Location Choice, Commuting, and School Choice*

Minseon Park
University of Wisconsin-Madison
Job Market Paper

Dong Woo Hahm
University of Southern California

January, 2023

Click here for the latest version

Abstract

We explore the impact of public school assignment reforms by building a households' school choice model with two key features—(1) endogenous residential location choice and (2) opt-out to outside schooling options. Households decide where to live taking into account that locations determine access to schools—admissions probabilities and commuting distances to schools. Households are heterogeneous both in observed and unobserved characteristics. We estimate the model using administrative data from New York City's middle school choice system. Variation from a boundary discontinuity design separately identifies access-to-school preferences from other location amenities. Residential sorting based on access-to-school preference explains 30% of the gap in test scores of schools attended by minority students versus their peers. If households' residential locations were fixed, a reform that introduces purely lottery-based admissions to schools in lower- and mid-Manhattan would reduce the cross-racial gap by 7%. However, households' endogenous location choices dampen the effect by half.

Keywords: Centralized School Choice, Neighborhood Sorting, School Segregation, Commuting to School
JEL Codes: D12, I21, I28, J15, R23

*Park: mpark88@wisc.edu; Hahm: dongwooh@usc.edu. Park is deeply grateful to her committee members John Kennan, Christopher Taber, Chao Fu, and Jesse Gregory. We also thank Milena Almagro, Naoki Aizawa, Panle Jia Barwick, Pierre-Andre Chiappori, Andrew Bongjune Choi, Rajeev Darolia, Jean-François Houde, Amrita Kulka, Rasmus Lentz, Margaux Luflade, Jeff Smith, and Matt Wiswall for their useful feedback. This research has benefited from comments from seminar participants at the labor workshop at UW-Madison, the brown bag at the Federal Reserve Bank of Richmond, and the applied micro theory colloquium at Columbia University. Thanks also go to the New York City Department of Education, particularly to Benjamin Cosman, Broggini Matthew, Stewart Wade, and Lianna Wright.

1 Introduction

Across the U.S. in 2016, 33% of K-12 students lived in a school district where they could choose, to some degree, which public school to attend.[1] Public school choice systems provide students with multiple options beyond the nearest school to their home, standing in contrast to classic settings where students are automatically assigned to schools based on their home addresses. Such contrast has raised hope that centralized school assignments could decouple educational disparities from spatial disparities at scale. However, many popular schools under a choice system give admissions priority to students from residential locations nearby, even when they accept applications from a broader set of students (Dur, Kominers, Pathak, and Sönmez, 2013). Such location-based admissions rules have triggered debate over the design of admissions rules, motivated by a concern that these contribute to the continued school segregation observed in many school choice settings (Cohen, 2021).

How effectively can we desegregate schools with reforms on the location-based admissions rules in a public school choice system? We answer this question by developing a households' school choice model that considers two important margins through which households may respond: residential location choice and opt-out to outside schooling options.

The key feature of the model is households' endogenous location choice. While previous work has documented that residential location explains half the racial gap in test scores of schools attended by students under centralized school choice (Laverde, 2020), *how* households make the residential location decision has received little attention in the school choice literature (e.g., Abdulkadiroğlu, Agarwal, and Pathak, 2017). In our model, households choose residential locations by considering *access to schools*, which refers to both admissions probabilities and commuting distances to schools that vary across locations.

We set up a multi-stage discrete choice model where households sequentially choose (1) which location to live in, (2) which school to apply to, and (3) whether to enroll in the assigned school or opt out to outside schooling options. Households have observed and unobserved heterogeneous preferences over a set of location and school characteristics, which leads to rich residential and school sorting patterns.

We start by providing causal evidence that households consider location-based ad-

[1] Based on authors' calculation using The National Center for Education Statistics 2019 National Household Education Surveys: Parent and Family Involvement in Education Survey.

missions rules when deciding where to live. Our empirical context is the middle school choice system in New York City (NYC), where 70,000 students and 700 middle schools are matched each year. Each student has over 30 public school options to apply to, and which Community School District (CSD)—a subdivision of the city—they reside in largely determines the choice set and admissions probabilities. Leveraging this institutional aspect, we apply a boundary discontinuity design (BDD) to compare Census blocks that are close to one another but located on opposite sides of a CSD boundary. By doing so, we deal with the endogeneity concern that locations with higher admissions chances to high-achieving schools might have amenities unobserved to researchers but are observed and valued by households. Estimates indicate that Census blocks within a CSD with one standard deviation higher school test scores have 22% more households with middle school applicants.

We use the variation from the BDD to identify how much households value access to school relative to other location amenities. We estimate our structural model using an extension of the expectation-maximization algorithm with a sequential maximization step (ESM, Arcidiacono and Jones, 2003). This keeps the estimation tractable while enabling us to jointly estimate all stages of the model to account for households' selection into locations.

The results show that endogenizing households' residential choice has important implications for (1) understanding the source of school segregation under the status quo, (2) obtaining an unbiased commuting cost estimate, and (3) predicting the implications of the counterfactual policy.

First, our estimates illustrate that households' location choices based on location's access to schools play a large role in explaining which students are matched to which schools. To show this, we shut down each part of the model in a decomposition exercise. 30% of the gap in test scores of schools attended by minorities versus non-minorities is explained by households' residential sorting based on locations' access to schools. Households' heterogeneous preference over other location characteristics and school characteristics explains 45% and 18% of the cross-racial gap, respectively.

Second, we find that a model that does not account for endogenous location choice overestimates commuting costs by 15%. Our model estimates show that a median household is willing to pay $19 per school day to reduce commuting time to school by 50 minutes. Commuting cost is an important parameter that governs the degree to which students take advantage of school choice options rather than applying to schools nearest to their residential locations. The reason for the overestimation is that households choose locations near schools they prefer since it increases their admissions probabilities. This leads to a spurious result in which they apply to schools nearby not because they care about distance but because in their location choice they cared about admission probability. Without correcting for households' selection into locations, the model would misinterpret households' applying to schools nearby as solely due to commuting costs. Due to residential sorting based on unobserved school preference, this is still true when one controls for households' observed characteristics.[2]

Finally, we describe how households' spatial reshuffling in response to a school desegregation reform can affect the effectiveness of the policy. We consider a counterfactual policy that introduces purely lottery-based admissions to schools in District 2, the district with the highest test scores and housing costs. Covering lower- and mid-Manhattan, District 2 has been at the center of ongoing policy debates regarding the design of location-based admissions criteria.[3] When we fix households' residential locations, lottery-based admissions to District 2 schools would close the cross-racial gap in school test scores by 7%. This is because some minority students residing outside the district are assigned to District 2 schools, which pushes out non-minority District 2 residents to lower-achieving schools.

However, households' location choices in response to the policy dampen the equity impact by half. Two types of spatial reshuffling exert opposing forces. On the one hand, some minority households choose residential locations closer to District 2 in response to the reform. With shorter commuting distances to District 2, they are more likely to apply to District 2 schools. Spatial reshuffling of this sort amplifies the desegregation effect of the policy. On the other hand, most of the non-minority households who reside in District 2 under the status quo relocate out of the district. Since other districts still have location-based admissions in place, they seek other locations that assure higher admissions probabilities to high-achieving schools. Such spatial reshuffling dampens the equity effect of the policy.

The equilibrium force amplifies the second reshuffling while muting the first. This is because purely lottery-based admissions to District 2 schools induce more

[2]For example, a household that puts a higher value on school safety than other observably similar households will sort into locations that increase their child's admission chances into a safer school.

[3]Shapiro, Eliza, N.Y.C. to Change Many Selective Schools to Address Segregation, the *New York Times*, December 18, 2020.

applications, and thus the equilibrium admissions cutoffs of these schools increase. This weakens the incentive of minority households to relocate closer to District 2 but strengthens that of non-minority households to relocate farther from District 2. We find that households substitute between opting-out to outside schooling options and choosing different residential locations. But, overall, opt-out plays a smaller role in determining the effectiveness of the reform on reducing the cross-racial gap.

Related Literature We contribute to two strands of the literature. First, we extend the school choice literature by considering households' endogenous location choice. While it is well known that residential location is the main source of school segregation (Laverde, 2020), little is known about *how* households choose where to live in response to the design of centralized school choice. Previous studies have focused on assignment mechanisms (Abdulkadiroğlu, Che, and Yasuda, 2015; Abdulkadiroğlu, Agarwal, and Pathak, 2017; He, 2015; Agarwal and Somaini, 2018; Che and Tercieux, 2019; Calsamiglia, Fu, and Güell, 2020); information provision (Hastings and Weinstein, 2008; Hoxby and Turner, 2015; Luflade, 2018; Corcoran, Jennings, Cohodes, and Sattin-Bajaj, 2018; Chen and He, 2021; Fack, Grenet, and He, 2019; Allende, Gallego, and Neilson, 2019); limited attention (Ajayi and Sidibe, 2020; Son, 2020); and previously attended schools (Hahm and Park, 2022).

By modeling households' endogenous location choices, we first compare the implications of counterfactual policies when households' residential locations are fixed versus adjusted. This approach aligns with reduced-form evidence that access to school shapes the composition of residents and housing costs of locations (Black, 1999; Reback, 2005; Brunner, Cho, and Reback, 2012; Schwartz, Voicu, and Horn, 2014; Billings, Brunner, and Ross, 2018). Moreover, we correct selection into locations in estimating school preference to obtain an unbiased estimate of commuting costs; we take a departure from the standard assumption in the literature that distances to schools are uncorrelated with households' unobserved school tastes conditional on their observable characteristics.[4]

Second, this paper adds to a large body of studies on within-city residential sorting, by studying households' location choice in a newly relevant setting of centralized school assignments. Among many papers in this literature, more closely related are

[4]This assumption is often found in the broader economics of education literature, which uses distance to schools as an instrumental variable for school application and attendance (Card, 1993; Schwartz, Stiefel, and Wiswall, 2013; Walters, 2018; Mountjoy, 2022)

those that give special attention to schools compared with other location amenities.[5] Earlier studies have focused on classical settings where each residential location is zoned to one public school (Bayer, Ferreira, and McMillan, 2007) while incorporating limited forms of school choices such as private school vouchers or inter-district transfers (Manski, 1992; Nechyba, 2000; Epple and Romano, 2003; Ferreyra, 2007).

Under centralized school assignments, households choose among many public schools from a given location. This enables us to study households' heterogeneous values over a set of school characteristics, including commuting distance. Indeed, this two-way heterogeneity is one of the main sources of school segregation under school choice settings (Idoux, 2022; Hahm and Park, 2022, e.g.). In contrast, frameworks in the location choice literature (Bayer, Ferreira, and McMillan, 2007) have considered one-dimensional school characteristics, usually mean test scores, due to lack of variation coming from their setting where each location is zoned to one public school.

With the recent popularity of centralized school assignments, there have been a few papers proposing a unified framework of location choice and school choice. These include theoretical models (Xu, 2019; Avery and Pathak, 2021; Grigoryan, 2021) and a quantitative model (Agostinelli, Luflade, and Martellini, 2021). Our paper complements theoretical models by estimating our model using data.

The closest paper to ours is by Agostinelli, Luflade, and Martellini (2021), from which we differentiate in two respects. First, our model features rich heterogeneity in households' location and school preferences. For example, Grigoryan (2021) shows that preference heterogeneity is crucial in determining the welfare implication of a school choice design.[6] We depart from the assumption that households have the same ordinal preferences over schools. We also consider location sorting based on unobserved school preferences to obtain unbiased commuting costs. Second, we model outside schooling options, another margin that some households use with the introduction of a more extensive school choice system.

Organization The remainder of the paper is organized as follows. Section 2 describes the public middle school choice system in NYC and the data. Section 3 presents

[5] Broader set of papers have studied how residential sorting is determined by other factors such as access to work (Ahlfeldt, Redding, Sturm, and Wolf, 2015), ease of commuting (Barwick, Li, Waxman, Wu, and Xia, 2021), consumption amenities (Almagro and Domınguez-Iino, 2019; Miyauchi, Nakajima, and Redding, 2022), or neighborhood composition (Davis, Gregory, and Hartley, 2019).

[6] See Almagro and Domınguez-Iino (2019) for a similar discussion in a model where households have heterogeneous preferences over a set of urban amenities.

motivating evidence on the interaction between residential location choice and school choice. Section 4 describes the model. Section 5 describes the empirical strategy and presents estimation results. Section 6 investigates the source of school segregation. Section 7 studies the equity impacts of a school desegregation reform.

2 Institutional Background and Data

2.1 Public Middle School Choice in NYC

Each year, about 70,000 entering students and 700 middle school programs participate in the NYC citywide middle school choice system. There are about 500 middle schools. Multiple programs with separate curriculum can be offered by one school, and students apply to each program. In the following, we use the terms "program" and "school" interchangeably when there is no confusion. Schools that are part of the centralized choice system are governed by the city. The property tax rate is constant within the city and the city allocates the pooled funding to schools directly, largely based on the number of students.[7]

The main round of the school choice process starts in December of students' last year of elementary school. Students are given a customized list of programs they are eligible for and submit a rank-ordered list (henceforth, ROL) by designating their preference rankings over schools. In 2014-15, the average student had about 30 choice options. There is no list-length restriction, and students can list as many schools as they like (an example of an ROL is in Appendix A). The city uses the student-proposing deferred acceptance (SPDA) algorithm, which takes students' applications, schools' ranking over students, and the number of seats as main inputs and produces at most one assignment for each student (Gale and Shapley, 1962).[8]

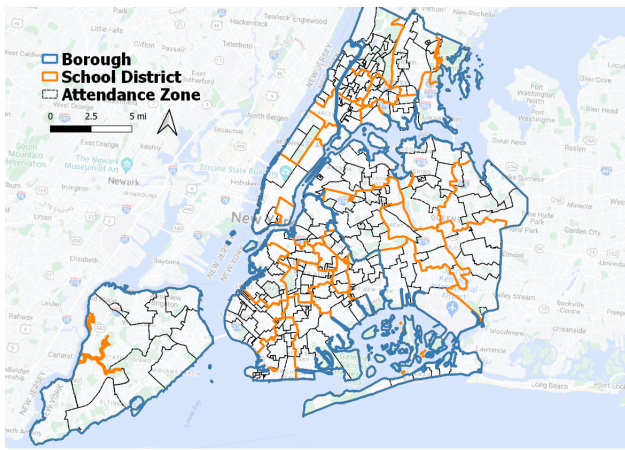
Schools rank students by pre-announced admission rules, which consist of three layers. The first is eligibility, which determines students' school choice sets. If a student is not eligible for a program, she is never considered by the program, even when

[7]In 2002, Chapter 91 (Bill A.11627/S.7456-B) was enacted to reorganize the education system and has established centralized power. Since then, the public education system has been governed by the Panel for Educational Policy (PEP), which has 15 members; 9 of which are nominated by the mayor. The citywide school choice system was introduced in 2004 as part of this effort (Abdulkadiroğlu, Pathak, and Roth, 2005).

[8]90% of students are assigned to a program on their list. The rest are matched to their fall-back option, which is usually a school in their attendance zone. See Appendix A for details on the timeline and SPDA.

there are remaining seats. Second, eligible applicants are classified into a small number of priority groups. A program considers all students in the higher-priority group before considering any student in a lower priority group. Henceforth, we use the term “priorities" to refer to eligibility *and* priority groups, if not specified. Lastly, tie-breaking rules determine which students to admit among applicants of the same priority group. Some programs use a *nonrandom tie-breaker*, which is a school-specific function of the student's previous year's GPA, statewide standardized test scores, and punctuality. The rest use a *lottery* system in which each student receives one lottery number that applies to all such programs. See Appendix A for more details on the admissions rules.

Figure 1: Geographic Divisions



Note: The city is divided into five boroughs (=counties), which are further divided into 32 school districts and 300 middle school attendance zones.

Students' residential locations are the main criterion for the eligibility and priority of schools. Figure 1 depicts different levels of geographic subdivisions that determine location-based admissions rules. The city is split into 5 boroughs, 32 Community School Districts (districts, henceforth), and more than 300 attendance zones.

Depending on their eligibility criteria, middle schools are classified into zoned programs, district programs, borough programs, and citywide programs. A student's residence or the location of her elementary school decides her eligibility for each type of school.[9] Of 669 programs in academic year 2014-2015, 14 were citywide programs, 27 were borough programs, 478 were district programs, and the rest (150) were zoned programs. Schools can further assign priority based on finer geographic divisions. For

[9]87% of 2014-15 middle school applicants attended an elementary school in their residential district.

instance, 81 of 478 district programs gave top priority to students from a particular attendance zone.

2.2 Data

Student and School Data Student-level data from the NYC Department of Education (DOE) cover middle school applicants in academic year 2014-15. The data have two crucial components for the purpose of this paper—students' school applications and residential Census block. The data also contain students' enrollment decisions, demographic characteristics, and statewide standardized test scores.[10]

We construct school characteristics by digitizing the *Directory of Public Middle Schools*.[11] It covers each program's admissions criteria, address, performance measures, previous year's capacity, and number of applicants. Students, parents, and guidance counselors use this as their primary information source during the middle school application process (Sattin-Bajaj, Jennings, Corcoran, Baker-Smith, and Hailey, 2018). We augment this data by adding the number of crime incidents of different categories in each school building from a NYC Police Department's School Safety Report.

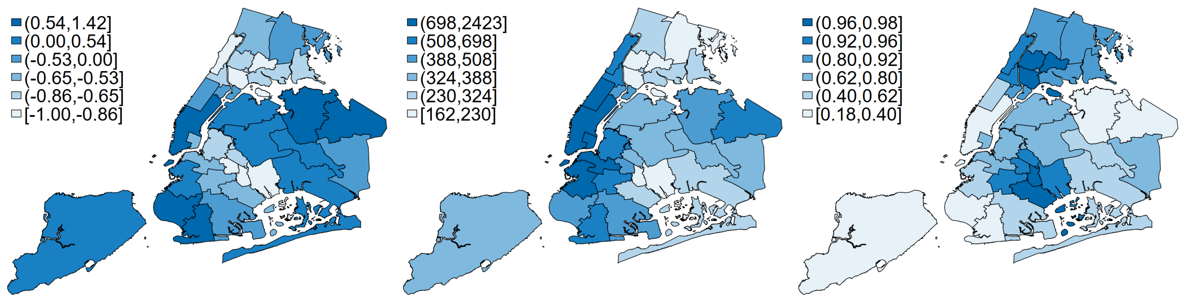
Housing Cost and Structure Housing cost and housing characteristics are from the NYC Department of Finance's (DOF) Rolling Sales files. The data include the exact address of each sold property, which is granular enough for us to observe on which side of a school district boundary the property is located. We describe the cleaning process of the DOF Rolling Sales files in detail in Appendix B.

Amenities of Residential Location We construct location amenities from various sources. Land use comes from the Primary Land Use Tax Lot Output. We also obtain consumption amenities such as the number of cafes from business licenses published by NYC Consumer and Worker Protection. Next, we collect information on bus stops, metro stations, and park areas using NYC OpenData GIS data files. We aggregate variables to Census block level. Finally, the demographic composition of each Census block group, such as ethnicity, age, education, and income, comes from the American Community Survey (ACS) 5-year estimates.

[10] We focus on academic year 2014-15 because students can list only up to 12 middle schools in more recent years. With this list length restriction, students have the incentive to list less preferred schools with higher admissions chances (see Section 4).

[11] The city began publishing a digitized version in academic year 2017-18.

Figure 2: Main Variables by District



(a) School Test Score (b) Sales Price ($1,000) (c) Minority Share in Schools

Note: In panels (a) and (c), we take the average of the variables across schools within each district. The school test score is the average NYS standardized test scores of enrolled students. In panel (b), we present the average unit sales price of residential properties in each district.

Figure 2 presents the average characteristics for each district, which demonstrates a strong correlation among school achievement, housing cost, and share of minorities in schools. Summary statistics of main variables are in Table B.2.

3 Motivating Data Pattern

3.1 Effect of Admissions Probability on Residential Sorting

This section presents evidence that households choose where to live by considering location-based admissions probabilities. Specifically, we show that locations with higher admissions chances to high-achieving schools have greater number of households with middle school applicants and higher housing costs. This makes the main motivation to model endogenous location choices under centralized school choice. Moreover, we show that these locations also have a lower minority share among households with middle school applicants, which implies that households have heterogeneous rates of substitution between housing cost and higher admissions chances to high-achieving schools. The main challenge to credibly show these patterns is that locations with higher admissions chances to high-achieving schools may have amenities unobserved to the econometrician but observed by and desirable to households, such as a well-kept playground.

