Distance to Schools and Equal Access in School Choice Systems

Mariana Laverde*†

June 2024

Most recent version

Abstract

This paper studies the impact of geography on cross-racial access to schools under school choice systems. Using data from Boston Public Schools, I show that white prekindergarteners are assigned to schools that are rated higher using measures of test-score levels, test-score growth, and race-balanced growth, than Black students; and that cross-race school-rating gaps under choice are no lower than would be generated by a neighborhood assignment rule. I find that longer commutes to high-rated schools reduce access for Black students. Consistent with a more favorable geography; Hispanic students, on the other hand, sort toward high-growth and race-balanced growth schools under choice.

Keywords: School choice, distance to schools, racial opportunity gaps

JEL: D47, I20, I24

*I am grateful to Scott Ashworth, Kerwin Charles, Steven Durlauf, and Seth Zimmerman for continuous advice and support. I also thank Santiago Caicedo, Michael Dinerstein, Margaux Luflade, Peng Shi, Charlie Murry, Parag Pathak and seminar participants at the University of Chicago, Boston College, University of Wisconsin-Milwaukee, Southern Methodist University, Clemson University, The Boston Fed, ISU Market Design Conference, Boston University, Rice University, Universidad de la Republica, FGV-Rio, Universidad de los Andes, Universidad del Rosario, University of New Hampshire, and others for feedback. I thank Lisa Harvey and Apryl Clarkson from Boston Public Schools for providing data and guidance. I gratefully acknowledge support by the Successful Pathway from School to Work initiative of the University of Chicago, funded by the Hymen Milgrom Supporting Organization. The conclusions of this research do not necessarily reflect the opinions or official position of Boston Public Schools or the Massachusetts Department of Education. All errors are my own. This paper previously circulated under the name "Unequal Assignments to Public Schools and the Limits of School Choice."

†Department of Economics, Boston College. Email: laverdem@bc.edu

1 Introduction

Since the late 1980s, many cities across the United States have adopted centralized school choice systems.[1] These systems give families a choice among public schools, unlike neighborhood assignments, in which school districts assign students to schools based on proximity to residences. Neighborhood assignments replicate residential segregation and can sustain educational inequality across racial and income groups. Proponents of choice argue that by decoupling residences and schools, choice systems can reduce school segregation, improve student-school match quality, and equalize access to high-quality schools. In doing so, these systems generate competitive pressures that drive ineffective schools to improve. School districts emphasize that a guiding principle of school choice is creating equitable access to high-quality schools.[2]

This paper asks how effectively choice systems reduce cross-racial gaps in access to high-quality schools. Using assignment data from Boston Public Schools (BPS), I begin by showing that under Boston's choice system, white prekindergartners are assigned to higher-achieving schools than their Black and Hispanic peers; and more effective schools, measured by test-score growth and a racially balanced measure of growth, than Black students. Moreover, I document that Black-white gaps under choice are no lower than would be generated by an assignment based on proximity between residences and schools under current residential choices. This suggests that there are limits to the effectiveness of school choice systems in equalizing access to educational resources.

An effective policy response to the above requires an understanding of why the effects of

[1] According to the nonprofit Education Commission of the States, 47 states plus the District of Columbia have passed laws to allow or mandate a version of school choice. School districts that have implemented open enrollment include New York City, Boston, Cambridge, Charlotte, and New Haven.

[2] Boston Public Schools superintendent expressed this in the proposal for the 1988 choice plan: "My overall goal is to create a student assignment plan that provides all Boston students with high-quality desegregated education" (Boston Desegregation Project 1988). Other examples include the Charlotte-Mecklenburg School District (https://www.cms.k12.nc.us/boe/Pages/2010%20GuidingPrinciplesforStudentAssignment.aspx).

choice are limited. Cross-race differences in choice-based assignments are explained by differences in the demand for schools or by assignment rules that generate different probabilities of assignment conditional on submitted preferences. Under choice systems, both parental demand and assignment probabilities are crucially affected by a family's residential location.

Regarding demand, a large body of evidence shows that parents value proximity to home when choosing a school (see Agarwal and Somaini 2019 for a summary). Parents may care about long commutes to schools because their schedules are not flexible enough to allow them to get their children to schools that are farther away, or they might worry about longer commutes on school buses. If so, the benefit of attending high-quality schools may be lower than the costs for families who live far from them.

Turning to assignment rules, those that link assignments to a student's residential location may contribute to the gap. Two assignment rules in Boston, which are prominent in many choice systems, may do this. First, students are prioritized for assignment based on proximity to schools. This means that students who live closer to high-quality schools are more likely to get assigned to these schools.[3] The second rule limits the menu of schools a student can apply to based on their residential location.[4] Differences in these menus can result in differential access to educational resources.

Understanding and quantifying how location matters for choice-based school assignments is critical to evaluating such widely implemented policies. While choice systems are touted as equity enhancing for weakening the link between residences and schools, this paper shows that geography remains a crucial barrier to equalizing access to high-quality schools. Moreover, a weaker demand for higher-quality schools explained by geography can erode competitive pressures on ineffective schools, limiting the potential for system-wide improvements in quality.

[3] Examples of other cities that implement proximity priorities include New York City and Barcelona (Abdulkadiroğlu et al. 2005, Calsamiglia and Güell 2018).

[4] BPS has had this type of restriction since the early 1990s, and it modified these menus in 2014, after the end of my study period.

To disentangle the contribution of distance and assignment rules, I estimate a model of school demand using data on the rankings submitted by all applicants to prekindergarten in BPS between 2010 and 2013. The model captures preference heterogeneity by exploiting choice data linked to the residential location of families, allowing estimation by clusters as in Abdulkadiroğlu et al. (2020), Hastings et al. (2017), and Langer (2016). Under assumptions described in Section 4, I estimate parental preferences for proximity and parental mean school values net of travel costs. I then use the model estimates to generate counterfactual assignments that quantify the contribution of assignment rules and distance to cross-race differences in access to high-rated schools in Boston.

I find that differences in distance to high test-score schools explain between 33% and 44% of the Black-white gap and 11% and 22% of the Hispanic-white gap in test-score levels. Turning to effectiveness ratings, I find that if Black families faced the menu of distances to effective schools of white families, access for Black students would improve beyond that of white students. Eliminating proximity priorities and menu restrictions has a negligible impact on the distribution of school ratings by race. The salience of travel costs for the resulting school choice assignments has important policy implications. It shows that even though choice systems give all families the option to apply to the best schools in the city, if they face different costs of accessing those schools, gaps in access to high-quality schools will not vanish through school choice alone.

The empirical analysis comes with two challenges. The first is measuring school quality for the schools that prekindergarteners attend.[5] Data on standardized test scores are only available starting in grade 3, and estimation of school effectiveness also requires student-level test scores in two or more grades. In the absence of student-level data, and to minimize the number of schools dropped from the sample, I mainly restrict the analysis to schools that offer at least third and fifth grade. These schools represent 89% of the schools that

[5] I am interested in a measure of school value-added for schools that offer a prekindergarten program, not prekindergarten value-added, since many students will continue with their prekindergarten assignment in the years to follow.

enroll prekindergarteners and 90% of prekindergarten admissions. Since test scores might not be good predictors of school quality if higher-income and nonminority students tend to have better outcomes for reasons unrelated to the quality of the schools they attend, I use two measures that predict school effectiveness. The first, test-score growth, measures improvements in average math and reading achievement percentiles of the students enrolled at each school. This measure is produced by the school district and is publicly available. While growth takes care of selection by controlling for prior student scores, it is correlated with the share of white students, which could signal some degree of bias. The second measure, race-balanced test-score growth, is built following Angrist et al. (2022). The authors show that this measure predicts value-added at least as well as unadjusted measures in New York and Denver, and is uncorrelated with the share of white students by construction. I complement the analysis with measures of test-score levels. In this case, I use all schools that offer at least third grade.

The second challenge is identifying parental preferences over distance. Since families may be choosing to live close to their preferred schools, the distance to a school is potentially correlated with unobserved school tastes. In this case, the estimated distance parameters would be biased downward and we would conclude that families care about distance more than they really do. This issue has been emphasized by Agostinelli et al. (2021) and Park and Hahm (2023). To make progress with this challenge, I run sensibility analyses where I estimate counterfactuals assuming various degrees of bias. I start with the assumption that preferences for distance are overestimated by 15%, as in Park and Hahm (2023), and also consider these to be overestimated by as much as 40%. A larger bias implies the estimated impact of location is smaller, but results are not highly sensitive within the range considered. I conclude bias would need to be substantial for results to vanish.

Related Literature. This paper contributes to the literature that examines the effectiveness of school choice policies. A first set of papers studies the impact of choice in generating system-wide improvements in school effectiveness (Campos and Kearns 2021, Hoxby 2003, Chubb and Moe 1990, Friedman 1982) and the limits of choice systems in doing so if parents do not rank schools on the basis of school effectiveness (Abdulkadiroğlu et al. 2020, Hastings et al. 2009, Barseghyan et al. 2014, Borghans et al. 2015). Adding to this literature, our paper shows that if effective schools are concentrated in some areas of a city, parents may rank nearby lower-quality schools over more effective ones that are farther away, lowering demand pressures for quality improvement.[6]

A second set of papers studies the impact of school choice on student sorting and school integration. In this strand of literature, a first group of papers focuses on studying the effects of voucher policies on the composition of the student body by achievement and income, in both the public and private sectors (Epple and Romano 1998, Epple et al. 2004, Hsieh and Urquiola 2006, Altonji et al. 2015). More recently, a set of papers has studied the impact of open-enrollment choice policies on school integration. These papers evaluate the impact of admission criteria, residential sorting, limited consideration, beliefs about admission chances, or willingness to submit risky applications on school segregation (Oosterbeek et al. 2021, Lee and Son 2022, Idoux 2022, Calsamiglia et al. 2021). Relatedly, Angrist et al. (2021) use an instrumental variables approach to study the impact of being assigned to a non-neighborhood school on integration and student achievement in Boston and New York City. While students traveling to a non-neighborhood school experienced increased integration, there was little or no effect on academic outcomes. Finally, Sartain and Barrow (2022) study the mechanisms that explain lower access to high-achieving schools for Black and low-income students in Chicago. Distinctively, this paper identifies the effect of families' residential location on their access to educational resources under choice systems.

This paper also contributes to the literature on neighborhood effects. Growing up in low-opportunity areas has been found to be related to low adult earnings and educational achievement (Chetty et al. 2014, Chetty et al. 2016, Chetty and Hendren 2018, Chetty et al. 2018), and some of these effects may be explained by the quality of public education in these ar-

[6] Neilson (2013) and Allende (2019) find that horizontal differentiation across schools explained by distance contributes to reduced competition in Chile and Peru—systems with public and private supply and choice in the form of vouchers.

eas (Biasi 2019, Laliberte 2018). This paper shows a first-order effect of location on access to educational resources. The salience of travel costs is consistent with results that show substantial spatial variation of place-based effects for geographic areas as small as census tracts.

The rest of the paper is organized as follows. Section 2 discusses the institutional context and the data. Section 3 summarizes the main observed differences in application behavior and discusses evidence for the mechanisms. Section 4 presents the model used to recover demand parameters, discusses the assumptions and estimation, and analyzes the results. Section 5 describes the methodology and assumptions required for the counterfactual exercises in this section and discusses the results. I conclude in Section 6.

