School Choice, Student Sorting and Academic Performance

Andrei Munteanu[†]

Université du Québec à Montréal

Abstract

This study examines the impact of school choice on academic achievement. I use differences in the number of schools across similar Romanian towns, generating variation in school choice for local students, who compete for seats via test scores. I find that more school choice results in increased sorting of students by admission scores across different schools. Sorting widens achievement gaps between high- and low-admission score students. High-scorers having access to better teachers and peer effects are the primary factors explaining these widening gaps. Lastly, between-school competition via school choice does not increase average achievement levels.

Keywords: education, school choice, sorting, inequality, peers, teachers

JEL Classification Codes : I24, I21, I28

[†]Email: munteanu.andrei@uqam.ca. I am grateful to Fabian Lange, Francesco Amodio and Nicolas Gendron-Carrier for excellent supervision and advice, as well as Michael Gilraine, Andra Hirișcău, Brian Jacob, Lawrence Katz, John Klopfer, Andreea Mitruț, Daniela Morar, Ioana Neamțu, Cristian Pop-Eleches, Laëtitia Renée, Fernando Saltiel, Kevin Thom, Miguel Urquiola, Xian Zhang and three anonymous referees. A previous version of this paper circulated under the title "Student Sorting and Academic Performance".

1 Introduction

School choice remains a contentious issue in education policy. Supporters argue that introducing competition through school choice would improve public schools, benefiting all students. However, critics assert that school choice could exacerbate disparities by ability, race, and socioeconomic factors, particularly when geographic constraints, limited mobility, and incomplete information are present. The challenge in effectively analyzing the impact of school choice, especially the distributional, general-equilibrium effects, is exacerbated by a lack of data and a paucity of quasi-experimental variation, hindering a comprehensive understanding of its consequences on students.

In this paper, I ask whether school choice is indeed a "rising tide that lifts all boats" or merely exacerbates existing educational inequalities. To do this, I study the Romanian high school system. In Romania, prospective high school students compete for high school seats of their choice via a national, standardized exam. Students then fill out preferences over high schools. A centralized serial dictatorship rule, prioritizing students with high scores on a standardized admission exam, assigns them to their most preferred available seat.

I exploit differences in the number of high schools across Romanian towns with similar populations and characteristics. These differences engender variation in school choice for local students. For example, in a small town with only one high school, there is no school choice or student sorting, as all students must attend the same high school (unless they migrate or commute). In a similar town with two high schools, students face school choice and, with high-scoring students receiving priority, can potentially sort by admission scores across schools.

Using these differences in the number of schools, and more than two million linked student and graduation records, I first study how choice impacts student sorting across schools. Second, I explore the effects of this sorting on academic outcomes. Lastly, using novel data on teacher placement scores and school spending, that I match to the student records, I explore three plausible channels through which school choice-induced sorting affects academic outcomes: peer effects, access to better teachers and school spending.

I argue that the variation in the number of schools across similar towns is plausibly exogenous for three reasons. First, student baseline characteristics (including test scores) across towns with similar populations but different numbers of schools are statistically indistinguishable. Second, geographic mobility and household sorting across locations in Romania is limited, partially due to its record-high homeownership rate. Third, the number of schools is highly invariant over time. This number is largely predetermined by the last communist-regime schooling expansion in the 1960s and 1970s and is uncorrelated with current demographic and economic realities. To summarize, the number of high schools across towns is invariant and divorced from current socioeconomic and student population characteristics, while household sorting across towns is severely restricted.

Three main findings emerge from this study. I first find that locations with more high schools experience more segregation by test scores across schools. This sorting is much more pronounced than across middle schools in the same towns, where a neighborhood assignment rule makes allocations. The mechanism underlying this process is simple: the more high schools there are in a given market, the more competitive admissions to very selective schools are. High-score students can sort into these high schools, while their low-scoring counterparts are increasingly relegated to low-quality schools.

Second, I show that the sorting ultimately leads to increased inequalities in test scores at graduation without affecting mean achievement. Low-admission score students fall farther and farther behind their high-score counterparts in locations with more high schools. Meanwhile, there is no evidence that mean achievement increases with school choice, despite strong competitive pressures on schools to improve their services and attract better students.

I confirm these findings using several robustness checks. I first study school openings. After a new school opens, sorting increases markedly, and the graduation exam score gap between high- and low-admission score students increases. I then use propensity score matching to pair control students to treated students with similar admission scores and living in similar-sized towns with one additional high school. Treated students sort more heavily on admission scores, and wider graduation test score gaps form between high- and low-admission score-treated students. The results are also robust to excluding students attending high school in a different town or to defining education markets based on student attendance patterns rather than at the town level.

Lastly, I explore three potential mechanisms explaining how student sorting affects academic outcomes: peer effects, access to high-ability teachers, and access to schools with more financial resources. I first use year-to-year variation in admission scores within high school tracks to show that students who benefit from higher-scoring peers score higher on the graduation exam.

Next, I visit the relationship between student performance and access to teachers, using teacher hiring records that include placement scores on subject-specific exams designed to allocate teachers to schools. I find that high-admission score students have better access to high-placement score teachers, particularly in towns with more high schools. This implies that in areas with more high schools, top students are not only more likely to attend more selective schools but also more likely to have access to high-ability teachers. Access to high-ability teachers translates into higher test scores at graduation, after controlling for student and peer baseline characteristics. In contrast, other teacher characteristics like college GPA, experience, and education do not seem to have as strong an effect. This suggests that teacher subject-specific expertise may be a key factor in determining student success. Meanwhile, school spending only weakly varies with student test scores, with low-achieving students attending schools that spend marginally more per student. In terms of achievement, access to higher-spending schools is not predictive of higher test scores.

I then estimate the relative contribution of each of these channels on student performance. To achieve this, I estimate a model of student graduation scores as a function of high school selectivity, to which I sequentially add measures of teacher ability, peer admission score variation and school spending. Teacher sorting and peer effects explain the lion's share of the effect on test scores of attending a more selective school. School spending explains very little of the effect of attending a more selective school on graduation scores.

The unique Romanian setup and data allow me to make three main contributions. First, I am able to tease out the distributional, general equilibrium implications of school choice on student sorting patterns and student achievement. This is difficult to achieve in other setups, because of i) a lack of baseline and endline test scores, ii) no plausibly exogenous variation in school choice and iii) a lack of variation in school choice affecting all students in a given market. Thus, I am able to go beyond analyzing the impact of school choice only on some marginal students or only on mean scores.

Second, I show how access to good schools and, in particular in the Romanian context, the mechanism allocating students to schools is perhaps even more important than school choice itself. Suppose policymakers are not able to ensure access to quality schools for all students. In that case, school choice can serve as an engine that exacerbates educational inequalities, rather than a tool to improve school efficiency.

Third, data on teacher hiring and school spending allows me to disentangle exactly how the school choice and admission policies affect student outcomes. This is particularly policy-relevant, since decision makers can allocate these resources across schools to complement school choice policies.

This paper speaks to three main strands of literature. The first one is the literature on school choice. A large body of literature discusses the efficiency gains generated by school choice, with some studies finding evidence for increases in mean test scores (e.g., Hoxby 2000, Campos and Kearns 2022). In contrast, others find no effects of school choice on achievement and explore the theoretical reasons why, in the presence of peer effects, asymmetric information and other market frictions, such gains are modest (Hsieh and Urquiola 2003, Rothstein 2006). In light of this, other studies have focused on the effects of school choice on student sorting and stratification (MacLeod and Urquiola 2015, Altonji et al. 2015, Barseghyan et al. 2019, MacLeod and Urquiola 2019, Abdulkadiroğlu et al. 2020, Laverde 2021 and Machado and Szerman 2021), especially since a large body of theoretical literature predicts the reallocation of students across schools to be the main effect of school choice, rather than efficiency gains.

The rich Romanian data allows me to link these two strands of literature. Indeed, I first find that school choice accentuates student sorting on scores and then proceed to show its effects on student learning outcomes. While I do not find any differences in mean scores across localities with differences in the degree of school choice students have, school choice increases the variance of test scores, accentuating the differences between high- and low-ability students.

Another relevant strand of literature is the literature on the impact of attending a better school on student achievement. Attending a better school improves academic outcomes (for example, Deming et al. 2014 and Pop-Eleches and Urquiola 2013 in the Romanian context, among many others). I contribute to this literature by providing general-equilibrium evidence that high-ability students progressively improve their test scores when they sort into better schools.

Lastly, I contribute to the literature on tracking in education. The Romanian admission mechanism essentially creates tracking by ability between different high schools. Duflo et al. (2011) suggest that tracking may benefit everyone in contexts where students are of very different abilities. My results suggest that tracking, especially when associated with a potential disparity in access to teaching resources, may lead to a widening gap between high- and low-scoring students. This different result can be rationalized by the fact that, unlike in Duflo et al. (2011), teaching and other school resources are allocated endogenously. Moreover, in the Kenyan context, students are of very heterogeneous ability levels and even ages which might drive the effectiveness of tracking by ability. In contrast, the Romanian high school student population is much more homogeneous. Peer effects and access to good teachers might dominate the benefits of attending school with similar-ability peers.

2 Setup and Data

High School Admission Exam Each summer, Romanian middle schoolers completing 8^{th} grade are assigned to high schools by a centralized system. Students first write a standardized, national high school admission exam. This exam covers mathematics, Romanian, and, in some years, a choice between history and geography components. Each student is then assigned an admission score comprised of a weighted average of their admission exam score and their middle school (grades 5-8) GPA. While the weights on each component vary across time, at least half of the total score is attributed to the admission exam.

High School Tracks Romanian high schools offer multiple tracks that broadly fall into one of two streams: theory tracks (e.g., science, humanities) and technical tracks (e.g., industrial, tourism). These tracks offer very uniform curricula across different schools. Although tracks can be thought of as schools-within-schools to a certain extent, as students apply to each track separately and there is segregation of students across tracks, students in different tracks do interact in some classes (for example, foreign language classes and possibly electives) and outside of classes. Additionally, all students benefit from having the same teaching faculty and school facilities. Lastly, anecdotally, parents have strong preferences over schools, with some parents preferring any track in a very prestigious school to better-suited tracks in slightly less renowned ones. For these reasons, I conduct the main analyses at the track and school levels separately.

High School Allocation Mechanism After receiving their admission scores, students fill out an almost untruncated list of ranked preferences over combinations of high schools and tracks they wish to attend. For example, a student can rank the humanities track in high school A as their first choice, the humanities track in high school B as their second choice and the science track in high school A as their third choice. A centralized algorithm then allocates students to high school seats via a serial dictatorship rule that gives high-scoring students absolute priority over lower-admission score students. Students have no incentive to strategically manipulate their preferences over tracks in the hope of a better assignment. Brochures provided by the Ministry of Education explain these features of the allocation and head teachers help parents and students fill out the preference lists.

High School Graduation Exam After completing four years of high school (grades 9-12), Romanian high school students register to take a national, standardized high school exit exam. It consists of several subject components, including Romanian, mathematics, and other track-specific subjects. Like the admission exam, the graduation exam is high-stakes. Passing it (by obtaining a grade of at least 50% on every component) provides students with an additional certification often seen as more valuable than a high school diploma. Moreover, results on the graduation exam can be used as an admission requirement for students planning to attend postsecondary schooling. The admission and graduation exams thus allow me to measure student learning progress in high school.

Linked Admission and Graduation Record Data I match more than two million student admissions (2004-2015) and graduation records (2008-2019). The data include middle school attended by students, middle school GPA, admission and exit exam grades (by subject component) and high school and track the student was assigned to. Student middle schools, assigned using a neighborhood assignment rule, provide a good proxy for student residential location during middle school. In a later section, I study how students migrating across towns affects my results concerning student sorting and differences in outcomes. I restrict my analysis to students with matched admission and graduation records. Notes on matching are available in the online appendix A.2. I complement these data with yearly data on town population, high-school dropout rates, unemployment rates, and county-level wage rates.

Teacher Hiring Data I link the student records data with scraped data containing the universe of new teacher hires in Romania in 2015-2019. These include placement test results, placement outcomes, and individual characteristics of close to 200,000 prospective teachers competing for 40,000 teacher jobs across Romanian high schools. Characteristics include college GPA, experience, education level, and results on other qualification exams used by teachers to achieve higher pay grades.

Prospective teachers in Romania are assigned to vacant high school teaching jobs via an annual standardized subject-specific examination consisting of an oral and a written component. Although the teacher allocation mechanism is slightly more complicated than the student one, high-scoring teachers generally have priority over low-scoring ones in choosing the school where they work. For example, priority is given to teachers who want to return to their hometown and to teachers occupying temporary positions who want to apply for permanent jobs in their current schools and meet some minimum criteria. Salaries are standardized for all teachers, so teacher preferences do not reflect salary considerations.

These data have two limitations. First, I do not observe the stock of teachers in schools, only new hires over a period of five years. Second, I do not observe teacher classroom allocations.

However, incoming cohorts in Romanian high schools are relatively small (approximately 122 students per high school per year). Moreover, high schools typically offer multiple tracks with very different curricula. As such, newly-hired teachers teaching track-specific subjects (such as chemistry or philosophy) are very likely to teach students admitted in those tracks.

School Spending Data I complement the student and teacher data with more than one million school expenditure items scraped from the Romanian Electronic Purchase System (SEAP), where public institutions, including schools, must publicly post their expenditures on goods and services. There are two types of spending items.

Direct spending covers any item or service purchased by the school whose value is under a certain threshold. The threshold changes from year to year, but it was around €30,000 for goods and services and €100,000 for renovations in 2016-2017. Indirect (or auction) spending includes items with costs above the thresholds. These require the organization of a formal bidding process on the SEAP website.

Typically, direct spending includes day-to-day items, like classroom materials, food for the cafeteria, utility bills, etc. Indirect spending includes major renovations to classrooms, gyms, science and computer science labs and contracts with private firms to ensure school security. Spending data does not include teaching and admin staff wages. However, these wages are uniform across all schools and vary only according to seniority, education and rank.