To this end, we adopt a boundary discontinuity design (BDD) (Black, 1999; Bayer, Ferreira, and McMillan, 2007). Ideally, we would compare two locations with the same amenities but with different admissions probabilities to schools. BDD mimics the ideal design by comparing locations that are within a narrow buffer around a school district boundary but on opposite sides. The identification assumption is that unobserved amenities are as good as random within a narrow buffer around a boundary. This assumption likely holds if other amenities are continuous in geography.[12]

We consider a narrow buffer that covers locations within 0.25 miles from a border at which a pair of school districts meet. Figure 5 illustrates this idea. Tables in Appendix B present estimates with a 0.2-mile buffer. Table B.2 presents summary statistics of student, housing, and Census block group characteristics of all sample in comparison to sample included in the BDD analysis. The differences in characteristics largely come from the fact that we exclude Staten Island since it consists of one school district. For example, Staten Island has larger number of White student, thus BDD sample has smaller share of White students (8.5%) than the full sample (12.5%).

The baseline regression is as follows.

$$y_i = \beta \underbrace{Q_{d(i)}}_{\text{district school quality}} + \theta_{b(i)} + f(r_i) + \epsilon_{id} \,. \tag{1}$$

The unit of observation i is a housing transaction record when y_i is the log house sales price. The unit of observation i is a Census block when y_i is the number and characteristics of middle-school-applying residents in the block. $b(i)$ is the boundary region fixed effect in which i is located. $f(r_i)$ is a local cubic control for distance to the boundary $b(i)$, which we allow to differ by whether the district in which i is included has higher school quality than the bordering district.

$Q_{d(i)}$ is district school quality, measured by the mean NYS standardized test score of students enrolled in middle schools (previous cohorts) in the district. In the model estimation, we consider multidimensional school "quality" measures and allow students to have heterogeneous preferences over measures. In this section, we use a one-dimensional measure for simplicity. Our choice of the mean test score is motivated by Abdulkadiroğlu, Pathak, Schellenberg, and Walters (2020)'s finding that parents

[12] We do not apply BDD on attendance zone boundaries, because there is a concern about these boundaries' being determined by residents themselves. School district boundaries can be redrawn only every 10 years, and the decision is made at city level (New York State Law 2590-B). Meanwhile, attendance zone boundaries can be redrawn every year by the district council, whose members include parents and representative students. We still consider that admissions probability chances vary across attendance zones in the model estimation (Section 5).

of high school students in NYC do not value school effectiveness beyond the average test scores of students enrolled in a school.

The identification assumption is unlikely to hold if school district boundaries were drawn to divide already divided neighborhoods; even if they were exogenously drawn in the beginning, location amenities might have evolved differently over time on opposite sides of a boundary (Baum-Snow and Ferreira, 2015). We do two things to address these and to render causal interpretation of our estimates more plausible (Bayer, Ferreira, and McMillan, 2007; Kulka, 2019; Zheng, 2022).

First, we drop boundaries in which locations on opposite sides are likely to differ in access to amenities other than schools. Thus, we exclude boundaries aligned with a river, creek, park, highway, or borough boundary. Second, in Appendix B, we show that neither housing characteristics nor urban amenities change sharply at school district boundaries, which suggests that the identification assumption is plausible in our context.

Figure 3 presents estimates $\hat{\beta}$ for various outcomes. Tables in Appendix B present coefficients plotted in Figure 3. For each outcome, we start from a simple BDD specification (Equation 1). Then we present coefficients from specifications where we control for various covariates. In each panel, the coefficient from our preferred specification is in the rightmost.

District school quality increases the quality of schools to which residents are assigned The top left panel of Figure 3 reports $\hat{\beta}$ for the mean score of schools to which middle-school-applying residents in a Census block are assigned. A one standard-deviation increase in district school quality increases assigned schools' test scores of residents by 0.26 student-level standard-deviation (p-value < 0.01); we control for the resident's ethnicity, FRL status, and test score, to absorb differences in school applications and admissions probabilities explained by applicants' observable characteristics. This result implies that school district boundaries determine admissions probabilities to high-achieving schools, which establishes the first stage of the BDD.

District school quality increases housing prices The top right panel of Figure 3 reports $\hat{\beta}$ for the log sales price of a residential unit. Including this panel, we plot coefficients from specifications where we sequentially add housing characteristics, neighbor characteristics, and urban amenities for the rest of the panels. Given that housing characteristics and urban amenities do not change at a boundary (Appendix B), we control for those to increase the precision of our estimates. Meanwhile, we control for neighbor characteristics to account for the fact that households might have preferences over neighbors' ethnicity or median income. We interpret the estimate from a model with full controls to describe the effect of district school quality.

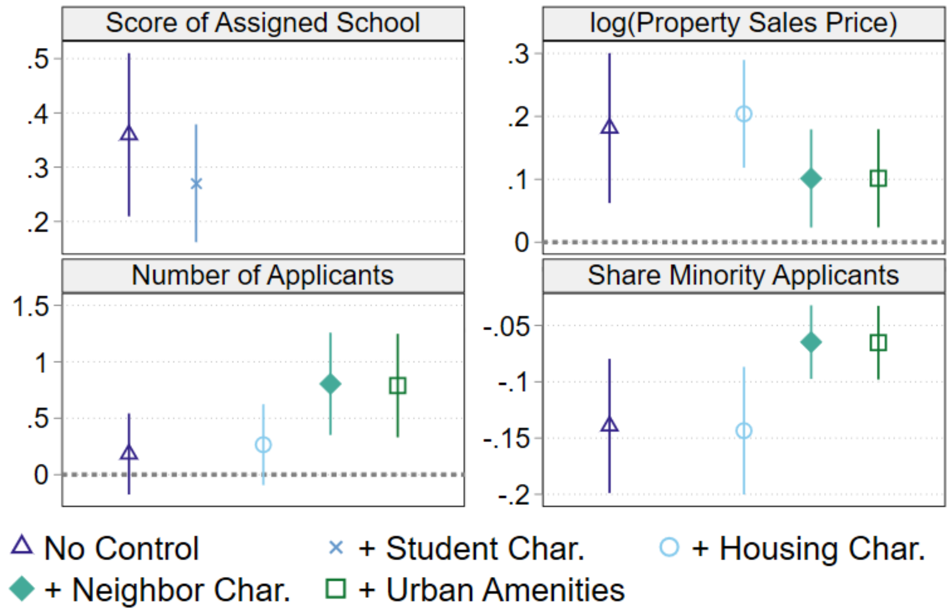
A one standard-deviation increase in district school quality increases housing sales price by 10% (p-value < 0.05). This implies that there is a higher demand for locations with higher admissions probability to better-performing schools.

We present $\hat{\beta}$ for house value and median gross rent from the ACS 5-year estimates in Appendix B. Estimates are 5.8% for both house value and median rent, although the estimate is only significant for median rent (p-value < 0.1). While sold properties might not be representative of all properties, we prefer sales prices to these two alternatives because the ACS 5-year estimates are at Census block group level, which is too coarse to study a change of housing costs at boundaries. In Appendix B, we explain how we use the distribution of total population and houses across Census blocks within each block group to weigh Census block groups in obtaining $\hat{\beta}$.

District school quality attracts households with middle school applicants The bottom left panel of Figure 3 reports $\hat{\beta}$ for the number of middle school applicants residing in a Census block. A one standard-deviation increase in district school quality increases the number of middle-school-applying residents, with $\hat{\beta} = 0.79$ (p-value < 0.01). An average Census block has 3.5 middle-school-applying residents, and thus this is a 22% increase from the average. This result is robust to controlling for the total number of population in Census Block Group ($\hat{\beta} = 0.81$, p-value < 0.01) from the ACS 5-year estimate. Thus, we exclude an explanation that Census blocks with higher district school quality have a greater number of households with middle school applicants merely because those blocks have more houses. Estimates are presented in Table B.5

District school quality attracts non-minority households more The two bottom panels of Figure 3 report $\hat{\beta}$ for the share of Black and Hispanic applicants among middle-school-applying residents in a Census block. A one standard-deviation increase in district school quality decreases the share of minority applicants by 6 percentage points (p-value < 0.01). An average Census block has 62% Black or Hispanic residents among middle-school-applying residents, and thus this is a 10% decrease from the average.

Figure 3: Estimated Effects of District School Quality on Residential Sorting



Note: The figure depicts the estimates (dots) and 95% confidence intervals (lines) of the coefficients of district school quality on various outcomes (β in Equation 1). The dependent variable in each panel is as follows (clockwise): (1) the mean score of the schools middle-school-applying residents in a Census block are assigned to, (2) the log sales price of a residential unit, (3) the number of middle-school-applying residents in a Census block, and (4) the share of Black and Hispanic applicants among those residents. In all panels, we plot the coefficient from a simple BDD specification (Equation 1) and coefficients from specifications that control for other variables. In the top left panel, we control for middle-school-applying residents' ethnicity, FRL status, and test score. In the rest of the panels, we sequentially add housing characteristics, neighborhood characteristics, and urban amenities. Standard errors are clustered at school district level. Housing characteristics include the space of the unit, land use of the tax lot, number of floors, age, renovation status, and storage area of the building, all of which we interact with a dummy if the property is coop. Neighbor characteristics include % minority, median household income, % college-or-more-educated, and median commuting time to work at Census block group. Urban amenities include the number of bus stops, subway stations, laundries, cafes, and crime incidents of different categories at Census block.

3.2 The Role of Commuting Distance in School Applications

Next, we show that while students apply to geographically proximate schools, the patterns are heterogeneous by students' characteristics and by the achievement level of schools near their residential locations. We run the following linear probability model:

$$100 * \mathbb{1}(Top3)_{ij} = \alpha d_{\ell_i j} + d_{\ell_i j} Z_i \beta + \delta \mathcal{J}_i + \epsilon_{ij} \,. \quad (2)$$

$\mathbb{1}(\text{Top3})_{ij}$ is an indicator for whether student i lists school j in her top three choices. j is a school for which student i is eligible. We multiply $\mathbb{1}(\text{Top3})_{ij}$ by 100 to interpret coefficients as percentage point changes. $d_{\ell_i j}$ is the driving distance in miles between school j and student i's residential census block ℓ_i. Z_i is a vector of student characteristics. α represents the association between distance to a school and the propensity of students to list the school as their top choice. β shows how that association changes by students' characteristics. To account for the fact that the probability of choosing a specific school as the top choice mechanically decreases when the number of eligible options increases, we control for the total number of schools for which i is eligible ($\mathcal{J}_i$). We cluster standard errors at student level.

Table 1: Commuting Distances and the Propensity of Listing as Top 3

	(1)	(2)
$d_{\ell_i j}$	-2.460	-2.411
	(0.013)	(0.012)
$d_{\ell_i j} \times \mathbb{1}(Minority)_i$	0.262	0.083
	(0.011)	(0.011)
$d_{\ell_i j} \times$ Quality of the three closest schools$_i$		-0.286
		(0.007)
N	1,745,513	1,745,513
R2	0.062	0.063
Dep. var mean	7.895	

Note: The dependent variable is a dummy if student i listed school j as one of their top three choices, multiplied by 100 for ease of interpretation. Pairs of a student and an eligible school within 10 miles from the student's residential Census block are included. The fastest driving distance between a school and a Census block is calculated using Open Route Services. A student is a minority if she is Black or Hispanic. Column (2) controls for the mean test score of the three closest schools from i's residential Census block ℓ_i. All columns control for the total number of schools a student is eligible for and the interaction of distance and student's standardized test scores. Standard errors in parentheses are clustered at the student level.

Columns (1) and (2) in Table 1 demonstrate that students are 2.4 percentage points less likely to rank a school that is 1 mile farther away as their top 3 choices (p-value < 0.01). Minority students seem to be less responsive to distance in column (1) ($\beta = 0.262$, p-value < 0.01). In column (2), we further control for the mean test score of the three closest schools from student i's residential Census block. Students are even less likely to apply to schools farther away when nearby schools have higher quality (column (2), $\beta = -0.286$, p-value < 0.01). Importantly, controlling for the quality of nearby schools reduces the coefficient of the minority dummy by two-thirds ($\beta = 0.082$, p-value < 0.01). This pattern, whereby students from disadvantaged location travel farther to schools that are higher performing than schools in their residential location, coincides with what has been reported in previous studies (Burdick-Will, 2017; Corcoran, 2018).

Motivated by these patterns, we model households as considering not only commuting distances but also other school characteristics. We also allow households to have heterogeneous commuting costs.[13]

4 A Model of Location Choice, School Choice, and Enrollment Decision

We model households' sequential decisions of residential locations, school applications, and enrollment decisions. Location choices affect school applications and assignments through two channels. First, distances to schools vary by residential location, which affects students' school applications. Second, applicants are ranked based on location-based priority rules, and thus two students from different locations who are otherwise similar face different admissions probabilities. Households take these two channels into account when choosing which location to reside in.

The model is guided by two key parameters. The first is access-to-school preference α^u. It is the weight households put on access-to-school utility that captures both commuting distances and admissions probability, relative to other location amenities. It also governs the extent to which counterfactual location-based priority rules would induce households to resort across locations. The second is commuting cost β^d, which affects to what extent students apply to schools that are farther away given their location choices as opposed to applying to only nearby schools. Together with α^u, it shapes the spatial distribution of households; for example, with infinite commuting costs and strictly positive access-to-school utility, households would choose locations closer to the schools they would apply to.

[13]This pattern is not explained by the difference in the number of schools proximate to their residential location. Students whose proximate schools are lower achieving have more schools proximate to their residential location. We present a histogram showing this result in Appendix B.

Next, we discuss our model in greater detail.

4.1 Household Preference, School Assignment, and Timeline

Household Heterogeneity and Preferences We use "household, applicant," and "student" interchangeably and model the unitary decision of a household. Household i is heterogeneous in both observable and unobservable (to the researcher) characteristics, denoted as Z_i and γ_i, respectively. Observable characteristics Z_i include students' race/ethnicity, poverty status (proxied by free and reduced-price lunch eligibility), and test score prior to their middle school application. Unobserved (discrete) type γ_i (Heckman and Singer, 1984) captures the fact that school characteristics may be valued differently by observably similar households.

i's utility from living in location ℓ and attending school j is

$$\underbrace{V_i(\ell;\eta_{i\ell})}_{\text{utility from location}} + \alpha^u \underbrace{U_i(j,\ell;\varepsilon_{ij})}_{\text{utility from school} \times \text{location}} . \tag{3}$$

We parameterize each component as follows:

$$V_i(\ell;\eta_{i\ell}) = \underbrace{W'_\ell}_{\text{location char.}} \alpha_i^W + \underbrace{p_\ell}_{\text{housing cost}} \alpha_i^p + \underbrace{\xi_\ell}_{\text{unobserved amenities}} + \underbrace{\eta_{i\ell}}_{\text{i.i.d. EVT1}} , \tag{4}$$

$$\text{where } \alpha_i^k = \alpha^{k0} + \underbrace{Z'_i}_{\text{student char.}} \alpha^{kz}, \text{ for } k = p, W .$$

$$U_i(j,\ell;\varepsilon_{ij}) = \underbrace{X'_j}_{\text{school char.}} \beta_i^X + \underbrace{d_{\ell j}\beta_i^d}_{\text{commuting cost}} + \underbrace{\varepsilon_{ij}}_{\text{i.i.d. EVT1}} , \tag{5}$$

$$\text{where } \beta_i^k = \beta^{k0} + \underbrace{Z'_i}_{\text{student char.}} \beta^{kz} + \underbrace{\gamma_i^k}_{\text{unobserved type}} , \text{ for } k = d, X .$$

W_ℓ is a vector of location observable characteristics, p_ℓ is the housing cost, Z_i is the vector of student observable characteristics, X_j is the vector of school characteristics, and $d_{\ell j}$ is the fastest driving distance between location ℓ and school j.

In addition, ξ_ℓ represents unobservable location amenities that are shared across households. $\gamma_i = (\gamma_i^X, \gamma_i^d)$ is the vector of student i's unobserved tastes over school characteristics and distance to schools, and $\eta_{i\ell}$, ε_{ij} are idiosyncratic preferences shocks over locations and schools. $\eta_{i\ell}$ and ε_{ij} are mutually independent and follow i.i.d extreme value type 1 distribution.

i's utility from living in location ℓ and attending an outside option ϑ is

$$\underbrace{V_i(\ell;\eta_{i\ell})}_{\text{utility from location}} + \underbrace{U_i^{\vartheta}(\vartheta;\varepsilon_{i\vartheta})}_{\text{utility from outside option}} . \tag{6}$$

We consider two outside options, non-public schools ϑ^{np} and public charter schools ϑ^{c}.[14] Non-public schools ϑ^{np} includes private schools, homeschooling, or moving out of NYC.[15] We further allow students to have heterogeneous preferences for outside options based on their observable characteristics. For example, non-minority students might assign higher value to non-public schools than their peers. Mathematically,

$$\begin{aligned} U_i(\vartheta;\varepsilon_{i\vartheta}) &= \beta_i^{\vartheta} + \underbrace{\varepsilon_{i\vartheta}}_{\text{i.i.d. EVT1}} \\ \beta_i^{\vartheta} &= \beta_0^{\vartheta} + \underbrace{Z_i}_{\text{student char.}} \beta_z^{\vartheta}, \text{ where } \vartheta = \vartheta^c, \vartheta^{np}. \end{aligned} \tag{7}$$

$\varepsilon_{i\vartheta}$ follows an i.i.d extreme value type 1 distribution.

School Assignment Next, we briefly discuss how schools rank applicants. As discussed in Section 2, priority groups are largely determined by students' residential location ℓ. The tie-breaker within priority groups is either a lottery or a school-specific aggregation of students' pre-middle-school academic measures. We capture programs' ranking over students with a priority score, $c_{ij}(\ell)$. This is the sum of an integer $g_{ij}(\ell)$ that corresponds to priority groups and decimal point $\tau_{ij} \in [0,1]$ that corresponds to tie-breakers.[16] Tie-breakers are either school-specific aggregation of students' academic measures or random lottery numbers ρ. The higher a student's $c_{ij}(\ell)$, the higher her admissions chance.

[14] Public charter schools are not parts of the centralized school choice system and they have separate admissions processes.

[15] Although we observe that a student is not enrolled in a public school in NYC, we do not know which non-public option a student chooses.

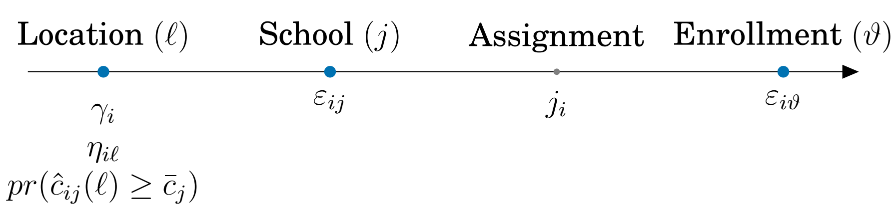
[16] For a program j with three priority groups, students in the first priority group have $g_{ij} = 3$. The second and the third priority groups' students have $g_{ij} = 2$ and 1, respectively. If a student is ineligible for program j, $g_{ij} = -\infty$

How priority groups are determined is public information. When it comes to school-specific aggregation of students' academic measures, we know which inputs a school uses—such as GPA, statewide test score, and punctuality—and the aggregated scores among its applicants. However, the exact function that schools use to construct these measures are unknown. We estimate school-specific linear functions of measures using a latent model and assume households form expectations ($\hat{c}_{ij}(\ell)$) in the same way; details are in Appendix C. Given students' rank-ordered list and priority scores, the city assigns students to at most one program using the SPDA algorithm. See Appendix A for a detailed explanation of the SPDA procedure.

Cutoff for each school $\bar{c}_j$ is given as the $\min\{c_{ij} : i \in \mathcal{I}_j\}$, where $\mathcal{I}_j$ is a set of students admitted to program j if the capacity of j is filled, and $-\infty$ otherwise. We assume the market is large enough (70,000 students) that an individual student considers the cutoffs as given (Fack, Grenet, and He, 2019; Agarwal and Somaini, 2020; Calsamiglia, Fu, and Güell, 2020).

Timing Figure 4 summarizes the timeline of the model and households' information at each stage. Households make choices on blue dots. They have full information on their own observable characteristics (Z_i) and those of schools (X_j) and residential locations (W_ℓ, p_ℓ), as well as locations' unobserved amenities (ξ_ℓ), throughout all stages.

Figure 4: Timeline and Information Set



Note: Households make choices on blue dots. They have full information on their own observable characteristics (Z_i) and those of schools (X_j) and residential location (W_ℓ, p_ℓ), as well as the shared neighborhood unobserved amenities (ξ_ℓ) throughout all stages. γ_i is the vector of student i's unobserved tastes over school characteristics. $\eta_{i\ell}$, ε_{ij}, and $\varepsilon_{i\vartheta}$ are idiosyncratic preferences shocks over locations, schools, and outside options respectively. j_i is the assignment result. $pr(\hat{c}_{ij}(\ell) \geq \bar{c}_j)$ is the predicted admissions probability.