2 School Choice in Boston

2.1 The Assignment Mechanism

Parents who wish to apply for a prekindergarten seat in a school within BPS are required to submit to the school district a ranking of programs and schools ordered by preference. A school typically offers a couple of general education programs and programs for language learners.[7] Students can rank any number of programs subject to the condition that the programs are housed in a school the student is eligible for. Eligibility is determined by the student's residential location. During the study period, Boston was divided into three zones: the north, east, and west (Figure A.1a). Students were eligible for any general education program in their residence zone plus any within a mile of their home. Geographic restrictions that determine eligibility for language programs are similar to those for general education programs; although these restrictions are not always binding (Pathak and Shi 2013a). I

[7] At the prekindergarten level general education programs are typically referred to as inclusion programs. I exclude from my analysis students applying to substantially separate programs since assignments for these students do not always follow the assignment rules and allow for exceptions when needed.

assume, as Pathak and Shi (2013a) do, that English-language-learner students can apply to any program across the city. A handful of citywide schools accept applications from students all over the city. I refer to the set of schools a parent can apply to as the parent's choice menu. Figure A.1b shows a partition of the city that groups families with the same choice menu.

Although parents in Boston apply to programs within schools, I make the simplifying assumption that parents rank schools. I transform school-program rankings into school rankings by eliminating instances in which different programs in the same school are ranked and keeping only the first time a school appears in the ranking (a similar assumption is made in Abdulkadiroğlu et al. 2020). Going forward when speaking of preferences I refer to parental preferences for schools.

Students are assigned to schools following a priority structure—defined by the school district—that is common across schools. Under this priority structure, students who have a sibling at a school have a higher priority at that school than students who do not. Also, students who live within a mile of a school—called the walk zone—have priority at that school over students that live farther away. Students who both have a sibling at a school and live in the walk zone have the highest priority. These are followed by students who have a sibling, then those who live in the walk zone. The remaining students have the lowest priority. Ties within each group are broken with a random number assigned to each applicant. This guarantees that priorities generate a strict ordering of students.[8] School districts also determine school capacities—that is, the number of seats available in each program. Preferences, priorities, and capacities feed into the assignment algorithm, which is a version of Gale and Shapley's (1962) student-proposing Deferred Acceptance (DA) algorithm (Balinski and Sönmez 1999; Abdulkadiroğlu and Sönmez 2003).

The DA algorithm guarantees that parents do not have incentives to misrepresent their true

[8] This priority structure is typically used in half of the seats in each school, while the remaining seats ignore walk-zone priorities altogether. A more detailed description of the algorithm is given in Appendix C. Dur et al. (2018) and Sönmez et al. (2019) discuss this design and its properties.

preferences when submitting rankings (Dubins and Freedman 1981, Roth 1982). This holds under the assumption that students are allowed to rank all desirable schools. Instances in which school authorities restrict the length of submitted rankings might not generate truthful reports, even under the DA algorithm (Haeringer and Klijn 2009, Calsamiglia et al. 2010). BPS is one of the few districts that do not restrict the length of the submitted rankings. These properties make Boston a good setting for studying parental school demand.

Students assigned to a school farther than a mile from their homes are eligible for free bus transportation to and from school. The pickup and drop-off location is set by the district to a location within a mile of a student's home. BPS estimates that the majority of riders are in elementary school and attend a school with a high population of low-income families. Among prekindergarten students, around half opted in for school transportation.

2.2 Data

I use two main data sources. The first is data from BPS that cover the universe of first-round applicants to prekindergarten between 2010 and 2013. For each applicant I observe the rank-ordered list submitted, the school assigned or an indicator for whether the student was unassigned, and the priority that generated the assignment.[9] I also observe the residential location and demographic information of the student including their race.[10] First-round applicants represent over 80% of admitted students (Pathak and Shi 2017); the rest apply in the second round and are assigned after first-round applicants.

Using the location of each school and the geocode of each student's residence, I measure the

[9] A student will be unassigned if they are rejected from every school on their submitted rank list. Students who are unassigned in the first round can reapply in the second round or search for options outside the school district.

[10] Residential locations are coded by the school district at the geocode level. Geocodes partition the city into 868 polygons of average area of 0.1 square miles. The assignment algorithm is built using such geocodes; hence that level of aggregation does not represent any loss of information for the purposes of the assignment algorithm. I remove from my sample students with an invalid geocode, representing around 2% of the sample.

distance between students and schools in one of two ways: as the walking distance between the geocode's centroid and the school or as the linear distance between the two points. The former is obtained using Google Maps travel estimates. Using these locations, I also generate the walk-zone priority status for each student-school pair and the choice menu of each student, recreating the procedure used by BPS.[11]

Ideally, I would have the sibling priority status of every student at every school. In reality, I only observe the sibling priority status of student i at school j if i was assigned to j with this priority. Throughout the analysis, I assume that all students that are not assigned with a sibling priority do not have a sibling at any school and that students assigned with a sibling priority at j do not have a sibling at another school. Using data on the priorities that generated each assignment, I find evidence in support of this assumption: in most schools every student who applied with a sibling priority was admitted. This means that for the set of schools each student ranks, I observe sibling status when it indeed exists, with the exception of students who have a sibling priority at multiple schools or those who rank the sibling school sufficiently low and are assigned to a school ranked higher. In the first case, I would only be able to account for the sibling status at the sibling's school that is ranked higher.[12]

The second data source has yearly school characteristics and comes from the Massachusetts Department of Education. Using this data I build three measures of school ratings: test score levels, test score growth, and race-balanced growth. *Levels* ratings are built from the

[11]Student i is in the walk zone of school j if a one-mile radius from school j intersects the geocode of residence of i. Similarly, I define the choice menu of each student using data on the zone in which each school and geocode lies.

[12]If the following conditions are satisfied, a school did not reject a student with a sibling priority: First, if there are fewer assigned students than available seats, then no student was rejected. Second, if a school accepted a student with either the walk-zone priority or with no priority, then that school did not reject anyone with a sibling priority. Otherwise, the resulting match would not be stable. The number of schools that do not satisfy either of these conditions in 2010 is three, in 2011 it is two, and in 2012 it is six. For these schools I cannot rule out that they rejected a student with a sibling priority.

share of enrolled students scoring advanced or proficient at each school on the Massachusetts Comprehensive Assessment System (MCAS) tests, averaged across math and reading, and standardized to have zero mean and a standard deviation of one in each year. MCAS scores are available for every school that offers at least third grade. *Growth* ratings are based on district-built measures of year-to-year improvement in enrolled students' achievement percentiles. I average measures across math and reading, available for schools that offer at least fifth grade, and standardize them to match the distribution of level ratings. Growth measures predict school value-added better than raw achievement measures do, with modest bias (Angrist et al. 2017, Angrist et al. 2022).

The measure of *race-balanced growth* builds on work by Angrist et al. (2022) and builds the measure as the residual of a regression of growth ratings on the share of white students at each school. Angrist et al. use data from Denver and New York City to show that this measure has larger predictive accuracy to school value-added than growth and raw achievement. Since schools in my sample offer different grade levels, in the final step I residualize all three measures with dummies for the maximum grade at each school.

Students. Table 1 describes the student sample. The sample has $8,869$ applicants to prekindergarten schools between 2010 and 2013. Close to half of the applicants are Hispanic, while Black and white students constitute around one-fifth of the sample each. Asian and other minority families make up around 10% of the applicant pool. This composition is in contrast to Boston's residential makeup, as white residents account for about half of Boston's population. Selection of the white student population outside the BPS sample, is likely to be higher income and use private prekindergarten options. Gaps in school ratings in the BPS population potentially underestimate those of the overall Boston population. We will focus on the BPS population.

Families who apply for a prekindergarten seat in BPS can choose from a set of 25 schools on average. This contrasts with other school choice settings, such as New York City's high school system, in which families choose from about 700 options (Lee and Son 2022). Out

Table 1: Student Descriptive Statistics

	All	Black	Hispanic	White	Asian	Other
Applicants	8,869	22.9%	42.8%	22.8%	7.8%	3.6%
Tract Income	55,551	43,705	49,873	76,753	55,166	63,660
	(25,429)	(19,205)	(21,711)	(24,850)	(22,875)	(27,363)
Applications						
Size of Choice Menu	24.8	26.0	24.8	23.5	25.0	24.4
	(2.4)	(2.2)	(2.4)	(1.9)	(1.9)	(2.3)
Distance in Choice Menu	2.6	2.4	2.7	2.7	2.6	2.5
	(0.8)	(0.7)	(0.9)	(0.8)	(0.8)	(0.8)
Maximum Distance in Choice Menu	5.6	5.5	5.8	5.3	5.9	5.3
	(1.3)	(1.1)	(1.3)	(1.5)	(1.2)	(1.4)
Length of Submitted List	5.0	5.5	5.0	4.8	4.1	5.7
	(3.1)	(3.4)	(3.0)	(2.8)	(2.7)	(3.6)
% English Language Learners	37.5	19.4	58.2	11.4	64.7	11.7
Assignments						
Assigned Rank	1.8	1.9	1.7	1.7	1.6	2.3
	(2.2)	(2.1)	(1.8)	(2.7)	(1.6)	(3.3)
Distance to Assigned School	1.2	1.3	1.3	1.0	1.1	1.2
	(1.3)	(1.3)	(1.3)	(1.0)	(1.1)	(1.2)
% Assigned with Sibling Priority	36.0	31.3	34.4	43.8	40.0	33.9
% Assigned with Walk-Zone Priority	48.4	47.4	46.6	53.5	48.1	49.1
% Unassigned	26.1	23.0	24.2	33.2	22.7	30.8

Note: The first row shows the total number of applicants and their racial makeup. The second row shows the average tract-level household income from the five-year 2010 ACS (I match the geocode of each applicant to a census tract by overlaying both geographies and keep the tract with the largest share of each geocode's area). I show the mean and the standard deviation in parentheses, except for variables marked as percentages. Distances are measured in miles. The length of the submitted list and the rank of the assigned school are computed using school rankings transformed from school-program rankings. Students assigned with a sibling or walkzone priority are expressed as a percentage of assigned students. Students assigned with a sibling and walkzone priority are included in both categories.

of the options in Boston, families typically rank 5. Black students submit longer lists, while white students submit shorter lists, possibly because outside options of white families are ranked higher among public schools.[13] Students who are unassigned after running the assignment algorithm may apply in a subsequent round. Since prekindergarten attendance is not mandatory, some applicants are not assigned to any school and need to search for options outside of the public school district. About a quarter of the students that apply in the first round are unassigned, and out of all unassigned students nearly 75% do not enroll in any public school.

Figure A.2 shows the spatial distribution of students by race. Although there are clear sorting patterns, students of all races can be found across the city. One way to quantify this is to zoom in on each school and see the racial distribution of residents within a close buffer. If I consider a radius of 3.8 miles around every school, which is the average distance to the farthest ranked school, I find that on average there are about two hundred students of each race and income group who can apply to each school; and for all schools there are students of all races. Similarly, looking at applications, I find that the average school has applicants from each race and income group (Table A.1).

Schools. Between 2010 and 2013, a total of 67 public schools offered a prekindergarten program, and not all schools had prekindergarten seats in all years. There is substantial variation in students' demographic characteristics and school performance measures. On average 39% of students at each school scored advanced or proficient in math and reading; the school with the poorest achievement had 10% of students scoring advanced or proficient, while for the highest performing school, the percentage was close to 80%. In terms of growth, the school with the lowest growth has its average student around the 11th percentile for English and 14th for math; these numbers go close and above the 80th percentile for the higher-growth schools. On average, schools have 32% Black students and 21% white students, while both groups represent about 20% of applicants. The distribution across

[13] Lee and Son (2022) documents something similar in the case of New York City's high school choice system.