School Incentives Romanian high schools operate in a highly competitive environment, facing substantial pressures to enhance their services, making them well-suited for studying the effects of school choice. School choice in Romania is transparent, universal, and easily accessible to all students, with no local monopolies based on school districts. This allows students to apply to their preferred schools and introduces competitive pressures. School rankings based on admission scores and graduation exam performance are widely known, further contributing to competition.

Failing to attract high-achieving students has severe implications for schools. As schools receive funding per student, a decrease in enrollment directly affects their budgets. Additionally, Romanian high schools compete for a shrinking student population due to low birth rates and em-

Table 1: Summary of Data

	Number of High Schools in Town				
	1	2	3	4-15	16+
Towns (Yearly)	344.2	56.1	26.2	53.6	19.3
High Schools	1.0	2.0	3.0	8.7	47.9
Middle Schools	4.0	4.5	4.9	7.9	49.5
Tracks (per High School)	1.5	1.7	1.8	1.8	1.7
Matched Students	180,191	94,893	82,769	370,795	487,573
Yearly Exit Exam Students (per Town)	46	141	264	577	2102
Yearly Exit Exam Students (per School)	46	76	99	97	96
Yearly Exit Exam Students (per Track)	18	22	24	25	28
Admission Exam Score (Percentile)	49	57	59	61	64
Graduation Exam Score (Percentile)	48	56	57	59	62
Graduation Exam Pass Rate (%)	60	72	73	75	80
Hired Teachers (Total)	5,050	1,680	969	5,339	6,957
Yearly Hired Teachers (per Town)	3.1	6.3	9.1	21.8	133.8
Yearly Hired Teachers (per School)	3.1	3.1	3.0	3.0	4.5
Teacher Score (Percentile)	47	47	49	49	54
Total Town Spending (€1,000/year)	79	154	98	359	2,418
Total School Spending (€1,000/year)	79	92	37	56	91
Total Spending per Student (€/year)	533	950	199	548	1,774
Direct Spending per Student (€/year)	461	681	128	482	250

Summary statistics for high school admission, graduation, teachers and school spending. Student statistics are for students with matched admission and graduation records who do not change schools. School spending can be direct (for day-to-day items) or indirect and requiring an auction (for larger projects, e.g. renovations).

igration. Between 2004 and 2017, the number of students writing the high school admission exam decreased by a staggering 43%, from around 197,000 to under 113,000. This is due to low birth rates and emigration. With fixed classroom sizes and the number of seats determined by demand in previous years, undersubscribed schools may need to downsize and reduce teacher working hours.

The inability to attract strong students also impacts teacher recruitment, increasing the likelihood of temporary teachers with high turnover rates. Given that teacher hiring is competitive, teachers tend to prefer applying to high-admission score schools. Furthermore, school principals, who serve fixed-length terms, have strong incentives to improve their schools along these dimensions.

The competitive pressures on Romanian high schools are exemplified by notable scandals. In response to reports of teachers leaking exam solutions to students, the government installed CCTV cameras in all exam rooms (Borcan et al. 2017). There have also been reports of teachers discouraging students from registering for the graduation exam in low-achieving schools to artificially inflate pass rates.

In summary, Romanian high schools and their administrators face significant competitive pressures to improve school quality. The absence of barriers to school choice, coupled with the consequences of failing to attract high-achieving students, highlights the responsiveness of schools to competition. This context provides an ideal setting for examining the competitive effects of school choice.

Summary Statistics The data are summarized in Table 1. Since I will later use variation in the number of high schools across different locations, the summary statistics are broken down by towns with different numbers of high schools. These statistics account for variation within localities across time. For example, if a new high school opens, a town may switch from having one high school to two high schools.

Generally speaking, the schools in one-high school towns are smaller than those in towns with more high schools. However, class sizes are typically fixed at around 28 students, so class size will not vary with town size (except in heavily undersubscribed schools). At the same time, high school admission, graduation, and teacher placement scores are increasing in the number of high schools in a town, which probably captures socioeconomic differences between rural and urban areas. Schools typically offer the same number of tracks and hire the same number of new teachers regardless of town size, except for towns with more than 16 high schools, where more teachers per school are hired. Lastly, schools spend more per capita in places with many high schools, but this is mainly driven by large contract items, such as renovations. Spending on everyday items is similar across the different types of towns.

3 School Choice, Student Sorting and Achievement Gaps

In this section, I first show that the average Romanian student has strong preferences for high-admission score schools. Secondly, these preferences, combined with competitive admissions processes, result in the segregation of students based on their admission scores across high schools. Moreover, localities with more high schools experience more student sorting. Thirdly, these patterns of student segregation are reflected in inequalities in graduation exam scores: in locations with a larger number of high schools, students with high admission scores tend to perform better, while those with low admission scores tend to perform worse, compared to their counterparts in towns with fewer high schools.

Characterizing Student Preferences The average Romanian student has strong preferences for high-admission score high schools in their proximity, and these preferences lead to the sorting patterns observed at high school admissions. To show this, I begin by reconstructing how students rank schools. While I do not have access to student preference lists filled out by students during the high school application process, I can partially reconstruct these rankings using i) student admission scores, ii) the high school they were admitted to, and iii) the admission score cutoffs of high schools in their choice sets. For example, suppose that a student has a higher admission score than the admission cutoff to both schools A and B. If they are allocated to school A, it means they must have ranked A over B.

I restrict my attention to schools within a 10 km radius of each student's location. I start by geocoding high schools and middle schools in Romania, using middle schools as proxies for student homes. I was able to geolocate the middle schools of over 85% of admitted students and the high schools of over 97% of them. This is likely a good proxy because Romania is relatively densely populated, has many schools, and students are assigned to middle schools using a neighborhood assignment rule. For each student, I thus create a feasible choice set of schools that i) are located within a 10 km radius of the middle school the student attended and ii) have a lower required admission score than the student's admission score.

Using these choice sets and rankings, I characterize student preferences over high schools and tracks in two ways in Table 2. First, I present descriptive statistics of student preferences in the bottom panel of the table. When faced with a maximum of two viable choices within 10 km of their middle school, 68% of students opt for the highest-admission score school. This percentage drops as the number of high schools in their choice set increases, but even in urban areas where students have access to an average of 21.6 high schools, they still choose the highest-admission score school 16% of the time, while choosing a top-3 admission score school 50% of the time. Moreover, when ranking students' attainable high schools from 0 (lowest admission score) to 1 (highest admission score), they generally select schools with scores ranging from an average of 0.68 to 0.80. Thus, Romanian student has strong preferences for the schools that end up attracting high-admission score students.

I then estimate a model of preferences that includes distances to each high school and differences in average admission scores between the various schools in the choice set. This model shows that the average student strongly prefers the highest-admission score school even when controlling for distance and score differences between schools in their choice set. This rules out the descriptive results presented above being driven by students disproportionately living closer to their top-admission score schools. The model is:

$$\text{attend}_i = \beta_0 + \beta_d \text{Distance}_{\text{Top}-\text{Closest}} + \beta_a \text{Admission Score}_{\text{Top}-\text{Closest}} + \varepsilon_i \tag{1}$$

For each student i, the outcome (attend_i) is 1 if they attend the high school with the highest mean

Table 2: Student Preferences over High Schools

	P(Attend Top School) (p.p.)			
Number of Eligible High Schools within 10 km Radius	2	3	4-15	16+
Intercept	72.7^{***}	57.4^{***}	36.6^{***}	20.7^{***}
	(1.5)	(1.5)	(1.3)	(2.8)
$\text{Distance}_{\text{Top}-\text{Closest}}$ (km)	-6.8^{***}	-5.5^{***}	-3.9^{***}	-2.0^{***}
	(0.4)	(0.5)	(0.4)	(0.5)
$\text{Admission Score}_{\text{Top}-\text{Closest}}$ (percentile)	0.13^{*}	0.09	-0.06^{**}	0.04
	(0.07)	(0.06)	(0.03)	(0.03)
Observations	150,318	109,616	506,500	76,201
R^2	0.08	0.04	0.03	0.02
Attend Highest Admission Score School	0.68	0.54	0.29	0.16
Attend Top-2 Admission Score School	1.00	0.85	0.54	0.35
Attend Top-3 Admission Score School	1.00	1.00	0.71	0.50
Average Rank of School Attended (0-low, 1-high)	0.68	0.69	0.71	0.80
Mean Number of High Schools in Choice Set	2.0	3.0	7.9	21.6
Mean Distance to HS in Choice Set (km)	3.0	2.8	2.5	3.5
Mean Distance to HS Attended (km)	2.5	2.4	2.1	2.9

This table characterizes student preferences over high schools. The top panel shows estimates of equation 1. The dependent variable is an indicator that takes value 1 if the student attends a particular school and 0 otherwise. Controls include town and student and year fixed effects. Each column represents different subsamples, defined by the number of eligible high schools within a 10 km radius of a student's middle school ranging from 2 schools in Column (1) to 16 schools and more in Column (4). The bottom panel shows summary statistics for each subsample. Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Standard errors are clustered at the county level.

admission score in their choice set and 0 if they attend any other high school. The controls are i) how much farther the top-scoring high school is to student *i*'s middle school compared to the closest high school ($\text{Distance}_{\text{Top}-\text{Closest}}$), ii) how much higher the top-scoring high school's mean admission score is compared to to the closest high school ($\text{Admission Score}_{\text{Top}-\text{Closest}}$) and iii) a constant.

The top panel of Table 2 presents the model estimates. When the closest high school is also the highest-admission score school (i.e., $\text{Distance}_{\text{Top}-\text{Closest}}$ and $\text{Admission Score}_{\text{Top}-\text{Closest}}$ are both equal 0), 73% of students having two choices prefer to attend this school. Conversely, if the highest-admission score school is not the nearest to the student's middle school, then the likelihood of the student attending it decreases by 6.8 pp for each additional kilometer that separates the high school from the middle school relative to the distance between the middle school and the nearest high school. Meanwhile, differences in admission scores between the top-ranked and closest high schools have little incidence over choices.

The findings remain consistent for students who have a larger number of high schools to choose from. Conditional on the distance and score controls, 58%, 37%, and 21% of students choose the top-admission score high school when faced with 3, 4 to 15, and 16 or more choices, respectively. The estimated effects become smaller as the choice set expands, given that more variables may affect the student's decision, and the top high school faces competition from not only the nearest high school but also numerous other schools. Nonetheless, the evidence suggests that students consistently prefer enrolling in the top-admission score school. In Table A.9, I show similar results for preferences over tracks.

Student Sorting In Figure 1, I show that students' preferences for high-admission score schools, coupled with the serial dictatorship assignment mechanism, lead to pronounced sorting patterns by test scores across high schools.

I first show that immediately prior to high school admissions, there is little sorting by scores across middle schools. The left panel shows the average scores of students' *middle school* peers (measured as admission score percentiles at the *national-cohort* level), split by i) students' own admission score ranks in their town (measured as admission score decile at the *town-cohort* level) and ii) number of high schools in the town.

Students in towns with only one high school and belonging to the lowest admission score decile within their town-cohort attend schools where the average high school admission score is in the 40th percentile. In contrast, their highest-decile students counterparts attend middle schools where the average admission score percentile is 46, a mere 6 percentile difference. In locations with more high schools, average scores are higher, and sorting patterns are more pronounced, but still relatively mild. These relatively mild sorting patterns are a result of middle schools employing a neighborhood assignment rule. Furthermore, geographic sorting based on socioeconomic status, particularly within towns, is highly limited in Romania. This limited sorting can be attributed to the egalitarian housing assignment policies of the communist regime and the persistent low levels of wealth inequality in the country, as explained in more detail in section 4.1.

By contrast, the reshuffling of students occurring as a result of high school admissions leads to a dramatic increase in sorting by scores (middle panel). Indeed, while in one-high school towns there is almost no high-school level sorting (as everyone attends the same high school), these sorting patterns become increasingly pronounced in locations with a greater number of high schools. Put simply, as the number of high schools available for selection increases, low-scoring students tend to attend high schools with progressively lower admission scores, while high-scoring students tend to enroll in high schools with higher admission scores. Thus, there is a significant rise in the degree of sorting based on scores across schools as the number of local high school choices expands. In metropolitan areas, where sorting is most extreme, top-decile students opt for high schools with admission scores ranking in the 91st percentile, whereas the lowest decile of students are relegated to schools where the average entrance score falls in the 24th percentile.

Lastly, in the rightmost panel, it becomes evident that the sorting of students across high school tracks is even more substantial compared to sorting across schools, as students are able to further sort by admission scores across tracks of the same schools, which have different entry cutoffs. Even in towns with only one high school, where students are unable to choose more selective high

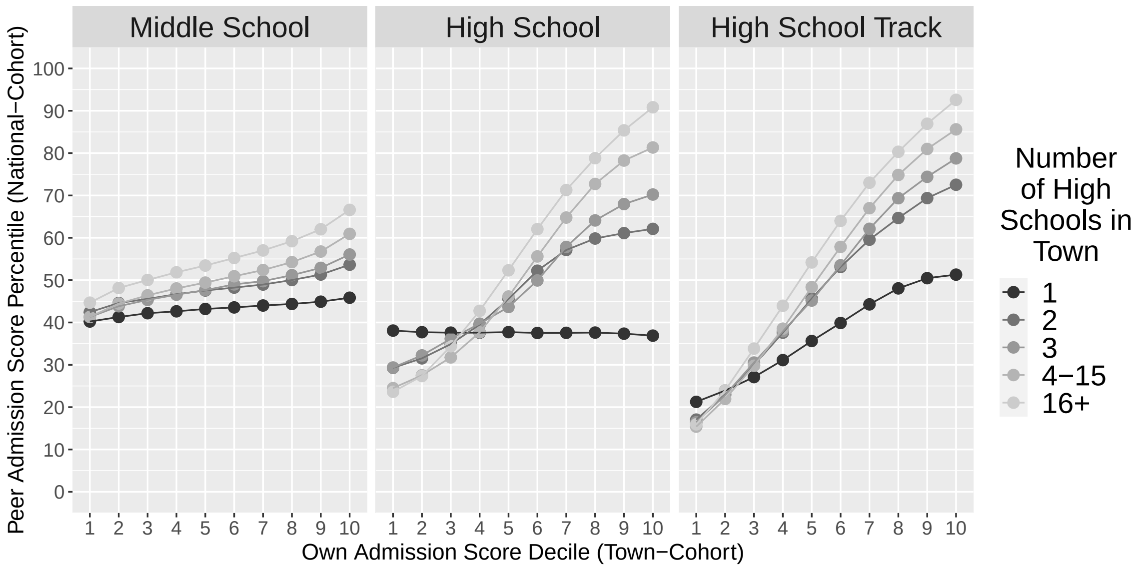

Figure 1: This figure illustrates how students sort by admission scores across schools and tracks. Three panels plot average admission scores of 1) middle school peers, 2) high school peers and 3) high school track peers, by their own admission score decile and number of high schools in the town of high school attendance.

schools, they can partially make up for this by sorting into more selective tracks. However, sorting remains much more pronounced in locations with a greater number of high schools.