Households know their unobserved tastes over school characteristics (γ_i) from the beginning of the location choice stage, so these unobserved preferences influence their residential choice. This becomes a source of bias in estimating commuting costs if we estimate school preference without correcting for the selection into locations. For example, a household that values school safety more than other observably similar households would choose locations that assure higher location-based admissions probability for safer schools. This household would apply to only nearby schools because it already lives near its safer schools, but a model that does not correct this selection would mistakenly justify such behavior with a high commuting cost. Together with unobserved taste, households observe their idiosyncratic preference shocks over locations ($\eta_{i\ell}$) and form predictions on admissions probabilities to schools, $pr(\hat{c}_{ij}(\ell) \geq \bar{c}_j)$.[17]

At the beginning of the school choice stage, student i observes her preference shock over programs, ε_{ij}. To sum up, households know know unobserved tastes γ_i but not idiosyncratic shock ε_{ij} when deciding where to live. Once the assignment is realized, they know the exact assignment result j_i. The preference shock over outside options $\varepsilon_{i\vartheta}$ is realized at the enrollment-decision stage to rationalize the fact that 7.74% of students assigned to their top choice enroll in outside options. $\varepsilon_{i\vartheta}$ is either an income shock that affects households' affordability for private schools (Calsamiglia, Fu, and Güell, 2020) or charter school lotteries realized after the application stage is complete. The idiosyncratic shock ε_{ij} over assigned school does not change in the enrollment-decision stage.[18]

4.2 Household's Problem

Next, we describe the household's problem corresponding to the blue dots in Figure 4, which we solve backward.

Stage 4: Enrollment Residential locations and assignment results are set in previous stages. Given those, students decide whether to enroll in their assigned school, or the non-public option, or a public charter school to maximize their utility:

$$U_i^*(\ell_i) \equiv max\Big\{U_i(j_i(\ell_i), \ell_i; \varepsilon_{ij}), U_i^{\vartheta}(\vartheta^{np}; \varepsilon_{i\vartheta}), U_i^{\vartheta}(\vartheta^{c}; \varepsilon_{i\vartheta})\Big\}, \quad (8)$$

[17] The admissions cutoffs $\bar{c}_j$ households use at this stage are calculated using observed school application, which is a function of students' preference shocks over programs ε_{ij} that are realized in the next period. The large market assumption establishes the internal consistency—i.e., the admissions cutoffs are determined in the large market, and $\bar{c}_j$ are consistent estimators of those.

[18] An alternative model is such that students draw new shocks on assigned schools at the enrollment stage and the final shock is a weighted sum of the old and the new shock. However, it is impossible to tell to what extent the idiosyncratic shock ε_{ij} is time-invariant, since all other choices from the application stage are forgone except for j_i. That is, this alternative model would generate the same school application list and enrollment decision. Such model would have been possible if students had more than two options that are relevant at both the application stage and enrollment stage.

where j_i is the assignment outcome from the assignment stage and ℓ_i is the location chosen in the previous stage.

Stage 3: Assignment Students are passive as their assigned school j_i is determined by their priority score at each program and admissions cutoffs, given their ROLs from the previous stage. Mathematically,

$$j_i(\ell_i) \equiv f(\underbrace{\mathcal{ROL}_i(\ell_i)}_{\text{application list}}, \underbrace{c_{ij}(\ell_i;\rho)}_{\text{priority score}}, \underbrace{\bar{c}_j}_{\text{cutoff vector}}). \tag{9}$$

Stage 2: Application We assume that students submit an ROL following their true preference order up to their fallback options. The fallback option is the school a student is assigned to when rejected by all programs on her ROL—either pre-designated zoned school or an undersubscribed school in her school district. The middle school choice system in NYC uses the Deferred Acceptance algorithm, in which students can list as many schools as they want, which jointly renders truth-telling—ranking schools based on one's true preference order—a weakly dominant strategy (Gale and Shapley, 1962).

Stage 1: Residential Location Choice Given the solution in the subsequent period, household i chooses the location that solves

$$\max_{\ell} \underbrace{V_i(\ell;\eta_{i\ell})}_{\text{utility from location}} + \underbrace{\alpha^u \mathbb{E}_{\varepsilon_{ij},\rho,\varepsilon_{i\vartheta}} U_i^*(\ell)}_{\text{expected utility from enrolled school given location}}, \tag{10}$$

where $U_i^*(\ell)$ is the utility from enrolled school (Equation 8). This is is location dependent because locations decide commuting costs and admissions probabilities, and as a result, which school student i enrolls in. Households form an expectation over $U_i^*(\ell)$, since they do not know their idiosyncratic preference shocks over schools and outside options (ε_{ij} and $\varepsilon_{i\vartheta}$) as well as their lottery number ρ_i.

4.3 Equilibrium

To define the school assignment equilibrium, we extend the supply and demand characterization of Azevedo and Leshno (2016).

Definition 4.1. *An equilibrium is a pair of decisions $\{\ell_i, \mathcal{ROL}_i\}$ for each i and a vector of admissions cutoffs $\{\bar{c}_j\}_{j=1}^J$ where*

1. *Given cutoffs $\{\bar{c}_j\}_{j=1}^J$ and the first-stage choice ℓ_i, $\{\mathcal{ROL}_i\}$ is the school application list based on i's true preference order up to their fallback option.*

2. *Given $\{\bar{c}_j\}_{j=1}^J$, ℓ_i solves i's problem Equation 10 for each i.*

3. *Admissions cutoffs clear the market; i.e., $S_j \geq D_j(\{\bar{c}_{j'}\}_{j'=1}^J)$ for each $j \in J$. S_j is capacity of school j and $D_j(\{\bar{c}_{j'}\}_{j'=1}^J)$ is the aggregate demand for school j given the cutoffs $\{\bar{c}_{j'}\}_{j'=1}^J$.*

Aggregate demand for schools can be further simplified by using the fact that when students are truth-telling, the realized matching is stable—i.e., each student is matched to her favorite feasible school. Details are in Appendix C.

4.4 Discussion

Truth-telling We consider truth-telling a reasonable assumption in our context. There are well-known factors that make this assumption less plausible: (1) list-length restriction (2) limited consideration set (3) application cost. First, there is no list-length restriction in our setting. In a setting with this restriction (Luflade, 2018; Son, 2020), truth-telling is no longer a weakly dominant strategy when students really want to be assigned to some school (Haeringer and Klijn, 2009).[19] Second, students are given a customized list of all eligible schools with an average of about 30 schools. This stands in contrast to settings in which they have to construct a consideration set out of hundreds of options, where they are unlikely to consider all options in their choice set when deciding which school to apply to (Ajayi and Sidibe, 2020; Son, 2020). Third, both monetary and psychological application cost are relatively low. There is no application fee. Also, they can add one more school to their application list just by marking the ranking to the customized list that they received.

Even though assuming truth-telling is reasonable in our context, it is still a weakly dominant strategy (Artemov, Che, and He, 2017; Che, Hahm, and He, 2022). For example, "skipping the impossible" yields the same assignment results, and detecting impossible options is feasible given that each school's capacity and the number of previous year's applicants are public information. Instead of imposing truth-telling

[19] List-length restrictions were introduced in NYC's middle school choice in years more recent than our setting.

assumption, one can estimate the model based on stability (Fack, Grenet, and He, 2019; Agarwal and Somaini, 2020; Hahm and Park, 2022), which rely on assignment results rather than ranking strategies, in Appendix D. While imposing a weaker assumption, this estimation strategy loses the precision of estimates by focusing only on the assignment outcome instead of the full list.

Asymmetry in Utility from Location and School In our model, utility from locations includes unobserved amenities shared by households but no household-specific unobserved tastes. Meanwhile, utility from schools includes household-specific unobserved taste but no unobserved quality shared by households (Equation 4).

These modeling choices are largely driven by the motivation to obtain unbiased estimates of two key parameters—access-to-school preference α^u and commuting cost β^d—while keeping the estimation tractable. Access-to-school preference α^u will be biased if unobserved location amenities are correlated with access-to-school utility.[20] Meanwhile, households' sorting into locations based on household-specific unobserved school tastes biases commuting cost β^d.[21]

Utility from the Outside Option Utility from the outside option is not a function of school characteristics because of a lack of data on schools outside the system, especially non-public options. It can also be a function of location. By abstracting away from it, our location demand estimates might capture the unequal geographic distribution of outside schooling options. For example, locations with higher median household income would have more private schools nearby, and the estimated preference over neighbors' income in Section 5 might capture households' preference over geographic proximity to non-public options. We assume the geographical distribution of outside options does not change under the counterfactual scenario.[22]

[20] Moreover, we lack variation to identify household-specific unobserved tastes. For example, Bayer, McMillan, Murphy, and Timmins (2016) sets up a dynamic location choice model and uses the panel structure of the data to identify households' unobserved attachment to a specific location. Or, Barwick, Li, Waxman, Wu, and Xia (2021) constructs household-specific location choice set by leveraging that they observe when each household bought the house.

[21] For example, Allende (2019) and Abdulkadiroğlu, Pathak, Schellenberg, and Walters (2020) estimate unobserved school quality to obtain causal estimates on how much households value peer quality aside from other school factors such as the building quality.

[22] See, for example, Dinerstein and Smith (2021) to see how private schools' entry and exit decisions can be affected by public school policies.

5 Estimation Procedure and Results

5.1 Identification of Key Parameters

We discuss the identification of the two key parameters of the model: access-to-school preference α^u and commuting cost β^d. We also discuss the identification of price coefficient α^p.

The biggest concern regarding credibly identifying access-to-school preference α^u is to distinguish it from preferences on unobserved location amenities (ξ_ℓ). To this end, we use variation from our boundary discontinuity design. Similar to Section 3, the identification assumption is that the unobserved amenities are as good as random within a narrow buffer around a boundary. Meanwhile, the access-to-school utility sharply changes at a boundary, since 70% of schools give eligibility or higher priority to students from the same school district (Section 2), and there is marked heterogeneity in school characteristics across districts. So intuitively, seeing that households are more likely to live in the side of a district boundary with higher admissions probabilities to schools whose characteristics are more desirable (Section 3) would lead to a larger value of α^u.

To obtain an unbiased estimate of commuting cost β_i^d, we need to account for the fact that households choose locations based on their unobserved school demand γ_i. When students apply to nearby schools, we need to identify to what extent this is explained by commuting costs as opposed to households' residential sorting in order to be assigned higher priority by the schools they prefer. If residential sorting arises only from households' observed characteristics, we can obtain unbiased commuting costs by controlling for those characteristics in estimating school preference without fully modeling residential sorting. Thus, previous papers have assumed that idiosyncratic preference shocks and unobserved tastes over schools are independent of distances to school conditional on student observable characteristics—i.e., $(\varepsilon_{ij}, \gamma_i) \perp d_{l_ij}|Z_i$ (e.g., Agarwal and Somaini, 2018; Laverde, 2020). By modeling and jointly estimating location and school choice, we relax this assumption and allow an individual's unobserved type γ_i to be correlated with distances to schools—i.e., $\varepsilon_{ij} \perp d_{l_ij}|Z_i$.

Moreover, we need to identify unobserved type γ_i to correct for households' selection into locations based on it. Whereas the different applications of two observably identical students can be explained by either unobserved tastes (γ_i) or idiosyncratic preference shock (ε_{ij}), these two components can be disentangled for two reasons.

First, unobserved tastes are student-specific but the preference shock is independent across schools within each student. We observe students' full application lists. To what extent characteristics among the schools on a student's ROL are correlated helps to identify the unobserved type separate from the idiosyncratic shock (Bhat, 2000; Berry, Levinsohn, and Pakes, 2004). Second, while households choose residential locations knowing their unobserved taste, the idiosyncratic shock is realized after location choice. Among observably similar students, variation across residents of different locations pins down unobserved taste, while variation among residents from the same location are captured by the idiosyncratic preference shock.[23]

The final parameter to identify is the price coefficient, α_p, since housing cost (p_ℓ) is likely to be correlated with location unobserved amenities (ξ_ℓ). We instrument for housing cost of a location with the land use of other locations that are (1) 2 miles away from the location (2) but within 3 miles (Bayer, Ferreira, and McMillan, 2007; Barwick, Li, Waxman, Wu, and Xia, 2021; Davis, Gregory, Hartley, and Tan, 2021). Given a location, other locations that are far away are unlikely to share its unobserved amenities (exclusion restriction). However, the land use of other locations that are near enough to the location could affect its housing cost if people decide where to live among those locations (relevance restriction).

5.2 Estimation Procedure

Challenge 1: Granularity of Location With over 38,000 Census blocks in NYC, estimating the model at block level might decrease the precision of estimates by having too many parameters relative to the data (Dingel and Tintelnot, 2020).[24] But we still aim to estimate the mean utility of location ($\delta_\ell = W_\ell \alpha^{W0} + p_\ell \alpha^{p0} + \xi_\ell$) to account for the endogeneity of access-to-school and housing price (Berry, Levinsohn, and Pakes, 1995). To this end, we define neighborhoods—a unit of residential location—by merging Census blocks.

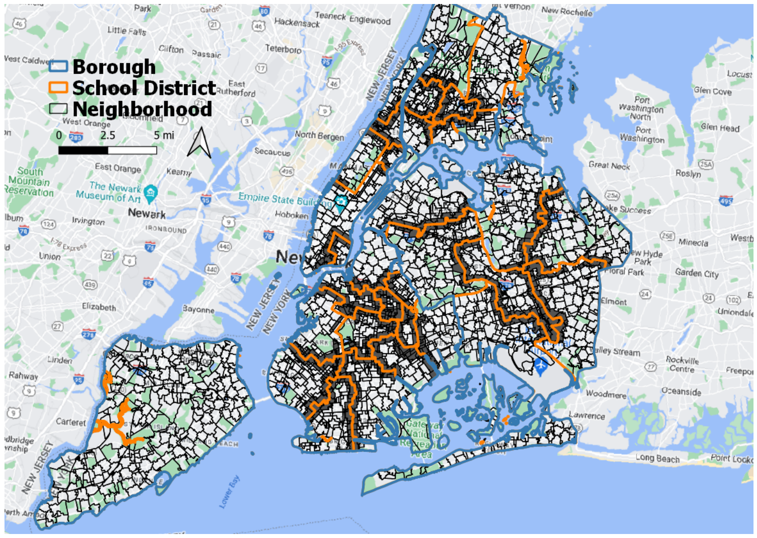
Two Census blocks are in the same neighborhood if they satisfy the following criteria. First, they are in the same cluster when we group Census blocks based on the distance to all schools using k-mean clustering, with k of 1,000. Second, they share the same location-based admissions probability to all schools. Third, they are either

[23] In principle, we can even divide unobserved taste into two components: individual-specific and realized at the location choice stage versus individual-specific and realized at the application stage.

[24] There are also many Census blocks with no middle-school-applying residents or housing transaction records during the time of the study.

both within the 0.2-mile buffer of a school district boundary or neither is.

Figure 5: Defined Neighborhoods



Note: We aggregate 38,798 Census blocks into 2,778 neighborhoods using the procedure described in Subsection 5.2. The darker shaded neighborhoods along school district boundaries (in orange) are those to which we apply BDD to identify access-to-school preference α^u.

With this procedure, we aggregate 38,798 Census blocks into 2,778 neighborhoods. In comparison, there are 2,165 Census tracts in NYC. Figure 5 shows the map of neighborhoods defined by this procedure. The darker shaded neighborhoods along school district boundaries (in orange) are those to which we apply BDD to identify access-to-school preference α^u.

Challenge 2: Computational Burden from Joint Estimation We aim to jointly estimate all stages of the model to address the selection into locations. Full information maximum likelihood (FIML) involves calculating a large Hessian matrix (Train, 2009), which renders the computation infeasible. See Appendix D for details on FIML.

To circumvent the computational burden, we employ the expectation-maximization algorithm with a sequential maximization step (ESM) proposed by Arcidiacono and Jones (2003).[25] In summary, the idea is to (1) reformulate the full information likelihood function into additive separable terms, each of which represents the likelihood of each stage; (2) update estimates of each stage; and (3) iterate the procedure until convergence.

[25] Dempster, Laird, and Rubin (1977) and Train (2009) show that solving the EM algorithm is identical to solving maximum likelihood estimation (MLE), and Arcidiacono and Jones (2003) prove the consistency of estimates with a multi-stage model.

The expectation function (reformulation) for the household i is a sum of the log of the likelihood for each stage weighted by the conditional probability of its being each unobserved type, given the school application and location choice observed in the data. Then we take the sum across i's expectation function.

$$\begin{aligned}\mathcal{E}(p,\gamma,\theta|\hat{q},\hat{\gamma},\hat{\theta}) &= \Sigma_i\Sigma_k q(k|x_i;\hat{q},\hat{\gamma},\hat{\theta}) log q_k \\ &+ \Sigma_i\Sigma_k q(k|x_i;\hat{q},\hat{\gamma},\hat{\theta}) log P^{LC}(x_i;\theta^{EC},\theta^{SC},\theta^{LC},\gamma_k) \\ &+ \Sigma_i\Sigma_k q(k|x_i;\hat{q},\hat{\gamma},\hat{\theta}) log P^{SC}(x_i;\theta^{SC},\gamma_k) \\ &+ \Sigma_i\Sigma_k q(k|x_i;\hat{q},\hat{\gamma},\hat{\theta}) log P^{EC}(x_i;\theta^{EC},\theta^{SC},\gamma_k)\,.\end{aligned} \tag{11}$$

$q(k|x_i;\hat{q},\hat{\gamma},\hat{\theta})$ is the conditional probability of being type k given data x_i, calculated using Bayes' rule. θ^{LC}, θ^{SC}, θ^{EC} are the set of location, school, and outside option preference parameters, respectively. γ_i is the unobserved taste, and q_k is the unconditional probability of each type k. $\theta = \{(\theta^{LC},\theta^{SC},\theta^{EC}),\{\gamma_k,q_k\}_k\}$ is the full set of parameters to be estimated. P^{LC}, P^{SC}, P^{EC} are the likelihood of location choice, school choice, and enrollment choice, respectively. Likelihood functions are presented in Appendix D.

Then we update the guess on each element of θ sequentially by maximizing each line of the expectation function. Starting from an initial guess, we iterate the updating process until the guess of θ converges. We used squared extrapolation methods (see Varadhan and Roland, 2008) to make convergence faster. See Appendix D for the cookbook of the iteration process.

We update θ^{SC}, θ^{EC}, and γ using maximum likelihood estimation to obtain efficient estimates of γ by exploiting full information in application lists.

Meanwhile, we update θ^{LC} using method of moments estimation to deal with the endogeneity of price and access-to-school utility (Berry, Levinsohn, and Pakes, 1995). Location preference parameters include those that govern heterogeneous preferences $(\alpha^{Wz},\alpha^{pz})$; common preferences $(\alpha^{W0},\alpha^{p0})$; and the access-to-school preference α^u. To estimate $(\alpha^{Wz},\alpha^{pz})$, we match the first-order condition of location choice likelihood P^{LC} (presented in Appendix D) with respect to α^{Wz} and α^{pz},

$$\underbrace{\Sigma_i W_{\ell_i} z_i^r}_{\text{cov. of } W \text{ and } z \text{ in the data}} = \underbrace{\Sigma_i \Sigma_k q(k|x_i; \hat{q}, \hat{\gamma}, \hat{\theta}) \Sigma_\ell P^{LC}(\ell; \hat{\theta}^{SC}, \hat{\theta}^{EC}, \theta^{LC}, \hat{\gamma}_k) W_\ell z_i^r}_{\text{predicted cov. of } W \text{ and } z}, \tag{12}$$

$$\underbrace{\Sigma_i p_{\ell_i} z_i^r}_{\text{cov. of } p \text{ and } z \text{ in the data}} = \underbrace{\Sigma_i \Sigma_k q(k|x_i; \hat{q}, \hat{\gamma}, \hat{\theta}) \Sigma_\ell P^{LC}(\ell; \hat{\theta}^{SC}, \hat{\theta}^{EC}, \theta^{LC}, \hat{\gamma}_k) p_\ell z_i^r}_{\text{predicted cov. of } p \text{ and } z}, \tag{13}$$

where z^r is each element of observed household characteristics Z—e.g., the minority dummy.