Table 2: Descriptive Statistics: Schools

	Mean	Std. Dev.	Min.	Max.
Prekindergarten Capacity	30.9	15.7	6	108
% Advanced or Proficient in English	39.2	14.0	10.0	80.0
% Advanced or Proficient in Math	38.9	14.2	10.0	82.0
English Growth Percentile	48.5	11.2	18.0	78.5
Math Growth Percentile	53.5	13.5	20.0	85.0
% Black Students	32.0	19.3	2.1	79.7
% Hispanic Students	44.2	19.3	14.3	91.1
% White Students	20.7	17.6	0.6	71.6
% Low-Income K Students	67.5	19.8	7.7	100.0
Observations		258 (67 distinct schools)		

Note: Statistics for enrolled students in all grades, except for the first and last row. There are 22 (8.5%) school-year observations with missing achievement and 28 (10.8%) school-year observations with missing growth. These are schools that do not offer third and/or fifth grade.

schools is uneven, with some schools having as little as 1% and 2% of white and Black students, and others having close to 80% and 72%, in each case. Each school has close to 70% low-income students, and the school with the lowest percentage of low-income students has 8% (Table 2).

3 The Gap in School Ratings and the Possible Explanations

In this section I describe the main facts concerning school-ratings gaps that motivate the paper. Then, I discuss the mechanisms that might explain the gaps. Finally, I provide some evidence on the relevance of each mechanism.

3.1 The Racial Gap in School Ratings

Between 2010 and 2013, white prekindergarteners in Boston were assigned to schools with higher test score levels and a smaller fraction of low-income students and non-white students than their Black and Hispanic peers. Measures of achievement might not be good predictors of school quality if nonminority students tend to have better achievement for reasons unrelated to the quality of the schools they attend. Estimating the causal impact of a school on student test scores requires rich data unavailable for the majority of schools in my sample. In this paper, I use growth and race-balanced growth ratings to measure school effectiveness, and I complement the analysis using data on level ratings. Section 2.2 discusses details about how these are constructed and the rationale for using them.

I find that Black and Hispanic prekindergarteners in Boston are assigned to schools with an average of 0.8 and 0.6 standard deviations (SD) lower achievement than white students as shown in the first row, columns (5) and (6), of Table 3. As was pointed out before, some of these racial differences are likely explained by differences in peer composition; with white students being assigned to peers who are higher-achieving for reasons partly unrelated to school quality. Looking into growth measures, Black students access schools with an average of 0.2 SD lower growth than white students, and a small gap in favor of Hispanic students is found. Finally, using race-balanced growth, we find a larger gap in favor of Hispanic students and a smaller gap for Black students who access schools with 0.05 SD lower race-balanced growth than white students.

Comparing school assignments under the DA algorithm to hypothetical neighborhood assignments shows choice may have a differential impact across racial groups.[14] While white families appear to sort toward higher achieving peers and not more effective schools under choice, the contrary appears to be the story for Hispanic students who get assigned to schools

[14] I generate the neighborhood assignment by running the DA algorithm taking the set of students assigned via the choice system and redefining their preferences and priorities to be determined exclusively by proximity: students prefer schools closer to home, and schools prioritize students that live closer to schools.

Table 3: School Ratings by Group in DA and Neighborhood Assignments

	All	Black	Hispanic	White	BW Gap	HW Gap
	(1)	(2)	(3)	(4)	(5)	(6)
Test score levels - DA	-0.061	-0.361	-0.136	0.427	0.788	0.563
Test score levels - Neighborhood	-0.074	-0.359	-0.110	0.312	0.671	0.422
Test score growth - DA	0.011	-0.160	0.085	0.051	0.211	-0.034
Test score growth - Neighborhood	0.008	-0.123	0.053	0.058	0.181	0.005
Race-balanced score growth - DA	0.021	-0.105	0.123	-0.051	0.054	-0.174
Race-balanced score growth - Neighborhood	0.024	-0.049	0.087	-0.023	0.026	-0.110
Distance - DA	1.218	1.322	1.262	1.003	-0.319	-0.259
Distance - Neighborhood	0.327	0.294	0.315	0.389	0.095	0.074
% Same race - DA	0.463	0.427	0.542	0.338	-0.089	-0.204
% Same race - Neighborhood	0.431	0.425	0.493	0.309	-0.116	-0.184

Note: The table shows average school ratings, distance, and share of same-race students assigned under the DA and a hypothetical neighborhood assignment, for all students and for students in each group. Columns 5 and 6 show gaps relative to white students for each measure.

with higher growth and race-balanced growth under choice, but not higher test scores. Black families, on the other hand, don't appear to be effectively sorting either on high-achieving peers or school effectiveness even though they are traveling farther than white and Hispanic students do.

These differential forces imply that the Black-white gaps in levels, growth, and race-balanced growth under choice are not lower than those generated under the hypothetical neighborhood assignment. This is not the case for Hispanic students who gain access to more effective schools relative to white students under choice. Here I do not interpret the hypothetical neighborhood assignment as a likely counterfactual under a neighborhood assignment rule, but instead as a benchmark that shows how assignments would look like if students enrolled in their neighborhood schools under current residential choices.[15]

3.2 The Mechanisms

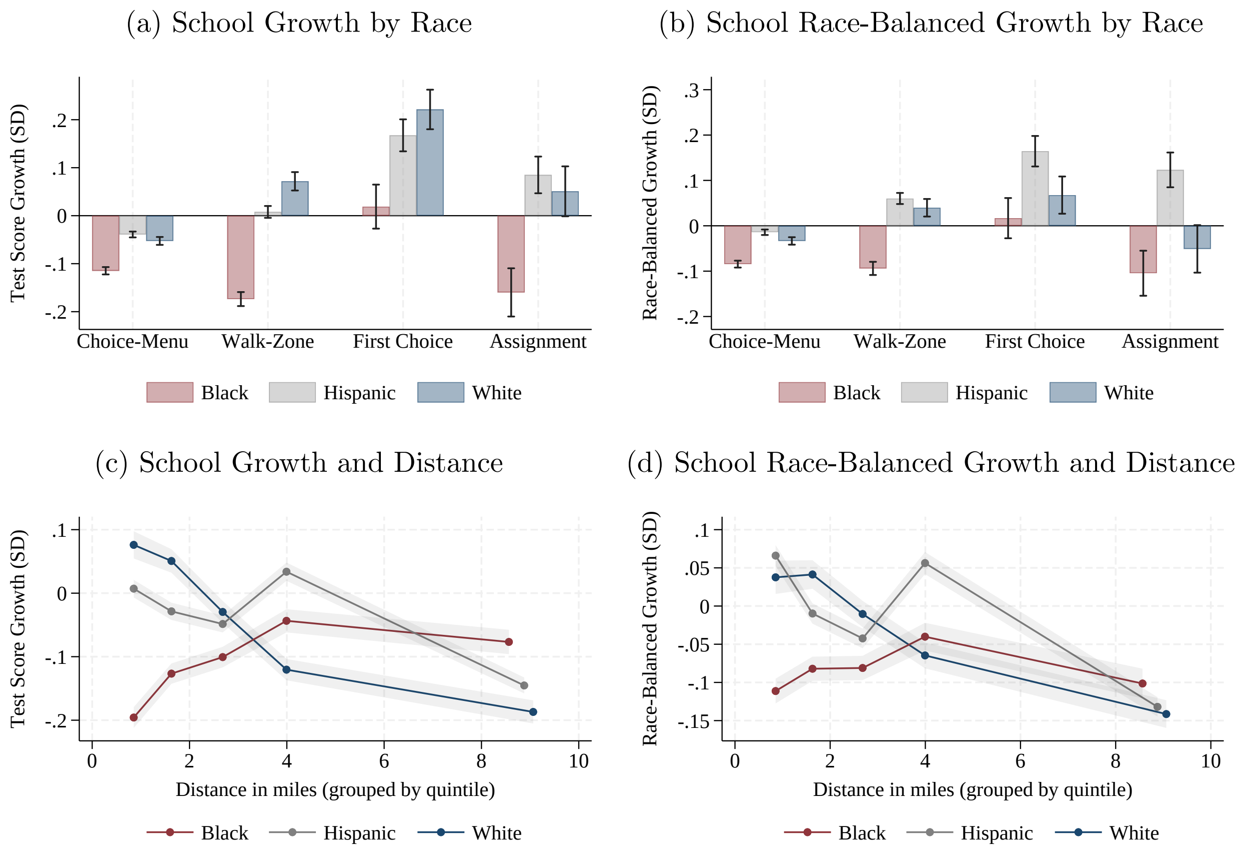
Understanding how location matters for choice-based school assignments is critical to evaluate the equity consequences of such widely implemented policies. Even in a choice-based system in which the link between residences and schools is weakened, the residential location of families may play a crucial role in their school assignments. If parents value proximity, travel costs may offset the benefit of attending high-value-added schools. Also, assignment rules that constrain the choices of families or that prioritize students based on proximity to schools can generate inequities in access even in the absence of travel costs.

Figure 1 shows evidence that location-specific assignment rules and distance may all be relevant to explain the cross-race gap in assigned-school ratings for Black students. To the left of Panels 1a and 1b, differences in average growth and race-balanced growth ratings across the choice menus of students by race suggest the availability of schools may contribute to the gap. Notably, the choice menus of Black students have lower average growth and race-balanced growth than those of their white and Hispanic peers. Gaps are more salient when we consider schools in families' walk zones. This suggests that proximity priorities may affect the gap and that white families tend to be closer to schools that are rated higher. Crucially, this may affect parental demand even if the valuation for all school attributes is the same across groups. Gaps in first-ranked schools show that demand differences may contribute to the gap. Of course, cross-race differences in school demand can be explained by heterogeneity in the valuation of location-independent school attributes or by longer distances for Black families.

Panels 1c and 1d show the rating gradient as the distance between schools and students grows. Two facts stand out. First, Black students live on average farther from high growth

[15] To generate a counterfactual neighborhood assignment a researcher may need to include the endogenous decision of housing location that plausibly follows a change in the assignment rules.

Figure 1: Mechanisms: Rules, Distance, and Preference Heterogeneity



Note: The top graph shows average growth and race-balanced growth weighted by capacity at the schools in the choice menu and walk zone of applicants by race, and the average of these measures for the schools ranked first and the schools assigned to students by race. The bottom graph shows the average of these at schools in the choice menu by quintiles of distance to families' residences. Distance, in miles, is plotted on the x-axis.

and race-balanced growth schools than white and Hispanic families do. Second, as the distance grows, the average rating of schools drops for white families. For Hispanic families, this is true in the most relevant 4-mile radius around their home. The opposite is true for Black families who trade off travel times and ratings, having to travel on average farther to access high-rated schools.

4 Estimating Parental Preferences

In this section, I present the model and assumptions used to recover parental preferences for schools. At the end of the section, I discuss the estimated parameters.