Student Achievement Gaps Figure 2 shows that the achievement gaps in graduation scores are consistent with the student sorting patterns across admission scores. First, low-scoring students score higher on the graduation exam when they attend high school in locations with fewer high schools. There is less potential for sorting in these locations, making these students less likely to be relegated to poor-quality high schools. More specifically, students in the lowest admission score decile in one-high school towns score 2.8 percentiles higher on the graduation exam than their counterparts in localities with sixteen or more high schools.

Second, high-scoring students seem to benefit from having more high schools to choose from, as they can enroll in more selective schools, resulting in higher performance on the graduation exam. In towns with sixteen or more high schools, students in the top admission score decile score an average of 2.9 percentiles higher than their counterparts in one-high school towns. The results suggest that the gap in graduation scores widens in areas where students can more easily sort themselves by admission scores across schools. These findings are not driven by within-decile differences in scores across the different types of towns. Figure A.8 in the appendix shows that the same pattern holds when zeroing in on students with the same admission score *percentile*.

In the remainder of the paper, I establish a causal link between the number of school choices students have and increasing inequalities in educational outcomes. Having just found that student sorting patterns comove with the number of high schools, I show that there are plausible reasons why the number of high schools across towns with similar populations varies quasi-exogenously. I then exploit this plausibly exogenous variation in the number of high schools across locations to tease out differences in student sorting and, ultimately, in student learning outcomes.

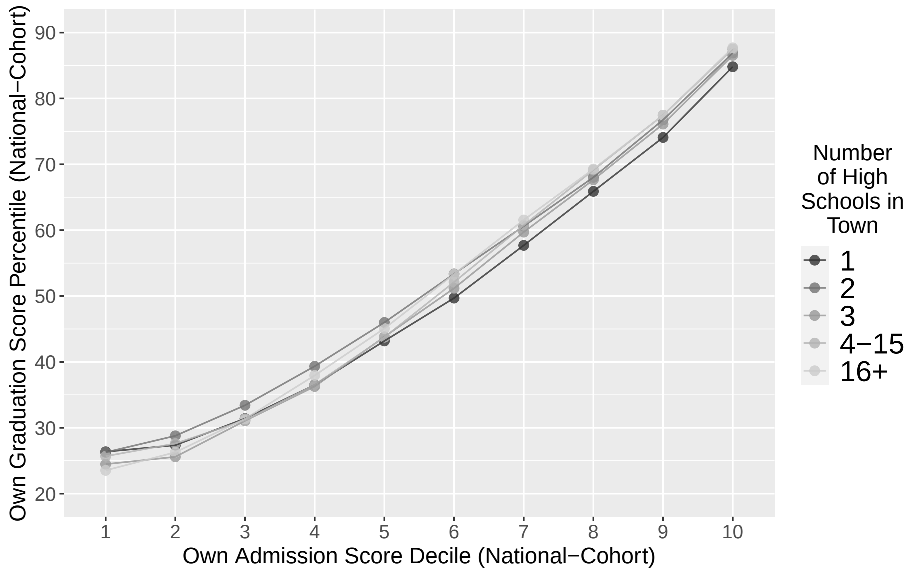

Figure 2: This figure shows how high school graduation exam scores vary with i) the rank of student high school admission scores within their (national) admission cohort and ii) the number of high schools in the town of high school attendance.

4 Model and Identification

4.1 Quasi-Random Variation in the Number of High Schools

Differences in the number of high schools across Romanian towns with similar student populations are quasi-random. Three key points support this assertion. First, the number of schools has remained essentially unchanged since the 1970s, despite significant demographic and economic shifts affecting distinct regions differently, particularly following the downfall of the communist regime in 1989. Secondly, Romanian households exhibit relative geographic immobility and are not markedly sorted by socioeconomic attributes within or across similar towns, a consequence of the communist legacy. Lastly, I show that across towns with similar student populations, but different numbers of high schools, there are no systematic differences in baseline student characteristics. This suggests that I can use the number of high schools as a quasi-random shifter of student sorting patterns without worrying that this instrument may reflect differences in underlying student characteristics.

High school numbers across localities: Despite significant demographic and economic changes that have impacted different Romanian regions differently since the fall of the communist regime in 1989, the number of high schools across Romanian towns has remained largely fixed and is divorced from town socioeconomic conditions in the sample period. In Table A.10 of the online appendix, I show several pieces of evidence in this regard. Roughly 94% of the variation in the number of high schools across towns is explained by the town population measured immediately after the fall of the communist regime in the 1992 census.

Local changes in population, average wages and unemployment between 1992 and the sample period do not predict changes in the number of high schools. Between 1992 and 2019, the last year of the sample, the Romanian population officially decreased by 15%. In addition, The post-communist period in Romania has seen a host of economic changes across towns. Small towns, often centered around one state-owned enterprise, have seen significant changes in their fortunes as most such enterprises have faced layoffs, privatization, or bankruptcy. Nonetheless, the changes in high school numbers across towns are minimal.

Lastly, although I lack the data to confirm this, the number of high schools across towns was probably determined even earlier, in the early 1970s. The Romanian high school expansion essentially ended during this period after the visit of dictator Nicolae Ceausescu to China and North Korea. The number of high schools and university graduates was reduced to ensure a higher supply of manual laborers and reduce political dissent. The decades preceding the fall of the regime were punctuated by severe recessions in the context of the Oil and Energy Crises. They were exacerbated by a policy of austerity that aimed to eliminate foreign government debt. In this climate, expansions in upper-secondary and tertiary education were halted. Meanwhile, in post-1989 Romania, the generally good school coverage, lack of funding, and a rapidly decreasing school-age population made school openings a rarity.

Romanian household (im)mobility: Second, Romanian households are relatively geographically immobile, making it unlikely that households sort across locations to attend better schools. Using the Romanian 2011 census, I find that only 11% of Romanian enrolled high school-aged children lived in a locality that was neither their parents' birthplace nor their place of residence in 1990. A particular institutional and historical context engenders this reality. Before the fall of the communist regime in 1989, individuals entering the labor market were assigned jobs across the country based on their qualifications and centrally-determined labor demand. Also, since the private property of land or homes was essentially abolished, housing was assigned almost randomly based on availability. At the fall of the communist regime, households were allowed to purchase the homes they were occupying and were given access to mortgages. Due to rampant hyperinflation in the early 1990s, most households could then quickly pay off these mortgages. At the same time, relatively weak economic conditions in the post-communist economic recovery made for a relatively thin real estate market. Today, Romania continues to boast the highest homeownership rate in the world: 96.4% in 2018(Eurostat, 2021). Lastly, household wealth inequality in Romania is very low, with household wealth Gini ranked among the lowest in the world(Credit Suisse Group, 2019). As a result of all these factors, while significant urban-rural sorting exists in Romania, Romanian households today continue to be weakly geographically sorted by socioeconomic characteristics across similar towns.

Baseline student characteristics: Lastly, in Table 3, I show that the number of high schools in a town is uncorrelated to student baseline characteristics. Mean admission scores (column 1) and the distribution of these scores (i.e. the proportion of students in different admission score quartiles - columns 2 to 5) are only weakly correlated with the number of high schools once I control for the number of admitted students in a town. Thus, there is little evidence of students sorting by ability into towns with more high schools, ceteris paribus.

This means that once I condition for student populations, I can use the number of high schools (interacted with student scores) as an instrument for student sorting patterns across towns without worrying that the instrument is picking up differences in student baseline scores. In the appendix Figure A.9, I also show that there is a significant amount of exploitable variation in the number of high schools across town-cohorts with similar numbers of admitted students students.

4.2 Instrumental Variable Approach

Having argued that the number of schools across locations with similar student populations is quasi-random, I now use this variation as an instrument for the type of schools students attend. Specifically, in the first stage, I use the number of high schools in a student's town, interacted with their admission score, as a shifter of how selective a school a student can attend, conditional on student and town baseline characteristics. Each high school's selectivity is proxied by its average admission score, which leads to the following first-stage equation:

$$\mu^{a}_{-iht} = \gamma_0 + \gamma_a a_i + \gamma_X X_{ct} + \gamma_{d\times X}(d_i \times X_{ct}) + \delta_d + \delta_n + \delta_{d\times n} + \delta_m + \delta_c + \delta_t + \xi_i \tag{2}$$

Here, μ^{a}_{-iht} is the mean admission score in high school h (excluding student i) in year t, a_i is student i's admission score, which controls for student learning at baseline and the instrument set $\delta_{d\times n}$ is a series of fixed effects capturing the interaction between the number of high schools n in a given town and a student's admission score decile d at the town-cohort level, capturing the student's within-town-cohort admission score rank. This allows the model to reflect differences

Table 3: Admission Scores vs Number of High Schools, Conditional on Town Population

	Admission Score Percentile (0-100)	Percentage of Students in Admission Score Quartile 1	2	3	4
2 HS-Town	1.71	-2.12	-1.05	1.82	1.35
	(1.47)	(1.99)	(0.85)	(1.16)	(1.36)
3 HS-Town	0.69	-0.54	-1.07	0.95	0.66
	(2.03)	(2.68)	(1.05)	(1.52)	(1.88)
4-15 HS-Town	1.35	-0.76	-1.79	0.91	1.64
	(2.66)	(3.55)	(1.13)	(1.94)	(2.37)
16+ HS-Town	2.95	-1.67	-3.62**	0.97	4.33
	(3.54)	(4.46)	(1.58)	(2.16)	(3.55)
N	2,131,565	2,131,565	2,131,565	2,131,565	2,131,565
R^2	0.07	0.05	0.01	0.01	0.04
DV Mean	50	25	25	25	25

This table shows the relationship between student admission scores and the number of high schools in the town students attend high school. Column 1 shows estimates of a regression of admission scores on the number of high schools and columns 2-4 show estimates of regressions of indicator variables for student admission score quartiles on the number of high schools in the town students attend high school. Controls include the number of admitted students (grouped into 50 student bins) and county fixed effects. Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered at the county level.

in student sorting across the admission score distribution and towns with varying numbers of high schools.

Other controls include: X_{ct}, a matrix of town or city (c) characteristics, including unemployment rate, high school dropout rate, average wage at the county level and the size of the high school entering cohort at time t and $d_i \times X_{ct}$, an interaction between student admission score decile and town characteristics that captures differences in student sorting for students with comparable admission scores, but living in towns with different characteristics. Lastly, δ_d, δ_n, δ_m, δ_c and δ_t are admission score decile, number of high schools in town, middle school, town (city), and year fixed effects, respectively and ξ is an error term. Since I separately control for δ_n and δ_d, there are 36 $\delta_{d\times n}$ instruments in total (9 deciles and 4 number of high school categories). Overidentification is addressed in the robustness checks.

It is important to keep in mind that μ^a_{-ih}, the average peer admission score, should be interpreted as a measure of school selectivity rather than a measure of peer effects. Peer effects, teacher characteristics, school resources or the efficiency of school management may all drive the correlation between school selectivity and school value-added.

The second stage describes how sorting into schools with varying degrees of selectivity affects student graduation scores. The second-stage equation is:

$$g_i = \beta_0 + \beta_a a_i + \beta_\mu \hat{\mu}^a_{-iht} + \beta_X X_{ct} + \beta_{d\times X}(d_i \times X_{ct}) + \eta_d + \eta_n + \eta_m + \eta_c + \eta_t + \varepsilon_i \tag{3}$$

Here, g_i is the graduation score grade of student i who entered high school h in year t from middle school m in town or city c and $\hat{\mu}^a_{-ih}$ is the first stage estimate. Lastly, η_d, η_n, η_m, η_c and η_t are admission score decile, number of high schools in town, middle school, town (city) and year fixed effects, respectively and ε is an error term.

5 Results

5.1 School Choice and Sorting

I begin by showing that even after controlling for town and student characteristics, the number of high schools in a location leads to more student sorting by admission scores across schools. This aligns with the observed descriptive sorting pattern illustrated in Figure 1.

Figure 3 (and Table A.11 of the appendix) shows the first stage estimates. I plot mean admission scores in schools attended by students in different categories. These categories are defined by: i) student admission score rank (decile) in their town-cohort and ii) the number of high schools in their town. I adjust these estimates for the number of high schools (δ_n), admission score decile (δ_d), and differences in mean admission scores (a_i) within each category. I then plot all estimates relative to close-to-median students in towns with only one high school who have an admission score in the sixth decile in their town-cohort.

In the first panel of the figure, I show that compared to this reference category, all other students in one high school towns, as well as other close-to-median students in locations with more high schools, attend schools with similar mean admission scores. In contrast, high-scoring students attend progressively more selective schools, the more high schools there are in their town. Conversely, low-admission score students attend progressively less selective schools, the more high schools there are in their town. Towns with sixteen or more high schools thus exhibit substantial sorting based on admission scores, with lowest-decile students being admitted to schools with average admission scores 45 percentiles lower than their top-decile counterparts.