To obtain the remaining parameters $(\alpha^{W0}, \alpha^{p0})$, α^u, we first search the mean utility of location $\delta_\ell = W_\ell \alpha^{W0} + p_\ell \alpha^{p0} + \xi_\ell$ that satisfies the first-order condition of P^{LC} with respect to δ_ℓ, [26]

$$\underbrace{\Sigma_i \mathbb{1}(\ell_i = \ell)}_{\text{observed share}} = \underbrace{\Sigma_i \Sigma_k q(k|x_i; \hat{q}, \hat{\gamma}, \hat{\theta}) P^{LC}(\ell; \hat{\theta}^{SC}, \hat{\theta}^{EC}, \theta^{LC}, \hat{\gamma}_k)}_{\text{predicted share}}, \ \forall \ell. \tag{14}$$

Finally, we get $(\alpha^{W0}, \alpha^{p0}, \alpha^u)$ by targeting the following conditions"

$$E(\xi_\ell \hat{p}_\ell^{IV}) = 0 \tag{Price IV}$$

$$E(\xi_\ell \mathbb{1}(\text{right side of BD})_\ell | \text{BD}_\ell, \mathbb{1}(\ell \in \mathcal{B}(\text{BD}_\ell; 0.25\text{mi.})) = 0. \tag{BDD IV}$$

where $\hat{p}_\ell^{IV}$ is a vector of other observed location characteristics W_ℓ and price IV. $\mathcal{B}(BD_\ell; 0.25)$ is the buffer around each boundary BD with a radius of 0.25 mile. The procedure consists of the outer loop that searches parameters that satisfy Equation 12, 13, Price IV, and BDD IV and the inner loop that searches δ_ℓ that satisfy Equation 14. We present price IV regression results in Appendix D.

5.3 Estimation Results

Demand Estimates Estimates in Table 2 have expected signs. Households prefer locations with higher access-to-school utility (EU, 1.419) and lower housing costs. They prefer schools that are higher achieving and safer. There is homophily (i.e., pref-

[26]The process is accelerated by Newton's nonlinear root-finding algorithm. We thank Jean-François Houde for sharing his code.

erence for one's same race and FRL status) in both location and school preferences.

Willingness to pay To better interpret estimates, we calculate households' willingness to pay (WTP) for school and location characteristics in Table 3. WTP for a one-unit increase in location characteristics for household i is given by $\frac{\alpha_i^W}{\alpha_i^p}$. WTP for a one-unit increase in school characteristics of *all* schools in a school district is the sum of $\frac{\alpha^u}{\alpha_{p_i}} \frac{\partial EU_{i\ell}}{\partial X_j}$ across school js in a school district of location ℓ. This is household WTP to ensure the increase in characteristics of the assigned school[27] and is a function of location; thus we take the average across locations. We further convert WTPs in monetary terms by multiplying them by \$1,366, the mean of median gross monthly rent at the Census tract from 2014 5-year ACS estimates.

For some characteristics, households uniformly agree on what makes a location or a school more desirable. Both the 25th and 75th percentiles of households are willing to pay a positive amount for an increase in the median income of neighbors, mean test score of schools, and safety of schools.[28] For other characteristics, there is marked heterogeneity in preferences. For an increase in the minority share among neighbors or school peers, some households are willing to pay a positive amount while others must be compensated to stay indifferent.[29]

Next, we present households' WTP for a reduction in commuting time to school. Commuting time has been used as the numéraire in previous studies on public-school choice (Agarwal and Somaini, 2018), and we convert it into the monetary term using housing cost. A median household is willing to pay \$19 (=384/20 days) per school day to reduce commuting to school by 50 minutes a day.

We view our WTP estimate to capture various challenges that middle school students face during school commuting. For example, parents answer a survey by Sattin-Bajaj and Jennings (2022) that safety on the journey to a school is a main consid-

[27] $\frac{\partial EU_{i\ell}}{\partial X_j}$ can be simplified to $\beta_i^X Prob_{ij}(\ell) Prob(j \in \mathcal{J}_i(\ell;\rho_i))$ where $Prob_{ij}(\ell)$ is the probability of j's being the most preferred feasible option for student i when she lives in ℓ. $Prob(j \in \mathcal{J}_i(\ell;\rho_i))$ is the probability of j being i's feasible choice when she lives in ℓ.

[28] WTP for a one-standard-deviation increase in schools' test score is 11.3%. Ours is slightly higher than the range reported by previous papers (3%-10%) that study households' WTP for a test score increase in *one* school such as a zoned school or a charter school (Black, 1999; Bayer, Ferreira, and McMillan, 2007; Zheng, 2022). In contrast, we consider a test score increase for *all* schools in a district.

[29] Bayer, McMillan, Murphy, and Timmins (2016) estimate that for a 10-percentage-point increase in the fraction of White neighbors, an average White family in the San Francisco Bay Area is willing to pay \$2,428 annually in 2000 dollars from their dynamic location choice model, and \$1,901 from their static model. Our estimate for a similar scenario is \$1,740. ($= 0.7 \times 704 \times 12 \times (0.1/0.34)$), with a 0.7 adjustment to 2000 dollars using CPI (source: BLS CPI New York-Newark-Jersey City area)

Table 2: Demand Estimates

	Main	Additional Effects				
	Type1	Type2	Type3	Black/Hisp	FRL	Low-achieving
Panel A: Neighborhood Demand						
log(SalesPrice)	-2.039	-	-	-0.015	-0.048	-0.035
	(1.383)			(0.283)	(0.360)	(0.272)
Frac. Black or Hisp.	-3.459	-	-	3.933	-0.119	0.208
	(2.025)			(1.727)	(1.671)	(1.672)
log(Med. HH Income)	2.848	-	-	-0.161	-1.201	-0.303
	(1.991)			(1.545)	(1.891)	(2.141)
Med. Time to Work (hr)	17.676	-	-	-27.016	13.242	15.394
	(359.156)			(278.872)	(402.963)	(474.360)
Med. Time to Work^2 (hr)	-13.456	-	-	17.269	-8.257	-10.207
	(247.912)			(191.749)	(278.871)	(323.379)
EU	1.419	-	-	-	-	-
	(1.276)	-	-	-	-	-
Panel B: School Demand						
Mean test score	0.121	0.256	0.187	0.134	-	-0.253
	(0.052)	(0.074)	(0.469)	(0.029)	-	(0.028)
Frac. Black or Hisp.	-1.722	-0.501	0.159	1.958	0.117	0.216
	(0.612)	(1.118)	(5.442)	(0.417)	(0.281)	(0.223)
Frac. FRL	-0.771	-0.356	1.020	-0.540	0.882	-0.162
	(0.865)	(1.041)	(5.097)	(0.429)	(0.220)	(0.213)
Non-safety	-0.059	0.003	0.027	0.018	-	-
	(0.008)	(0.012)	(0.057)	(0.010)	-	-
Commuting Cost (mi.)	0.221	1.085	9.136	-0.084	-0.010	-0.046
	(0.032)	(0.054)	(1.757)	(0.045)	(0.021)	(0.014)
Prob.	0.352	0.631	0.017	-	-	-
	(0.127)	(0.083)	(0.055)	-	-	-
Panel C: Outside Option						
Non-public	-1.225	-	-	-0.136	-0.800	-0.200
	(0.174)			(0.135)	(0.125)	(0.134)
Public Charter	-3.767	-	-	1.883	0.173	-0.059
	(0.256)			(0.207)	(0.152)	(0.124)

Note: Standard errors in parentheses are calculated from 75 bootstrapped samples. Columns are for students' heterogeneity and rows are for school and neighborhood characteristics. FRL stands for free or reduced-price lunch eligibility. The fastest driving distance to a school is calculated using the Open Route Service. School non-safety measure is constructed by running a principal component analysis on crime incidence of different categories at each school building. See Appendix D for details.

Table 3: Willingness to Pay

	Std of Var.	WTP		
		p25	median	p75
Panel A: Neighborhood Characteristics				
Frac. Black or Hisp.	0.341	-704	32	71
log(Med. HH Income)	0.429	303	379	595
Med. Time to Work (min, daily)	11.52	-236	-181	-114
Panel B: School Characteristics				
Time to School (min, daily)	51	-752	-384	-171
Mean test score	1.002	89	154	238
Frac. Black or Hisp.	0.28	-355	-25	11
Frac. FRL	0.173	-98	-83	-53
Safety	4.228	94	121	149

Note: The unit of willingness to pay is the mean of median gross monthly rent at the Census tract from 2014 5-year ACS estimates, $1,366. We use the standard deviation of distance to the assigned schools across students. For other school characteristics, we calculate the standard deviation across schools. FRL stands for free or reduced lunch eligibility. The fastest driving time to school is calculated using the Open Route Service. School safety is constructed by running a principal component analysis (PCA) on crime incidence of different categories at each school building. Appendix D has more details on the PCA result.

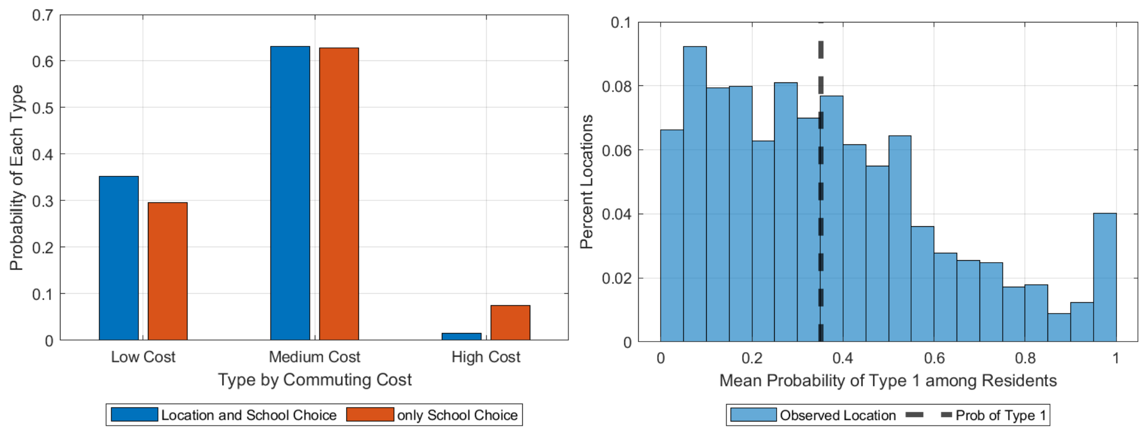
eration factor for school application. Such concern of parents arises because many students commute to schools by themselves by public transportation or on foot. Middle school students are eligible for school bus service only in their first year if their schools offer any. Moreover, we calculate from the 2017 National Household Travel Survey that at least 70% of students in our sample commute to schools without any adult accompanied.[30] Finally, our WTP estimate is high to be interpreted as the forgone earning of middle school students. Adult commuters' value of commuting time is known to be 50%-70% of their hourly wage (Parry and Small, 2009; Purevjav, 2022), and the minimum wage in New York in 2015 was $9.

Overestimation of Commuting Cost Commuting cost is overestimated when we ignore households' residential sorting. We estimate a different version of the model without location choice (estimates are presented in Appendix D), and find that commuting cost is overestimated by 15% on average (mean β_i^d is -0.97 in a model with both location and school choice, and -1.11 with only school choice). Figure 6 describes what leads the model without endogenous location choice to overestimate commuting

[30]To be accurate, 70% of middle school students residing in the NY-NJ-PA area, which is the finest geography available, commute to schools without any adult accompanied.

cost.

Figure 6: Residential Sorting on γ and the Overestimation of Commuting Cost



(a) Mean Probability of Unobserved Type (b) Location Sorting on Unobserved Type

Note: Panel (a) presents the probability of each unobserved type under each model. Panel (b) shows the mean probability of residents' being type 1 across residential locations.

Panel (a) shows probabilities of types (q_k) from a model with both location and school choice compared with those from a model with only school choice. In the latter, we over-classify students into the high commuting-cost type, since we rationalize households' applying to schools nearby as due only to high commuting costs, as opposed to households' residential sorting based on unobserved school taste γ_i. Panel (b) plots the mean probability of being type 1 among residents across locations. For each student i, we calculate the probability of her being type 1 based on how well her location and school choices can be justified by being type 1 relative to other types (Bayes' rule). In the absence of residential sorting based on unobserved type, the mean probability of being type 1 among residents of a location should be similar across all locations. In contrast to this, some locations have zero type-1 students while others have many type-1 students, which implies sorting based on unobserved type.[31]

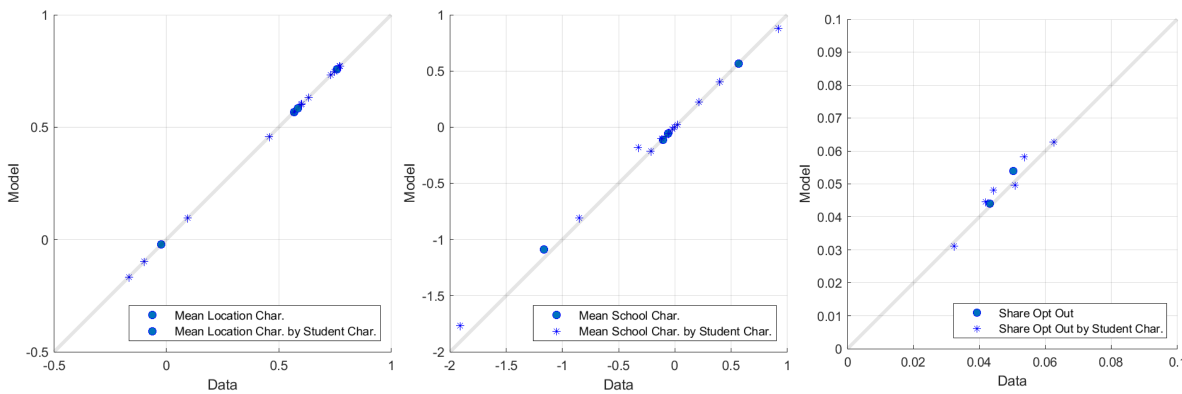
Model Fit We simulate choices using our estimates to validate whether our model can replicate the data patterns. To minimize the idiosyncrasies coming from preference shocks and the lottery number, we present the average over 100 simulations.

Figure 7 plots the simulated and observed moments from school and location

[31] There is also an idiosyncrasy coming from a finite sample. Figure D.4 compares the distribution from the data and that from a simulation in which we allocate households randomly across locations.

choice. Moments include the mean observable characteristics of chosen options and the correlation between students' characteristics and those of their chosen options. Unsurprisingly, our targeted moments from location choice are well aligned with the 45-degree line. Meanwhile, even though we do not target school choice moments, and rather estimate school preference parameters via MLE, our simulated moments of school choices are close to data moments. Table D.14 presents the numbers plotted in Figure 7.

Figure 7: Model Fit



(a) Location Choice (b) School Choice (c) Enrollment

Note: We take the average over 100 simulations with draws of η, ϵ, and the lottery number. Moments include the mean observable characteristics of chosen options and the correlation between students' characteristics and those of their chosen options. In panel (b), we focus on students' first choice. In panel (c), we present the fraction of students who choose each outside option.

6 Source of School Segregation

In this section, we use model estimates to identify the sources of school segregation. Even with an extensive school choice system in place, NYC middle schools are highly segregated.[32] There are also large differences in academic achievement across these segregated schools. In the 2014-15 academic year, classmates of minority students (in their assigned schools) had standardized test scores than were one standard-deviation lower than the classmates of non-minority students. In this section, we explore which components of the model explain the cross-racial gap in the test scores of students' peers in their assigned schools

[32]In terms of racial composition, 77% of Black and Hispanic students attend schools that enroll less than 10% of White students, while only 11% of White students and 43% of Asian students attend schools that enroll less than 10% of White students (Cohen, 2021).

In Table 4, we investigate to what extent the cross-racial gap in test scores is explained by the following components of the model: access-to-school preference, heterogeneity in preference over location characteristics, and that over school characteristics. Column (1) in Table 4 presents the cross-racial gap under the status quo; minority students attend schools with lower test scores by one student-level standard deviation than their non-minority peers.

Table 4: cross-racial Gap in Coassigned Peers' Test Score

Status Quo Racial Gap	(1) Racial Gap Explained by: Access-to-school Preference	(2) Heterogeneous Preference over: Other Location Characteristics	(3) Heterogeneous Preference over: School Characteristics
-1.048	-0.312	-0.466	-0.182

Note: The cross-racial gap is the difference in test scores of the schools students attend for minority and non-minority students. We shut down each channel for one household one at a time. In column (2), we impose $\alpha^u = 0$. In column (3), we impose $\alpha^{WZ} = \alpha^{pZ} = 0$, In column (4), we impose $\beta^{XZ} = \beta^{dZ} = 0$.

Next, columns (2) and (3) demonstrate that residential sorting is the main driver of school segregation, which is in line with past studies (Laverde, 2020; Monarrez, 2020). We further break down what part of the gap is explained by residential sorting based on access to school (column (2)) versus sorting based on other location amenities (column (3)). Column (2) demonstrates that 31% of the gap observed in the data is explained by residential sorting based on access-to-school utility. In this scenario, we shut down residential sorting based on access to school ($\alpha^u = 0$). Thus, households choose locations as if they do not know that the locations chosen determine commuting costs and location-based admissions probabilities to schools. The cross-racial gap in this scenario comes from households' heterogeneous preferences over location characteristics other than access to school and those over school characteristics.

In columns (3) and (4), we investigate the role of preference heterogeneity. Column (3) shows that households' heterogeneous location preferences play a key role in generating school sorting. We shut down heterogeneous preferences over location characteristics and price by setting $\alpha^{WZ} = \alpha^{pZ} = 0$; thus residential sorting is only based on access to school. This scenario explains 46% of the gap.[33] In column (4), we shut down heterogeneous preferences over school characteristics by setting $\beta^{XZ} = \beta^{dZ} = 0$, so that households choose locations and schools as if they have perfect consensus over

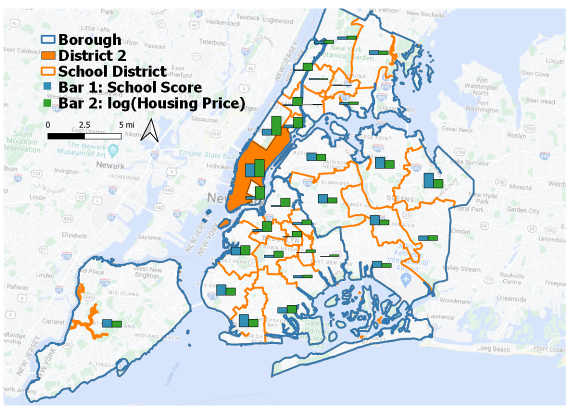
[33] This largely comes from heterogeneous preferences over location characteristics rather than price. Shutting down only the heterogeneous preference over housing costs reduces the gap by only 0.009.

what makes a school desirable, even though they disagree on what makes a location desirable. This explains 18% of the cross-racial gap.

7 Citywide Access to Highest-achieving Schools

NYC middle schools are intensely segregated (Cohen, 2021; Idoux, 2022), and many believe that location-based priorities are the main cause. The city has long acknowledged this issue and proposed plans to relax location-based priorities, many of which have triggered heated debate among parents, students, and educators.[34]

Figure 8: District 2 Characteristics



(a) Location of District 2

	(1) District 2	(2) Others
Panel A: School Characteristics		
Mean z-score	1.124	-0.155
Safety	-0.087	-1.145
Share Minority	0.320	0.723
N of Schools	24	646
Panel B: Neighborhood Housing Price		
Unit Price (1K)	2,606	464
N of Neighborhoods	81	2,057

(b) District 2 Characteristics

Note: District 2 is the shaded area in the figure. School score is the mean of z-scores among enrolled students from the NYS standardized Math and Language test. Housing price is the mean price of residential units sold in 2013-14 located in each school district. Safety is a composite of crime incidences of different categories at the school building. Minorities include Black and Hispanic.

We evaluate a scenario in which we introduce purely lottery-based admissions to schools in School District 2. The district is located in lower Manhattan and has been at the center of ongoing policy debate regarding whether to retain location-based admissions rules.[35] In District 2, the average housing cost is about six times higher

[34] For example, a plan to scrap all location-based priorities for high schools was canceled due to pushbacks from parents (Russo, Barbara , Zoned High School Options for NYC Students Will Remain in Place, NY Metro Parents, December 14 2021). Meanwhile, smaller plans have been implemented; For example, starting in the 2019-2020 academic year, Bronx middle schools have been open to all students in the Bronx (Zingmond, Laura, Bronx Middle School Best Tets, InsideSchools, October 20 2020).

[35] Shapiro, Eliza, N.Y.C. to Change Many Selective Schools to Address Segregation, the *New York Times*, December 18, 2020

than other districts, and the mean test score for its schools is more than 1.2 standard deviations higher than those in other districts (Figure 8).

We compare the status quo, in which we simulate households' location and school choices under the current admissions rules, with the following scenarios in which we scrap all admissions criteria—both location-based priority rules and academic screening—for schools in District 2.