4.1 Model

I model preferences using a random-utility model where $i \in \mathcal{I}$ index a student and $j \in \mathcal{J}$ index a school. To capture rich heterogeneity in preferences, I estimate separate models for six subgroups of students defined by the intersections of students' covariates. The covariate clusters are defined by the intersections of the students' race and students' census-tract income, where students are grouped together if their census-tract income is above or below applicants' median-tract income. This strategy follows that of Abdulkadiroğlu et al. (2020) in a school choice setting and Hastings et al. (2017) and Langer (2016) in other settings. For each cluster c, I use data on individual choices to estimate the model.

For applicant i in cluster c, I model the indirect utility from being assigned to school j as

$$u_{ij} = \beta_c d_{ij} + \gamma_c z_{ij} + \delta_{cj} + \varepsilon_{ij}. \tag{1}$$

The variable d_{ij} denotes the walking distance from i's residence to school j, and z_{ij} is a vector of observable characteristics that vary by student and school within clusters.[16] The parameter β_c summarizes preferences for proximity for parents in cluster c, and δ_{cj} summarizes the location-independent attractiveness of school j for parents in the cluster. This includes parents' assessment of school characteristics that are observable or unobservable to the econometrician. Finally, ε_{ij} represents i's idiosyncratic taste for school j. I assume values of ε_{ij} are independent and distributed type-1 extreme value with scale normalization. I consider a model where the utility of the outside option, u_{i0}, is normalized to zero for all students, and I also consider a model that allows for neighborhood-specific values of the out-

[16] Here I include an indicator of whether school j offers a program in student i's first language.

side option. The latter model captures differences in quality, price, or availability of private childcare options in different neighborhoods.

To model location-specific outside options, I consider a partition of the city into 12 neighborhoods. This partition builds on neighborhood boundaries defined by the city and thus groups areas that have a similar amenity supply. I assume that the residents of neighborhood n that belong to the same cluster share a common value for the outside option. To guarantee the model is identified, I normalize to zero the value of the outside option in a reference neighborhood that intersects the three zones of choice in the city. I estimate the value of the outside option in the other neighborhoods relative to the value of the outside option in the reference neighborhood. Then, if $n(i)$ is i's neighborhood and $c(i)$ is i's cluster, $u_{i0_{n(i)}} = 0$ if $n(i)$ is the reference neighborhood and $u_{i0_{n(i)}} = \kappa_{n(i)c(i)}$ otherwise, where $\kappa_{n(i)c(i)}$ is a parameter of the model. Identification of outside-option values follows from the fact that these are connected strict substitutes of the reference outside option (Berry et al. 2013).

In robustness checks, I also consider a model of utility in which preferences for distance are not linear but quadratic. This model captures the possibility that the first miles traveled are marginally more costly. This would happen if families only considered taking the train or the school bus if their travel exceeded some distance and if the marginal cost of each mile traveled was different across transportation modes.

Identification. Two distinct sources of variation separate school mean utilities and parental sensibility to distance. Rankings of students who are equidistant to any pair of schools generate the variation used to identify school mean utilities, while parents who rank schools farther from their residence above closer schools help identify parents' sensibility to distance. Ranking behavior presents sizable identifying variation, as 84% of the students in the sample rank at least one farther school over a closer school, and on average each student does this 1.9 times.

The main challenge to identification is identifying the distance parameter. If there is within-cluster heterogeneity in school valuations that is correlated with distance to schools—because families choose to live near preferred schools—the distance parameters will be biased away from zero. In this case, we will conclude that families care about distance more than they really do. The more families are able to sort close to their preferred schools the larger the bias.

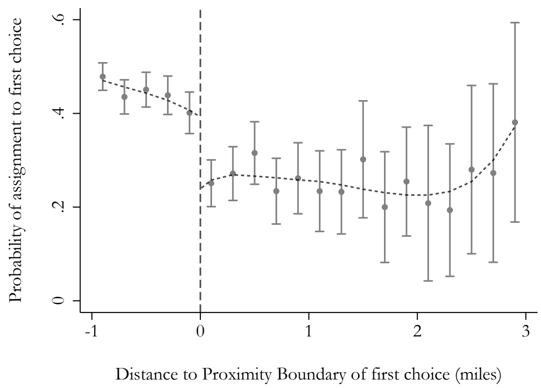
While sorting can't be ruled out, submitted rankings suggest it is not prominent. Out of all families in the sample, 76% rank a school that is not the closest as their first choice.[17] Moreover, 44% of students do not rank the school closest to their homes. This fraction is the smallest for white families in the higher-income group (27%) and is the largest for Black families in the lower-income group (59%). Additional evidence suggesting sorting is not prominent is obtained by studying the geographic discontinuities in admission probabilities generated by walk-zone priorities. Figure 2a shows the discontinuity in the assignment probability to the first-ranked school at the proximity boundary. The graph shows that families benefit from choosing their residence 0.9 miles from a preferred school over 1.1 miles from it. If families are sorting on these boundaries, we may see parents who choose to live less than a mile from a school ranking it in the first position more often than parents who live just over a mile from it. Figure 2, shows the probability of ranking a school in the first position as the distance to the proximity boundary of that school changes. The zero in the x-axis represents the one-mile threshold. To the left of zero, students live within the walk-zone. The downward trend shows that parents value proximity while a discontinuity at zero is evidence of sorting. The plots show no evidence of sorting on these boundaries for students in any group, and for a sample restricted to competitive schools.

While this evidence is reassuring, I do not assume sorting is not present in the data. Instead, I assume some level of bias in the distance parameter may exist and run counterfactuals considering several scenarios. I generate results assuming sorting leads to distance parameters that are overestimated by 15%, following work by Park and Hahm 2023, or 40%.[18] Results

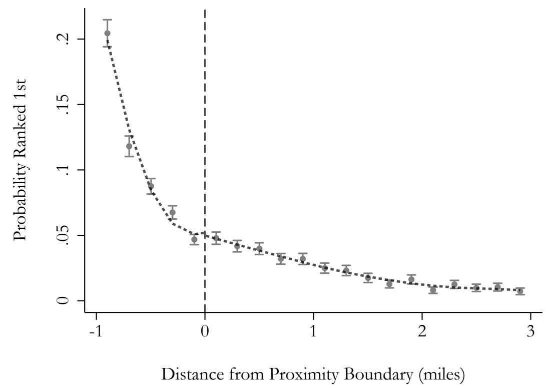
[17] For white families, the share who rank the closest school as the first choice is 33%, which would be consistent with more sorting, or a stronger preference for proximity. For Black families, this share is 17%

[18] I multiply the distance parameter by a distance factor of 0.85 and 0.6 in each case.

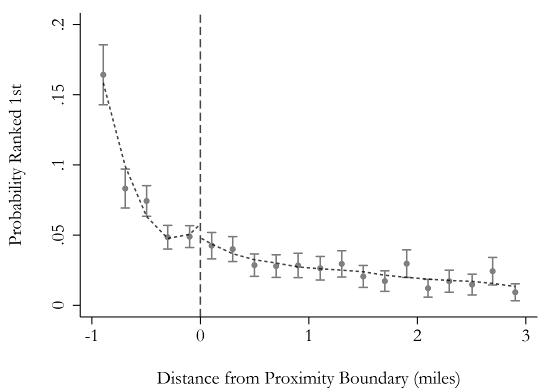
Figure 2: Proximity Priority and Ranking Behaviour



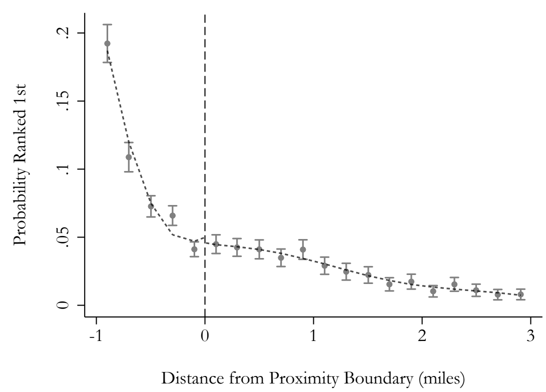
(a) Assignment to First Choice



(b) All Students - First Rank



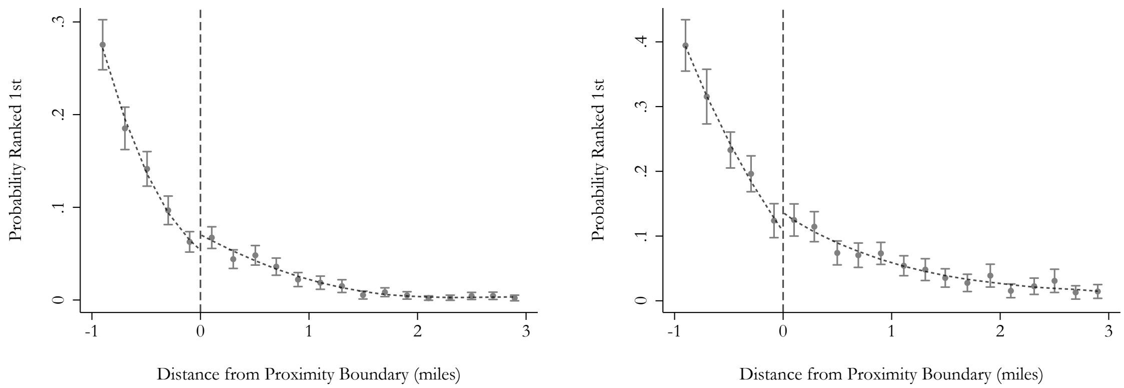
(c) Black students - First Rank



(d) Hispanic students - First Rank

Note: Graphs show bin-scatter plots with equally sized distance bins on each side of the boundary. For every student-school pair, I construct the linear distance between that student and the proximity boundary of the school. Negative values indicate that a student lives within the walkzone zone. Panel A plots the probability of getting assigned to the first choice as the distance to the proximity boundary to the first choice changes. Panels B - F plots the probability of ranking a school in the first choice. Range plots show 95% confidence intervals, while the dashed line represents a local linear fit estimated on each side of the boundary. Competitive schools in panel F are the five schools that are ranked in the first position more often. Results are similar if I instead consider the schools that accepted the least number of students without any priority.

Figure 2: Proximity Priority and Ranking Behaviour (Continued)



(e) White students - First Rank

(f) Competitive Schools - First Rank

are compared to the benchmark that assumes the distance parameter is not biased. This exercise is intended as a sensibility check to measure how results change as the distance parameter varies.

Truth Telling. I assume that the submitted rankings are truthful. This means that parents rank all acceptable schools in true preference order. A school is acceptable if it is preferred to the outside option, which is the best option parents can find if unassigned in the first round. This assumption is motivated by the algorithm's incentive compatibility and the property that the number of schools parents can rank is unrestricted. Such restrictions, even under the DA mechanism, can generate reports that are not truthful (Haeringer and Klijn 2009, Calsamiglia et al. 2010, Luflade 2018).