Sorting across tracks (in the right panel of Figure 3) is similar to sorting across high schools. The main difference is that significant across-track sorting occurs even in one high school towns. Top-decile students attend tracks with 14-percentile higher average admission scores than their lowest-decile counterparts. However, this effect is more pronounced in many-high school towns, where this gap stands at 39 percentiles.

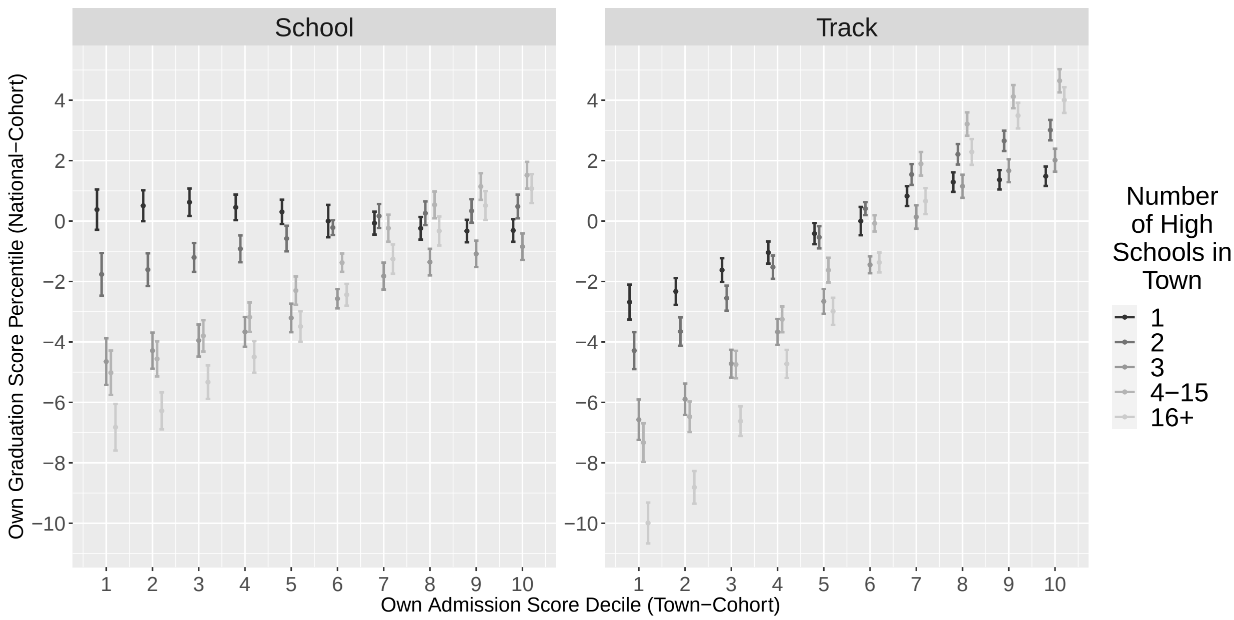

Figure 3: This figure shows a visualization of the first stage estimation (equation 2). It shows how peer admission scores vary by i) own admission score (a_i), ii) within town-cohort admission score decile d, and iii) number of high schools in town of high school attendance n in the model, once I control for town, year and middle school characteristics. Each point represents the mean estimated $\gamma_a a_i + \delta_d + \delta_n + \delta_{d\times n}$ within each admission score decile and town type. The reference category is students attending high school in towns with one high school, whose admission score is in the sixth decile. The left panel shows peer admission scores in the student's high school cohort. The right panel shows peer admission scores in the student's high school track cohort.

Table 4: Effect of Attending a More Selective School (Stage 2 IV)

	Graduation Score Percentile				
	School Level		Track Level		Town Mean
	OLS	IV	OLS	IV	OLS
	(1)	(2)	(3)	(4)	(5)
Admission Score (a_i)	0.54^{***}	0.53^{***}	0.51^{***}	0.46^{***}	0.56^{***}
	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
Peer Admission Score (μ^a_{-i})	0.14^{***}	0.18^{***}	0.21^{***}	0.36^{***}	
	(0.02)	(0.02)	(0.01)	(0.05)	
2 HS-Town	-0.55	-0.56	-0.58	-0.62	-0.52
	(0.82)	(0.82)	(0.75)	(0.75)	(0.84)
3 HS-Town	-2.74^{*}	-2.73^{*}	-2.70^{**}	-2.65^{**}	-3.20^{**}
	(1.39)	(1.39)	(1.33)	(1.31)	(1.39)
4-15 HS-Town	-1.41	-1.39	-1.37	-1.30	-1.39
	(1.55)	(1.54)	(1.50)	(1.48)	(1.56)
16+ HS-Town	-2.51	-2.46	-2.48	-2.35	-2.17
	(1.94)	(1.92)	(1.93)	(1.89)	(2.08)
Observations	1,161,358	1,161,358	1,161,135	1,161,135	1,161,358
Adjusted R^2	0.64	0.64	0.64	0.64	0.64

This table shows how graduation scores are affected by attending a more selective school (1-2) and track (3-4), as per equation 3. Column 5 shows mean graduation scores across towns with different numbers of high schools. Controls include own admission score decile (within town-cohort) fixed effects and interactions of these decile indicators with: the number of students admitted in towns in a given year, town unemployment level and county high school dropout levels and average wage levels. Additional controls include the number of high schools in a town, town fixed effects, year fixed effects, track type fixed effects, and middle school fixed effects. The reference levels are students in the sixth admission score decile and one high school towns. Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Standard errors are clustered at the town level.

5.2 School Choice and Student Scores

Moving forward, I investigate the impact of school choice on student outcomes across the admission score distribution. In Table 4, I show that sorting into more selective schools (or tracks) leads to higher graduation scores, even after controlling for student and town baseline characteristics. An increase of one percentile in the average admission scores of the chosen school (track) leads to a corresponding increase of 0.18 (0.36) percentiles in the graduation score. These results align with the findings of Pop-Eleches and Urquiola (2013), who also observed score improvements among students who were marginally able to attend more selective schools. I rule out weak instruments using an F-test as per Sanderson and Windmeijer (2016).

I also show that there is little evidence of school choice increasing average attainment via between-school competition. Table 4 uses close-to-median students as a reference group. Indeed, conditional on student and town baseline characteristics, graduation scores for students in the sixth decile in their town are largely invariant with the number of schools in their town, as seen by fixed effects estimates for the number of high schools. In fact, there is evidence that close-to-median students in towns with many schools underperform their counterparts in one-high school towns, though these effects are poorly estimated. Even without taking into account high school characteristics that may absorb score differences between the different types of towns (column 5), there is no evidence students at the median are performing better in towns with more high schools. While these findings could in theory be influenced by attrition or self-selection into graduation that can occur at different rates in different towns, I find that this is not the case (see section 5.4).

Subsequently, I present Figure 4 to illustrate how the effects of school choice act across the admission score distribution, ultimately leading to a widening gap in academic achievement. The graph plots the effects on graduation scores of student sorting. Specifically, I plot the causal effect of attending a more selective school (β) multiplied by school selectivity as predicted by the number of high schools in students' towns ($\hat{\mu}$), adjusted for fixed effects for the numbers of schools (η_n). These effects are plotted separately for different student subgroups, defined by their admission score ranks in their town-cohort and the number of high schools in their town. The estimates are plotted relative to near-median (sixth decile) admission score students in towns with one high school.

Firstly, the sorting across schools (or tracks) contributes to a significant widening of the graduation score gap between high- and low-admission score students in towns with more than sixteen high schools. In such locations, the graduation score gap is amplified by nearly 8 (14) percentiles. Compared to median students, low-scoring students experience learning setbacks that are not entirely offset by the learning gains of high-admission score students. Indeed, students in areas with a higher number of high schools in the former group have the opportunity to select exceptionally selective schools, allowing them to outperform similar students in one-high school towns by 1 (3) percentiles. Meanwhile, in many-high school towns, low-achieving students are confined to less desirable schools (or tracks), resulting in a performance difference of 7 (7) percentiles compared to their peers in one-high school towns, who are not subject to this type of school segregation.

Secondly, after taking school and track sorting into account, the additional competition between high schools does not seem to have a positive impact on student graduation scores. If anything, the median-student achievement is slightly lower (1-2 percentiles) in towns with many high schools than in one-high school towns. Moreover, the negative impact of sorting on low-achieving students seems to drown out any positive effects at the bottom of the distribution.

5.3 Other Outcomes

I explore several other outcomes beyond graduation scores in Table A.12 of the online appendix. First, I examine the probability of matching a student's admission and graduation records. The match rate between admission and graduation records is increasing in student admission scores. Inability to match a student's admission records to their graduation records could occur if the student drops out, does not register for the graduation exam or changes schools and I am unable to match them. I then examine how school changes vary with student admission scores and the average admission score in their school and find that students in lower-admission score schools are more likely to change schools. Finally, I study the probability of passing the exit exam. Higher-

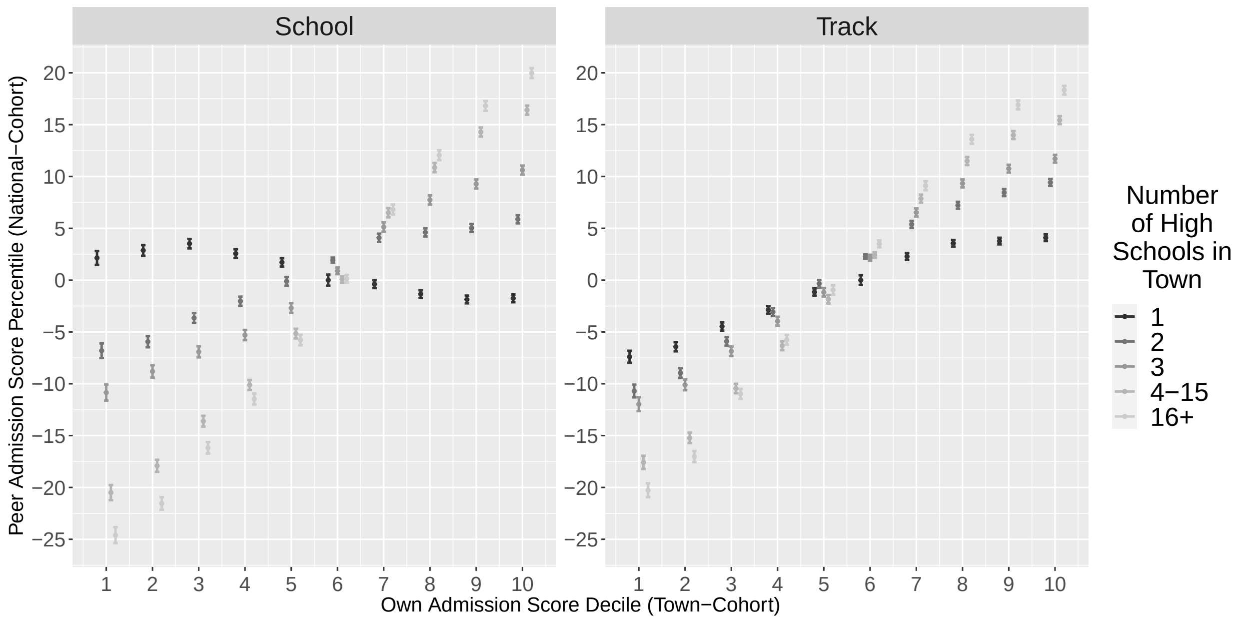

Figure 4: This figure shows a visualization of the effect of sorting on graduation scores by i) own admission score (a_i), ii) within town-cohort admission score decile d, and iii) number of high schools in town of high school attendance n. Each point represents the average value of predicted peer admission scores (shown in Figure 3) multiplied by β_μ (from equation 3), and adjusted for the number of high school fixed effects (η_n). The reference level is students attending high school in towns with one high school, whose admission scores are in the sixth decile in their town-cohort. The left panel shows the effect of sorting across high schools. The right panel shows the effect of sorting across high schools and tracks.

ability students are more likely to pass the exit exam, but attending a better school does not affect the exam pass rate, conditional on student and town characteristics.

5.4 Robustness Checks

School Openings: Triple Difference I use the opening of new schools in small towns to confirm that the number of schools greatly affects sorting patterns. I study these openings using a triple difference framework. I compare graduation scores of i) high- and low-admission score students across ii) treatment towns (where a new high school opens) and control towns (similar in size but without a new high school) iii) before and after the high school opening. In Section A.3.1 of the online appendix, this model is estimated, accounting for various town and county characteristics (such as unemployment, high school dropout rates, and wage levels), as well as numerous fixed effects (year, locality, within-town admission score ranking, and their interactions).

I find that high school openings exacerbate the sorting of students by admission scores across schools. Following the opening of a new high school, students with admission scores above the median tend to sort into high schools attended by peers with admission scores 5 to 6 percentiles higher. Moreover, the average achievement of students, considering their admission scores, does not improve. Lastly, the graduation score gaps between high- and low-admission score students widen in towns where a new high school opens.

Propensity Score Matching I next use a nearest neighbor matching algorithm (Stuart et al., 2011) to validate the main results. I pair control students with treated students from the same admission cohort, having similar admission scores and living in towns with similar numbers of high school-bound students, but where there is one extra high school. I then compare sorting patterns by admission scores across the treatment and control groups.

Table A.18 of the online appendix contains the results of this exercise. Compared to control students in towns with n high schools, treatment students in towns with n+1 high schools display more pronounced sorting patterns by admission scores across high schools, larger graduation score gaps between high- and low-admission score students, and have similar graduation exam scores on average. These results are confirmed by zeroing in on two- and one-high school towns. Moreover, ceteris paribus, the average graduation exam scores of treated students are less than one percentile higher than those of control students. Yet again, there is very little evidence of school choice being a “rising tide”.