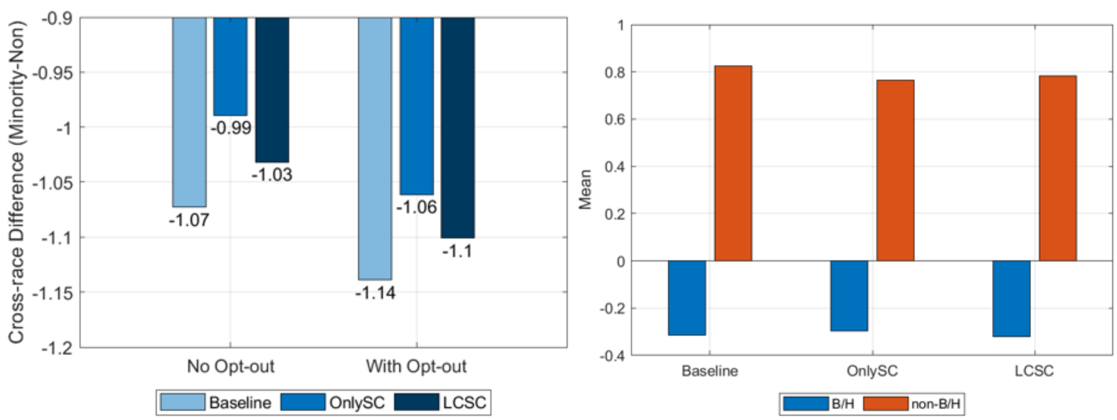
1. **OnlySC + No Opt-out**: Residential locations under status quo are fixed. We report the characteristics of the schools students are assigned to.
2. **LCSC + No Opt-out**: Households reoptimize residential locations. We report the characteristics of the schools students are assigned to.
3. **OnlySC + With Opt-out**: Residential locations under status quo are fixed. We report the characteristics of the schools students are enrolled in, excluding those who opt out.
4. **LCSC + With Opt-out**: Households reoptimize residential locations. We report the characteristics of the schools students are enrolled in, excluding those who opt out.

We solve the new equilibrium admissions cutoffs under the policy to address oversubscription to popular schools, especially District 2 schools. The main outcome of interest is the cross-racial gap in the characteristics of coassigned or coenrolled school peers, which we interpret as the measure of inequity or school segregation. In **With Opt-out** cases, we calculate the mean characteristics of students who enroll in each school, excluding those who choose outside options.

We predict the effects of a policy that targets only one cohort of middle school applicants. Thus, we assume that the housing market can absorb changes in the demand of households with middle school applicants, who account for only 3% of the population. We also assume that school characteristics are invariant under a new policy. Furthermore, we compare the distribution of households across schools and residential locations in a steady state, since we do not model moving costs.

Cross-racial Gap in Peer Characteristics Figure 9 shows the gap in coassigned or coenrolled peers' test scores between minority and non-minority students in each scenario. While the reform narrows the cross-racial gap in peer test scores, households' location choices dampen such effect. The y-axis in panel (a) is the cross-racial difference school peers' standardized test scores. The policy closes the cross-racial gap in coassigned peers' test scores from 1.07 to 0.99, thus approximately 7%, if households' residential locations were fixed (**No Opt-out, OnlySC**).[36] However, when households reshuffle across locations, the effect reduces to 3.3% (**No Opt-out, LCSC**). The cross-racial gap in assigned schools (**No Opt-out**) is always smaller than that in enrolled schools (**With Opt-out**). But the effects of the policy and households' endogenous location choices on the cross-racial gap in enrolled schools are similar to those on the cross-racial gap in assigned schools. In panel (b), we present the mean of coenrolled peers' test scores by minority and non-minority students. It shows that the policy closes the cross-racial gap both because non-minority students enroll with lower-achieving peers and minority students enroll with higher-achieving peers.

Figure 9: Cross-racial Gap in School Characteristics



(a) Gap in Peers' Test Score (1 std) (b) Mean of Peers' Test Score (1 std)

Note: Panel (a) shows the difference in mean test scores of coassigned/coenrolled peers between Black/Hispanic and other students. Panel (b) shows the mean test score of coenrolled peers for each group separately. We use z-scores from the NYS standardized Math and Language test.

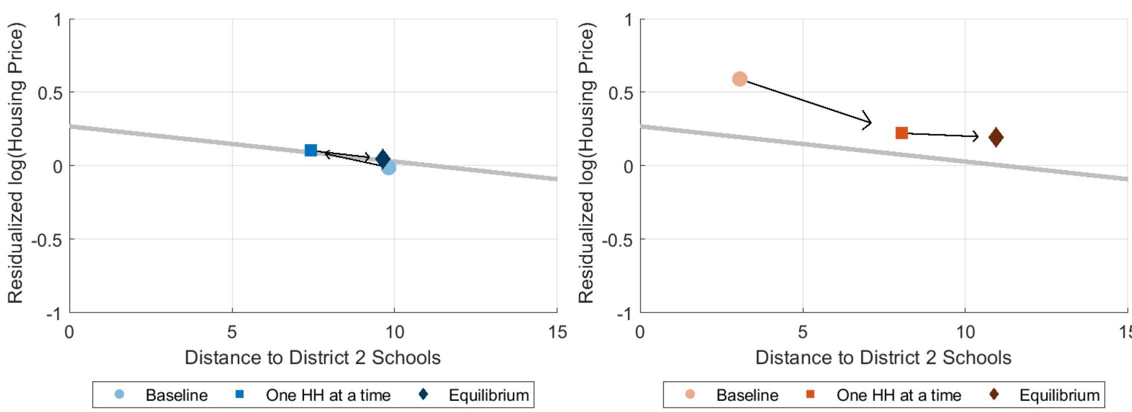
Location Choice Patterns Next, we delve into households' location choices to understand how those dampen the equity impact of the policy. The key is that locations decide on commuting costs as well as location-based priorities, which together determine access-to-school utility of locations. Under the status quo, households have positive admission chances to District 2 schools only when they reside in District 2. Hence, utility from access to District 2 schools differs only by whether a location is either within or outside the district. On the other hand, the policy equalizes admissions

[36]Zooming in District 2 schools, the cross-racial gap reduces from 0.35 standard deviation to 0.15 standard deviation, thus the gap reduces by 57% among students who are assigned to District 2 schools.

probability to District 2 schools across locations. Hence, locations differ in utility from access to District 2 schools by their proximity to District 2. Standing in contrast to the status quo, locations outside District 2 have different levels of utility from access to District 2 schools from one another.

These changes in access-to-school utility result in a different reoptimization in location choice patterns among households who live in District 2 under the status quo (=*D2 residents*) and others (=*Non-D2 residents*). We first present the location choice patterns of these two groups in a scenario in which we introduce purely lottery-based admissions to District 2 schools to *one household at a time*. In this scenario of *one household at a time*, a given household does not expect other households to modify their behavior in response to the policy change.

Figure 10: Location Choice Patterns



(a) Non-D2 Residents

(b) D2 Residents

Note: Panel (a) illustrates the location choices of *Non-D2 residents*, and (b) of *D2 residents*. The x-axis is the average distance to schools in District 2 from locations. The y-axis is the residualized log sales price. Each dot shows the median characteristics of locations chosen by *Non-D2 residents* and *D2 residents* who change locations under the policy, respectively. In each panel, we plot location choice patterns when we grant citywide access to District 2 schools to one household at a time and when we grant citywide access to all households. For the latter, we solve equilibrium admissions cutoffs.

In Figure 10, the y-axis is the demeaned log of housing price and the x-axis is the average distance to schools in District 2 from each location. Each dot describes the mean characteristics of locations chosen by *Non-D2 residents* (panel (a)) and *D2 residents* (panel (b)).

In the *one household at a time* scenario, *Non-D2 residents* relocate closer to District 2 at the expense of higher housing costs (panel (a)). While the policy makes them eligible to apply to and enroll in District 2 schools, such an option is not attractive when they stay in their baseline locations due to the high commuting cost to District 2 schools. Meanwhile, *D2 residents* choose locations with lower housing costs, but farther from District 2 schools (panel (b)). Purely lottery-based admissions make it no longer necessary to live in District 2 to ensure positive admissions probabilities to District 2 schools.

In equilibrium, the location choice behaviors of *Non-D2 residents* are largely muted, while those of *D2 residents* are reinforced. This is because citywide access to District 2 schools induces applicants from a broader area, and thus the admissions chances to District 2 schools are lower from each household's point of view. This makes choosing locations nearer to District 2 by *Non-D2 residents* less attractive and choosing locations farther from District 2 by *D2 residents* more attractive.

Connection between Location Choice and Peer Characteristics Table 5 reveals the link between households' location choice and the school desegregation effect of the policy. *Non-D2 residents*' spatial reshuffling narrows the cross-racial gap in school characteristics. For example, by relocating, minority *Non-D2 residents* are assigned to schools with a 13.7-percentage-point lower minority share. This largely comes from their choosing locations nearer to District 2 schools and more actively applying to and enrolling in those schools.

Table 5: From Location Choice To School Assignment

	(1)	(2)	(3)	(4)
	Non-D2 Resident		D2 Resident	
	Non-Minority	Minority	Non-Minority	Minority
Share	33.53%	62.81%	2.72%	0.94%
Panel A: Location Choice when Granting Citywide Access to One HH at a Time				
Change Location under New Policy?	0.220	0.176	0.592	0.511
Conditional on Changing Location:				
Δ Frac. Minority of Assigned School	0.107	-0.137	-0.150	-0.085
Δ Mean Score of Assigned School	-0.204	0.181	0.273	0.095
Δ Frac. Minority of Neighborhood	0.047	-0.216	0.147	0.344
Panel B: Location Choice in Equilibrium				
Change Location under New Policy?	0.136	0.092	0.974	0.936
Conditional on Changing Location:				
Δ Frac. Minority of Assigned School	0.022	-0.048	-0.186	-0.023
Δ Mean Score of Assigned School	-0.211	0.030	0.281	-0.099
Δ Frac. Minority of Neighborhood	0.004	-0.042	0.130	0.416

Note: Minority includes Black or Hispanic. *D2 residents* are those who reside in one of the locations in District 2 under the status quo. Each column shows the mean of variables for each group.

The location choice patterns of *D2 residents* stand in contrast to those of *Non-*

D2 residents; their spatial reshuffling dampens the equity impact of the policy. They seek locations that come with a secured seat in higher-achieving and lower-minority schools, and the purely lottery-based admissions to District 2 schools make locations within District 2 less attractive. Instead, they choose locations where location-based admissions are kept in place. By doing so, they are assigned to schools with a 15 percentage point lower minority share.

Previously, we have shown while relocation motives of *Non-D2 residents* are muted in equilibrium those of *D2 residents* are reinforced (Figure 10). Then, Table 5 shows relocation of *Non-D2 residents* amplifies the equity impact of the policy, while that of *D2 residents* dampens the impact. Combined together, endogenous location choices dampen the effect of the policy on the cross-racial test score gap, as depicted in Figure 9.

Commuting Distance to School and Welfare Another widespread concern about relaxing the importance of location-based admissions priority is that students would have to commute longer distances. In Table 6, we present average commuting distances to schools under each scenario by minority and non-minority students. We also present the change in welfare, which is a number that summarizes various changes in outcome induced by the counterfactual policy.[37]

Table 6: Effect on Commuting Distance and Welfare

	(1)	(2)	(4)	(5)	(6)	(7)	(8)	(9)
	Distance to school		Δ Welfare (% Housing Cost)					
	Non-Minority	Minority	Non-Minority	Minority	Overall			
	Mean	Mean	Mean	Mean	Mean	p50	Sum Loss	Sum Gain
Baseline	1.450	1.595	-	-	-	-	-	-
OnlySC	1.728	1.645	-0.074	-0.006	-0.030	0.012	894.620	591.450
LCSC	1.738	1.769	0.005	0.045	0.031	0.029	264.270	572.680

Note: Minority includes Black or Hispanic. *D2 residents* are those who reside in one of the locations in District 2 under the status quo. The fastest driving distance between a school and a Census block is calculated using Open Route Services. Welfare is measured by exante utility (Equation 10 at the chosen location), which we we convert into log housing cost. We present the difference in welfare under the policy relative to the baseline scenario.

Commuting distances increase for both minority and non-minority students, so

[37] While the welfare measure is a good summary of various changes, we might want to use caution in interpreting this. This is because our demand estimates might not represent households' true preferences, even though they capture how households make location, school, and enrollment choices. Former studies have documented various types of friction in location and school choices such as limited information, limited attention, and even discrimination by landlords or schools (Christensen and Timmins, 2018; Luflade, 2018; Allende, Gallego, and Neilson, 2019; Son, 2020; Christensen, Sarmiento-Barbieri, and Timmins, 2020; Ferreira and Wong, 2020).

the gap is decided by which group experiences a larger increase. While the policy narrows the cross-racial gap in commuting distance under **OnlySC**, households' endogenous location choice partially undoes this effect (**LCSC**).

Next, we calculate the change in welfare, which we measure with the ex-ante utility (Equation 10) at each household's optimal location. We convert the utility into percentage housing cost for ease of interpretation. Since we restrict households from reoptimizing locations, **OnlySC** mechanically gives lower welfare in comparison to the other two scenarios.

Under **OnlySC**, both minority and non-minority students experience a decrease in welfare on average. This comes from a large decrease in welfare among *D2 residents*, who cause the distribution of welfare change to skewed to the left. Indeed, a median household experiences an increase in welfare by 1.2% of housing cost. The sum of welfare losses is greater than the sum of gains by 300% of housing cost. This suggests while the policy might be approved by the voting among these households, it might face harder pushback from households who lose.

In the long run, where households adjust their locations (**LCSC**), both the average and the median household experience welfare gain, by 3.1% and 2.9%, respectively. The benefit is largely concentrated among minority students (a 4.5% increase), who obtain eligibility to District 2 schools while living in affordable locations. We consider this welfare gain an upper bound given that we assume away from moving cost.

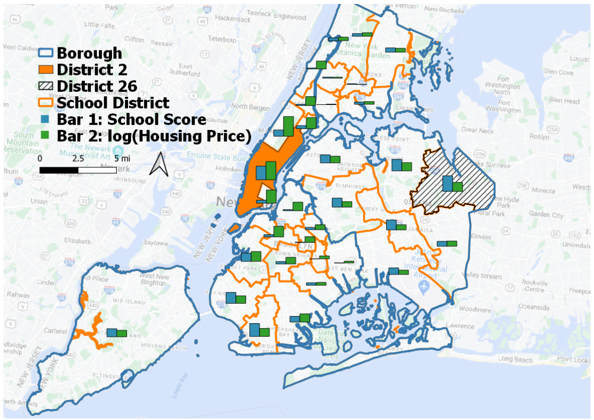
Other Margins The counterfactual policy has effects that go beyond changing the cross-racial test score gap, which we briefly discuss here. First, residential segregation, which we measure with an entropy-based segregation measure (Appendix E) decreases by 1.5% by race and 10% by income. Second, the policy increases the mean score of peers even among the lowest-performing minority students, an effect that is also dampened by households' location choices. There is a 6.5% increase in coassigned peers' test scores in **OnlySC** and a smaller effect (4% increase) in **LCSC**.

Other Policy Plans Next, we discuss the impact of another plan to introduce purely-lottery based admissions to District 26 schools. District 26 is located in upper Queens and features the highest mean test score of schools. The average housing unit price was $614,000, which is about a quarter of the housing price in District 2.

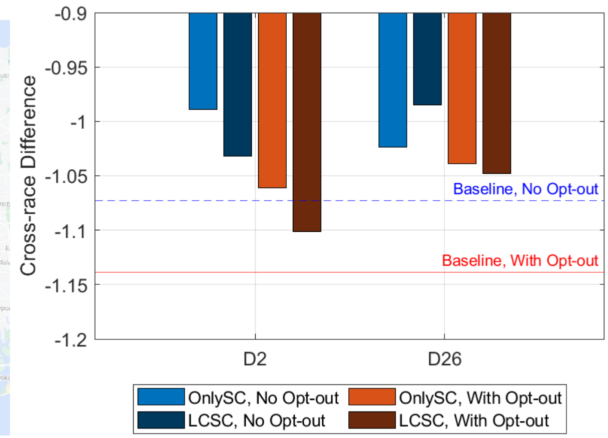
Figure 11 presents the characteristics of District 26, and how the cross-racial gap

Figure 11: Lottery-based Admissions to District 26 Schools

(a) Location of District 26



(b) Gap in Peers' Test Score (1 std)



Note: Panel (a) shows the location of District 26. Panel (b) shows the difference in mean test scores of coassigned/coenrolled peers between Black/Hispanic and other students. We use z-scores from the NYS standardized Math and Language test. Under **D2**, we introduce purely-lottery based admissions to schools in school District 2. Under **D26**, we target District 26.

changes across scenarios. First, whether and how households' endogenous location choices change the equity impact of the policy varies across policies. Focusing on **No Opt-out** scenarios, while endogenous location choices dampen the effect of the lottery-based admissions to District 2 schools by half, it amplifies the effect of the policy targeting District 26. This is because, in the latter scenario, minority households' endogenous location choice responses to shorten commuting distances to District 26 schools dominate non-minority households' location choices to get away from the policy. The lower housing cost of District 26 relative to District 2 is the key reason.

Second, the comparison across policies changes depending on if we consider households' endogenous location choice or not. Focusing on **No Opt-out** scenarios, while targeting District 2 schools seems more effective in reducing the cross-racial gap when we take residential locations as given, households' endogenous location choices in response make targeting District 26 more effective.

Lastly, households substitute between opting out to outside schooling options and reoptimizing their residential locations. While opt-out plays a minor role when we target District 2 schools, its role is more pronounced when we target District 26 schools. This is because households who live in District 26 cannot afford all other school districts as District 2 do. Thus, they take advantage of outside options to enroll their kids in a more preferable school when they lose the advantage in admissions changes to District 26 schools.

We also present results when we scrap location-based admissions to District 2 schools but with academic screening in place in Figure E.5. The policy does not close the cross-racial gap in school peers' test scores when households' residential locations are fixed. Instead, households' spatial reshuffling rather widens the gap by 2.6%. This is because non-minority, higher-achieving students who reside outside of District 2 are largely motivated to move locations nearby District 2, which pushes minority students in District 2 out to schools with a higher proportion of minority peers.

8 Conclusion

Increasingly more school districts have adopted centralized school choice systems in the hope that they can break the tie between spatial disparities and educational disparities. Whether they can achieve these goals, however, crucially depends on the extent to which students are willing to take advantage of school choice options as well as how households respond to the policy by (1) reshuffling across locations and (2) opt-out to other schooling options.

This paper develops a unified framework of households' residential location choice and school choice under a centralized school choice system. By doing so, we extend empirical school choice literature that has studied many factors for students' school applications and assignments but has given little attention to endogenous residential location choices. Residential locations determine location-based admissions probabilities and commuting distances to schools, which motivates households to choose locations by considering such ties. Our framework captures this as well as the possibility of opting out to outside schooling options. Rich heterogeneity in households' observed and unobserved preferences over various school and location characteristics generates sorting into locations and schools. We map the framework to New York City's middle school choice context, which is the largest unified district with a centralized school choice system.

Our policy analysis shows how a radical school desegregation effort might have a minimal effect, largely because of households' choosing locations that can undo the policy. The policy grants citywide access to the school district that covers lower- and mid-Manhattan. We find that households' spatial reshuffling dampens the policy effect by half. Some minority households choose locations from which the commute to top district schools is easier, which amplifies the desegregation effect. However, other non-minority households choose locations that come with secured seats in higher-achieving schools outside of the affected district, which undoes the effect of the policy.

Several lines of inquiry are left for future work. First, such work might quantify the complementary effect between school desegregation policies and housing market policies. We find that 45% of school segregation is explained by households' heterogeneous preference over location characteristics other than price and access to school. Recent evidence shows that such heterogeneity stems from information frictions (Ellen, Horn, and Schwartz, 2016; Ferreira and Wong, 2020) or housing market discrimination (Christensen and Timmins, 2018), which suggests that policy interventions can change how households choose locations. Second, although we take the location of schools as given in the paper, future work can consider where to open a new school or how to allocate resources to schools in different locations. Such work informs policymakers' ongoing efforts to design school choice systems that could benefit a larger number of students.

References

ABDULKADIROĞLU, A., N. AGARWAL, AND P. A. PATHAK (2017): "The Welfare Effects of Coordinated Assignment: Evidence from the New York City High School Match," American Economic Review, 107(12), 3635–89.

ABDULKADIROĞLU, A., Y.-K. CHE, AND Y. YASUDA (2015): "Expanding "Choice" in School Choice," American Economic Journal: Microeconomics, 7(1), 1–42.

ABDULKADIROĞLU, A., P. A. PATHAK, AND A. E. ROTH (2005): "The New York City High School Match," American Economic Review, 95(2), 364–367.