The truth-telling assumption may be violated if admission outcomes are largely predictable or parents think they are. In this case, parents may misrepresent their preferences by not ranking schools that are desirable but where they perceive a low probability of admission or by not ranking all acceptable schools if they perceive they have high admission chances in a subset of these. Skipping schools that are perceived as impossible to be admitted to is more likely in settings where an applicant knows their own priority and the distribution of priorities before applying—for example, a college choice system where priorities are determined by a test score and historical cutoffs are observable to applicants. In the case of Boston, although parents can observe the category in which they lie in the priority ladder, meaning they know their sibling and walk-zone statuses, they do not observe the random number that determines their actual priority ranking, nor do they observe historical cutoffs used to predict the fraction of admitted students with sibling or walk-zone priority at each school. Moreover, even if parents were able to predict these probabilities with accuracy, an analysis of the admissions data reveals that all schools admitted students without any priority during at least one year in the sample, meaning that across the years the probability of being accepted without a sibling and without walk-zone status was not zero. On the other side, families who believe they have high admission chances may stop adding acceptable schools to the bottom of their ranking. Arteaga et al. (2022) find evidence in Chile and New Haven of families that are too optimistic about their admission chances and skip acceptable schools. In my sample, I find that students in the highest priority group at a school—who have a sibling and walk-zone priority—rank on average three schools below the school they will get assigned to with certainty, suggesting parents do not behave as if they observed true admission chances and do not stop ranking schools as admission chances approach one.

Consideration Set. I assume that students consider all schools in their choice set. This means families can process information about all the schools they are eligible for and can rank all those options. The assumption is motivated by the relatively small size of choice sets in this setting: families have an average of 25 schools to choose from. This is in contrast with assumptions made in Lee and Son (2022), in which families are asked to choose from around 430 high school programs in New York City. Lee and Son estimates that in this context families are aware of about 65 programs.

Formally, the assumptions about consideration sets and truth telling imply that if $R_i =$

$(R_{i1}, \cdots, R_{il_i})$ is the rank-ordered list submitted by i and $\mathcal{J}_i$ is the choice menu of i, then

$$R_{i1} = \underset{j \in \mathcal{J}_i}{\arg\max} \quad u_{ij} \tag{2}$$

$$R_{ik} = \underset{j \in \mathcal{J}_i \setminus \{R_{im} : m < k\}}{\arg\max} \quad u_{ij}. \tag{3}$$

Moreover, if u_{i0} is the utility of the outside option, then

$$u_{ij} > u_{i0} \quad \forall \quad j \in R_i \tag{4}$$

$$u_{i0} > u_{ij} \quad \forall \quad j \in \mathcal{J}_i \setminus R_i. \tag{5}$$

Utility u_{i0} represents the expected utility at the time of the application associated with the best alternative if unassigned in the first round. This may include private childcare options, parochial schools, and other private pre-kindergarten programs. Typically these programs announce admission decisions simultaneously to BPS. While we do observe students who are unassigned after the first round and who enroll in a school in the district in the second round, and students who get admitted to a school in the first round but who don't enroll; we assume these actions are rationalized by shocks to the utility of outside option between rounds one and two. This happens, for instance, when a family overestimates the probability of admission into their first round outside option. In that case, their outside option value receives a negative shock and families may reconsider some inside options in the second round. I do not model these dynamic considerations; instead, I use first round applications and interpret the parameters of the model as a summary of parents' preferences and expectations in the first round.[19]

Estimation and Inference. I estimate preferences for the subsamples of Black, Hispanic, and white students. I do not estimate preferences for Asian students and other racial minorities because their sample size is small.[20] In consequence, there are a total of six covariate clusters, with students spread across the city.

[19] Kapor et al. (2020) estimate interim beliefs in a similar setting.

[20] For these groups I use the submitted rankings instead of simulated rankings in the counterfactuals.

I estimate utility parameters for each cluster by maximum likelihood. Bootstrapped standard errors are obtained by sampling the data by student with replacement, keeping the application profile submitted by each student, and reestimating the model in each of 100 samples.

4.2 Parameter Estimates

Table 4 shows estimated model parameters. Negative distance parameters imply a disutility for traveling, and positive language match parameters show that within clusters, English language learners value being assigned to a school that offers a program in their first language. School mean utilities, δ_{cj}, summarize the cluster-specific average attractiveness of a school after discounting the effect of distance and the language match. Variation in school mean utilities, measured by the standard deviation of the estimated δ's, is largest for families of white students and lowest for families of Black students. The same is true for the neighborhood-level outside option values.

Looking at correlations in the δ's across clusters, higher-income families have preferences that are closer to other higher-income groups, and preferences also tend to be correlated within race and across income. School values for low-income white families are more dissimilar to those of other groups. Low- and high-income Black and Hispanic families have correlated preferences, while higher-income white families behave closest to higher-income Hispanic families.

Table A.3 shows results from independent regressions between standardized values of δ_{cj}/β_c and school observable characteristics. Considering the ratio between school mean utilities and the distance parameter facilitates comparison between clusters as the variable is expressed in terms of standard deviations of miles traveled. Estimates of δ_{cj} and β_c are from a model that assumes a common outside option value across space for each c. School mean utilities have a positive association with test-score levels and with the percentage of white students for all clusters. For example, for lower-income Black families, increasing test-score levels by one

Table 4: Estimated Model Parameters

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
	Panel A. Distance and Language Match					
Distance	−0.46	−0.45	−0.60	−0.67	−0.78	−0.91
	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)	(0.02)
Language Match	0.41	0.41	0.66	0.73	0.98	1.25
	(0.10)	(0.22)	(0.05)	(0.06)	(0.25)	(0.26)
	Panel B. School Mean Utilities					
Standard Deviation	0.76	0.77	0.86	0.85	2.66	2.22
Correlations						
Cluster 1	1.00	0.88	0.86	0.75	0.29	0.38
Cluster 2	0.88	1.00	0.77	0.82	0.28	0.47
Cluster 3	0.86	0.77	1.00	0.87	0.47	0.54
Cluster 4	0.75	0.82	0.87	1.00	0.54	0.65
Cluster 5	0.29	0.28	0.47	0.54	1.00	0.45
Cluster 6	0.38	0.47	0.54	0.65	0.45	1.00
	Panel C. Neighborhood-Level Outside Options					
Standard Deviation	0.34	0.37	0.49	0.51	1.38	1.86
Race	B	B	H	H	W	W
Income	Q1	Q2	Q1	Q2	Q1	Q2

Note: Panel A shows estimated distance and language match parameters. Bootstrapped standard errors are shown in parentheses. Panel B shows the standard deviation of the point estimates of school mean utilities and the correlation matrix of δ_{cj}/β_c across clusters. Panel C shows the standard deviation of the point estimates of neighborhood-level outside option values.

standard deviation is associated with increased school mean utilities, valued as a reduction in travel distance of 0.28 SD. For this same group, increasing the share of white students is associated with a reduction of 0.37 SD in travel distance. The equivalent quantities for higher-income white families are 0.54 SD and 0.6 SD, showing a stronger association between school mean utilities, and test scores and the share of white students. Test-score growth and race-balanced growth show a positive but weaker association with school mean utilities than test-score levels. The share of Black students and the share of low-income kindergarten students are negatively related to school mean utilities for all clusters.

The effect of the racial composition tends to be stronger than that of achievement levels for most clusters, suggesting demographics are a big driver of parental preferences and a source of preference heterogeneity. Table A.4 shows results from a regression that includes both test scores and demographic controls. While the percentage of white students remains largely associated with school mean utilities, test-score levels become less important and only significant for higher-income Black families and white families. Non-monotonicity in parental preferences for demographics, captured by quadratic terms in the percentage of Black and white students, highlights preference heterogeneity across clusters.

5 Counterfactual Assignments

In this section, I describe under what assumptions the counterfactual assignments are generated, and then I discuss the results. I simulate counterfactual assignments for all students in each year and aggregate results across years. The counterfactuals are then used to estimate the contribution of the mechanisms described in Section 3.

5.1 Change the Location of a Student

To estimate how much of the cross-race gap in school ratings can be attributed to differences in distance-related costs to high-rated schools, I first evaluate how the assignments of Black and Hispanic students change if their residential locations were randomly drawn from the set of white students' locations. For the main results, I restrict counterfactual locations to be in the same income group as the original location of each student, hence isolating the effect of race. After drawing a new residential location for a single Black or Hispanic student, I use the demand model to generate the ranking that the student would have submitted at the new location. I further assume that the length of counterfactual rankings is determined by the position of the outside option—in other words, parents rank every acceptable school in the new location. The change in distance to all schools shifts travel costs. Also, choice-menu restrictions may change parents' available choices, and changes in proximity priorities change assignment probabilities.

While a change in the residential location of families is not modeled after a known implementable policy, the counterfactual is interpreted as a means to quantify the impact of the residential location of families over school assignments. I argue quantifying this impact is relevant for policy even when the counterfactual proposed is not intended as an implementable policy.

I consider the relocation of one student at a time. Changing the residential location of a single student guarantees that schools are unchanged across counterfactuals, then preference parameters are the same. If the locations of all students changed simultaneously, we would expect, among other things, the demographic composition of schools to change. This means that the results estimate the average impact of relocating a single student as opposed to the equilibrium effect of relocating a large mass of students.

Considering the relocation of a larger mass of students is interesting but may not be the best-suited counterfactual for answering the proposed question. If the mass of relocated students is large enough to cause changes in parental preferences for schools, potentially cascading into residential choices, there is no reason to believe the resulting equilibrium would be substantially different from what is observed; namely, segregated neighborhoods and unequal access to high-quality schools across groups. A partial equilibrium exercise, on the contrary, allows to quantify the impact of location on access, holding all else equal. While this counterfactual is not tied to an implementable policy it serves as a way to identify the impact of location on access.

About a third of Black and Hispanic students had a sibling priority at the time they applied to BPS. I make three distinct assumptions for treating sibling priorities under counterfactual locations. Results are not sensitive to these assumptions, so I conclude that while no single assumption may be realistic, together they are informative about the impact of location changes. The main results assume that in the new location, students lose any sibling priority they previously held. This is the case for a family that relocates and searches for a new school for both siblings. Two other assumptions are made, and the results are presented in the appendix. First, I assume that families keep the sibling priority they had, meaning that the older sibling holds their seat and the younger searches for one in the new location. Second, I assume every student with a sibling has a sibling priority at their first-ranked school in the new location. This is the extreme case in which the older sibling is assigned to the first-ranked school in the new location had the family lived there when the older sibling applied to BPS, and the younger sibling is in a high-priority group at that school. The main results that assume students lose their sibling priority may underestimate the impact of location as this assumption reduced their chances of getting assigned to their listed schools relative to white peers in the new location. This happens because more than 40% of white applicants have a sibling priority at one of their ranked schools, putting them in a high-priority group in at least one school.

Finally, to study the sensibility of the results to changes in the distance parameter due to potential bias, I generate rankings in the new location assuming the estimated distance parameter is over-estimated by 15% and 40% (distance factor of 0.85 and 0.6), and compare this to the case where the estimated distance parameter is unbiased (distance factor of 1).

Results from each counterfactual assignment are compared to simulated benchmark assignments. The benchmark is generated using model parameters and the same realization of the random parameters that are used for counterfactuals and considering all families' original residential locations. Benchmark assignments use a distance factor of 1.

I draw a single counterfactual location for each Black and Hispanic student in the sample, and for each location, I simulate counterfactual assignments for each of three random draws of ε_{ij} from a type-1 EV distribution. Taste shock values are not further restricted using observed rankings given that for most counterfactual locations choice menus differ from those at original locations.

