***	-0.036***	-0.016***
	(0.008)	(0.006)	(0.009)	(0.090)	(0.004)	(0.007)	(0.004)
Mid-Priced	-0.057	-0.103***	0.023	-1.506***	0.029*	-0.042	0.022
× SAS Level	(0.037)	(0.024)	(0.046)	(0.390)	(0.017)	(0.030)	(0.014)
N	575,892	1,004,228	1,001,505	1,004,230	1,384,665	698,350	728,999
				Panel C: Free-Tuition Schools			
SAS Level	0.019**	0.110***	0.119***	0.790***	0.039***	-0.038***	-0.017***
	(0.008)	(0.006)	(0.009)	(0.091)	(0.004)	(0.007)	(0.004)
Free Tuition	0.078***	0.010	-0.028	0.081	-0.030**	0.019	0.024
× SAS Level	(0.029)	(0.020)	(0.033)	(0.335)	(0.014)	(0.023)	(0.015)
N	575,892	1,004,228	1,001,505	1,004,230	1,384,665	698,350	728,999

Standard errors clustered at classroom level in parentheses. *** p<0.01 ** p<0.05 * p<0.1

Each observation corresponds to a student who switched schools at a given year to compare trend changes in students enrolled at each school. Outcomes correspond to the background characteristics of the incoming students enrolled at each school. Regressions include year times school fixed effects to compare outcome within a school-year application period. Robust standard errors clustered at the school level in parenthesis. *** p<0.01 ** p<0.05 * p<0.1.

Table 14: Peers' Background Effect: Income

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Take College	Enroll in	Coll. Exam	College Admission Exam Subject			
	Adm. Exam	College	Percentile	Reading	Math	History	Science
Panel A: Proportion of Low-SES Classmates							
Prop. Low-SES	0.189	0.106	-55.833*	-183.766	-197.773*	11.683	-66.145
	(0.516)	(0.457)	(31.265)	(119.872)	(120.151)	(181.992)	(178.185)
N	209,069	209,069	142,699	142,699	142,699	142,699	142,699
Panel B: Proportion of Classmates' Mother High School Degree							
Prop. High School Deg.	0.213	-0.005	17.870	131.237	96.144	53.540	184.513
	(0.495)	(0.561)	(43.817)	(157.181)	(153.269)	(233.568)	(242.327)
N	160,507	160,507	115,259	115,259	115,259	115,259	115,259
Panel C: Proportion of Classmates' Mother College Degree							
Prop. Mother College Deg.	0.053	0.831	90.167*	237.203	433.197*	547.940*	144.829
	(0.449)	(0.602)	(51.499)	(162.998)	(222.989)	(304.876)	(233.068)
N	160,507	160,507	115,259	115,259	115,259	115,259	115,259
OutcomeMean	0.749	0.298	45.744	476.786	474.114	270.173	320.414
SD	0.434	0.458	28.093	122.191	127.771	253.378	241.765
IncludesLags	No	No	No	No	No	No	No

Each observation corresponds to a student-year. Regressions control for fixed effects by grade-year, decile of applicants' heterogeneity, region, school dependence, and SEP program adherence. Actual incoming students characteristics are instrumented using deviation in each school-grade admitted set of students relative to its applicants. Robust standard errors clustered at the school-grade level in parenthesis. *** p<0.01 ** p<0.05 * p<0.1.

A.3 Peer Effects Instrument Validity

To verify the validity of this instrument, let z_i be the characteristics' vector of a student i and $Z_j = (z_1, \ldots, z_q)$ be the vector of q applicants to school j. Then the allocation mechanism will assign a subset $\mu(X_j) = [\mu_1^j, \ldots, \mu_q^j]$, where $\mu_i^j = 1$ indicates that student i was assigned to school j and $\mu_i^j = 0$ otherwise. The function μ depends on X_j to reflect that the allocations could depend on students' characteristics X under alternative assignment mechanisms. Given the limited number of spots, it has to hold that $\sum_i \mu_1^j \leq k_j \ \forall j$. Note that we can then rewrite the potential outcomes as follows:

$$Y_{ij} = \alpha_j + \beta x_i + \gamma_{(i,-i)} \mu(Z_j) Z_j^{-i} + \epsilon_{ij} \tag{12}$$

Given that applicants self-select when applying to schools, we have that $\mu(Z_j)Z_j^{-i} \not\perp \epsilon_{ij}$ even when spots in schools are randomly allocated among applicants. This is because the applicants' set differs for every school due to characteristics possibly related to unobservables. Consequently, a direct OLS measurement of $\gamma_{(i,-i)}$ would yield biased estimates. To overcome this, define instead $W(Z_j) = \mu(Z_j)Z_j - E[\mu(Z_j)Z_j|Z_j]$, which we refer to as classroom shocks. First, note that this is uncorrelated with the error term in the structural equation once we condition on applicants' characteristics:

$$\begin{aligned}
E[W(Z_j)'\epsilon_j|Z_j] &= E[Z_j'\mu(Z_j)'\epsilon_j - E[Z_j'\mu(Z_j)'|Z_j]\epsilon_j|Z_j] \\
&= Z_j'E[\mu(Z_j)'\epsilon_j|Z_j'] - Z_j'E[E[\mu(Z_j)'|Z_j]\epsilon_j|Z_j] \\
&= Z_j'E[\mu(Z_j)'\epsilon_j|Z_j'] - Z_j'E[\mu(Z_j)'\epsilon_j|Z_j] \\
&= 0
\end{aligned}$$

Where the first part is zero because the random assignment from DA guarantees $\mu(Z_j) \perp \epsilon_{ij}$. However, we would not expect this correlation to be zero if spots were not randomly allocated or if schools performed screening practices.

To analyze the relevance of $W(Z_j)$ as an instrument for Z_j^{-i}, we can rewrite this in the following manner:

$$\begin{aligned}
E[W(Z_j)Z_j^{-i}|Z_j] &= E[\mu(Z_j)Z_jZ_j^{-i} - E[\mu(Z_j)Z_j]Z_j^{-i}|Z_j] \\
&= E[\mu(Z_j) - E[\mu(Z_j)]|Z_j]Z_jZ_j^{-i} \\
&= V[\mu(Z_j)]|Z_j]Z_jZ_j^{-i}
\end{aligned}$$

From here, we can conclude that there are three conditions are necessary for $\mu(Z_j)$ to be a relevant instrument for classroom composition: i) that the school is oversubscribed, so that $V[\mu(Z_j)] \neq 0$; ii) that there is variation among the applicants themselves, so that $Z_j Z_j^{-i} \neq 0$; and iii) that the proportion of randomized spots is large enough.

In an empirical setting, we rarely observe schools with identical characteristics and an equal number of spots and applicants. However, our estimation of the empirical distribution of new students allows us to identify schools with equivalent shock distributions. In practice, the definition of similarity will depend on the specific functional form used in the analysis. For example, in a linear-in-means model, a similar applicant pool would be one with a similar distribution of the average of admitted students. Therefore, we control by applicants' average characteristics and bins of standard deviation to produce schools with comparable admitted students distributions.