A Dynamic Framework of School Choice: Effects of Middle Schools on High School Choice*

Dong Woo Hahm† Minseon Park‡

July 19, 2022

Click here for the latest version

Abstract

We explore how students' previously attended schools influence their subsequent school choices and how this relationship affects school segregation. Using administrative data from New York City, we document the causal effects of the middle school a student attends on her high school application/assignment. Motivated by this finding, we estimate a dynamic model of middle and high school choices. We find that the middle schools' effects mainly operate by changing how students rank high schools rather than how high schools rank their applications. Counterfactual analysis shows that policymakers can design more effective policies by exploiting the dynamic relationship of school choices.

JEL Classification Numbers: D12, D47, D63, H75, I21, I24, I28.
Keywords: School choice, dynamic models, deferred acceptance, segregation by race.

*We are grateful to Sandra E. Black, Michael Best, Yeon-Koo Che, Pierre-André Chiappori, Chao Fu, Jesse Gregory, Gautam Gowrisankaran, YingHua He, Seungki Hong, Jean-François Houde, John Kennan, Simon Lee, W. Bentley MacLeod, Bernard Salanié, Jeffrey Smith, Suk Joon Son, Christopher Taber, Olivier Tercieux, Miguel Urquiola, Matthew Wiswall for valuable comments and suggestions. We would also like to thank the seminar participants at Columbia University and the University of Wisconsin-Madison. Thanks also go to the New York City Department of Education, particularly to Benjamin Cosman, Stewart Wade, and Lianna Wright. We acknowledge computing resources from Columbia University's Shared Research Computing Facility project, which is supported by NIH Research Facility Improvement Grant 1G20RR030893-01, and associated funds from the New York State Empire State Development, Division of Science Technology and Innovation (NYSTAR) Contract C090171, both awarded April 15, 2010. Hahm acknowledges funding from the Program for Economic Research (PER) at Columbia University. All errors are our own.

†Department of Economics, University of Southern California, USA. Email: dongwooh@usc.edu.

‡Department of Economics, University of Wisconsin-Madison, USA. Email: mpark88@wisc.edu.

1 Introduction

Worldwide, numerous jurisdictions employ a centralized, market design-based student assignment system called school choice, which assigns students to schools based on families' preferences. Notably, many such jurisdictions offer school choice *at multiple levels of education*. For example, New York City (NYC) uses the student-proposing Deferred Acceptance (DA) mechanism (Gale and Shapley, 1962) to assign students to public schools at all levels of education, starting as early as 3-K to high school, and Boston Public Schools also use DA to assign students starting from kindergarten to high school.

Naturally, the students' choices at different levels are closely related. The school a student attends determines her learning environment, the set of teachers, and her peers and friends, all of which influence the student's subsequent choices. For example, a middle school student with many high-performing classmates may aspire to attend a high-quality high school. Not surprisingly, anecdotal evidence often suggests that parents are well-aware that a carefully chosen middle school could lead their children to good high schools and consequently to good colleges.[1]

Nevertheless, the sequentiality of school choice has been neglected by both scholars and policymakers. Existing studies have not examined the school choices at multiple levels together but focused on one level at a time. As a result, we do not have an appropriate framework or empirical evidence of such sequentiality of school choice. Also, policymakers neglected the sequentiality of school choice in designing a policy to tackle a conspicuous problem, racial segregation across schools in large urban school districts. Admissions reforms to desegregate schools have targeted each level of education separately, without carefully examining how a reform at one level may influence the school choice outcomes at other levels. For example, suppose middle schools are already segregated. In that case, minority students' middle school experience may differ from that of White students, resulting in different high school application patterns that lead to racial segregation across high schools. Any high school-only admissions reform that neglects the sequentiality of school choice may fail to address this aspect of segregation.

This paper explores the dynamic relationship of school choices at different educational

[1] The administration often offers relevant information to assist families' choices in this aspect. For example in NYC, families can obtain information on which high schools are frequently attended by the graduates of each middle school (see https://www.myschools.nyc/en/).

levels and how it affects racial segregation across schools. To our knowledge, this is the first attempt to study the relationship between students' school choices at different education levels. To be specific, we ask two questions:

1. Does a student's previously attended school affect the subsequent school choices, and if so, how?
2. How can one address racial segregation across schools using this relationship?