Migration Migration across towns may bias the results. In particular, there is potentially selection on unobservables. For example, more motivated students may disproportionately apply to high schools in locations with higher student sorting, where they can attend a more selective school. To the extent that these unobservables (rather than attending a more selective school) affect performance, the above estimates are upward biased.

I address this issue in the online appendix section A.3.3. Instead of using towns as separate educational markets, I define these markets endogenously. For example, if sufficiently many students from other middle schools in town A enroll in high schools in both towns A and B, then I consider both towns A and B to be part of the same market. Using this approach, I reestimate the first and second stages of the instrumental variable model at the school and town levels. I additionally exclude cross-market migrants from the analysis. The second stage results are presented in Tables A.19 to A.22 of the online appendix. The results are very robust to these different specifications.

Sample Selection Next, I address sample selection. The main results of the paper use matched admission and graduation records. Students who appear in the admission records may drop out from the sample for several reasons, including changing schools, deferring the graduation exam, not registering for the graduation exam, or dropping out of high school. I replicate the main results using the methodology from Ainsworth et al. (2023), who include all students with admission records, but no matching graduation records, by coding their graduation scores as zeros.

The results - presented in Table A.23 of the online appendix - confirm our main findings. Sorting into a more selective school increases graduation scores. Median scores continue to show a decreasing trend in the number of high schools, so there is no evidence of competition between schools yielding score gains. Lastly, mean graduation scores are decreasing in the number of high schools, which is likely due to a higher proportion of urban students enrolling in high school only to drop out later.

First Stage Overidentification Lastly, I address a first stage overidentification concern. Indeed, there are 36 instruments (4 high schools/town dummies and 9 decile dummies) used in the first stage for only one endogenous regressor. The Sargan-Hansen overidentification test yields a p-value close to 0. Although this is not necessarily a concern, given the large sample size and precise estimates, I reestimate the IV model using only 1 instrument: an indicator of above/below median admission score interacted with an indicator of a town with more than 3 high schools. I present the results in section A.3.5 of the online appendix. This alternative specification yields very similar results to those shown above while producing an insignificant Sargan-Hansen statistic.

6 Mechanisms

Below, I provide descriptive evidence regarding the relative importance of peer effects, teacher ability, and school spending in explaining the effects of sorting on student scores.

6.1 Peer Effects

To the extent that school choice is responsible for increasing student sorting, peer effects can significantly increase educational outcome inequalities. High-achieving students who are admitted to selective schools may benefit from positive peer effects and succeed academically. I identify peer effects using variation in admission scores of students admitted to the same track across different years. Students admitted to the same high school track in different years are exposed to peers with slightly different abilities.

As before, I consider that peer admission scores indicate how selective a school is. This measure of selectivity is endogenous because it results from student choices and is plausibly correlated with student unobservable characteristics. For example, families attaching a higher value to education could be more likely to send their children to more selective schools, ceteris paribus. I instrument this regressor using the interaction between the number of high schools and students' admission score rank. Whereas in the baseline mode, this measure of selectivity was measured at

Table 5: Peer Effects

	Graduation Score Percentile	
	OLS	IV
	(1)	(2)
Admission Score	0.46***	0.39***
	(0.01)	(0.03)
Change in Mean Track Admission Score ($\mu^a_{-ihst} - \mu^a_{-ihs}$)	0.06***	0.12***
	(0.01)	(0.02)
Mean Track Admission Score (μ^a_{-ihs})	0.23***	0.40***
	(0.01)	(0.05)
Observations	507,692	507,692
Adjusted R^2	0.65	0.65

This table shows how cohort-to-cohort changes in peer admission scores in the same high school track correlate with graduation scores (as per equation 4). Controls include own admission score decile (within town-cohort) fixed effects and interactions of these decile indicators with: the number of students admitted in towns in a given year, town unemployment level and county high school dropout levels and average wage levels. Additional controls include the number of high schools in the town (grouped into 5 bins), town fixed effects, year fixed effects, track type fixed effects and middle school fixed effects. Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Standard errors are clustered at the town level.

the high school (or track)-cohort level (μ^a_{-ihst}), I now use μ^a_{-ihs}, the peer average across all cohorts admitted to the same track over the years.

This allows me to use the variation in peer admission scores different cohorts (t) of the same high school (h) track (s), $\mu_{-ihst} - \mu_{-ihs}$, to identify peer effects. A high value indicates that the year t cohort was particularly strong. I assume that this year-to-year variation in incoming cohort scores within a track results from random deviations in the scores of the small sample of students who apply to a high school track in a particular year. If this is true, then this difference is exogenous and no additional source of bias is introduced - despite the endogenous regressor appearing in the difference. This would not be true, if, for example students have inside information about changes in a school's quality over time which would also impact their enrollment decisions.

The model I estimate is similar to the main specification in equation 3:

$$g_i = \theta_0 + \theta_a a_i + \theta_\mu \hat{\mu}^a_{-ihs} + \theta_{\mu_t}(\mu^a_{-ihst} - \mu^a_{-ihs}) + \theta_X X_{ct} + \theta_{d\times X}(d_i \times X_{ct}) + \eta_d + \eta_n + \eta_m + \eta_c + \eta_t + \varepsilon_i$$

The parameter capturing the peer effects is θ_{μ_t}. I assume that these this relative cohort quality directly impacts a student's graduation score performance through peers, allowing me to interpret the coefficient θ_{μ_t} as a peer effect. I restrict my attention to tracks with fewer than 25 students, where all admitted students are guaranteed to be placed in the same classroom to circumvent the issues of class allocation in multiple-class tracks.

Table 5 shows that, conditional on town, school, track, and individual controls, being admitted to a given track in a year when peer admission scores are 1 percentile higher is associated with a 0.12 percentile increase in student graduation scores. The results suggest that the magnitude of peer effects is roughly one-third of the size of the effect of sorting into a better track.

6.2 School Resources: Teacher Ability and School Spending

Access to schools with better resources is a significant factor through which school choice-driven student sorting can influence their outcomes. I examine two types of resources: teacher ability and school spending. I focus on the subsample of high schools reporting full-time teacher hirings (only available for 2015-2019) and reporting positive spending. This subsample contains more urban,

Table 6: Access to School Resources

	Teacher Placement Score (Percentile)	School Spending (Euro per Admitted Student per Year)
Own Admission Score	-0.02	-2.63
	(0.04)	(2.09)
Peer Admission Score ($\hat{\mu}^{a}_{-i}$)	0.19***	-4.70**
	(0.06)	(1.93)
2 HS-Town	1.37	45.63
	(1.30)	(559.16)
3 HS-Town	3.54	-54.47
	(3.59)	(553.62)
4-15 HS-Town	4.41	121.66
	(4.20)	(555.32)
16+ HS-Town	17.23***	249.78
	(5.55)	(564.11)
N	404,258	273,353
R^2	0.51	0.51

This table shows how access to school resources varies with one's own admission score, peer admission score and town characteristics. Controls include within town-cohort own admission score decile fixed effects and interactions of these decile indicators with: the number of students admitted in towns in a given year, town unemployment level and county high school dropout levels and average wage levels (all computed as averages of 2010-2019 yearly data). Additional controls include the number of high schools in town (grouped into 5 bins), town fixed effects, year fixed effects, track type fixed effects, and middle school fixed effects. The mean peer admission score is instrumented by the interaction between the student admission score decile and the number of high schools in their town. Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Standard errors are clustered at the town level.

in-demand schools and more recent years than the full sample.

I match subject-specific performance on the high school graduation exam to subject-specific teacher ability. To do this, for each subject and high school, I calculate the weighted average of newly hired teachers' placement scores, where the weights are based on the number of teaching hours specified in each teacher's contract. Teacher-classroom matches are not observed. However, the relatively small number of admitted students in each cohort (122 per high school on average) and the fact that there is significant variation in subjects taught across tracks within the same high school allows for a relatively granular teacher-student match by subject.

I focus specifically on teacher placement test scores as an indicator of teacher ability. In Table A.13 of the appendix, I demonstrate that this measure has the strongest correlation with student achievement, conditional on student admission scores and school and town characteristics. Conversely, more traditional measures such as experience, college GPA, years of schooling, and teacher rank (determined by passing rank exams and associated with a pay increase) do not exhibit a strong correlation with student achievement. School spending is measured as the total expenditure in Euros per admitted student per year.

Using the instrumental variable approach outlined in equation 3, with teacher ability and spending as dependent variables, I find that high-admission score schools tend to hire teachers with higher placement test scores while also spending slightly less per student on average. The results are summarized in Table 6. Attending a school with a 1 percentile higher mean admission score is associated with being taught by teachers who score 0.19 percentiles higher on teacher placement tests. Furthermore, teachers placed in towns with sixteen or more high schools have an average score that is 17 percentiles higher than teachers placed in one-high school towns. This suggests that teachers, who compete for job positions based on their placement scores, prefer teaching in high-achieving schools in larger cities. Additionally, I show that attending a school with a 1 percentile higher mean admission score is linked to attending a school that spends 4.7 Euros less per student per year. There are three plausible explanations for this finding: low-admission score schools may have a lower (unobserved) infrastructure stock and require additional expenditure to catch up, high-admission score schools tend to be larger and more urban on average, benefiting from economies of scale and require less spending per student, and lastly, the government may be willing to allocate more funds to schools with low-achieving students from disadvantaged backgrounds in order to narrow the learning gap between them and more privileged students. However, disentangling these explanations falls outside the scope of this paper.

In Table 7, I provide evidence that access to high-scoring teachers has a positive impact on student graduation scores, independent of student admission scores and town and school characteristics. Conversely, there is no correlation between spending and student achievement. Note that I use i) subject-specific scores at baseline and at graduation and ii) a selected sample of 2015-2019 schools reporting teacher hires and spending, which contains more urban, high-achieving schools. This means that these estimates, and those in the following sections, are not directly comparable to the main results presented in Table 4.

I find that a 1 percentile higher teacher placement score is associated with a 0.08 percentile higher score in the subject taught by the teacher, even after accounting for the school's mean admission score, student's baseline scores on the admission exam, and other town and school characteristics (column 1). However, I find no relationship between school spending and student achievement (column 2).

In Table A.14 of the appendix, I delve further into the relationship between teacher and student scores. I observe that teachers with higher placement scores disproportionately benefit students with strong numeracy skills. Additionally, teacher placement scores are more indicative of student performance in math-intensive fields such as Computer Science, Economics, and Physics.

6.3 Decomposition

I assess the importance of teacher ability, peer effects, and school spending in explaining the effect of school choice-driven student sorting on academic performance. Controlling for these factors reduces the estimated effect of sorting into a more selective school on student scores. I employ a Gelbach decomposition (Gelbach, 2016) to quantify the importance of each of the three channels.

Table 7: Student Graduation Scores vs School Resources

	Graduation Score Percentile	
	(1)	(2)
Own Admission Score		0.702***
		(0.024)
Own Admission Score (Romanian)	0.354***	
	(0.017)	
Own Admission Score (Math)	0.344***	
	(0.019)	
Teacher Placement Score	0.083***	
	(0.011)	
Per Student Spending (Euro, thousands)		−0.059
		(0.122)
Peer Admission Score ($\hat{\mu}^a_{-i}$)	0.101	0.053
	(0.063)	(0.043)
Observations	436,969	266,918
R^2	0.51	0.69

This table shows the correlation between school resources (teachers - column 1 - and school spending - column 2) and student scores on the graduation exam. Controls include track type, year fixed effects, town fixed effects, and middle school fixed effects. Additional controls for (1) include teacher college GPA, experience, rank and years of schooling. The sample includes high schools who hired teachers between 2015 and 2019 (1) and schools providing spending information between 2008 and 2019 (2). The mean peer admission score is instrumented by the interaction between the student admission score decile and the number of high schools in their town. Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered at the town level.

This mediation approach consists of sequentially adding different controls to a baseline specification and quantifying how the coefficient of interest changes as these controls are included. It has the added advantage of providing results that are invariant to the order in which the mediators are added to the model. I first estimate a version of the two-stage least squares model in equation 3:

$$g_i = \beta_0 + \beta_a a_i + \beta_\mu \hat{\mu}^a_{-ih} + \beta_X X_{ct} + \beta_{d\times X}(d_i \times X_{ct}) + \eta_d + \eta_n + \eta_m + \eta_c + \eta_t + \xi_i \tag{4}$$

There are several differences between this and the main specifications. First, the choice of proxy for school quality (μ) is different. The main specification uses the mean admission score within a given high school (h) track (s) cohort (t) μ^a_{-iht} as a proxy for school selectivity. The specification estimated in this section uses the average admission score across all sample years for each high school track (μ^a_{-ih}). In this way, I can estimate the effect of attending a more selective school using μ^a_{-ih}, while using the across-year variation in admission scores to estimate peer effects. Second, the sample used in this section includes only 2015-2019 schools with reported spending and teacher hires. And third, I use subject-specific admission and graduation scores, rather than average scores, to better capture the effect of teachers on scores in the subjects they teach. This also means that I can only include student-subject pairs for high schools hiring teachers for that given subject in 2015-2019. For these reasons, the results presented in this section are not directly comparable to the main results.