ABDULKADIROĞLU, A., P. A. PATHAK, J. SCHELLENBERG, AND C. R. WALTERS (2020): "Do Parents Value School Effectiveness?," American Economic Review, 110(5), 1502–39.

AGARWAL, N., AND P. SOMAINI (2018): "Demand Analysis using Strategic Reports: An Application to a School Choice Mechanism," Econometrica, 86(2), 391–444.

——— (2020): "Revealed Preference Analysis of School Choice Models," Annual Review of Economics, 12, 471–501.

AGOSTINELLI, F., M. LUFLADE, AND P. MARTELLINI (2021): "On the Spatial Determinants of Educational Access," HCEO Working Paper.

AHLFELDT, G. M., S. J. REDDING, D. M. STURM, AND N. WOLF (2015): "The Economics of Density: Evidence from the Berlin Wall," Econometrica, 83(6), 2127–2189.

AJAYI, K., AND M. SIDIBE (2020): "School Choice Under Imperfect Information," Economic Research Initiatives at Duke (ERID) Working Paper, (294).

ALLENDE, C. (2019): "Competition under Social Interactions and the Design of Education Policies," Working Paper.

ALLENDE, C., F. GALLEGO, AND C. NEILSON (2019): "Approximating the Equilibrium Effects of Informed School Choice," Work. Pap., Princeton Univ., Princeton, NJ.

ALMAGRO, M., AND T. DOMINGUEZ-IINO (2019): "Location Sorting and Endogenous Amenities: Evidence from Amsterdam," Working Paper.

ARCIDIACONO, P., AND J. B. JONES (2003): "Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm," Econometrica, 71(3), 933–946.

ARTEMOV, G., Y.-K. CHE, AND Y. HE (2017): "Strategic 'Mistakes': Implications for Market Design Research," Working Paper.

AVERY, C., AND P. A. PATHAK (2021): "The Distributional Consequences of Public School Choice," American Economic Review, 111(1), 129–52.

AZEVEDO, E. M., AND J. D. LESHNO (2016): "A Supply and Demand Framework for Two-sided Matching Markets," Journal of Political Economy, 124(5), 1235–1268.

BARWICK, P. J., S. LI, A. R. WAXMAN, J. WU, AND T. XIA (2021): "Efficiency and Equity Impacts of Urban Transportation Policies with Equilibrium Sorting," Discussion paper, National Bureau of Economic Research.

BAUM-SNOW, N., AND F. FERREIRA (2015): "Causal Inference in Urban and Regional Economics," in Handbook of Regional and Urban Economics, vol. 5, pp. 3–68. Elsevier.

BAYER, P., F. FERREIRA, AND R. MCMILLAN (2007): "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," Journal of Political Economy, 115(4), 588–638.

BAYER, P., R. MCMILLAN, A. MURPHY, AND C. TIMMINS (2016): "A Dynamic Model of Demand for Houses and Neighborhoods," Econometrica, 84(3), 893–942.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," Econometrica: Journal of the Econometric Society, pp. 841–890.

——— (2004): "Differentiated Products Demand Systems from a Combination of Micro and Macro data: The New Car Market," Journal of Political Economy, 112(1), 68–105.

BHAT, C. R. (2000): "Incorporating Observed and Unobserved Heterogeneity in Urban Work Travel Mode Choice Modeling," Transportation Science, 34(2), 228–238.

BILLINGS, S. B., E. J. BRUNNER, AND S. L. ROSS (2018): "Gentrification and Failing Schools: The Unintended Consequences of School Choice under NCLB," Review of Economics and Statistics, 100(1), 65–77.

BLACK, S. E. (1999): "Do Better Schools matter? Parental Valuation of Elementary Education," The Quarterly Journal of Economics, 114(2), 577–599.

BRUNNER, E. J., S.-W. CHO, AND R. REBACK (2012): "Mobility, Housing Markets, and Schools: Estimating the Effects of Inter-district Choice Programs," Journal of Public Economics, 96(7-8), 604–614.

BURDICK-WILL, J. (2017): "Neighbors but not Classmates: Neighborhood Disadvantage, Local Violent Crime, and the Heterogeneity of Educational Experiences in Chicago," American Journal of Education, 124(1), 37–65.

CALSAMIGLIA, C., C. FU, AND M. GÜELL (2020): "Structural Estimation of a Model of School Choices: The Boston Mechanism versus its Alternatives," Journal of Political Economy, 128(2), 642–680.

CARD, D. (1993): "Using Geographic Variation in College Proximity to Estimate the Return to Schooling," National Bureau of Economic Research Working Paper.

CHE, Y.-K., D. W. HAHM, AND Y. HE (2022): "Leveraging Uncertainties to Infer Preferences: Robust Analysis of School Choice," Working Paper.

CHE, Y.-K., AND O. TERCIEUX (2019): "Efficiency and Stability in Large Matching Markets," Journal of Political Economy, 127(5), 2301–2342.

CHEN, Y., AND Y. HE (2021): "Information Acquisition and Provision in School Choice: An Experimental Study," Journal of Economic Theory, 197, 105345.

CHRISTENSEN, P., I. SARMIENTO-BARBIERI, AND C. TIMMINS (2020): "Housing Discrimination and the Toxics Exposure Gap in the United States: Evidence from the Rental Market," The Review of Economics and Statistics, pp. 1–37.

CHRISTENSEN, P., AND C. TIMMINS (2018): "Sorting or steering: Experimental Evidence on the Economic Effects of Housing Discrimination," Discussion paper, National Bureau of Economic Research.

COHEN, D. (2021): "NYC School Segregation Report Card: Still Last, Action Needed Now!," UCLA Civil Rights Project.

CORCORAN, S., AND H. M. LEVIN (2011): "School Choice and Competition in the New York City Schools," in Education Reform in New York City: An Ambitious Change in the Nation's Most Complex School System. Harvard University Press.

CORCORAN, S. P. (2018): "School Choice and Commuting," .

CORCORAN, S. P., J. L. JENNINGS, S. R. COHODES, AND C. SATTIN-BAJAJ (2018): "Leveling the Playing Field for High School Choice: Results from a Field Experiment of Informational interventions," Discussion paper, National Bureau of Economic Research.

DAVIS, M. A., J. GREGORY, AND D. A. HARTLEY (2019): "The Long-Run Effects of Low-Income Housing on Neighborhood Composition," .

DAVIS, M. A., J. GREGORY, D. A. HARTLEY, AND K. T. TAN (2021): "Neighborhood Effects and Housing Vouchers," Quantitative Economics, 12(4), 1307–1346.

DEMPSTER, A. P., N. M. LAIRD, AND D. B. RUBIN (1977): "Maximum Likelihood from Incomplete Data via the EM algorithm," Journal of the Royal Statistical Society: Series B (Methodological), 39(1), 1–22.

DINERSTEIN, M., AND T. D. SMITH (2021): "Quantifying the supply response of private schools to public policies," American Economic Review, 111(10), 3376–3417.

DINGEL, J. I., AND F. TINTELNOT (2020): "Spatial Economics for Granular Settings," Discussion paper, National Bureau of Economic Research.

DUR, U. M., S. D. KOMINERS, P. A. PATHAK, AND T. SÖNMEZ (2013): "The Demise of Walk Zones in Boston: Priorities vs. Precedence in School Choice," Discussion paper, National Bureau of Economic Research.

ELLEN, I. G., K. M. HORN, AND A. E. SCHWARTZ (2016): "Why don't Housing Choice Voucher Recipients Live Near Better Schools? Insights from Big Data," Journal of Policy Analysis and Management, 35(4), 884–905.

EPPLE, D. N., AND R. ROMANO (2003): "Neighborhood Schools, Choice, and the Distribution of Educational Benefits," in The Economics of School Choice, pp. 227–286. University of Chicago Press.

FACK, G., J. GRENET, AND Y. HE (2019): "Beyond Truth-telling: Preference Estimation with Centralized School Choice and College Admissions," American Economic Review, 109(4), 1486–1529.

FERREIRA, F. V., AND M. WONG (2020): "Estimating Preferences for Neighborhood amenities under Imperfect Information," Discussion paper, National Bureau of Economic Research.

FERREYRA, M. M. (2007): "Estimating the Effects of Private School Vouchers in Multi District Economies," American Economic Review, 97(3), 789–817.

GALE, D., AND L. S. SHAPLEY (1962): "College Admissions and the Stability of Marriage," The American Mathematical Monthly, 69(1), 9–15.

GRIGORYAN, A. (2021): "School Choice and the Housing Market," Available at SSRN 3848180.

HAERINGER, G., AND F. KLIJN (2009): "Constrained School Choice," Journal of Economic theory, 144(5), 1921–1947.

HAHM, D. W., AND M. PARK (2022): "A Dynamic Framework of School Choice: Effects of Middle Schools on High School Choice," in Proceedings of the 23rd ACM Conference on Economics and Computation, pp. 292–293.

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

HE, Y. (2015): "Gaming the Boston School Choice Mechanism in Beijing," Working Paper.

HECKMAN, J., AND B. SINGER (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," Econometrica: Journal of the Econometric Society, pp. 271–320.

HOXBY, C. M., AND S. TURNER (2015): "What High-achieving Low-income Students Know about College," American Economic Review, 105(5), 514–17.

ICELAND, J. (2004): "The Multigroup Entropy Index (also known as Theil's H or the Information theory index)," US Census Bureau. Retrieved July, 31, 2006.

IDOUX, C. (2022): "Integrating New York City schools: The role of admission criteria and family preferences," Discussion paper, Technical Report.

KULKA, A. (2019): "Sorting into Neighborhoods: The Role of Minimum Lot Sizes," Working Paper.

LAVERDE, M. (2020): "Unequal Assignments to Public Schools and the Limits of School Choice," Working Paper.

LUFLADE, M. (2018): "The Value of Information in Centralized School Choice Systems," Working Paper.

MANSKI, C. F. (1992): "Educational Choice (vouchers) and Social Mobility," Economics of Education Review, 11(4), 351–369.

MIYAUCHI, Y., K. NAKAJIMA, AND S. J. REDDING (2022): "The Economics of Spatial Mobility: Theory and Evidence Using Smartphone Data," Discussion paper.

MONARREZ, T. E. (2020): "School Attendance Boundaries and the Segregation of Public Schools in the U.S.," American Economic Journal: Applied Economics.

MOUNTJOY, J. (2022): "Community Colleges and Upward Mobility," American Economic Review, 112(8), 2580–2630.

NECHYBA, T. J. (2000): "Mobility, Targeting, and Private-school Vouchers," American Economic Review, 90(1), 130–146.

PARRY, I. W., AND K. A. SMALL (2009): "Should Urban Transit Subsidies be Reduced?," American Economic Review, 99(3), 700–724.

PUREVJAV, A.-O. (2022): "The Welfare Effects of Congestion-Reduction Policies," Working Paper.

REBACK, R. (2005): "House Prices and the Provision of Local Public Services: Capitalization under School Choice Programs," Journal of Urban Economics, 2(57), 275–301.

RUST, J. (1994): "Structural Estimation of Markov Decision Processes," Handbook of Econometrics, 4, 3081–3143.

SATTIN-BAJAJ, C., AND J. JENNINGS (2022): "Understanding Safety As a Driver of School Selection for Parents and Students: Lessons from New York City," in 2021 APPAM Fall Research Conference. APPAM.

SATTIN-BAJAJ, C., J. L. JENNINGS, S. P. CORCORAN, E. C. BAKER-SMITH, AND C. HAILEY (2018): "Surviving at the Street Level: How Counselors' Implementation of School Choice Policy Shapes Students' High School Destinations," Sociology of Education, 91(1), 46–71.

SCHWARTZ, A. E., L. STIEFEL, AND M. WISWALL (2013): "Do Small Schools Improve Performance in Large, Urban Districts? Causal Evidence from New York City," Journal of Urban Economics, 77, 27–40.

SCHWARTZ, A. E., I. VOICU, AND K. M. HORN (2014): "Do Choice Schools Break the

Link Between Public Schools and Property Values? Evidence from House Prices in New York City," Regional Science and Urban Economics, 49, 1–10.

SON, S. J. (2020): "Distributional Impacts of Centralized School Choice," Working Paper.

TRAIN, K. E. (2009): Discrete Choice Methods with Simulation. Cambridge university press.

VARADHAN, R., AND C. ROLAND (2008): "Simple and Globally Convergent Methods for Accelerating the Convergence of any EM Algorithm," Scandinavian Journal of Statistics, 35(2), 335–353.

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

XU, J. (2019): "Housing Choices, Sorting, and the Distribution of Educational Benefits under Deferred Acceptance," Journal of Public Economic Theory, 21(3), 558–595.

ZHENG, A. (2022): "The Valuation of Local School Quality under School Choice," American Economic Journal: Economic Policy, 14(2), 509–37.

9 List of Appendices

A Details of NYC School Choice Process

A.1 Student-Proposing Deferred Acceptance Algorithm

In detail, DA works as follows (Gale and Shapley, 1962; Abdulkadiroğlu, Pathak, and Roth, 2005):

- **Step 1**
  Each student proposes to her first choice. Each program tentatively assigns seats to its proposers one at a time, following their priority order. The student is rejected if no seats are available at the time of consideration.

- **Step $k \geq 2$**
  Each student who was rejected in the previous step proposes to her next best choice. Each program considers the students it has tentatively assigned together with its new proposers and tentatively assigns its seats to these students one at a time following the program's priority order. The student is rejected if no seats are available when she is considered.

- The algorithm terminates either when there are no new proposals or equally when all rejected students have exhausted their preference lists.

DA produces the student-optimal stable matching and is strategy-proof i.e., truth-telling is a weakly dominant strategy for students.

A.2 NYC School Admission Methods

Middle school programs use a variety of admission methods—Unscreened, Limited Unscreened, Screened, Screened: Language, Zoned and Talent Test. Unscreened programs admit students by a random lottery number, and Limited Unscreened programs use rules that give priority to those who attend information sessions or open houses. Screened programs as well as Screened: Language programs select students by program-specific measures such as elementary school GPA, statewide test scores, punctuality and interviews. Zoned programs guarantee admissions or give priority to students who reside in the school's zone, and Talent Test programs use auditions.

A.3 The Timeline of Admission Process

The timeline of the admission process is as follows (Corcoran and Levin (2011), Directory of NYC Public High Schools). By December, students are required to submit their ROLs. By March, DA algorithms are run and determine students' assignments. Students who accept their offer finalize, and if a student rejects an offer, then she goes to the next round. This describes the main round of the entire system. A majority of students finalize in the main round (about 90% each year). Students who are not assigned in the main round or rejected the assignment go to the Supplementary round which is similarly organized as the main round and includes programs that did not fill up their capacities in the main round, or programs that are newly opened. Finally, there is an administrative round in which students who are not assigned a school even after the second round are administratively assigned to a school.

A.4 Example of ROL

Figure A.1: Example of Customized List and Rank-Ordered List

Rank your choices in order of preference here.

Choice Number	District	Program Code	School Name/Program Name
2	1	M378L	School for Global Leaders
3	21	K239CM	Mark Twain (I.S. 239) Magnet Program - Computer/Math
1	21	K239VO	Mark Twain (I.S. 239) Magnet Program - Vocal
	1	M292S	Henry Street School for International Studies
4	1	M140S	The Nathan Straus Preparatory School of Humanities (P.S. 140)
	1	M301S	Technology, Arts, and Sciences Studio

SAMPLE

Source: NYC DOE Middle School Directory 2014-15

B Supplementary Materials for Section 3

B.1 Cleaning Procedure of DOF Annualized Selling Record

First, we drop non-residential properties such as industrial buildings, commercial buildings, and vacant land, based on both the tax class and building class. Then we merge the selling record with the Primary Land Use Tax Lot Output (PLUTO) to recover the exact location of each sold property.[38] Lastly, we exclude transactions that

[38]Two data sets are merged based on the identifiable tax lot number. One complication is in merging condos. The selling record has a unique id for each unit, while PLUTO for each condo. We use the

are unlikely at arm's-length. We drop transactions of zero price that include transfers within a family. Also, we drop records of significantly low prices relative to other properties of similar characteristics. Specifically, we run a hedonic pricing model that includes tax class, assessment value, the interaction of the two, calendar time FE at each borough, month FE by borough, building type, land area, building area, total unit, odd shape, age, age square, garage area, the year of alteration, and commercial area (R square = 0.67). Then we drop observations of the predicted residual is less than 1 percentile. We run the regression separately for coops. Following Schwartz, Voicu, and Horn (2014), we lag housing cost by one year to take into account that there could be some time lag for school quality to be capitalized in the housing cost.

B.2 Housing Cost and Structure in ACS 5-year Estimates

While ACS 5-year estimates capture the price and characteristics of representative housings, the biggest limitation is that each observation is at the Census block group level, which could be too coarse to capture the change within a narrow bandwidth around the boundary. 1,944 of 7,506 Census block groups whose centroid is within 0.2 miles from a school district boundary overlay across a school district boundary.[39] Thus, the distance from the centroid of a Census block group to the closest boundary is a crude measure of proximity to the boundary.

Table B.1: Example of Census Block Groups

Census Block Group A:	Block A-1	Block A-2
Distance to Boundary (mi.)	0.15	0.28
The Number of Occupied Units	30	60
Census Block Group B:	Block B-1	Block B-2
Distance to Boundary (mi.)	0.15	0.28
The Number of Occupied Units	60	30

Note: Consider two Census Block Groups A and B with the same distance from their centroids to the nearest boundaries.

Therefore, we further exploit the variation of population density across Census blocks within a block group. Table B.1 illustrates two exemplary cases. Consider two census block groups A and B whose distances from their centroid to the closest boundaries are the same. Census block group A consists of two Census blocks, one of which is 0.15 miles away from the boundary and the other 0.28 miles away. Note that the

Department of City Planning Property Address Directory that lists unit ids to a matching condo id.

[39] On the contrary, only 489 out of 38,498 Census blocks overlay across a boundary.

two Census blocks differ in population density. Out of 90 occupied units in the block group, two-thirds are living in A-2. Census block group B has the opposite pattern.

To consider such differences in density, we weigh Census block groups with the percent of occupied units in Census blocks within 0.25 miles from the closest boundary when running Equation 1. In the example, Census block group A is given a weight of 0.33, while B has a weight of 0.66. We present the estimated effects of district school quality in Table B.8.

B.3 Evidence Supporting the Identification Assumption of BDD

The identification assumption of a boundary discontinuity design is that unobserved location amenities are as good as random within a narrow buffer around a boundary. While we cannot check this assumption directly, we present that other observed location characteristics are continuous in geography, which suggests that the assumption is plausible in this context.

Table B.9 reports estimates $\hat{\beta}$ (Equation 1) for various housing characteristics and urban amenities. $\hat{\beta}$s for most of the variables are not statistically significant. One exception is that sold properties located within a school district with higher school quality are more likely to have been renovated (p-value < 0.05). Thus, we use residualized prices when estimating the model to absorb variations coming from housing characteristics and urban amenities, including the renovation status.

While each variable is not the main driver of sharp change in sales price at boundaries (Figure 3 and Table B.4), a *set* of variables might. Table B.10 checks this further. First, we run a hedonic regression of log sales prices on various housing characteristics and urban amenities using transaction records within a 0.25-mile buffer around boundaries. Then we sum variables using coefficients from the hedonic regression, which capture the extent to which each variable explains the variation in sales prices. Finally, we run the BDD regression (Equation 1) using the predicted prices as dependent variables and check if estimates $\hat{\beta}$ are significant.

Columns (1)-(2) of Table B.10 describe that even a very extensive set of housing characteristics and urban amenities does not explain the sharp change in sales prices at boundaries—i.e., change in school quality is the main driver. Meanwhile, the estimate $\hat{\beta}$ is marginally significant (p-value < 0.1) when we include neighbors' composition (column (3)), which captures households' residential sorting at boundaries as well as their preference over neighbors' composition.