5.2 Eliminate Choice-Menu Restrictions and Walk-Zone Priorities

When a student changes locations, not only do her travel costs change but her choice menu and the set of schools in which she has a walk-zone priority change. This means the effects of location bundle the effects of location-specific assignment rules and changes to the menu of distances. To disentangle the effect of rules from that of travel costs, I run two additional counterfactuals. First, I eliminate choice-menu restrictions and allow parents to rank schools from across the city. Under this counterfactual, parents of a Black or Hispanic student can rank the same schools they would have ranked under the location counterfactual. The only reason why these rankings would not coincide is the differences in travel costs to these schools from both locations. In the second counterfactual, I run the assignment algorithm assuming no one has a walk-zone priority. Eliminating priorities does not change rankings, but it does change assignments via priorities. This counterfactual captures the effect of an assignment-probability change that is explained by the walk-zone priority. Notice that since these counterfactuals are about algorithm rules, I do not isolate the effect of the counterfactual on a single student; instead, I change assignment rules for all students.

A crucial assumption that is made when running walk zone priority counterfactuals is that families are not using their perceived assignment probabilities as input for making ranking decisions. In other words, families do not change their submitted rankings as a function of their priority at a school. This assumption is backed by the evidence presented in Figure 2 that shows families' ranking decisions, in particular their decision to rank a school in the first position, does not change at the walkzone discontinuity threshold.

5.3 Results

Column (1) of Table 5 shows cross-racial gaps in ratings under the simulated benchmark assignment predict well those generated with observed rankings (Table 3). Columns (2) - (4) in Panel A, show location change counterfactuals reduce the gap in test-score levels for Black students from 0.8 SD to 0.45 SD assuming no bias in the distance parameters, and to 0.54 SD when the parameters are over-estimated by 40%. This implies reductions in the gap of 44% and 33%. For Hispanic students, gaps in levels go from 0.52 SD to between 0.41 SD and 0.47 SD, representing reductions between 22% and 11% of the gap. Following the elimination of walkzone priorities and choice menu restrictions, the gaps marginally changed for both Black students and Hispanic students (columns (5) and (6)).[21] Taken together, salient gap reductions when location changes and a limited impact when rules change, suggest that the effect of location is mainly explained by changes in the cost of traveling to high test-score schools and not by location-specific assignment rules.

When looking into test-score growth (Panel B), the gap reverts or vanishes after a change in location for Black students. Assuming a distance factor of 1 or 0.85, the gap goes from 0.18 SD to -0.04 SD —a gap in favor of Black students; and to an indistinguishable-from-zero gap when the distance factor is 0.6. This implies Black students would access schools with between 0.18 SD and 0.22 SD higher average test-score growth following a location change. Eliminating walkzone priorities or choice menu restrictions does not impact the gap which implies the impact of location is mainly that of larger distance-related costs. For Hispanic students, the benchmark assignment shows a small gap in their favor that disappears with a location change but does not change with changes in walkzone priorities.

[21] Dropping walkzone priorities reduces by 0.03 SD the gap for Black students with no significant impact from eliminating choice menu restrictions. For Hispanic students, eliminating choice menu restrictions implies a marginal increase in the gap, with no changes after eliminating walkzone priorities. Across panels A through C dropping choice menu restrictions marginally reduces access to high-rated schools for Hispanic students, implying these boundaries reduce competition for school seats for this group.

Table 5: Summary of Counterfactuals: Main Specification

	Benchmark (1)	Location (df=1) (2)	Location (df=0.85) (3)	Location (df=0.60) (4)	Walkzone Priority (5)	Choice Menu (6)
			Panel A. Gap in Test Score Levels (SD)			
Black Students	0.799	0.449	0.477	0.537	0.768	0.788
	(0.013)	(0.018)	(0.018)	(0.018)	(0.013)	(0.013)
Hispanic Students	0.524	0.409	0.431	0.468	0.532	0.567
	(0.011)	(0.012)	(0.012)	(0.012)	(0.010)	(0.010)
			Panel B. Gap in Test Score Growth (SD)			
Black Students	0.178	-0.044	-0.037	-0.007	0.179	0.168
	(0.011)	(0.018)	(0.017)	(0.019)	(0.011)	(0.012)
Hispanic Students	-0.027	-0.004	0.002	0.007	-0.022	-0.011
	(0.010)	(0.011)	(0.010)	(0.011)	(0.009)	(0.009)
			Panel C. Gap in Race-Balanced Growth (SD)			
Black Students	0.020	-0.152	-0.150	-0.127	0.030	0.014
	(0.011)	(0.018)	(0.017)	(0.020)	(0.010)	(0.011)
Hispanic Students	-0.159	-0.106	-0.104	-0.104	-0.154	-0.147
	(0.010)	(0.011)	(0.010)	(0.011)	(0.009)	(0.009)
			Panel D. Gap in Share White (Share)			
Black Students	0.197	0.128	0.134	0.142	0.188	0.195
	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Hispanic Students	0.165	0.122	0.126	0.132	0.165	0.171
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)

Note: The table shows gaps in benchmark and counterfactual assignments for three measures of school ratings and the share of white students. Counterfactual locations are restricted to locations with the same income category as that of the original location of each family. Model parameters assume utilities are a linear function of distance and include neighborhood-level outside option values. Benchmark assignments, Choice Menu, and Walkzone Priority counterfactuals are generated using a distance factor of 1, and location counterfactuals are shown using distance factors of 1, 0.85, and 0.6. Bootstrapped standard errors are obtained by running counterfactual assignments with each of 100 vectors of preference parameters.

Panel C shows changes in the gap in race-balanced growth. Under the Benchmark, a small gap for Black students reverts with location changes, implying a location change would improve access for Black students to high race-balance growth schools beyond that of white students. Black students would access schools with between 0.15 SD and 0.17 SD higher race-balanced growth relative to their current locations. For Hispanic students, on the other hand, we find a Benchmark gap in their favor that shrinks with location changes implying location changes are associated with reduced access to high race-balanced growth schools by an average of 0.05 SD. For Black students, location-specific rules don't have an impact on the gap.

Panel D shows that higher access to race-balanced growth schools for Black students in counterfactual locations comes with a larger exposure to white peers, not less. A Black student would be assigned to a school with about 6% to 7% more students who identify as white on average. For Hispanic students, on the contrary, location changes reduce access to high race-balance growth schools while increasing the exposure to white peers by about 3% to 4%.

In summary, as shown in Figures 1c and 1d, Black families experience longer distances to high-rated schools than their white and Hispanic peers and must trade off distance and ratings. Those costs influence their school assignments resulting in cross-race gaps in access. Had a Black student faced the distance menu of white students in their same income group, they would experience increased access to schools with larger test-score levels (by 0.26 to 0.35 SD), higher test-score growth (by 0.02 to 0.03 SD), and higher race-balanced growth (by 0.05 to 0.06 SD). Increased access would reduce the gap between white and Black students, and in some cases revert it. Importantly, results are not meaningfully sensitive to the choice of distance factor between 1 and 0.6. The former corresponds to the case when the estimated distance parameters are unbiased, and the latter when the distance parameters are overestimated by 40%. A larger bias in the distance parameter implies the impact of location is overestimated. Estimates suggest bias must be significantly larger than 40% for results to vanish.

The salience of travel costs for Black families shows why neighborhoods matter, highlighting how the effective provision of public goods can be affected by geography at very granular levels. The results show that even in a generous choice environment in which parents face minimal restrictions on their choices and free transportation is provided, distance can contribute greatly to inequity and the design of the assignment algorithm can do little to break structural place-based inequities. This finding is not only relevant for the prekindergarten population. Not only can early investments have lasting impacts on adult outcomes, but choice systems are typically designed to grandfather students into subsequent grades within a school. So even if travel costs are lower for older children, early assignments are held for several years. In consequence, inequities in prekindergarten extend well after that period.

One potential cost-effective policy that can reduce the costs of travel to higher-rated schools concerns transportation. During my study period, BPS offered very generous transportation: it guaranteed school busing to all families assigned to a school farther than a mile from their homes and capped every student's travel time at an hour. Although school buses are effective in reducing travel costs to families (Trajkovski et al. 2021), the results in this paper show there are limits to their ability to equalize access to educational resources. Moreover, school transportation can represent high costs to school districts bearing to question their cost-effectiveness. At the time of this study, BPS spent about 10% of its budget on transportation, which constituted the highest per-student transportation cost in the US (Bertsimas et al. 2020).

Robustness Results. If counterfactual locations are not restricted to be in the same income group as benchmark locations, location counterfactuals would be associated with increased access to high test-score schools for Black and Hispanic students, but with lower access to high growth and high race-balanced growth schools (Table B.5). This highlights a correlation between income and school achievement that is not entirely mediated by school effectiveness.

Various robustness checks are presented in Appendix Section B. Table B.6 shows results are robust to alternative assumptions about sibling priorities for students who are relocated. This implies that sibling priorities, although important for assignment probabilities, are second-order when estimating the impact of location. Tables B.7 and B.8 show results if we assumed that the parents' utilities are quadratic rather than linear in distance, or if the value of the outside option was common across neighborhoods for each covariate cluster. Results are essentially unchanged.

6 Conclusion

Among other objectives, choice-based systems aim to enhance equity and diversity in public schools by decoupling families' residences and schools. Under these systems, residential choices, often influenced by family income and subject to documented discrimination, do not dictate a family's school assignment, but instead, families can access a variety of schooling options across the city, even if far from home. But how much can choice systems accomplish if longer distances between a school and a family's residence can impose high costs on families?

I find that differential travel costs reduce access to high-rated schools for Black families relative to white families. A Black family that faced the same distance menu as white families in their income group would experience increased access to schools with higher test-score levels, higher test-score growth, and higher race-balanced growth. Increased access would reduce the white-Black gap, and in some cases revert it. Hispanic students, on the other hand, experience reduced access to high-test score schools relative to white students under benchmark residential locations, but sort effectively toward higher growth and race-balanced growth schools under choice. This is consistent with reduced form evidence that shows a positive relation between distance to schools and ratings for Black families, but not for Hispanic families in Boston.

While I have shown that geography can explain gaps in access to high-rated schools for prekindergarteners, it remains important to quantify how geography influences access later in life. Families may be more prone to sort toward high-amenity schools when seeking admission in higher grades, and at the same time, travel costs may change as children grow. Also, quantifying the impact of pre-kindergarten assignments on student later-life outcomes remains an important question for future work.

References

Abdulkadiroğlu, A., Pathak, P. A., and Roth, A. E. (2005). The new york city high school match. *American Economic Review*, 95(2):364–367.

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

Abdulkadiroğlu, A. and Sönmez, T. (2003). School choice: A mechanism design approach. *American economic review*, 93(3):729–747.

Agarwal, N. and Somaini, P. J. (2019). Revealed preference analysis of school choice models. Technical report, National Bureau of Economic Research.

Agostinelli, F., Luflade, M., Martellini, P., et al. (2021). On the spatial determinants of educational access. Technical report.

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

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

Angrist, J., Gray-Lobe, G., Idoux, C., and Pathak, P. A. (2021). Is busing worth the trip? school travel effects in boston and new york. *Unpublished manuscript*.

Angrist, J., Hull, P., Pathak, P. A., and Walters, C. R. (2022). Race and the mismeasure of school quality. Technical report, National Bureau of Economic Research.

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

Arteaga, F., Kapor, A. J., Neilson, C. A., and Zimmerman, S. D. (2022). Smart matching platforms and heterogeneous beliefs in centralized school choice. *The Quarterly Journal of Economics*, 137(3):1791–1848.

Balinski, M. and Sönmez, T. (1999). A tale of two mechanisms: student placement. *Journal of Economic theory*, 84(1):73–94.

Barseghyan, L., Clark, D., and Coate, S. (2014). Public school choice: an economic analysis. Technical report, National Bureau of Economic Research.