To this model, I successively add controls for the three mediators: teacher ability (proxied by teacher placement scores), peer effects (proxied by the within-high school variation in admission scores across different cohorts) and school spending. The results, shown in Table 8, indicate that adding peer admission scores, particularly teacher placement scores, largely accounts for the causal effects of attending a better school on student outcomes. In contrast, differences in school spending have minimal explanatory power. The baseline model shows that attending a school with a 1 percentile higher average admission score improves graduation exam scores by 0.055 percentiles. Peer effects and teacher ability reduce this estimate to 0.033 and 0.040 percentiles, respectively. Including both teacher ability and peer admission scores further decreases the estimate to 0.015 percentiles. In Table A.15 of the appendix, I show similar results using track-level variation in

Table 8: Decomposition of Sorting Effects

	Graduation Exam Percentile (Subject-Specific)					
	Baseline	Teacher	Peer	Spending	T+P	T+P+S
Admission Score (Rom.)	0.156***	0.157***	0.145***	0.155***	0.146***	0.146***
	(0.011)	(0.011)	(0.010)	(0.011)	(0.010)	(0.010)
Admission Score (Math)	0.071***	0.074***	0.059***	0.071***	0.062***	0.062***
	(0.013)	(0.013)	(0.012)	(0.013)	(0.012)	(0.012)
Peer Admission Score	0.055**	0.033	0.040	0.052*	0.018	0.015
	(0.028)	(0.028)	(0.028)	(0.028)	(0.029)	(0.029)
Teacher Placement Score		0.096***			0.097***	0.096***
		(0.011)			(0.011)	(0.011)
Δ Peer Admission Score			0.177***		0.177***	0.182***
			(0.035)		(0.035)	(0.035)
School Spending				−1.479		−1.812
				(1.832)		(1.674)
N	309,942	309,942	309,942	309,942	309,942	309,942
R^2	0.51	0.51	0.51	0.51	0.51	0.51

This table shows the relationship between student scores at graduation and school quality with the following additional controls: teacher placement scores, peer effects (proxied by within-high school track changes in admission scores across cohorts) and school spending (in Euro per admitted student per year). Peer admission scores are instrumented by the interaction between the student admission score decile and the number of high schools in their town. Controls include within town-cohort own admission score decile and interactions of these decile dummies with the number of students admitted in towns in a given year, town unemployment levels, county high school dropout levels and average wage levels, the number of high schools in town, town fixed effects, year fixed effects, and middle school fixed effects. Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Standard errors are clustered at the town level.

peer quality. Lastly, the decomposition exercise reveals that, together, the three channels explain 72% of the total effect. Specifically, teacher ability accounts for 41% of the overall impact, while peer effects and school spending explain 27% and 5% of the effect, respectively.

7 Discussion and Conclusion

This paper investigates the effects of school choice on student sorting and educational outcomes, using administrative data from Romanian high schools, generating two main findings. First, school choice coupled with competitive admissions can lead to severe student sorting on test scores. This type of sorting is much higher than the one generated by neighborhood assignment rules.

Second, this sorting increases the test score gaps between high- and low-ability students without affecting mean achievement. Therefore, school choice is not a "tide that lifts all boats". Instead, it mainly exacerbates inequalities in access to good schools. In large cities, where students have many schools to choose from, school choice-induced sorting is severe. It widens the inequalities between high- and low-admission score students by roughly 8 percentiles. When considering sorting across different high school tracks, this figure stands at 14 percentiles. The two main channels underlying the widening of these score gaps are peer effects and access to high-ability teachers. When high-admission score students are able to sort into more selective schools, they are exposed to better teachers and other high-scoring pupils, which helps their academic performance.

In conclusion, combining school choice with exam-based admissions, the Romanian high school system serves as a cautionary tale. Even though schools face strong incentives to improve their quality, the primary outcome of school choice in Romania is very pronounced sorting by test scores across schools, which leads to polarization in outcomes. This extreme sorting results from the universal and centralized serial dictatorship rule that favors high-scoring students. While it is true that in other contexts and by designing more egalitarian school assignment rules, school choice will not necessarily lead to the same levels of sorting, the Romanian setup serves as a reminder that unless policymakers take concrete steps to ensure equitable access to good-quality schools, school choice could become a driver of inequality.

For example, in large US metropolitan areas, where wealth inequality and geographic segregation by socioeconomic status are much more pronounced, and travel times longer than in Romanian cities, school choice may fail to deliver on its promises to increase school quality, instead entrenching segregation. A mounting body of evidence shows that even school choice systems designed with alleviating inequality in mind can lead to as much segregation and widening test score gaps across schools as neighborhood assignment rules (Hastings et al. 2009, Hastings et al. 2009, Laverde 2021 and Park and Hahm 2022).

Thus, the potential benefits of school choice policies must be carefully weighed against their unintended effects, which are often difficult to anticipate ex-ante, only rearing their heads in general equilibrium. Policymakers must take into consideration a slew of local realities, ranging from the geographic sorting of households to levels of inequality, racial composition, and transportation times and costs, to ensure that school choice does not cause more harm than good.

References

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

Robert Ainsworth, Rajeev Dehejia, Cristian Pop-Eleches, and Miguel Urquiola. Why do households leave school value added on the table? the roles of information and preferences. *American Economic Review*, 113(4):1049–1082, 2023.

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

Levon Barseghyan, Damon Clark, and Stephen Coate. Peer preferences, school competition, and the effects of public school choice. *American Economic Journal: Economic Policy*, 11(4):124–58, 2019.

Oana Borcan, Mikael Lindahl, and Andreea Mitrut. Fighting corruption in education: What works and who benefits? *American Economic Journal: Economic Policy*, 9(1):180–209, 2017.

Christopher Campos and Caitlin Kearns. The impact of neighborhood school choice: Evidence from Los Angeles' Zones of Choice. Working paper, March 2022.

Credit Suisse Group. Global wealth databook 2019. Technical report, Credit Suisse, 2019.

David J Deming, Justine S Hastings, Thomas J Kane, and Douglas O Staiger. School choice, school quality, and postsecondary attainment. *American Economic Review*, 104(3):991–1013, 2014.

Esther Duflo, Pascaline Dupas, and Michael Kremer. Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review*, 101(5):1739–74, 2011.

Eurostat. Distribution of population by tenure status, type of household and income group - EU-SILC survey. Technical report, Eurostat, 2021.

Jonah B Gelbach. When do covariates matter? And which ones, and how much? *Journal of Labor Economics*, 34(2):509–543, 2016.

Justine Hastings, Thomas J Kane, and Douglas O Staiger. Heterogeneous preferences and the efficacy of public school choice. Working Paper 2145, National Bureau of Economic Research, 2009.

Caroline M Hoxby. Does competition among public schools benefit students and taxpayers? *American Economic Review*, 90(5):1209–1238, 2000.

Chang-Tai Hsieh and Miguel Urquiola. When schools compete, how do they compete? An assessment of Chile's nationwide school voucher program. Working Paper 10008, National Bureau of Economic Research, 2003.

Mariana Laverde. Unequal assignments to public schools and the limits of schools choice. Working paper, September 2021.

Cecilia Machado and Christiane Szerman. Centralized college admissions and student composition. *Economics of Education Review*, 85:102184, 2021.

W Bentley MacLeod and Miguel Urquiola. Reputation and school competition. *American Economic Review*, 105(11):3471–88, 2015.

W Bentley MacLeod and Miguel Urquiola. Is education consumption or investment? Implications for school competition. *Annual Review of Economics*, 11:563–589, 2019.

Minseon Park and Dong Woo Hahm. Location choice, commuting, and school choice. Working paper, November 2022.

Cristian Pop-Eleches and Miguel Urquiola. Going to a better school: Effects and behavioral responses. *American Economic Review*, 103(4):1289–1324, 2013.

Jesse Rothstein. Good principals or good peers? Parental valuation of school characteristics, Tiebout equilibrium, and the incentive effects of competition among jurisdictions. *American Economic Review*, 96(4):1333–1350, 2006.

Eleanor Sanderson and Frank Windmeijer. A weak instrument F-test in linear IV models with multiple endogenous variables. *Journal of Econometrics*, 190(2):212–221, 2016.

Elizabeth A Stuart, Gary King, Kosuke Imai, and Daniel Ho. Matchit: nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 2011.

A Online Appendix for "School Choice, Student Sorting and Academic Performance"

A.1 Additional Tables and Figures

A.1.1 Graduation Score Gaps

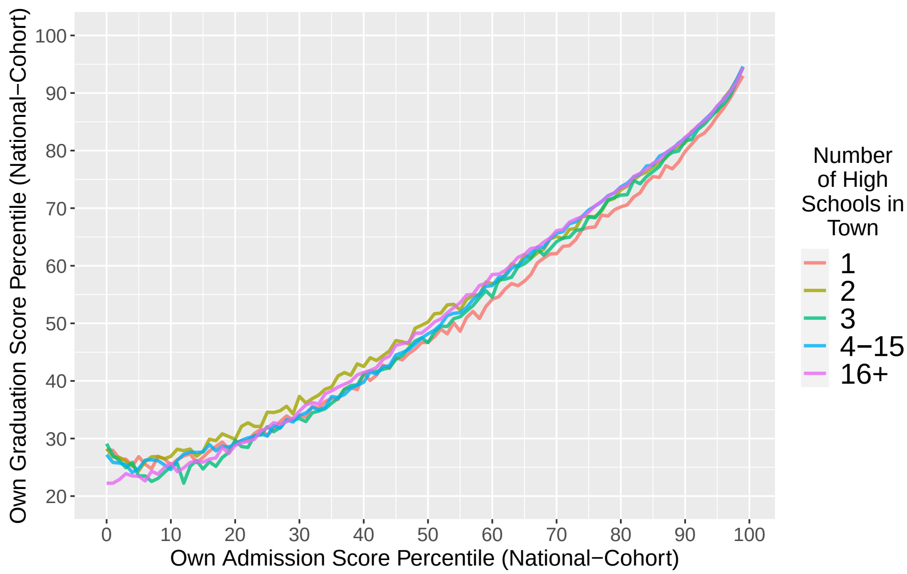

Figure A.8: This figure shows how high school graduation exam scores vary with i) the percentile rank of student high school admission scores within their (national) admission cohort and ii) the number of high schools in the town of high school attendance.

A.1.2 Student Preferences over Tracks

Table A.9: Student Preferences over Tracks

	P(Attend Top Track) (p.p.)			
Number of Eligible Tracks within 10 km Radius	2	3	4-15	16+
Intercept	64.4^{***}	53.1^{***}	39.0^{***}	23.7^{***}
	(1.1)	(1.0)	(1.0)	(1.0)
$\text{Distance}_{Top-Closest}$ (km)	-6.5^{***}	-4.6^{***}	-3.6^{***}	-2.3^{***}
	(0.3)	(0.3)	(0.2)	(0.2)
$\text{Admission Score}_{Top-Closest}$ (percentile)	-0.16	-0.29^{***}	-0.13^{***}	-0.03^{*}
	(0.14)	(0.09)	(0.04)	(0.02)
Observations	50,085	52,714	421,038	481,744
R^2	0.04	0.03	0.02	0.01
Attend Highest Admission Score Track	0.63	0.51	0.34	0.18
Attend Top-2 Admission Score Track	0.67	0.54	0.37	0.23
Attend Top-3 Admission Score Track	0.83	0.70	0.52	0.34
Average Rank of Track Attended (0-low, 1-high)	0.63	0.67	0.76	0.86
Mean Number of Tracks in Choice Set	2.0	3.0	8.8	34.5
Mean Distance to Track in Choice Set (km)	2.5	2.7	2.8	2.6
Mean Distance to Track Attended (km)	2.5	2.6	2.5	2.2

This table characterizes student preferences over high school tracks. The top panel shows estimates of equation 1. The dependent variable is an indicator taking a value of 1 if the student attends a particular school, and 0 otherwise. Each column represents different subsamples, defined by the number of eligible tracks within a 10 km radius of a student's middle school. The bottom panel shows summary statistics for each subsample. Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered at the county level.

A.1.3 Student Population vs Number of Schools

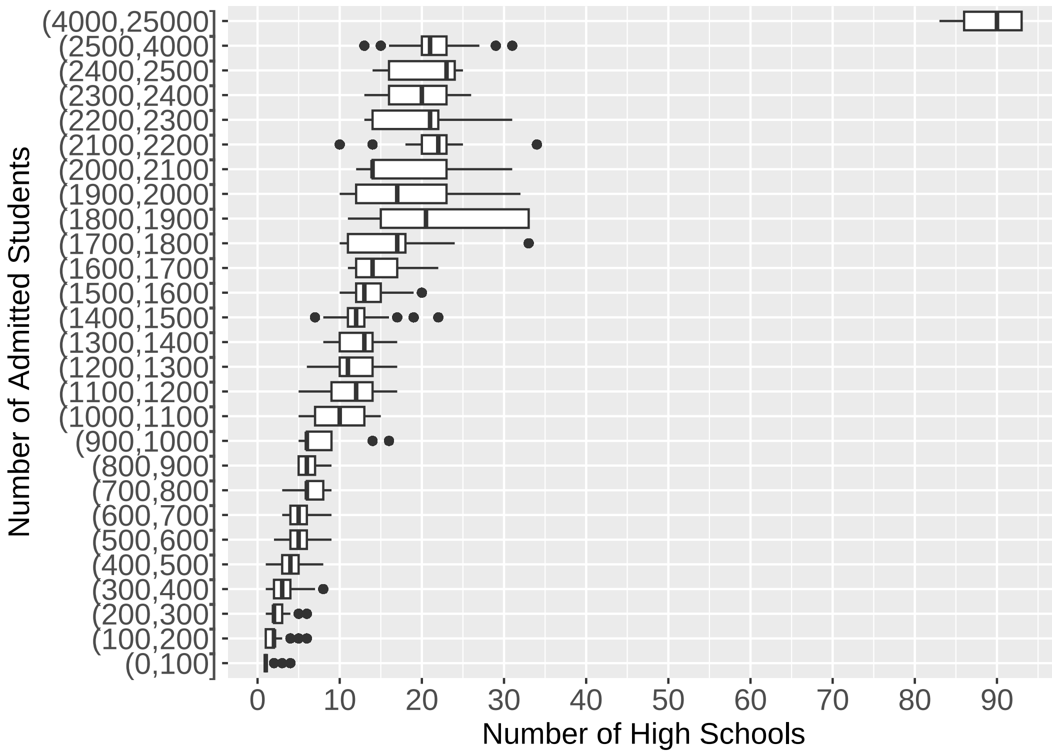

Figure A.9: This figure shows how the number of high schools in a given town-year varies with town-cohort size.