To answer those questions, we develop a novel, evidence-based dynamic framework of school choice and provide new insight into understanding and addressing the school segregation patterns at different educational levels.

Our empirical context is NYC public middle and high school choices, which provide a suitable setting to study our questions. NYC is the largest public school district in the United States. It utilizes centralized school choice at multiple levels of education, which generates panel data on middle and high school applications and assignments of the same students. No less importantly, NYC has been at the center of attention regarding racial segregation across public schools.

We start by providing empirical evidence of the causal effects of middle schools on high school applications and assignments. To overcome students' selection into middle schools based on unobservables, we adopt the research design in Abdulkadiroğlu, Angrist, Narita, and Pathak (2021). The design utilizes quasi-random assignments to middle schools generated by the tie-breaking rules that distinguish among applicants with the same applications and priorities. Our two-stage least squares (2SLS) estimates reveal that all else equal, students who attend high-achievement middle schools are more likely to *apply to* high schools of higher quality measured by graduation rate, college enrollment rate, and the proportion of high-performers. Furthermore, we find such students are more likely to be *assigned to* high schools of even higher quality.

Motivated by the empirical evidence, we proceed to develop and estimate a novel dynamic framework of school choice to decompose this effect into different channels and analyze the equilibrium consequences of counterfactual policies. To do so, we adapt the dynamic discrete choice framework and combine it with large market matching theory.

Our two-period model has three key features. First, it allows how students rank high schools to depend on the middle schools they attend (the *application* channel). Students' utilities that underlie their high school applications may depend on their attended middle schools. Second, how high schools rank students for admissions also depends on the middle school students attend (the *priority* channel). Attending different middle schools may affect students' admission chances at each high school. Third, students are forward-looking. They apply to middle schools considering these potential effects on their eventual high school assignments.

For tractability, we estimate the model using data from Staten Island, which is geographically separated from other boroughs of NYC. Most students do not commute outside of Staten Island, and hence, it can be considered an independent school district.

Our main findings are as follows. First, we find that the middle school a student attends affects her tastes for high schools that underlie her high school application. For example, all else equal (including test scores that may change by attending different middle schools), students who attend high-achievement middle schools are willing to travel an additional 0.11 miles for a 10 pp increase in the proportion of high-performers in high schools. Also, attending a middle school with many students of the same race makes students value the proportion of the same race students in high schools even more. Second, our decomposition exercise shows that the effect of middle schools on high school assignments mainly operates through the *application* channel. For example, on average, attending the highest-performing middle school instead of the lowest-performing middle school leads to higher-achieving high schools in terms of an increase in the proportion of high-performers by 9.7 pp, and two-thirds of the increase is explained alone by how students rank high schools, the *application* channel.

With regard to segregation, our findings reveal both the challenges facing policymakers and a potential solution. First, middle school segregation has a reinforcing effect on high school segregation. High-achieving students are over-represented in high-quality middle schools, and attending such middle schools makes them aspire to higher-quality high schools even more.[2] Similarly, attending middle schools with many students of the same race strengthens racial homophily. These taste changes impact students' high school applications, and hence their assignments and high school segregation. However, our findings also imply that desegregating middle schools will substantially address high school segregation, leveraging that middle schools have subsequent impacts on high school applications and assignments.

[2] Several papers in the economics of education literature have similar findings that early tracking in schools can aggravate inequality (see Betts (2011) for a survey). We add to this literature by showing that segregation in an earlier educational level can affect their tastes for schools in later levels toward more segregation, which may be one potential mechanism of such findings.

This is what we quantify in the last part of the paper. We evaluate the desegregation impacts of NYC's recently announced admissions reforms when they are implemented at alternative educational levels.[3] In particular, we ask if middle school admissions reform can desegregate *both* middle *and* high schools; this works through the *application* and *priority* channels. We find that the middle school-only reform can desegregate not only middle schools but also high schools by altering students' high school applications toward less segregation and hence their assignments. Furthermore, combining middle and high school reforms has a larger effect on desegregating high schools than reforming only high school admissions.