B.4 Supplementary Tables

Table B.2: Summary Statistics: All vs. Sample for BDD

	(1)	(2)	(3)	(4)
	All		< 0.25-mile Buffer	
Variables	mean	std	mean	std
Panel A: Student Characteristics				
Asian	0.185	0.388	0.162	0.368
Black	0.295	0.456	0.333	0.471
Hispanic	0.380	0.485	0.407	0.491
White	0.125	0.331	0.085	0.279
FRL	0.736	0.441	0.783	0.413
Standardized English Score	0.028	1.009	-0.077	0.997
Standardized Math Score	0.172	0.979	0.074	0.963
N	57,593		18,761	
Panel B: Sold Properties' Characteristics				
Unit Price ($1000)	802.8	2,063	707.9	1,496
Age	70.88	32.68	80.71	97.16
Number of Floors	7.150	9.083	6.493	8.105
Coop[a]	0.299	0.458	0.287	0.452
Manhattan	0.245	0.430	0.170	0.376
Bronx	0.079	0.270	0.088	0.283
Brooklyn	0.263	0.440	0.441	0.497
Queens	0.319	0.466	0.300	0.458
Staten Island	0.0934	0.291	0	0
N	106,040		23,836	
Panel C: Census Block Group Characteristics				
Median Rent	1,404	503.4	1,255	501.5
Median Value ($1000)	636.7	355.9	612.9	331.1
Median Age	70.83	13.45	72.67	13.08
% College and Higher Degree	0.352	0.237	0.298	0.229
% Minority	0.272	0.316	0.334	0.309
N	4,828		609	

Note: Source of each data set is NYC Department of Education, NYC Department of Finance Selling Record, and ACS 5-year estimates. All from 2013 to 2017.

[a]We control for properties' co-op status in our analyses. Coops take up a large proportion of the NYC housing market (35% of sold properties, 50% of the housing stock) with two unique features. First, they are more common in Manhattan compared to other boroughs, and second, are cheaper to buy but come with high monthly maintenance fees. (Susan Stellin, Co-op vs. Condo: The Differences Are Narrowing, The New York Times, Oct. 5, 2012) Ignoring the composition of co-op and other housing types understate housing cost in Manhattan because the Sales files cover only sold price.

Table B.3: Effects of District School Quality on Assigned Schools' Quality

	(1)	(2)	(3)	(4)
Bandwidth	< 0.25 mile		< 0.2 mile	
Boundary FEs	Yes	Yes	Yes	Yes
Local Cubic Control for Distance	Yes	Yes	Yes	Yes
Student Characteristics	No	Yes	No	Yes
District School Quality	0.360	0.270	0.324	0.256
	(0.074)	(0.053)	(0.072)	(0.054)
N	16576	15809	13261	12657
R2	0.249	0.394	0.245	0.389
$\bar{y}$	-0.078	-0.055	-0.092	-0.068
$std(y)$	0.800	0.779	0.796	0.773

Note: The dependent variable is the mean score of the schools middle-school-applying residents in a Census block are assigned to. Sample of 5th-grade students in academic year 2014-15 living in Census blocks within a buffer from the closest school district boundary. District school quality is measured by the mean NYS standardized test score of students enrolled in middle schools (previous cohorts) in the district. We use the 0.25-mile buffer in columns (1)-(2) and the 0.2-mile buffer in columns (3)-(4). Standard errors in parentheses are clustered at school district level. The local cubic control of distance differs at the opposite side of boundaries.

Table B.4: Effects of District School Quality on Housing Sales Price

	(1)	(2)	(3)	(4)
Boundary FEs	Yes	Yes	Yes	Yes
Local Cubic Control for Distance	Yes	Yes	Yes	Yes
Housing Characteristics	No	Yes	Yes	Yes
Neighborhood Characteristics	No	No	Yes	Yes
Urban Amenities	No	No	No	Yes
Panel A: 0.25-mile Buffer				
District School Quality	0.181	0.204	0.101	0.102
	(0.061)	(0.044)	(0.040)	(0.040)
N	23786	23786	23786	23786
R2	0.409	0.489	0.505	0.505
$\bar{y}$		12.88		
$std(y)$		1.112		
Panel B: 0.2-mile Buffer				
District School Quality	0.108	0.150	0.073	0.073
	(0.068)	(0.049)	(0.045)	(0.045)
N	19057	19057	19057	19057
R2	0.401	0.480	0.493	0.494
$\bar{y}$		12.84		
$std(y)$		1.086		

Note: The dependent variable is the log sales price of a residential unit. Sample of residential units sold within a bandwidth from the closest school district boundary. District school quality is measured by the mean NYS standardized test score of students enrolled in middle schools (previous cohorts) in the district. We use the 0.25-mile buffer in columns (1)-(2) and the 0.2-mile buffer in columns (3)-(4). Standard errors in parentheses are clustered at school district level. The local cubic control of distance differs at the opposite side of boundaries. Housing characteristics include the space of the unit, land use of the tax lot, number of floors, age, renovation status, and storage area of the building, all of which we interact with a dummy if the property is coop. Neighbor characteristics include % minority, median household income, % college-or-more-educated, and median commuting time to work at Census block group. Urban amenities include the number of bus stops, subway stations, laundries, cafes, and crime incidents of different categories at Census block.

Table B.5: Effects of District School Quality on the Number of Middle-school-applying Residents

	(1)	(2)	(3)	(4)	(5)
Boundary FEs	Yes	Yes	Yes	Yes	Yes
Local Cubic Control for Distance	Yes	Yes	Yes	Yes	Yes
Housing Characteristics	No	Yes	Yes	Yes	Yes
Neighborhood Characteristics	No	No	Yes	Yes	Yes
Urban Amenities	No	No	No	Yes	Yes
N of Population	No	No	No	No	Yes
Panel A: 0.25-mile Buffer					
District School Quality	0.183	0.266	0.805	0.789	0.808
	(0.176)	(0.175)	(0.222)	(0.225)	(0.219)
N	3755	3755	3755	3755	3755
R2	0.227	0.251	0.308	0.318	0.331
$\bar{y}$			3.515		
$std(y)$			4.122		
Panel B: 0.2-mile Buffer					
District School Quality	0.216	0.280	0.766	0.750	0.830
	(0.180)	(0.183)	(0.240)	(0.244)	(0.242)
N	2970	2970	2970	2970	2970
R2	0.223	0.246	0.295	0.303	0.313
$\bar{y}$			3.490		
$std(y)$			4.033		

Note: The dependent variable is the number of middle-school-applying residents in a Census block. Sample of Census blocks within a bandwidth from the closest school district boundary. District school quality is measured by the mean NYS standardized test score of students enrolled in middle schools (previous cohorts) in the district. We use the 0.25-mile buffer in panel A and the 0.2-mile buffer in panel B. Standard errors in parentheses are clustered at school district level. The local cubic control of distance differs at the opposite side of boundaries. Housing characteristics include the space of the unit, land use of the tax lot, number of floors, age, renovation status, and storage area of the building, all of which we interact with a dummy if the property is coop. Neighbor characteristics include % minority, median household income, % college-or-more-educated, and median commuting time to work at Census block group. Urban amenities include the number of bus stops, subway stations, laundries, cafes, and crime incidents of different categories at Census block. The number of population is at Census block group level, which we obtain from the ACS 5-year estimate.

Table B.6: Effects of District School Quality on Minority Share of Middle-school-applying Residents

	(1)	(2)	(3)	(4)
Boundary FEs	Yes	Yes	Yes	Yes
Local Cubic Control for Distance	Yes	Yes	Yes	Yes
Housing Characteristics	No	Yes	Yes	Yes
Neighborhood Characteristics	No	No	Yes	Yes
Urban Amenities	No	No	No	Yes
Panel A: 0.25-mile Buffer				
District School Quality	-0.139	-0.143	-0.065	-0.065
	(0.029)	(0.028)	(0.016)	(0.016)
N	2970	2970	2970	2970
R2	0.480	0.496	0.595	0.596
$\bar{y}$	0.620			
$std(y)$	0.417			
Panel B: 0.2-mile Buffer				
District School Quality	-0.132	-0.132	-0.066	-0.067
	(0.034)	(0.033)	(0.024)	(0.024)
N	2353	2353	2353	2353
R2	0.484	0.501	0.603	0.603
$\bar{y}$	0.633			
$std(y)$	0.415			

Note: The dependent variable is the share of Black and Hispanic applicants among middle-school-applying residents in a Census block. Sample of Census blocks within a bandwidth from the closest school district boundary. District school quality is measured by the mean NYS standardized test score of students enrolled in middle schools (previous cohorts) in the district. We use the 0.25-mile buffer in columns (1)-(2) and the 0.2-mile buffer in columns (3)-(4). Standard errors in parentheses are clustered at school district level. The local cubic control of distance differs at the opposite side of boundaries. Housing characteristics include the space of the unit, land use of the tax lot, number of floors, age, renovation status, and storage area of the building, all of which we interact with a dummy if the property is coop. Neighbor characteristics include % minority, median household income, % college-or-more-educated, and median commuting time to work at Census block group. Urban amenities include the number of bus stops, subway stations, laundries, cafes, and crime incidents of different categories at Census block.

Table B.7: Effects of District School Quality on Poverty Share of Middle-school-applying Residents

	(1)	(2)	(3)	(4)
Boundary FEs	Yes	Yes	Yes	Yes
Local Cubic Control for Distance	Yes	Yes	Yes	Yes
Housing Characteristics	No	Yes	Yes	Yes
Neighborhood Characteristics	No	No	Yes	Yes
Urban Amenities	No	No	No	Yes
Panel A: 0.25-mile Buffer				
District School Quality	-0.139	-0.142	-0.091	-0.094
	(0.038)	(0.034)	(0.033)	(0.033)
N	2970	2970	2970	2970
R2	0.203	0.218	0.261	0.263
$\bar{y}$	0.683			
$std(y)$	0.357			
Panel B: 0.2-mile Buffer				
District School Quality	-0.139	-0.141	-0.099	-0.102
	(0.042)	(0.038)	(0.037)	(0.037)
N	2353	2353	2353	2353
R2	0.173	0.192	0.238	0.240
$\bar{y}$	0.698			
$std(y)$	0.350			

Note: The dependent variable is the share of free or reduced lunch eligible applicants among middle-school-applying residents in a Census block. Sample of Census blocks within a bandwidth from the closest school district boundary. District school quality is measured by the mean NYS standardized test score of students enrolled in middle schools (previous cohorts) in the district. We use the 0.25-mile buffer in columns (1)-(2) and the 0.2-mile buffer in columns (3)-(4). Standard errors in parentheses are clustered at school district level. The local cubic control of distance differs at the opposite side of boundaries. Housing characteristics include the space of the unit, land use of the tax lot, number of floors, age, renovation status, and storage area of the building, all of which we interact with a dummy if the property is coop. Neighbor characteristics include % minority, median household income, % college-or-more-educated, and median commuting time to work at Census block group. Urban amenities include the number of bus stops, subway stations, laundries, cafes, and crime incidents of different categories at Census block.

Table B.8: Effects of District School Quality on Rent and House Value

	(1)	(2)	(3)
Boundary FEs	Yes	Yes	Yes
Local Cubic Control for Distance	Yes	Yes	Yes
Housing Characteristics	No	Yes	Yes
Neighborhood Characteristics	No	No	Yes
Urban Amenities	No	No	No
Panel A: Log(Median Gross Rent)			
District School Quality	0.163	0.159	0.058
	(0.047)	(0.047)	(0.035)
N	1875	1873	1873
R2	0.352	0.374	0.535
$\bar{y}$		7.119	
$std(y)$		0.377	
Panel B: Log(House Value)			
District School Quality	0.057	0.067	0.058
	(0.066)	(0.061)	(0.062)
N	1332	1331	1331
R2	0.478	0.540	0.558
$\bar{y}$		13.20	
$std(y)$		0.540	

Note: The dependent variable is the log median gross rent in panel A, and the log house value reported by homw owners in panel B. Unit of analysis is Census block group, where we weigh block groups by the share of occupied units in Census blocks within a 0.25-mile buffer from the closest school district boundary (See Subsection B.2). District school quality is measured by the mean NYS standardized test score of students enrolled in middle schools (previous cohorts) in the district. Standard errors in parentheses are clustered at school district level. The local cubic control of distance differs at the opposite side of boundaries. Housing characteristics include the space of the unit, land use of the tax lot, number of floors, age, renovation status, and storage area of the building, all of which we interact with a dummy if the property is coop. Neighbor characteristics include % minority, median household income, % college-or-more-educated, and median commuting time to work at Census block group. Urban amenities include the number of bus stops, subway stations, laundries, cafes, and crime incidents of different categories at Census block.

Table B.9: Housing Characteristics and Urban Amenities at School District Boundaries

	(1)	(2)	(3)	(4)
Panel A: Housing Characteristics of Sold Properties				
Dependent Variable:	N of Floors	Coop	Commercial Area (1K sqft)	Renovation
District School Quality	-1.55	0.076	-5.84	0.100
	(1.25)	(0.047)	(5.53)	(0.038)
N	23786	23786	23786	23786
R2	0.466	0.230	0.182	0.165
$\bar{y}$	6.498	0.288	11.19	0.809
$std(y)$	8.112	0.453	56.27	0.393
Panel B: Urban Amenities				
Dependent Variable:	Bus Stop	Subway Station	Laundries	Café
District School Quality	-0.014	-0.004	0.003	-0.004
	(0.024)	(0.008)	(0.008)	(0.007)
N	8091	8091	32340	32340
R2	0.025	0.020	0.068	0.087
$\bar{y}$	0.127	0.019	0.052	0.014
$std(y)$	0.413	0.140	0.252	0.157

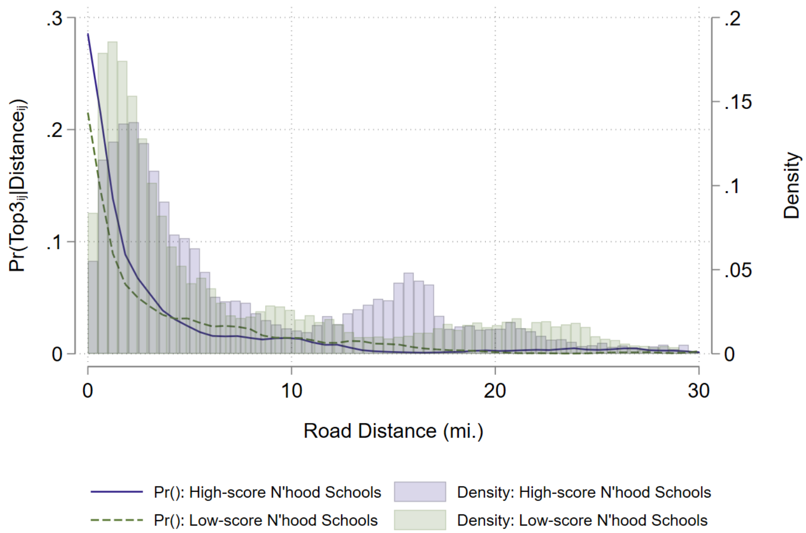
Note: Sample of residential properties sold within a bandwidth from the closest school district boundary in Panel A. Sample of Census block groups whose centroids are within a bandwidth from the closest school district boundary in Panel B. Sample of Census blocks whose centroids are within a bandwidth from the closest school district boundary in Panel C. We use 0.25 mile buffer. In panel B, Each Census block groups is weighted accroding to the procedure explained in the appendix. Standard errors in parentheses are clustered at each tax lot, Census block group, and Census block, respectively.We allow the local cubic control of distance to differ at the opposite side of boundaries.

Table B.10: Effects of District School Quality on Hedonic Sales Price

	(1)	(2)	(3)
Boundary FEs	Yes	Yes	Yes
Local Cubic Control for Distance	Yes	Yes	Yes
Housing Characteristics	Yes	Yes	Yes
Neighborhood Characteristics	No	No	Yes
Urban Amenities	No	Yes	Yes
Panel A: 0.25-mile Buffer			
District School Quality	-0.027	-0.018	0.094
	(0.069)	(0.065)	(0.052)
N	23786	23786	23786
R2	0.457	0.479	0.687
$\bar{y}$	12.88	12.88	12.88
$std(y)$	0.546	0.560	0.728
Panel B: 0.2-mile Buffer			
District School Quality	-0.021	-0.002	0.069
	(0.077)	(0.074)	(0.057)
N	19057	19057	19057
R2	0.443	0.467	0.686
$\bar{y}$	12.84	12.84	12.84
$std(y)$	0.509	0.526	0.696

Note: The dependent variable is the predicted log sales price of a residential unit, which we construct by running a hedonic regression of log sales price on covariates, as explained in Subsection B.3. Sample of residential units sold within 0.25-mile bandwidth from the closest school district boundary. District school quality is measured by the mean NYS standardized test score of students enrolled in middle schools (previous cohorts) in the district. Standard errors in parentheses are clustered at school district level. The local cubic control of distance differs at the opposite side of boundaries. Housing characteristics include the space of the unit, land use of the tax lot, number of floors, age, renovation status, and storage area of the building, all of which we interact with a dummy if the property is coop. Neighbor characteristics include % minority, median household income, % college-or-more-educated, and median commuting time to work at Census block group. Urban amenities include the number of bus stops, subway stations, laundries, cafes, and crime incidents of different categories at Census block.

Figure B.2: Probability of Listed as Top 3 Given Distance



Note: The fastest driving distance to school is calculated using the Open Route Service (ORS). the y-axis of the graph presents the probability of choosing school j as the top 3 choices among all schools a student is eligible for. Lines present the probability of student i's listing school j as one of her top 3 choices as a function of road distance between i and j, for all pairs of (i, j). We present the pattern separately for students whose neighborhood schools' mean test score is greater/smaller than the average, where we use the three closest schools as neighborhood schools.

C Supplementary Materials for Section 4

C.1 Stability of Matching and Aggregate Demand

Fixing location choice ℓ, the set of feasible schools are defined as $\mathcal{J}_i(\ell;\rho_i) = \{j | c_{ij}(\ell;\rho_i) \geq \bar{c}_j\}$, i.e. schools of which student i can clear the cutoffs. The set of feasible schools changes depending on which location ℓ student i chooses to reside in, and the lottery number.

Using the stability of matching and the distributional assumption on the idiosyncratic preference shock over locations and schools,

$$D_j(\{\bar{c}_{j'}\}_{j'=1}^{J}) = \int_i \Sigma_\ell \underbrace{\frac{exp(\tilde{V}_i(\ell))}{\Sigma_{\ell'} exp(\tilde{V}_i(\ell'))}}_{\text{Demand for location } \ell} \cdot \underbrace{\int_{\rho_i} \frac{\mathbb{1}(c_{ij}(\ell;\rho_i) \leq \bar{c}_j) exp(\tilde{U}_i(j,\ell))}{\Sigma_{j'} \mathbb{1}(r_{ij'}(\ell;\rho_i) \leq \bar{c}_{j'}) exp(\tilde{U}_i(j',\ell))} d\rho_i}_{\text{Demand for school } j \text{ given location } \ell} \tag{C.1}$$

where

$$\tilde{V}_i(\ell) = \underbrace{W_\ell}_{\text{location char.}} \alpha_i^w + \underbrace{p_\ell}_{\text{housing cost}} \alpha_i^p + \underbrace{\xi_\ell}_{\text{unobserved amenities}} + \alpha^u \underbrace{\mathbb{E}_{\varepsilon_{ij},\rho,\varepsilon_{i\vartheta}}}_{\text{expected utility from school given location}} \tag{C.2}$$

$$\tilde{U}_i(j,\ell) = \underbrace{X_j}_{\text{school char.}} \beta_i^X + \underbrace{d_{\ell j}\beta_i^d}_{\text{commuting cost}} \tag{C.3}$$

The second component of Equation C.1 means that a student has effective demand only when school j is feasible given location ℓ. For schools that use lottery number to break the tie, the feasibility depends on the lottery number ρ_i. We take the numerical integration over ρ_i. The existence and the uniqueness of the equilibrium follow from Azevedo and Leshno (2016). The key assumption is that the distribution of students' priority rank c_{ij} is continuous. This ensures a small change in the cutoff $\bar{c}_{j'}$ induces a small change in the demand for school j.

C.2 Stability of Matching and Expected Utility from School

In addition, based on the stability of assignment under DA with truth-telling, we can simplify the indirect utility from school choice stage $U_i^*(\ell)$:

$$U_i^*(\ell) = \max\{\max_{j\in\mathcal{J}_i(\ell;\rho_i)} U_i(j_i,\ell;\varepsilon_{ij}), U_i^\vartheta(\vartheta_p;\varepsilon_{i\vartheta}), U_i^\vartheta(\vartheta_c;\varepsilon_{i\vartheta})\Big\}) \tag{C.4}$$

With the assumption that $\varepsilon_{ij}, \varepsilon_{i\vartheta}$ follows i.i.d EVT1 distribution, the expected utility from school can be simplified as follows.

$$\mathbb{E}_{\varepsilon_{ij},\rho,\varepsilon_{i\vartheta}} = \mathbb{E}_\rho\left(\mu + log\left(\Sigma_{j\in\mathcal{J}_i(\ell;\rho_i)} exp(\tilde{U}_i(j,\ell;\varepsilon_{ij})) + exp(\tilde{U}_i^\vartheta(\vartheta_p;\varepsilon_{i\vartheta})) + exp(\tilde{U}_i^\vartheta(\vartheta_c;\varepsilon_{i\vartheta}))\right)\right) \tag{C.5}$$

where

$$\tilde{U}_i(j,\ell) = \underbrace{X_j}_{\text{school char.}} \beta_i^X + \underbrace{d_{\ell j}\beta_i^d}_{\text{commuting cost}} \tag{C.6}$$

$$\tilde{U}_i(\vartheta;\varepsilon_{i\vartheta}) = \beta_i^{\vartheta} \tag{C.7}$$

C.3 Estimating Program Preferences

s_{ij} is a weighted average of students' middle school GPA, NYS math and ELA score, and punctuality record, with the weights remaining as each program's private information.