A.1.4 Determinants of Number of Schools

Table A.10: Determinants of Number of Schools Across Towns

	Change in Number of Schools (2008-2019)		2019 Number of Schools		
	(1)	(2)	(3)	(4)	(5)
Population (2008, 1,000s)		-0.008^{***}			
		(0.002)			
Change in Population (2008-2019, 1,000s)	0.008	-0.005			
	(0.033)	(0.016)			
Population (1992, 1,000s)			0.161^{***}	0.165^{***}	-0.042^{**}
			(0.012)	(0.011)	(0.017)
Change in Population (1992-2019, 1,000s)				0.108	-0.013
				(0.066)	(0.111)
Observations	417	417	521	521	513
Adjusted R^2	0.04	0.33	0.94	0.94	0.91
Admitted Students (Town-Cohort)	No	No	No	No	Yes
High School Dropout Rate (County)	No	No	No	No	Yes
Wages (County, RON)	No	No	No	No	Yes
Town Unemployment Rate	No	No	No	No	Yes

This table shows the relationship between different town characteristics and the number of high schools. Column 1 shows estimates from a regression of the change in number of high schools across towns and the change in population, between 2008 and 2019. Column 2 shows the correlation between the number of schools in towns in 2019 and the town population in 1992. Column 3 shows estimates of a regression of the number of high schools in towns in 2019 and controls including 1992 population, change in population between 1992 and 2019, admitted students, unemployment rates and county-level wages and high school drop out rates. The strongest determinant of the number of schools remains the 1992 town population. Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Standard errors are clustered at the town level.

A.1.5 Instrumental Variable Additional Results

Table A.11: IV First Stage

	Average Peer Exam Score Percentile	
	School	Track
	(1)	(2)
Admission Score (Percentile)	0.262***	0.325***
d1 × n2	−8.911***	−3.115***
d2 × n2	−9.096***	−2.757***
d3 × n2	−7.927***	−2.239***
d4 × n2	−5.789***	−1.552***
d5 × n2	−3.441***	−1.075***
d7 × n2	2.287***	0.517*
d8 × n2	3.659***	0.943**
d9 × n2	4.546***	1.899***
d10 × n2	5.532***	2.881***
d1 × n3	−11.187***	−3.327***
d2 × n3	−10.353***	−3.057***
d3 × n3	−9.710***	−2.515***
d4 × n3	−7.740***	−1.972***
d5 × n3	−4.829***	−1.597***
d7 × n3	4.160***	1.523***
d8 × n3	7.418***	2.652***
d9 × n3	9.386***	3.767***
d10 × n3	11.176***	5.108***
d1 × n4-15	−18.538***	−7.391***
d2 × n4-15	−17.388***	−6.903***
d3 × n4-15	−14.675***	−5.267***
d4 × n4-15	−11.188***	−3.948***
d5 × n4-15	−6.213***	−2.216***
d7 × n4-15	6.245***	2.443***
d8 × n4-15	11.237***	4.387***
d9 × n4-15	15.134***	6.621***
d10 × n4-15	18.018***	8.846***
d1 × n16+	−21.340***	−9.499***
d2 × n16+	−20.107***	−8.591***
d3 × n16+	−16.688***	−6.107***
d4 × n16+	−12.254***	−4.012***
d5 × n16+	−6.802***	−2.284***
d7 × n16+	6.524***	2.652***
d8 × n16+	12.547***	5.639***
d9 × n16+	18.071***	9.076***
d10 × n16+	22.622***	12.036***
Observations	1,161,358	1,161,135
Adjusted R^2	0.82	0.86
F-stat	3,538.6	1,000.6
Weak Instrument Test p-value	0	0
Wu-Hausman p-value	0	0

This table shows first stage results form the equation 3. Peer admission scores (at the school level) are instrumented using an interaction between student's own admission score decile and the number of high schools int he town in which they attend high school. Controls include within town-cohort own admission score decile fixed effects and interactions of these decile indicators with: the number of students admitted in towns in a given year, town unemployment level and county high school dropout levels and average wage levels (all computed as averages of 2010-2019 yearly data). Additional controls include the number of high schools in town (grouped into 5 bins), town fixed effects, year fixed effects, track type fixed effects and middle school fixed effects. Note: $p<0.1$; $p<0.05$; $p<0.01$. Standard errors are clustered at the town level.

A.1.6 Other Outcomes

Table A.12: Other Outcomes

	Match Rate OLS (1)	School Change IV (2)	Pass Exit Exam IV (3)
Own Admission Score (a_i)	0.664^{***}	-0.027	0.647^{***}
	(0.023)	(0.017)	(0.044)
Mean Peer Admission Score (School - μ^a_{-i})	0.017	-0.301^{***}	-0.094
	(0.015)	(0.044)	(0.061)
Observations	1,864,431	1,341,775	1,341,775
Adjusted R^2	0.43	0.19	0.41
Wu-Hausman p-value	0	0	0

This table shows regression results for several outcomes: 1) admission-graduation match rate, 2) change of school rate and 3) exit exam pass rate. The dependent variables are: 1) an indicator for matched graduation and admission records, 2) a indicator for graduating from a different school than the one admitted to and 3) passing the exit exam (conditional on registering for it). For (2) and (3), the endogenous variable is students' mean peer admission scores (instrumented by student's own admission rank and number of high schools in their town). Controls include within town-cohort own admission score decile fixed effects and interactions of these decile indicators with: the number of students admitted in towns in a given year, town unemployment level and county high school dropout levels and average wage levels (all computed as averages of 2010-2019 yearly data). Additional controls include the number of high schools in town (grouped into 5 bins), town fixed effects, year fixed effects, track type fixed effects and middle school fixed effects. Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Standard errors are clustered at the town level.

A.1.7 School Resources

Table A.13: Teacher Characteristics and Student Achievement

	Own Graduation Score Percentile
Own Admission Score (Romanian)	0.3521***
	(0.011)
Own Admission Score (Math)	0.341***
	(0.011)
Teacher Placement Score	0.083***
	(0.011)
Teacher College GPA	−0.176
	(0.123)
Teacher Experience	−0.586***
	(0.174)
Teacher Rank	−0.026
	(0.175)
Teacher Years of Schooling	−0.091
	(0.142)
Peer Admission Score Mean	0.112***
	(0.034)
N	436,969
R^2	0.51

This table shows the correlation between graduation scores and teacher characteristics, conditional on own and peer admission scores. Controls include within town-cohort own admission score decile fixed effects and interactions of these decile indicators with: the number of students admitted in towns in a given year, town unemployment level and county high school dropout levels and average wage levels (all computed as averages of 2010-2019 yearly data). Additional controls include the number of high schools in town (grouped into 5 bins), town fixed effects, year fixed effects, track type fixed effects and middle school fixed effects. Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered at the town level.

Table A.14: Teacher Placement Score and Student Achievement Additional Results

	Graduation Score Percentile	
Own Admission Score (Romanian)	0.386***	0.355***
	(0.016)	(0.016)
Own Admission Score (Math)	0.243***	0.348***
	(0.033)	(0.018)
Teacher Placement Score	0.008	0.007
	(0.014)	(0.011)
Teacher Placement Score × Own Admission Score (Romanian)	−0.0005**	
	(0.0002)	
Teacher Placement Score × Own Admission Score (Math)	0.0017***	
	(0.0003)	
Teacher Placement Score × Chemistry		0.120***
		(0.019)
Teacher Placement Score × Computer Science		0.256***
		(0.029)
Teacher Placement Score × Economics		0.229***
		(0.050)
Teacher Placement Score × Geography		0.031
		(0.019)
Teacher Placement Score × Other Electives		0.069***
		(0.020)
Teacher Placement Score × Philosophy		0.020
		(0.017)
Teacher Placement Score × Physics		0.208***
		(0.018)
Teacher Placement Score × Psychology		0.033
		(0.026)
Peer Admission Score Mean	0.087	0.093
	(0.063)	(0.064)
Observations	436,969	436,969
R^2	0.51	0.52

This table shows additional results for the correlation between school resources and student achievement on the graduation exam. Column 1 shows interactions between student admission scores and teacher placement scores. Column 2 shows heterogeneity by subject. Controls for include, college GPA, experience in years, rank (i.e. whether or not the teacher passed rank exams, which come with a pay increase) and teacher years of schooling. Additional controls include track type, year fixed effects, towns fixed effects and middle school fixed effects. The sample includes students attending a high school and track where at least one teacher was hired during the 2015-2019 period. The mean peer admission score is instrumented by student admission score decile and the number of high schools in their town. The mean peer admission score is instrumented by student admission score decile and the number of high schools in their town. Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Standard errors are clustered at the town level.

Table A.15: Decomposition of Sorting Effects: Track Level

	Dependent variable:					
	Graduation Exam Subject Percentile					
	Baseline	Teacher	Peer	Expenditure	T+P	T+P+E
Own Admission Score (Romanian)	0.152***	0.155***	0.149***	0.152***	0.153***	0.153***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Own Admission Score (Math)	0.061***	0.070***	0.058***	0.062***	0.068***	0.068***
	(0.015)	(0.016)	(0.016)	(0.015)	(0.017)	(0.016)
Track Peer Admission Score ($\hat{\mu}^a_{-ihs}$)	0.098*	0.053	0.075	0.093	0.030	0.024
	(0.056)	(0.058)	(0.055)	(0.057)	(0.056)	(0.057)
Teacher Placement Score		0.096***			0.097***	0.096***
		(0.011)			(0.011)	(0.011)
Change in Mean Track Admission Score ($\mu^a_{-ihst} - \mu^a_{-ihs}$)			0.097***		0.082***	0.082***
			(0.027)		(0.027)	(0.026)
School Spending (Euro per admitted pupil per year, thousands)				-1.344		-1.622
				(1.885)		(1.728)
N	309,942	309,942	309,942	309,942	309,942	309,942
R^2	0.51	0.51	0.51	0.51	0.51	0.51

This table shows estimation results of Equation 4. It shows the relationship between graduation exam subject scores (e.g. math, Romanian) and school quality, proxied by average admission scores, conditional on own entrance grade. Each column represents a specification in which I add peer effects, teacher ability and/or school expenditure to the baseline model in column 1. Peer ability is identified by the differences in within-school average admission scores from year to year. Teacher ability is the average placement test percentile of teacher who worked during a student's four year stay, weighted by the number of weekly hours worked. Since I look at graduation exam components separately, only teachers teaching relevant subjects are included on the right-hand side (e.g. math and physics teachers for science component grades). Peer admission scores are instrumented by an interaction of town-cohort admission deciles and the number of high schools within a town. Controls include within town-cohort own admission score decile and interactions of these decile dummies with the number of students admitted in towns in a given year. Additional controls include town unemployment level, county high school dropout levels and average wage levels (all computed as averages of 2010-2019 yearly data), the number of high schools in town (grouped into 4 bins), town fixed effects, year fixed effects and middle school fixed effects. Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered at the town level.

A.2 Notes on Matching

I briefly discuss the matching of graduation records to admission records. I match records by student name (as no unique student identifier exists) within high schools. Thus, students who drop out or change high schools are not be matched. As a note, high school changes can occur for legitimate reasons (moving, for example) and, anecdotally, due to corruption. Students may be unmatched for several reasons. Some schools offer admissions using different criteria than the admission exam (for example, art schools), while students in other schools do not write the graduation exam. Students may also drop out, so, in this sense, the results are subject to sample selection. This issue is addressed as per Ainsworth et al. (2023), who code missing graduation scores as 0's, allowing me to include them in the analysis. Lastly, students may switch schools and have very common names, which makes it challenging to match them accurately.

Table A.16 shows summary statistics regarding yearly match rates. Several additional factors, aside from drop outs and transfers, negatively impact the match rate. Students in many programs write the high school graduation exam, but do not enter high school via the national admission exam. For example, students in religious, arts, teaching, sports and architecture, may be admitted based on other aptitudes, such as playing an instrument, sporting prowess, knowledge of the Bible, drawing skills, etc. Students in these tracks do not appear in the admission records. Moreover, some students may repeat a year, while others are former drop-outs who decide to complete their high school studies.

Likewise, not all high school students admitted to high schools complete the high school program write the graduation exam. This is the case with students in low-ranked schools and typically in non-academic or technical programs that aim to prepare students for tertiary education. Furthermore, students who do not feel confident of passing the exit exam sometimes do not register for it.

Regarding year to year variation in match rates, generally speaking, the match rate improves with time. This may be a sign of better data quality or of enforcement of school switching, as well as changes in drop-out rates.

Table A.16: Statistics Regarding Yearly Match Rates of Graduating and Entering Students

	Min	Q1	Median	Mean	Q3	Max
Graduating Students	56%	59%	64%	66%	71%	75%
Entering Students	48%	52%	59%	59%	66%	67%
Graduating Students (filtered)	58%	64%	79%	74%	84%	85%
Graduating Students (filtered and excluding technical tracks)	67%	73%	84%	80%	85%	86%

Note: This table contains summary statistics regarding yearly match rates between entering and graduating students. The third line shows match rates for graduating students, after filtering arts, music, education, architecture, sports and religious track students, who typically are not admitted through the regular admission exam, as well as graduating students from previous cohorts. The fourth line also excludes all technical track students, some of which (depending on their track), do not gain admission through the entrance exam.

A.3 Robustness Checks

A.3.1 Robustness Check: Triple Difference Using School Openings

I estimate the following model, which will capture how the graduation exam score gap between high- and low- achieving students is impacted by a high school opening:

$$g_i = \beta_0 + \beta_e e_i + \beta_X X_{ct} + \beta_{q\times X} q_t \times X_{ct} + \beta_T \mathrm{T}_{ct} + \delta_q + \delta_c + \delta_t + \delta_{qt} + \delta_{qc} + \delta_{Tq} + \varepsilon_i \tag{5}$$

Here, g_i is the graduation score grade of student i who entered high school h in year t from middle school m in town or city c, X_{ct} is a matrix of town characteristics, including unemployment rate, high school dropout rate, average wage at the county level and the size of the high school entering cohort at time t, δ_q is an admission score quartile fixed effect,[1] δ_c and δ_t are town (city) and year times effects, respectively and δ_{qt}, δ_{qc} are fixed effects representing the interaction of a student's admission score quartile and year and town dummies, respectively. Lastly, T is a treatment indicator which takes the value of 1 only for cohorts entering high school at the time or after the opening of a new high school in their town. Tq captures interaction effects between the treatment dummy and students in different quartiles of the admission score distribution, thus allowing the model to capture differences in the impact of a school opening across admission grades.