The policy implication of our paper is that the dynamic relationship of school choices at different education levels can be used to design more effective policies. More concretely, the *supply-side* reform on middle school admissions (e.g., changing priority rules of middle schools in admitting students) can influence students' *demand* for high schools—namely, which high schools students apply to. Policymakers often find it challenging to influence the demand-side, and as a result, most existing policies that seek to address school segregation have focused on reforming the supply-side (i.e., admission rule reform).[4] However, recent evidence shows that a large part of the school segregation under a centralized school choice setting comes from the demand-side, and the supply-side-only intervention has a limited role (Laverde, 2022). Our findings therefore uncover and suggest a novel channel to influence the students' demand-side behavior and enhance the effectiveness of existing approaches to desegregate schools.

The paper is primarily related to three strands of the literature. First, we add to the economics of education and labor economics literature on the effects of schools on students' future outcomes. Many researchers have studied the effects on outcomes such as academic performance, including test scores (Hastings and Weinstein, 2008; Jackson, 2010; Pop-Eleches and Urquiola, 2013; Abdulkadiroğlu, Angrist, and Pathak, 2014), graduation and college outcomes (Deming, Hastings, Kane, and Staiger, 2014; Dobbie and Fryer, 2014), or labor

[3] We combine the two recent affirmative action policies the NYC DOE announced in the academic year 2020-21: first, removing screening based on test scores, and second, removing geographic priority rules. Relatedly, Idoux (2022) studies the same policies in the context of NYC middle schools, focusing on comparing the desegregating effects of the two. Instead of comparing the two policies, we focus on how alternative timings of implementing such policies affect segregation at different educational levels.

[4] For example, Chicago exam schools (Ellison and Pathak, 2021) use an affirmative action policy that prioritizes students based on the socioeconomic status of the neighborhood they reside in. Recently, Boston exam schools also adopted a similar admission policy reform (Barry, Ellen, "Boston Overhauls Admissions to Exclusive Exam Schools", The New York Times, 15 July 2021).

market outcomes such as occupation or wages (Card and Krueger, 1992a,b; Betts, 1995; Hoekstra, 2009; Clark and Bono, 2016), among many others. To the best of our knowledge, we are among the first to evaluate the effects of schools on students' future academic choices in a K-12 context.[5] Given the importance of schools for future outcomes as past studies have found, it is crucial to understand what may impact the school attended itself, and we suggest that a student's previous schools may be one key factor.[6]

Second, we contribute to the school choice literature. Several papers have studied the factors that may influence the outcomes of school choice, such as the assignment mechanism (Abdulkadiroğlu, Che, and Yasuda, 2015; Abdulkadiroğlu, Agarwal, and Pathak, 2017; He, 2017; Agarwal and Somaini, 2018; Che and Tercieux, 2019; Calsamiglia, Fu, and Güell, 2020) or information provision (Hastings and Weinstein, 2008; Hoxby and Turner, 2015; Ajayi, Friedman, and Lucas, 2017; Luflade, 2017; Corcoran, Jennings, Cohodes, and Sattin-Bajaj, 2018; Chen and He, 2021a,b; Grenet, He, and Kübler, 2021). However, all these papers were in a static framework, and to our knowledge, we are the first to incorporate a dynamic framework into the school choice literature.[7] We add to the literature by explicitly studying the dynamic relationship between school choices at different educational levels.

Third, we relate to the literature that leverages the quasi-experimental features built in school assignments, which includes making use of lotteries in charter school admissions (Hoxby and Rockoff, 2004), the tie-breaking features of centralized assignments (Cullen, Jacob, and Levitt, 2006; Deming, Hastings, Kane, and Staiger, 2014; Abdulkadiroğlu, Angrist, Narita, and Pathak, 2017, 2021), and test score cutoffs (Hoekstra, 2009; Pop-Eleches and Urquiola, 2013; Abdulkadiroğlu, Angrist, and Pathak, 2014; Dobbie and Fryer, 2014), among many others. We adopt the methodology of Abdulkadiroğlu, Angrist, Narita, and Pathak

[5]Recently, Mark, Corcoran, and Jennings (2021) conducted a descriptive analysis and found a low similarity in high school application among students from the same middle school or neighborhood in NYC. By using a quasi-random experiment, we argue that attending different middle schools has a systematic effect on how students view high school characteristics, while the exact identities of high schools they apply to may differ.