Berry, S., Gandhi, A., and Haile, P. (2013). Connected substitutes and invertibility of demand. *Econometrica*, 81(5):2087–2111.

Bertsimas, D., Delarue, A., Eger, W., Hanlon, J., and Martin, S. (2020). Bus routing optimization helps boston public schools design better policies. *INFORMS Journal on Applied Analytics*, 50(1):37–49.

Biasi, B. (2019). School finance equalization increases intergenerational mobility: Evidence from a simulated-instruments approach. Technical report, National Bureau of Economic Research.

Borghans, L., Golsteyn, B. H., and Zolitz, U. (2015). Parental preferences for primary school characteristics. *The BE Journal of Economic Analysis & Policy*, 15(1):85–117.

Boston Desegregation Project (1988). (130). *Northeastern University Library, Archives and Special Collections*, 9(33).

Calsamiglia, C. and Güell, M. (2018). Priorities in school choice: The case of the boston mechanism in barcelona. *Journal of Public Economics*, 163:20–36.

Calsamiglia, C., Haeringer, G., and Klijn, F. (2010). Constrained school choice: An experimental study. *American Economic Review*, 100(4):1860–74.

Calsamiglia, C., Martínez-Mora, F., and Miralles, A. (2021). School choice design, risk aversion and cardinal segregation. *The Economic Journal*, 131(635):1081–1104.

Campos, C. and Kearns, C. (2021). The impacts of neighborhood school choice: Evidence from los angeles' zones of choice. *Available at SSRN 3830628*.

Chetty, R., Friedman, J. N., Hendren, N., Jones, M. R., and Porter, S. R. (2018). The opportunity atlas: Mapping the childhood roots of social mobility. Technical report, National Bureau of Economic Research.

Chetty, R. and Hendren, N. (2018). The impacts of neighborhoods on intergenerational mobility i: Childhood exposure effects. *The Quarterly Journal of Economics*, 133(3):1107–1162.

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

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics*, 129(4):1553–1623.

Chubb, J. E. and Moe, T. M. (1990). America's public schools: Choice is a panacea. *The Brookings Review*, 8(3):4–12.

Dubins, L. E. and Freedman, D. A. (1981). Machiavelli and the gale-shapley algorithm. *The American Mathematical Monthly*, 88(7):485–494.

Dur, U., Kominers, S. D., Pathak, P. A., and Sonmez, T. (2018). Reserve design: Unintended consequences and the demise of boston's walk zones. *Journal of Political Economy*, 126(6):2457–2479.

Epple, D., Figlio, D., and Romano, R. (2004). Competition between private and public schools: testing stratification and pricing predictions. *Journal of public Economics*, 88(7-8):1215–1245.

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

Friedman, M. (1982). Capitalism and freedom. 1962. *Whether government control over the provision of education through public schools is also efficient is a quite different matter.*

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

Haeringer, G. and Klijn, F. (2009). Constrained school choice. *Journal of Economic theory*, 144(5):1921–1947.

Hastings, J., Hortaçsu, A., and Syverson, C. (2017). Sales force and competition in financial product markets: the case of mexico's social security privatization. *Econometrica*, 85(6):1723–1761.

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

Hoxby, C. M. (2003). School choice and school productivity. could school choice be a tide that lifts all boats? In *The economics of school choice*, pages 287–342. University of Chicago Press.

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

Idoux, C. (2022). *Integrating new york city schools: The role of admission criteria and family preferences.* MIT Department of Economics.

Kapor, A. J., Neilson, C. A., and Zimmerman, S. D. (2020). Heterogeneous beliefs and school choice mechanisms. *American Economic Review*, 110(5):1274–1315.

Laliberte, J.-W. P. (2018). Long-term contextual effects in education: Schools and neighborhoods. *V manuscript.*

Langer, A. (2016). (dis)incentives for demographic price discrimination in the new vehicle market. *Working Paper*.

Lee, J. and Son, S. J. (2022). Distributional impacts of centralized school choice. Technical report, Working Paper.

Luflade, M. (2018). The value of information in centralized school choice systems. *Unpublished manuscript*.

Neilson, C. (2013). Targeted vouchers, competition among schools, and the academic achievement of poor students. *Documento de trabajo). Yale University. Recuperado de http://economics. sas. upenn. edu/system/files/event_papers/Neilson_2013_JMP_current. pdf*.

Oosterbeek, H., Sóvágó, S., and van der Klaauw, B. (2021). Preference heterogeneity and school segregation. *Journal of Public Economics*, 197:104400.

Park, M. and Hahm, D. W. (2023). Location choice, commuting, and school choice. *Commuting, and School Choice (January 18, 2023)*.

Pathak, P. A. and Shi, P. (2013a). Simulating alternative school choice options in boston-main report.

Pathak, P. A. and Shi, P. (2013b). Simulating alternative school choice options in boston-technical appendix.

Pathak, P. A. and Shi, P. (2017). How well do structural demand models work? counterfactual predictions in school choice. Technical report, National Bureau of Economic Research.

Roth, A. E. (1982). The economics of matching: Stability and incentives. *Mathematics of operations research*, 7(4):617–628.

Sartain, L. and Barrow, L. (2022). The pathway to enrolling in a high-performance high school: Understanding barriers to access. *Education Finance and Policy*, 17(3):379–407.

Sönmez, T., Yenmez, M. B., et al. (2019). Affirmative action with overlapping reserves. Technical report, Boston College Department of Economics.

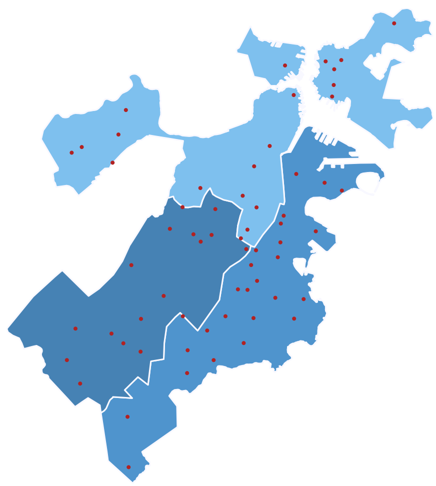
Trajkovski, S., Zabel, J., and Schwartz, A. E. (2021). Do school buses make school choice work? *Regional Science and Urban Economics*, 86:103607.

Appendix

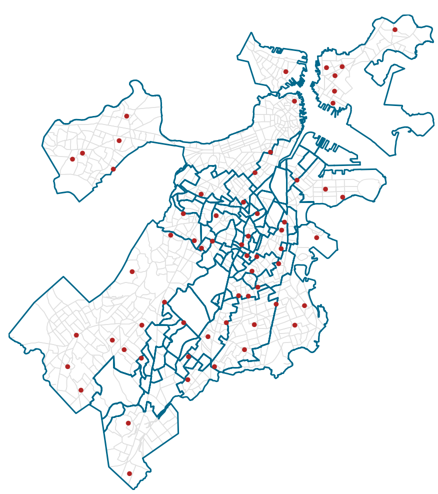
A Supplementary Tables and Figures

Figure A.1: Zoning and Choice Menus

(a) Zones

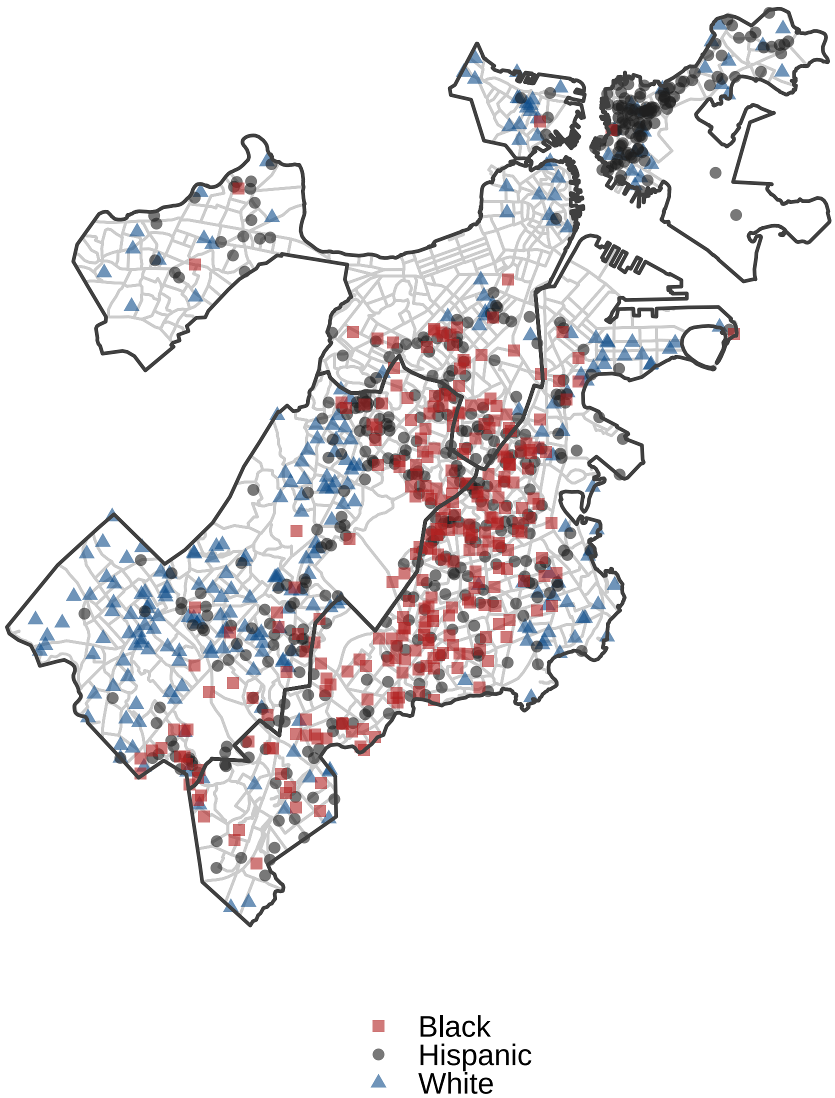


(b) Students with the same Choice Menu



Note: Red points are schools with a prekindergarten program in 2010. Choice menus are built using data on school and geocode coordinates.

Figure A.2: Spatial Distribution of Applicants by Race



Note: Each point represents 10 students from the 2010-2012 pooled data randomly located at the census tract level.

Table A.1: Number of applicants per school and potential applicants near each school

	Mean	St. Dev	Min	Max
Potential applicants within 3.8 miles of each school				
Cluster 1 - Black Q1	442.4	294.9	20	1,274
Cluster 2 - Black Q2	175.2	108.9	6	496
Cluster 3 - Hispanic Q1	618.9	297	116	1,515
Cluster 4 - Hispanic Q2	390.6	207	31	1,052
Cluster 5 - White Q1	60.5	32.4	10	144
Cluster 6 - White Q2	463.9	307.9	43	1,224
Applicants per school				
Cluster 1 - Black Q1	113.0	77.2	17	373
Cluster 2 - Black Q2	50.3	38.9	7	164
Cluster 3 - Hispanic Q1	171.0	128.7	13	686
Cluster 4 - Hispanic Q2	107.7	83.0	8	418
Cluster 5 - White Q1	17.5	16.8	0	71
Cluster 6 - White Q2	124.0	167.9	0	686

Note: The table shows statistics on the number of potential applicants and actual applicants per school, for students in each cluster.