The identifying assumption is that the timing of a high school opening in a given town is orthogonal to the change of the graduation score gap between high- and low-admission score students over time. In other words, the decision to open a new high school in a town and its precise timing is not correlated with an expected change in the relative performances of high- and low-entrance score students. I believe that this assumption is very plausible. Given that between the decision to consider opening a new school, approving the opening, earmarking the necessary funds, constructing the school, staffing it and finally opening it, a long period of time probably elapses. The following section demonstrates that a school opening immediately impacts town-level student

---

[1] Where admission score quartiles are computed at the town-cohort level.

sorting and student grades of the first entering cohort exposed to the opening. The opening itself plausibly causes these effects and not other confounding factors. The estimates are presented in Table A.17.

Table A.17: Effect of School Openings on Student Graduation Scores

	Peer Admission Score	Graduation Score	
	(1)	(2)	(3)
Admission Score	0.208***	0.667***	0.637***
	(0.016)	(0.009)	(0.018)
Post × Treated	−1.694	−4.944	−10.897*
	(1.923)	(3.686)	(5.985)
Q2 × Post × Treated	0.194		4.869***
	(2.471)		(1.545)
Q3 × Post × Treated	4.641*		8.871***
	(2.516)		(2.894)
Q4 × Post × Treated	5.376		6.828*
	(3.461)		(3.984)
Observations	219,338	211,528	211,528
Adjusted R^2	0.74	0.57	0.59

This table shows the impact of high school openings on student sorting (1), mean graduation scores (2) and the distribution of graduation scores (3). Data includes all one- and two-high school towns. I exclude towns where a non-general (e.g. religious denominational or private) high school opens. Admission and graduation scores are in national-cohort percentile ranks, quartiles (Q) are town-cohort admission score quartiles. All variables are interacted with the number of high schools in a town at baseline. I include year and town fixed effects and control for county high school dropout rates, unemployment rates, average wage rates and high school track type. For models (2) and (3), I interact the additional controls with the admission score quartiles (Q) to capture differences in the way students with different abilities perform in areas with different characteristics. Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered at the town level.

A.3.2 Robustiness Check: Propensity Score Matching

Table A.18: Propensity Score Matching Regression Results

	Peer Admission Score		Graduation Score			
	(1)	(2)	(3)	(4)	(5)	(6)
High Schools in Town	2 vs 1	n+1 vs n	2 vs 1	n+1 vs n	2 vs 1	n+1 vs n
Admission Score Percentile	8.18^{***}	22.61^{***}	72.24^{***}	73.57^{***}	57.48^{***}	54.13^{***}
	(0.44)	(0.36)	(0.78)	(0.49)	(1.15)	(0.67)
Treatment	-11.88^{***}	-6.15^{***}	-1.35^{*}	-2.26^{***}	0.39	0.10
	(0.35)	(0.27)	(0.81)	(0.52)	(0.27)	(0.12)
Treatment × d2	1.85^{***}	1.14^{***}	-0.78	0.52		
	(0.43)	(0.34)	(1.08)	(0.66)		
Treatment × d3	3.86^{***}	2.70^{***}	1.02	1.19^{*}		
	(0.41)	(0.33)	(1.02)	(0.62)		
Treatment × d4	6.11^{***}	4.07^{***}	1.73^{*}	0.62		
	(0.39)	(0.32)	(0.99)	(0.62)		
Treatment × d5	9.04^{***}	4.89^{***}	0.82	0.60		
	(0.38)	(0.32)	(0.97)	(0.61)		
Treatment × d6	11.90^{***}	6.01^{***}	1.75^{*}	1.13^{*}		
	(0.36)	(0.31)	(0.95)	(0.60)		
Treatment × d7	14.30^{***}	7.59^{***}	2.29^{**}	1.64^{***}		
	(0.36)	(0.31)	(0.95)	(0.59)		
Treatment × d8	15.58^{***}	8.43^{***}	1.67^{*}	1.82^{***}		
	(0.35)	(0.31)	(0.93)	(0.58)		
Treatment × d9	16.58^{***}	9.04^{***}	2.29^{**}	2.64^{***}		
	(0.35)	(0.30)	(0.91)	(0.57)		
Treatment × d10	16.97^{***}	9.14^{***}	1.73^{**}	2.66^{***}		
	(0.35)	(0.29)	(0.87)	(0.55)		
Observations	110,179	463,276	106,236	450,905	106,236	453,905
Adjusted R^2	0.78	0.78	0.52	0.57	0.65	0.64

This table shows results comparing students assigned to the control group (n high schools) and treatment groups (n+1 high schools) by nearest neighbour matching on admission scores, size of entering cohorts and admission year. The dependent variable for columns 1 and 2 is the admission score of peers admitted to the same high school. The dependent variable for columns 3 through 6 is the student's graduation score. Controls include within town-cohort own admission score decile fixed effects and interactions of these decile indicators with: the number of students admitted in towns in a given year, town unemployment level and county high school dropout levels and average wage levels (all computed as averages of 2010-2019 yearly data). Additional controls include the number of high schools in town (grouped into 5 bins), town fixed effects, year fixed effects, track type fixed effects and middle school fixed effects. Note: *p<0.1; **p<0.05; ***p<0.01.

A.3.3 Robustness Check: Migration

Table A.19: IV Second Stage (School): Endogenous Markets

	Graduation Exam Score Percentile		
	(1)	(2)	(3)
Admission Score (Percentile)	0.61***	0.61***	0.55***
	(0.04)	(0.03)	(0.03)
Instrumented Peer Admission Score (Percentile)	0.19***	0.13***	0.30***
	(0.02)	(0.02)	(0.07)
Observations	1,180,086	1,180,086	1,161,358
R^2	0.524	0.641	0.638

This table shows second stage results form the equation 3. Student graduation exams scores (in percentiles at the cohort-country level) are regressed on similar admission scores and peer admission scores. Peer admission scores (at the school level) are instrumented using an interaction between student's own admission score decile and the number of high schools in the education markets in which they attend high school. Education markets are defined endogenously based on student flows between middle schools and high schools across locations. The specification in Column (1) includes no controls, Column (2) includes year, school track type, middle school and market fixed effects, as well as within-market admission score decile and number of high schools within the market and Column (3) adds interactions of the decile dummies with: the number of students admitted in towns in a given year, town unemployment level and county high school dropout levels and average wage levels (all computed as averages of 2010-2019 yearly data). Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Standard errors are clustered at the county level.

Table A.20: IV Second Stage (School): Endogenous Markets, Excluding Migrants

	Graduation Exam Score Percentile		
	(1)	(2)	(3)
Admission Score (Percentile)	0.61***	0.61***	0.55***
	(0.04)	(0.03)	(0.03)
Instrumented Peer Admission Score (Percentile)	0.17***	0.11***	0.27***
	(0.02)	(0.02)	(0.07)
Observations	1,107,096	1,107,096	1,089,477
R2	0.524	0.642	0.640

This table shows second stage results form the equation 3. Student graduation exams scores (in percentiles at the cohort-country level) are regressed on similar admission scores and peer admission scores. Peer admission scores (at the school level) are instrumented using an interaction between student's own admission score decile and the number of high schools in the education markets in which they attend high school. Education markets are defined endogenously based on student flows between middle schools and high schools across locations. The specification in Column (1) includes no controls, Column (2) includes year, school track type, middle school and market fixed effects, as well as within-market admission score decile and number of high schools within the market and Column (3) adds interactions of the decile dummies with: the number of students admitted in towns in a given year, town unemployment level and county high school dropout levels and average wage levels (all computed as averages of 2010-2019 yearly data). Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered at the county level.

Table A.21: IV Second Stage (Track): Endogenous Markets

	Graduation Exam Score Percentile		
	(1)	(2)	(3)
Admission Score (Percentile)	0.53***	0.57***	0.39***
	(0.05)	(0.03)	(0.06)
Instrumented Peer Admission Score (Percentile)	0.33***	0.22***	0.70***
	(0.04)	(0.04)	(0.16)
Observations	1,179,854	1,179,854	1,179,854
R^2	0.52	0.64	0.62

This table shows second stage results form the equation 3. Student graduation exams scores (in percentiles at the cohort-country level) are regressed on similar admission scores and peer admission scores. Peer admission scores (at the track level) are instrumented using an interaction between student's own admission score decile and the number of high schools in the education markets in which they attend high school. Education markets are defined endogenously based on student flows between middle schools and high schools across locations. The specification in Column (1) includes no controls, Column (2) includes year, school track type, middle school and market fixed effects, as well as within-market admission score decile and number of high schools within the market and Column (3) adds interactions of the decile dummies with: the number of students admitted in towns in a given year, town unemployment level and county high school dropout levels and average wage levels (all computed as averages of 2010-2019 yearly data). Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Standard errors are clustered at the county level.

Table A.22: IV Second Stage (Track): Endogenous Markets, Excluding Migrants

	Graduation Exam Score Percentile		
	(1)	(2)	(3)
Admission Score (Percentile)	0.53***	0.58***	0.42***
	(0.05)	(0.03)	(0.06)
Instrumented Peer Admission Score (Percentile)	0.30***	0.19***	0.62***
	(0.04)	(0.04)	(0.16)
Observations	1,106,907	1,106,907	1,106,907
R^2	0.52	0.64	0.63

This table shows second stage results form the equation 3. Student graduation exams scores (in percentiles at the cohort-country level) are regressed on similar admission scores and peer admission scores. Peer admission scores (at the track level) are instrumented using an interaction between student's own admission score decile and the number of high schools int he town in which they attend high school. Controls include within town-cohort own admission score decile fixed effects and interactions of these decile indicators with: the number of students admitted in towns in a given year, town unemployment level and county high school dropout levels and average wage levels (all computed as averages of 2010-2019 yearly data). Additional controls include the number of high schools in town (grouped into 5 bins), town fixed effects, year fixed effects, track type fixed effects and middle school fixed effects. Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Standard errors are clustered at the county level.

A.3.4 Robustness Check: Sample Selection

Table A.23: Effect of Attending a More Selective School (Stage 2 IV)

	Graduation Score Percentile				
	School Level		Track Level		Town Mean
	OLS	IV	OLS	IV	OLS
	(1)	(2)	(3)	(4)	(5)
Admission Score (a_i)	0.47***	0.45***	0.45***	0.34***	0.51***
	(0.03)	(0.04)	(0.02)	(0.05)	(0.02)
Peer Admission Score (μ_{-i})	0.21***	0.27***	0.29***	0.58***	
	(0.02)	(0.06)	(0.02)	(0.12)	
2 HS Town	−1.19	−1.20	−1.25	−1.36	−0.67
	(1.05)	(1.05)	(1.00)	(1.02)	(1.23)
3 HS Town	−3.95**	−3.93**	−3.98**	−3.94**	−6.05**
	(1.80)	(1.79)	(1.78)	(1.75)	(2.63)
4-15 HS Town	−2.99	−2.96	−3.04	−2.96	−6.67**
	(2.28)	(2.25)	(2.25)	(2.20)	(3.32)
16+ HS Town	−4.53	−4.47	−4.63	−4.48	−11.53**
	(3.34)	(3.28)	(3.30)	(3.19)	(4.82)
Observations (millions)	1.4	1.4	1.4	1.4	1.4
Adjusted R^2	0.64	0.64	0.64	0.63	0.63

This table shows how graduation scores vary across towns with different numbers of high schools. Columns (1) and (3) show OLS regressions of graduation scores on admission scores and school-cohort and track-cohort admission scores respectively. Columns (2) and (4) instrument peer admission scores using an interaction between the within-town student admission score decile and the number of high schools in the town (as per equation 3). Column (5) highlights differences in mean graduation scores conditional on admission scores in towns with different numbers of high schools. Controls include within town-cohort own admission score decile fixed effects and interactions of these decile indicators with: the number of students admitted in towns in a given year, town unemployment level and county high school dropout levels and average wage levels (all computed as averages of 2010-2019 yearly data). Additional controls include the number of high schools in town (grouped into 5 bins), town fixed effects, year fixed effects, track type fixed effects and middle school fixed effects. Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Standard errors are clustered at the county level.

A.3.5 Robustness Check: Overidentification

Table A.24: IV Second Stage: Only Two Instruments

	Graduation Exam Score Percentile	
	School	Track
Instrumented Peer Admissiom Score (Percentile)	0.187***	0.416***
	(0.026)	(0.058)
Admission Score (Percentile)	0.522***	0.439***
	(0.027)	(0.035)
Observations	1,161,358	1,161,135
Adjusted R^2	0.637	0.636
Weak Instrument Test p-value	0	0
Sargan p-value (1st stage)	0.999779	0.999214
Wu-Hausman p-value	0	0

This table shows second stage results form the equation 3. Student graduation exams scores (in percentiles at the cohort-country level) are regressed on similar admission scores and peer admission scores. Peer admission scores (at the school level) are instrumented using an interaction between student's own admission score decile and the number of high schools in the education markets in which they attend high school. Education markets are defined endogenously based on student flows between middle schools and high schools across locations. Controls include within town-cohort own admission score decile fixed effects and interactions of these decile indicators with: the number of students admitted in towns in a given year, town unemployment level and county high school dropout levels and average wage levels (all computed as averages of 2010-2019 yearly data). Additional controls include the number of high schools in town (grouped into 5 bins), town fixed effects, year fixed effects, track type fixed effects and middle school fixed effects. Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Standard errors are clustered at the county level.