[6]Furthermore, without considering the sequentiality of school choice, the existing literature analyzing the impact of high school choice on future outcomes may potentially suffer from an omitted variable bias and may in particular overestimate the importance of high schools and therefore the effectiveness of any policy exclusively targeting high school assignments. Our findings suggest the importance of considering the sequentiality of school choice in studying such questions.

[7]In the broader empirical market design literature, several papers have considered dynamics such as in kidney waitlist (Zhang, 2010; Agarwal, Ashlagi, Somaini, and Waldinger, 2018; Agarwal, Ashlagi, Rees, Somaini, and Waldinger, 2021), public housing (Waldinger, 2021), or dynamic college admissions (Larroucau and Rios, 2020) among many others.

(2017, 2021) to obtain the 2SLS estimates of middle schools' causal effects on high school application/assignment. We use students' high school application patterns and assignment results as the outcome variables, departing from Abdulkadiroğlu, Angrist, Narita, and Pathak (2017, 2021), who study the effect of schools on students' achievement.

The rest of the paper is organized as follows. Section 2 provides the institutional background for NYC public school choice and describes our data. Section 3 analyzes the causal effects of middle schools on high school choice. Section 4 describes our structural model and provides the results of its estimation and the counterfactual analysis, and Section 5 concludes.

2 Institutional Background and Data

2.1 Public School Choice in NYC

NYC is one of the largest school districts worldwide that utilize centralized school choice to assign students to public schools. The school choice starts as early as for three-year-olds, and students/parents participate in the choice process of subsequent levels, including Pre-K, kindergarten, elementary, middle, and high school, as long as they wish to enroll in public schools. Schools that are part of the centralized choice system are governed and funded by the NYC Department of Education (DOE).

This paper focuses on middle and high school choices in NYC. The public middle school system consists of nearly 700 programs at around 500 middle schools. Multiple programs may be offered in one school. Similarly, the public high school system consists of nearly 800 programs at around 400 high schools.[8] Since the unit of admission is a program instead of a school, one may consider each program a separate school. In the following, we use the term 'program' and 'school' interchangeably when there is no confusion.

Both middle and high school systems use the student-proposing Deferred Acceptance (DA) algorithm (Gale and Shapley, 1962; Abdulkadiroğlu, Pathak, and Roth, 2005, 2009), which takes students' applications, schools' ranking over students, and the pre-announced

[8]Additionally, there are 9 specialized high schools in NYC such as Stuyvesant High School or Bronx High School of Science. We exclude these specialized high schools from our analyses since they use a separate admission process using a test called Specialized High Schools Admissions Test (SHSAT). Similarly, we exclude public charter schools because they have separate admission processes outside of the centralized school choice system.

number of seats as main inputs and produces at most one assignment for each student.[9]

Students apply to programs by submitting a rank-ordered list (ROL). In middle school choice, students can rank however many programs they are eligible for. In high school choice, students may rank up to 12 programs.[10]

Schools rank students by pre-announced admission rules, which consist of three layers. First, eligibility criteria determine the programs for which students are eligible. If a student is not eligible for a program, she is never considered by the program even when there is a remaining seat. Second, eligible applicants are classified into a small number of priority groups, for example, 'students or residents of Manhattan' or 'students who attended the information session'. A program considers all students in the higher priority group prior to any student in lower priority groups for admissions. For convenience, we use priority to denote both eligibility and priority groups when there is no confusion. Lastly, there are tie-breaking rules. Priority groups are often so coarse that the number of applicants from the same priority group exceeds the number of remaining seats. In this case, programs that actively screen students use *non-random tie-breakers* consisting of the previous year's GPA, statewide standardized test scores, attendance, and punctuality. Other programs break ties by a random *lottery* which is attached to each student and applies to all such programs in the same fashion (single tie-breaking rule).

Depending on the eligibility criteria, middle schools are classified into three types—district schools, borough schools, and citywide schools.[11] A student's residence or elementary school decides eligibility at each type of school. For example, among 670 programs in the academic year 2014-2015, 14 programs were citywide school programs, 39 programs were borough school programs, and the rest were district school programs. By contrast, the high school choice is fully citywide—students are eligible for almost all high school programs in NYC. Middle and high school programs can be further classified into subgroups depending on the details of the admission method, which is explained in detail in Appendix A.

[9] See Appendix A for details on how DA works.