Table A.2: Characteristics of schools with missing school achievement

	Not Missing Achievement	Missing Achievement
% Black	32.3	28.8
	(19.4)	(18.3)
% Hispanic	43.7	48.8
	(19.5)	(16.9)
% White	14.6	14.5
	(15.0)	(12.2)
% Low Income in K	67.4	68.7
	(19.9)	(19.4)
Observations	235	23

Note: Statistics of school year observations with and without observed school achievement.

Table A.3: School Mean Utilities and School Characteristics - Independent Regressions

	Standardized $\delta_{cj}/\lvert\beta_c\rvert$					
	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Test-Score Levels	0.28	0.43	0.34	0.46	0.34	0.54
	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)
	[236]	[236]	[236]	[236]	[236]	[236]
Test-Score Growth	0.06	0.11	0.11	0.10	0.15	0.15
	(0.06)	(0.06)	(0.06)	(0.07)	(0.06)	(0.06)
	[230]	[230]	[230]	[230]	[230]	[230]
Race-Balanced Growth	0.01	0.04	0.06	0.01	0.09	0.07
	(0.06)	(0.06)	(0.07)	(0.07)	(0.06)	(0.07)
	[230]	[230]	[230]	[230]	[230]	[230]
% Black Students	-0.25	-0.27	-0.65	-0.65	-0.48	-0.47
	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)
	[258]	[258]	[258]	[258]	[258]	[258]
% Hispanic Students	-0.07	-0.19	0.32	0.08	0.07	-0.09
	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.06)
	[258]	[258]	[258]	[258]	[258]	[258]
% White Students	0.37	0.50	0.38	0.63	0.44	0.60
	(0.06)	(0.05)	(0.06)	(0.05)	(0.05)	(0.05)
	[258]	[258]	[258]	[258]	[258]	[258]
% Low-Income Students in Kindergarten	-0.19	-0.35	-0.21	-0.39	-0.23	-0.34
	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)
	[256]	[256]	[256]	[256]	[256]	[256]
Race	B	B	H	H	W	W
Income	Q1	Q2	Q1	Q2	Q1	Q2

Note: Each coefficient is from an independent regression where the dependent variable is the standardized ratio $\delta_{cj}/|\beta_c|$, and the δ_{cj} are estimated assuming a common value of the outside option across neighborhoods. All the independent variables are standardized in each year. Standard errors are shown in parentheses and sample sizes are shown in square brackets.

Table A.4: School Mean Utilities and School Characteristics - Pooled Regressions

	Standardized $\delta_{cj}/\lvert\beta_c\rvert$					
	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Test-Score Levels	0.07	0.22	0.04	0.02	0.11	0.22
	(0.08)	(0.07)	(0.06)	(0.06)	(0.06)	(0.06)
% Black Students	-0.04	0.07	-0.58	-0.41	-0.25	-0.16
	(0.07)	(0.06)	(0.06)	(0.05)	(0.05)	(0.06)
% White Students	0.29	0.34	0.12	0.43	0.23	0.52
	(0.10)	(0.09)	(0.08)	(0.07)	(0.08)	(0.08)
% Black Students Squared	-0.02	-0.11	0.04	-0.01	-0.14	0.03
	(0.06)	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)
% White Students Squared	0.00	-0.02	-0.08	-0.08	-0.08	-0.16
	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)
% Low-Income Students in Kindergarten	-0.03	-0.13	-0.06	-0.14	-0.04	-0.08
	(0.07)	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)
Observations	234	234	234	234	234	234
Race	B	B	H	H	W	W
Income	Q1	Q2	Q1	Q2	Q1	Q2

Note: Coefficients from regressions of THE standardized ratio $\delta_{cj}/|\beta_c|$ and school characteristics, where the δ_{cj} are estimated assuming a common value of the outside option across neighborhoods. All the independent variables are standardized in each year.

B Counterfactuals

Table B.5: Location Changes when Counterfactual Locations are not Restricted by Income

	Main Specification + Counterfactual Location not Restricted by Income		
	Location (df=1)	Location (df=0.85)	Location (df=0.60)
	(1)	(2)	(3)
	Panel A. Gap in Test Score Levels (SD)		
Black Students	0.413	0.447	0.520
	(0.019)	(0.019)	(0.018)
Hispanic Students	0.338	0.362	0.416
	(0.015)	(0.015)	(0.014)
	Panel B. Gap in Test Score Growth (SD)		
Black Students	0.022	0.043	0.063
	(0.014)	(0.014)	(0.014)
Hispanic Students	-0.001	0.006	0.021
	(0.011)	(0.010)	(0.011)
	Panel C. Gap in Race-Balanced Growth (SD)		
Black Students	-0.075	-0.060	-0.051
	(0.014)	(0.014)	(0.014)
Hispanic Students	-0.088	-0.087	-0.08
	(0.012)	(0.011)	(0.011)
	Panel D. Gap in Share White (Share)		
Black Students	0.121	0.128	0.138
	(0.003)	(0.003)	(0.003)
Hispanic Students	0.108	0.115	0.125
	(0.003)	(0.003)	(0.003)

Note: The table shows gaps in benchmark and counterfactual assignments for three measures of school ratings and the share of white students. Counterfactual locations are not restricted to locations with the same income category as that of the original location of each family. Model parameters assume utilities are a linear function of distance and include neighborhood-level outside option values. Benchmark assignments are generated using a distance factor of 1, and location counterfactuals are shown using distance factors of 1, 0.85, and 0.6. Bootstrapped standard errors are obtained by running counterfactual assignments with each of 100 vectors of preference parameters.

Table B.6: Alternative Sibling Assumptions Under the Main Model Specification

	Main Specification + Alternative Sibling Assumptions		
	Location (df=1)	Location (df=1)	Location (df=1)
	Lose Sibling Priority	Keep Benchmark Sibling Priority	Sibling Priority at First-Ranked School
	(1)	(2)	(3)
	Panel A. Gap in Test Score Levels (SD)		
Black Students	0.449	0.446	0.428
	(0.018)	(0.018)	(0.018)
Hispanic Students	0.409	0.404	0.388
	(0.011)	(0.011)	(0.011)
	Panel B. Gap in Test Score Growth (SD)		
Black Students	-0.044	-0.044	-0.046
	(0.017)	(0.017)	(0.017)
Hispanic Students	-0.004	-0.006	-0.006
	(0.011)	(0.011)	(0.011)
	Panel C. Gap in Race-Balanced Growth (SD)		
Black Students	-0.152	-0.151	-0.151
	(0.018)	(0.018)	(0.018)
Hispanic Students	-0.106	-0.108	-0.106
	(0.011)	(0.011)	(0.011)
	Panel D. Gap in Share White (Share)		
Black Students	0.128	0.127	0.125
	(0.003)	(0.003)	(0.003)
Hispanic Students	0.122	0.121	0.118
	(0.002)	(0.002)	(0.002)

Note: The table shows gaps for three measures of school ratings and the share of white students after location change counterfactuals under three assumptions about the sibling priority status of families who are relocated. Counterfactual locations are restricted to locations with the same income category as that of the original location of each family. Model parameters assume utilities are a linear function of distance and include neighborhood-level outside option values. Bootstrapped standard errors are obtained by running counterfactual assignments with each of 100 vectors of preference parameters.

Table B.7: Results with Neighborhood-Level Outside Option and Quadratic Distance

	Quadratic in distance + Neighborhood-level outside option		
	Location (df=1)	Walkzone Priority	Choice Menu
	(1)	(2)	(3)
	Panel A. Gap in Test Score Levels (SD)		
Black Students	0.448	0.764	0.778
	(0.017)	(0.018)	(0.021)
Hispanic Students	0.398	0.527	0.560
	(0.01)	(0.02)	(0.017)
	Panel B. Gap in Test Score Growth (SD)		
Black Students	-0.034	0.180	0.170
	(0.019)	(0.014)	(0.020)
Hispanic Students	-0.001	-0.021	-0.010
	(0.01)	(0.08)	(0.053)
	Panel C. Gap in Race-Balanced Growth (SD)		
Black Students	-0.139	0.031	0.017
	(0.020)	(0.013)	(0.019)
Hispanic Students	-0.102	-0.153	-0.146
	(0.011)	(0.078)	(0.055)
	Panel D. Gap in Share White (Share)		
Black Students	0.125	0.189	0.193
	(0.003)	(0.004)	(0.003)
Hispanic Students	0.120	0.165	0.171
	(0.002)	(0.004)	(0.004)

Note: The table shows gaps for three measures of school ratings and the share of white students. Counterfactual locations are restricted to locations with the same income category as that of the original location of each family. Model parameters assume utilities are a quadratic function of distance and include neighborhood-level outside option values. Bootstrapped standard errors are obtained by running counterfactual assignments with each of 100 vectors of preference parameters.

Table B.8: Results with Common Outside Option and Linear Distance

	Linear in distance + Common outside option		
	Location (df=1)	Walkzone Priority	Choice Menu
	(1)	(2)	(3)
	Panel A. Gap in Test Score Levels (SD)		
Black Students	0.460	0.770	0.783
	(0.018)	(0.012)	(0.012)
Hispanic Students	0.408	0.531	0.567
	(0.011)	(0.011)	(0.012)
	Panel B. Gap in Test Score Growth (SD)		
Black Students	-0.033	0.175	0.157
	(0.018)	(0.011)	(0.012)
Hispanic Students	-0.011	-0.029	-0.014
	(0.01)	(0.009)	(0.01)
	Panel C. Gap in Race-Balanced Growth (SD)		
Black Students	-0.144	0.024	0.001
	(0.018)	(0.011)	(0.012)
Hispanic Students	-0.116	-0.163	-0.152
	(0.01)	(0.009)	(0.01)
	Panel D. Gap in Share White (Share)		
Black Students	0.133	0.191	0.197
	(0.003)	(0.002)	(0.002)
Hispanic Students	0.125	0.168	0.173
	(0.002)	(0.002)	(0.002)

The table shows gaps for three measures of school ratings and the share of white students. Counterfactual locations are restricted to locations with the same income category as that of the original location of each family. Model parameters assume utilities are a linear function of distance and include a common outside option value for each cluster across neighborhoods. Bootstrapped standard errors are obtained by running counterfactual assignments with each of 100 vectors of preference parameters.

C Assignment Algorithm

With the exception of a couple of schools, half of the seats at each school are assigned using the priority order explained in the main text. This includes sibling and walk-zone priorities. For the second half of seats, the priority does not include any walk-zone considerations. In consequence, students with a sibling have the first priority and the rest have the second priority. Ties between groups are broken using a unique random number drawn for each student.

Now, since a student may be eligible for seats in both halves at each school, a precedence order across halves is established. This is, the rule that determines whether a student is first considered for the first or second half of the seats at a school. A student with a walk-zone priority will be considered for the walk-half first while a student outside the walk zone is considered for the second half first. The DA algorithm, described below, is ran over school halves.

- *Step 1:* Applicants are sorted in priority order in their first-ranked schools and students over capacity are rejected. Those who are not rejected are provisionally admitted.
- *Step k:* For students rejected in step $k-1$, their next preferred option is considered. Each school ranks by priority order the set of provisionally admitted students jointly with those new students who are being considered in k. The program provisionally admits those with the highest priority and rejects students over capacity. The algorithm stops when every rank list has been exhausted or when there are no rejections.

More details about the assignment algorithm can be found in Pathak and Shi 2013a and Pathak and Shi 2013b.