However, a program j has to report the rank of s_{ij}s among its applicants for NYC DOE to implement the centralized DA. Therefore, given any pair of students i and i', we observe the value of $1(s_{ij} > s_{i'j})$, if both i and i' apply to the program j. Using this, we construct $\hat{s}_{ij}$ using a latent variable model by assuming

$$s_{ij} = Z_i\kappa_j + \eta_{ij},\ \eta_{ij} \sim \mathcal{N}(0,1) \tag{C.8}$$

where Z_i is student characteristics that are known to compose of s_{ij}, and κ_j a vector of weights which vary across js. η_{ij} is normalized to be $\mathcal{N}(0,1)$.

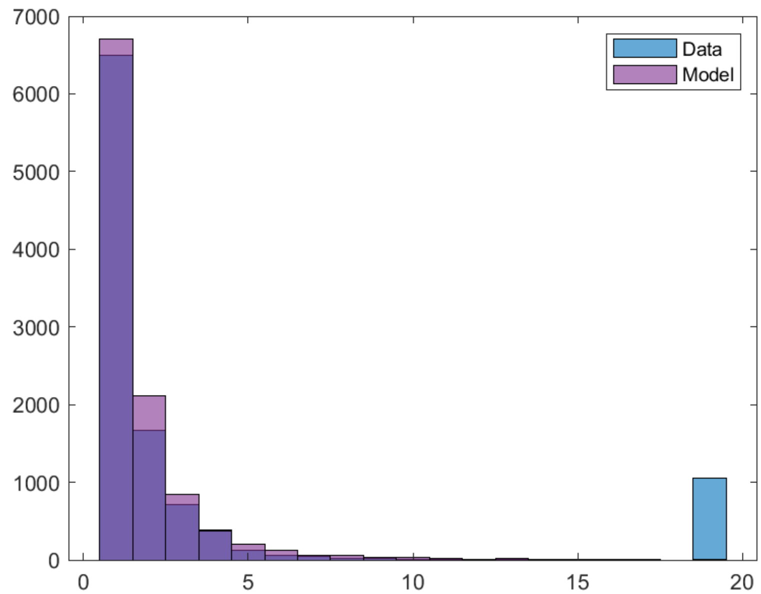
We estimate $(\kappa_j)_j$ using Maximum Likelihood Estimation (MLE) where the likelihood function is

$$\mathcal{LL}_c = \Pi_{i,i'\in\mathcal{A}_j} Pr(s_{ij} > c_{i'j})^{1(s_{ij}>c_{i'j})}(1 - Pr(s_{ij} > c_{i'j}))^{1(s_{ij}<c_{i'j})} \tag{C.9}$$

where $\mathcal{A}_j$ is the set of applicants to program j that uses a non-random tie-breaker.

Figure C.3 shows that our simulation recovers the distribution of the preference rank of the assigned program in the data. Both in simulation and data, around 63% of students are assigned to the 1st- or the 2nd- ranked programs while 8% do not get any offer from programs on the list.

Figure C.3: Rank of Assigned School: Model and Simulation



Note: Rank is the preference rank of the assigned program. Rank 19 in the data means that the student did not get any offer from programs included in the submitted list.

D Supplementary Materials for Section 5

D.1 Full Information Maximum Likelihood Estimation

Assuming idiosyncratic preference shocks over locations, schools, and outside options ($\eta_{i\ell}, \varepsilon_{ij},$ and $\varepsilon_{i\vartheta}$) are i.i.d, the full information log-likelihood function is as follows.

$$\mathcal{LL} = \underbrace{\Sigma_i}_{\text{3. sum over } i} \; log\Big(\underbrace{\Sigma_{k=1}^{K} q_k}_{\text{2. sum over type}} \; \underbrace{P^{LC}(x_i; \theta^{EC}, \theta^{SC}, \theta^{LC}, \gamma_k) P^{SC}(x_i; \theta^{SC}, \gamma_k) P^{EC}(x_i; \theta^{SC}, \theta^{EC}, \gamma_k)}_{\text{1. likelihood with fixed type}} \Big) \tag{D.10}$$

x_i is data, θ^{LC} is the set of location preference parameters, θ^{SC} is the set of school preference parameters, θ^{EC} is the set of outside option preference parameters, γ_i is the unobserved taste, and q_k the probability of each type k. $(\theta = (\theta^{LC}, \theta^{SC}, \theta^{EC}), \{\gamma_k, q_k\}_k)$ is the full set of parameters to be estimated.

The likelihood function for each step is not additive separable because of γ_i, making the maximization problem computationally very costly. Note that the sequential estimation strategy in Rust (1994) is also not applicable without additive separability of likelihoods.

D.2 EM Algorithm with Sequential Maximization

The conditional probability of type k given data x_i and current guesses of parameters, are derived using Bayes rule.

$$q(k|x_i;\hat{q},\hat{\gamma},\hat{\theta}) = \frac{\hat{q}_k q(x_i;k,\hat{q},\hat{\gamma},\hat{\theta})}{\Sigma_{k'}\hat{q}_{k'}q(x_i;k',\hat{q},\hat{\gamma},\hat{\theta})} \tag{D.11}$$

We estimate the model using the following iterative process.

1. Initial guess of q^0, γ^0, θ^0
2. Calculate conditional probability in Equation D.11 using the initial guess q^0, γ^0, θ^0
3. Solve maximization problem for q. It has a closed-form solution which is

$$q_k^1 = \frac{1}{I}\Sigma_i q(k|x_i;q^0,\gamma^0,\theta^0) \tag{D.12}$$

4. Taking other parameters as given, solve the maximization problem of **??**. Get θ_{LC}^1 using the Generalized Method of Moments (GMM) procedure.
5. Taking other parameters as given, solve maximization problem of **??** using MLE, get θ_{SC}^1 and γ^1.
6. Taking other parameters as given, solve the maximization problem of **??**. Get θ_{EC}^1 using MLE.
7. Repeat 2-6 until convergence

D.3 Likelihood Function

We presents the likelihood function P^{LC}, P^{SC}, and P^{EC} in this section.

For convenience, we introduced some notations.

$$\tilde{u}_{i\vartheta} = Z_i\beta_\vartheta \tag{D.13}$$

$$\tilde{u}_{ij} = X_j\beta_{Xi} + c(d_{\ell ij}, Z_i) \tag{D.14}$$

$$\tilde{v}_{i\ell} = W_\ell\alpha_{wi} + p_\ell\alpha_{pi} + \xi_\ell + \mathbb{E}_{\varepsilon_{ij},\rho_i,\varepsilon_{i\vartheta}}\left\{U_i^*(\ell)\middle|\ell\right\} \tag{D.15}$$

Each denotes utility from outside option ϑ, school j, and location ℓ at each decision stage net of idiosyncratic preference shocks.

With the distributional assumption on $\eta_{i\ell}$, the likelihood of location choice P^{LC} takes a simple form.

$$P^{LC} = \Pi_i \frac{exp(\tilde{v}_{i\ell_i})}{\Sigma_{\ell'} exp(\tilde{v}_{i\ell'})} \tag{D.16}$$

where ℓ_i is the observed location choice of household i.

We can extend the formula to construct the likelihood of school application P^{SC}. Assuming logit shock and weak truth-telling among eligible options,

$$P^{SC} = \Pi_i \underbrace{\frac{exp(\tilde{u}_{ij_1^i})}{\Sigma_{j' \in \tilde{\mathcal{J}}_i} exp(\tilde{u}_{i\ell'})}}_{\text{first-ranked}} \underbrace{\frac{exp(\tilde{u}_{ij_2^i})}{\Sigma_{j' \in \tilde{\mathcal{J}}_i / j_1^i} exp(\tilde{u}_{i\ell'})}}_{\text{second-ranked}} \cdots \underbrace{\frac{exp(\tilde{u}_{ij_{l_i}^i})}{\Sigma_{j' \in \tilde{\mathcal{J}}_i / \{j_1^i, j_2^i \cdots j_{l_i-1}^i\}} exp(\tilde{u}_{i\ell'})}}_{l_i^{th}\text{-ranked}} \tag{D.17}$$

where i's observed ranked-ordered list is $\{j_1^i, j_2^i, \cdots, j_{l_i}^i\}$, and $\tilde{\mathcal{J}}_i$ is the set of eligible options for i. Each term in the product represents the probability of choosing the option among eligible options that are not ranked higher.

Similarly, the likelihood of enrollment decision is a product of each i's likelihood, which takes different forms depending on the observed enrollment choice.

$$P^{EC} = \Pi_i \underbrace{\left(\frac{\Sigma_{j' \in \mathcal{J}_i} exp(\tilde{u}_{ij'})}{exp(\tilde{u}_{i\mu_i})}\right)^{\mathbb{1}(\vartheta_i = \mu_i)}}_{\text{adjustment when choose } \mu_i} \frac{exp(\tilde{u}_{\vartheta_i})}{exp(\tilde{u}_p) + exp(\tilde{u}_c) + \Sigma_{j' \in \mathcal{J}_i} exp(\tilde{u}_{ij'})} \tag{D.18}$$

The second term denotes the probability of choosing the enrollment option out of all available options including outside options. The first component adjusts that the distribution of $\varepsilon_{i\mu_i}$ conditional on being assigned to μ_i should be different from the marginal distribution of ε_{ij} and $\varepsilon_{i\vartheta}$. It is more realistic to have $\varepsilon_{i\mu_i}$ preserved the same in the application and enrollment decision stage, rather than assuming that a new $\varepsilon_{i\mu_i}$ is drawn from the Gumbel distribution with a location of zero and scale normalized to one. Without the first term, the value of private options should be underestimated.

The proof is available upon request.

D.4 Supplementary Tables

Table D.11: Demand Estimates from *OnlySC* Version

	Main	**Additional Effects**				
	Type1	Type2	Type3	Black/Hisp	FRL	High-achieving
Panel A: School Demand						
Mean test score	0.139	0.161	0.382	0.165	-	-0.245
Frac. Black or Hisp.	-2.002	-0.430	0.794	2.189	0.126	0.355
Frac. FRL	-0.330	-1.077	0.947	-0.590	0.860	-0.278
Non-safety	-0.036	-0.020	-0.007	0.010	-	-
Distance (mi.)	-0.190	-0.937	-5.331	0.096	0.012	0.029
Prob.	0.294	0.629	0.077	-	-	-
Panel B: Outside Option						
Non-public	-2.088	-	-	0.227	-0.879	0.505
Public Charter	-3.307	-	-	1.685	-0.373	-0.178

Note: Columns are for students' heterogeneity and rows are for school characteristics. FRL stands for Free-or-reduced Lunch eligibility. The fastest driving distance to school is calculated using the Open Route Service (ORS). School non-safety measure is constructed by running a Principal Component Analysis (PCA) on crime incidence of different categories at each school building.

Table D.12: IV Regression for Housing Cost

	(1) OLS	(2) 2SLS
Dep. var: Mean utility δ_ℓ		
Housing characteristics	Yes	Yes
Land use	Yes	Yes
Neighborhood characteristics	Yes	Yes
log(UnitHousingPrice)	-0.014	-2.441**
	(0.056)	(0.561)
N	1690	1690
First Stage F-stat		17.76
R2	0.641	0.233
ymean	-0.879	-0.879

Note: Instrument variable is the percent park area and the percent residential area of locations that are 2 miles away but within 3 miles from each location.

Table D.13: Principal Component of the School Safety Indices

	1st Component	Unexplained variance (percent)
Eigenvalue	3.246	
Total variance explained	64.920	
Eigenvectors:		
Major Crime	0.423	41.900
Other Crime	0.502	18.280
Non-Crime Incidents	0.413	44.520
Property Crime	0.444	36.110
Violent Crime	0.449	34.570

Note: Source - School Safety Report collected by the New York City Police Department which reports the number of crime cases at each school building. Major crimes include burglary, grand larceny, murder, rape, robbery, and felony assault. Other crimes include many crimes that range in severity such as arson, sale of marijuana, or sex offenses. Non-criminal incidents include actions that are not crimes but disruptive such as disorderly conduct, loitering, and possession of marijuana.

Figure D.4: Location Sorting on Unobserved Type

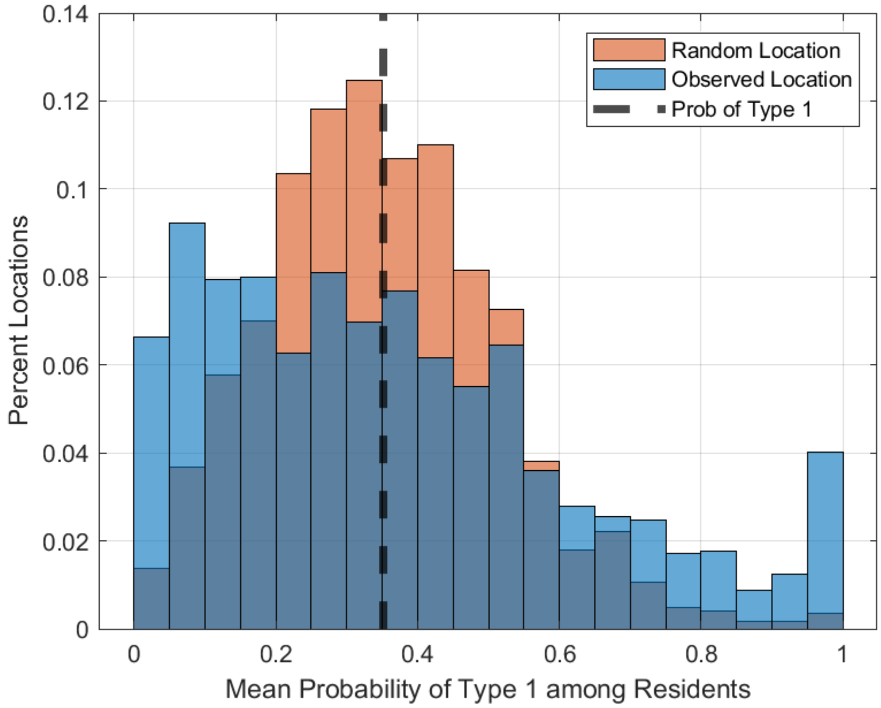


Table D.14: Model Fit

	All		Minority		FRL		High-achieving	
	data	model	data	model	data	model	data	model
Panel A1: Mean of Top 3 Ranked School Characteristics								
Mean test score	0.570	0.564	0.215	0.222	0.401	0.404	0.919	0.881
% Black or Hispanic	-0.107	-0.111	0.026	0.024	-0.054	-0.056	-0.212	-0.212
% FRL	-0.061	-0.054	-0.004	-0.001	-0.020	-0.018	-0.115	-0.104
Non-safety	-1.165	-1.091	-0.325	-0.183	-0.851	-0.810	-1.910	-1.766
Distance	1.965	2.032	1.975	2.052	1.949	1.976	1.974	2.065
Panel A2: Mean of Assigned School Characteristics								
Mean test score	0.311	0.352	-0.068	0.016	0.107	0.181	0.704	0.699
% Black or Hispanic	-0.098	-0.087	0.049	0.054	-0.039	-0.029	-0.204	-0.190
% FRL	-0.039	-0.040	0.023	0.017	0.006	-0.001	-0.097	-0.092
Non-safety	-1.037	-0.800	-0.095	0.174	-0.644	-0.493	-1.870	-1.546
Distance	1.382	1.867	1.388	1.940	1.352	1.834	1.455	1.878
Panel B: % Choosing an Outside Option								
Private	5.020	5.844	4.439	5.104	4.187	4.696	5.359	6.359
Charter	4.310	4.903	6.259	7.091	5.079	5.704	3.235	3.724
Panel C: Mean of Chosen Location Characteristics								
log(UnitPrice)	12.846	12.847	12.702	12.702	12.772	12.772	12.963	12.965
% Black or Hispanic	0.568	0.568	0.731	0.731	0.631	0.631	0.459	0.458
Median HH Income	10.696	10.696	10.562	10.561	10.604	10.603	10.814	10.814
Med. travel time to Work	0.759	0.759	0.772	0.772	0.770	0.770	0.747	0.747

E Supplementary Materials for Section 7

E.1 Segregation Measure: Theil's H Index

Theil's H Index is also known as the Information Theory Index or the Multigroup Entropy Index. In this paper, we closely follow the definition used by the United States Census Bureau to describe housing patterns (Iceland, 2004).[40]

First, the entropy score of the entire economy is calculated as:

$$E = \sum_{r=1}^{R} (\Pi_r) \log(1/\Pi_r)$$

where Π_r is a particular racial group r's proportion in the whole population in the economy. The entropy score measures the diversity in the economy, where higher number indicates higher diversity.

Next, for each school $j = 1, 2, \cdots, J$, the entropy score of j is calculated similarly:

$$E_j = \sum_{r=1}^{R} (\Pi_{r,j}) \log(1/\Pi_{r,j})$$

where $\Pi_{r,j}$ is a racial group r's proportion in the whole population in school j.

Finally, Theil's H index is calculated as the weighted average of deviation of each j's entropy from the entropy score of the entire economy, where the weight is the number of students at each school:

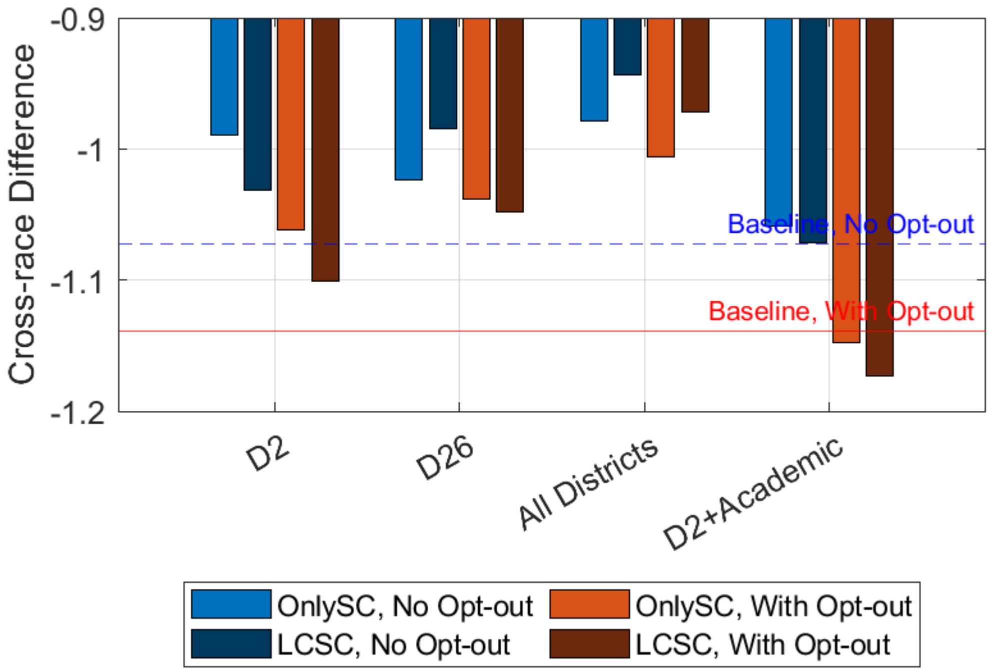
$$H = \sum_{j=1}^{J} \left[\frac{t_j(E - E_j)}{E \cdot T} \right]$$

where t_j is the total number of students in school j, and $T = \sum_{j=1}^{J} t_j$ is the total number of students in the economy. By construction, H is between 0 and 1 where 0 means maximum integration (i.e., all schools have the same racial composition as the whole economy), and 1 means maximum segregation.

E.2 Supplementary Figures

[40]See https://www.census.gov/topics/housing/housing-patterns/about/multi-group-entropy-index.html

Figure E.5: Cross-racial Gap in School Characteristics



Note: The figure shows the difference in mean test scores of coassigned/coenrolled peers between Black/Hispanic and other students across scenarios. There are 4 different counterfactual policies. Under **D2**, we introduce purely-lottery based admissions to schools in school District 2. Under **D26**, we target District 26. Under **All District**, we target all schools within the system. Under **D2+Academic**, we scrap location-based admissions rules among District 2 schools while keeping academic screening in place. The dotted line presents the cross-racial gap in coassigned peers' test scores under the status quo. The solid line presents the cross-racial gap in coenrolled peers' test scores under the status quo.