[10] In this regard, the algorithm used for high school assignment is a modified version of DA with a limit on the number of choices, which alters the nature of DA (Haeringer and Klijn, 2009; Calsamiglia, Haeringer, and Klijn, 2010). For example, strategyproofness does not hold. However, we do not rely on the strategyproofness of DA throughout this paper.

[11] The city is divided into 5 boroughs and 32 community school districts.

2.2 NYC School Choice Data

We focus on the main round application data of students who participated in the middle school (MS) application in the academic year 2014-15 and then participated in high school (HS) application in the academic year 2017-18. Appendix B provides more details on data sources and sample restrictions. Our main sample consists of 54,012 NYC students applying to 670 middle school programs (472 middle schools) in the academic year 2014-15 and 767 high school programs (426 high schools) in the academic year 2017-18.

In the following analysis, we focus on two types[12] of schools—1) high-achievement, and 2) high-minority, which are defined based on the characteristics of the previous cohort. A school is labeled 'high-achievement' if the average standardized test score of students belongs to the top 1/3 in the distribution across schools. Similarly, a school is labeled 'highly minority' if the proportion of Black and Hispanic students belongs to the top 1/3 in the distribution across schools.

We present summary statistics of baseline student characteristics in Table 1. Columns (1)-(2) present summary characteristics of all middle school applicants (whole sample, $N = 62,972$), and Columns (3)-(4) present those of middle school applicants net of attrition (main sample, $N = 54,012$). The demographic characteristics and middle school application behavior are very similar between the whole sample and the main sample. The majority of students are either Black (23%) or Hispanic (41%) and Free/Reduced-price Lunch (FRL) eligible (72%), and 53% of students ranked a high-achievement middle school as their first choice. While a student lists 1.7 high-achievement middle schools on average, there is a remarkable variation from one student to another, which is captured by the sizable standard deviation.

Table 2 shows summary statistics of programs on admissions criteria and enrolled students' characteristics, overall and by school type. While 94% of middle school programs are open only to students from a school district or an attendance zone, only 4% of high school programs are. Next, many middle and high schools employ non-random tie-breakers to admit high-achieving students. 59% of high-achievement middle schools adopt non-random tie-breakers relative to the average of 42%. The contrast is sharper at the high school level (61% vs. 38%).

[12]These types are neither exclusive nor exhaustive. Our main results in Sections 3 and 4 are not sensitive to a different definition of types, for example, using above median, 60th-, 70th-, and 75th-percentile in the respective distribution.

Table 1: Summary Statistics of Student Characteristics

Variables	(1) Mean	(2) Std	(3) Mean	(4) Std
	All MS Applicants (Whole Sample)		Both MS and HS Application (Main Sample)	
Panel A: Demographics				
5th Grade ELA score	300.2	35.4	300.6	35.0
5th Grade Math score	310.8	37.7	311.3	37.3
English Language Learner (ELL)	0.12	0.32	0.12	0.32
Free/Reduced-price Lunch (FRL)	0.72	0.45	0.73	0.45
Asian	0.18	0.39	0.19	0.39
Black	0.23	0.42	0.23	0.42
Hispanic	0.41	0.49	0.41	0.49
White	0.17	0.37	0.17	0.37
N	62,972		54,012	
Panel B: Middle School Application Behavior				
Ranked High-Achievement MS 1st?	0.53	0.50	0.53	0.50
# of High-Achievement MS Ranked	1.66	1.71	1.67	1.72
Ranked High-Minority MS 1st?	0.20	0.40	0.20	0.40
# of High-Minority MS Ranked	0.78	1.46	0.77	1.44
N	63,207		53,211	

Note: Summary statistics of student characteristics in 5th grade are presented. A middle school is 'high-achievement' (resp., 'high-minority') if the average standardized test score (resp., the percent of Black and Hispanic students) of the previous cohort is greater than the 66th percentile of that across all schools. 5th Grade ELA score ranges from 100 to 410, and 5th Grade Math score ranges from 130 to 420.

Table 2: Summary Statistics of Middle and High School Program Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
	All		High-Achievement		High-Minority	
Variables	Mean	Std	Mean	Std	Mean	Std
Panel A: Middle School Program Characteristics						
Open Only to District/Zoned Students?	0.94	(0.24)	0.89	(0.32)	0.96	(0.21)
Use Non-random Tie-breaker?	0.42	(0.49)	0.59	(0.49)	0.35	(0.48)
Average Test Score (6th Gr.)	297.3	(20.5)	313.0	(17.7)	282.5	(11.8)
% White	14.17	(20.88)	27.51	(25.27)	1.062	(1.78)
% Black/Hispanic	70.92	(30.51)	47.92	(30.82)	97.16	(2.74)
% Free/Reduced-price Lunch	76.09	(19.06)	66.70	(22.05)	87.48	(8.75)
Cohort Size	98.30	(90.67)	111.90	(103.00)	71.08	(37.07)
1(STEM)	0.14	(0.35)	0.13	(0.33)	0.17	(0.37)
N	670		253		198	
Panel B: High School Program Characteristics						
Open Only to District/Zoned Students?	0.04	(0.19)	0.05	(0.21)	0.01	(0.09)
Use Non-random Tie-breaker?	0.38	(0.49)	0.61	(0.49)	0.31	(0.46)
4yr Graduation Rate (%)	67.01	(17.34)	81.79	(10.81)	60.67	(16.19)
College Enrollment Rate (%)	58.38	(17.13)	73.96	(11.49)	52.23	(15.11)
Average Test Score (9th Gr.)	294.4	(17.49)	311.3	(14.36)	285.7	(13.37)
% White	10.41	(15.56)	19.47	(19.69)	1.629	(02.15)
% Black/Hispanic	76.58	(23.40)	57.98	(26.10)	95.49	(04.20)
% Free/Reduced-price Lunch	80.13	(15.33)	70.75	(18.37)	87.78	(08.55)
Cohort Size	83.04	(82.66)	114.50	(122.70)	65.36	(45.48)
1(STEM)	0.31	(0.46)	0.31	(0.46)	0.28	(0.45)
% From High-Achievement MS	29.89	(26.68)	58.99	(24.80)	15.69	(17.40)
% From High-Minority MS	32.96	(26.53)	15.62	(21.71)	52.78	(22.82)
N	767		254		249	

Note: A middle school is 'high-achievement' (resp., 'high-minority') if the average standardized test score (the percent of Black and Hispanic students) of the previous cohort is greater than the 66th percentile of that across all schools. Average Test Score is a mean of ELA (English Language Arts) and math test scores. Educational Option high school programs are not counted as programs which use non-random tie-breakers (See Appendix A). The average test score ranges from 110 to 410 (resp., 130 to 400) for middle school (resp., high school) programs.

Table 2 also illustrates that students' characteristics vary markedly depending on the school type, suggesting that students sort into different schools based on their characteristics. For example, the mean average test score among all middle schools is 297.3, while it is 313.0 among high-achievement middle schools and 282.5 among high-minority middle schools.

Importantly, the last two rows of Panel B of Table 2 show a correlation between the type of middle school a student graduated from and the type of high school she attends. While 30% of high school students on average graduated from a high-achievement middle school, the number is twice as large among high-achievement high schools. Similarly, high-minority high schools admit more students who graduated from high-minority middle schools than an average high school. These patterns suggest two possibilities. First, students may have consistent tastes over middle and high school program characteristics.[13] Second, which middle school a student attends may change how she applies and is assigned to high schools. We aim to explore these possibilities in the following sections.

In Appendix H.1, we present the average school characteristics by rank on students' ROLs of middle and high schools for interested readers.

3 Causal Effects of Middle School Attendance on High School Choice

In this section, we provide evidence that the middle school a student attends has a causal impact on the high school programs the student applies to and is assigned to. We focus on the effect of attending a high-achievement middle school in the main text and present the effects of attending a high-minority middle school in Appendix H.2.

3.1 Empirical Strategy

Our main identification concern is that students may sort into different middle schools based on unobserved factors (to the researcher), which could also affect how students choose high schools and where they are assigned to. For example, a student who prefers high-achievement

[13]For example, one reason for the consistency could be that geographically close middle and high schools have similar characteristics, and students usually have the same residential location when they apply to middle and high schools. We control for the borough of residence in Section 3 and the distance to each school in Section 4.

middle schools more than her peers of the same observable characteristics presumably also prefers high-achievement high schools. To deal with this selection issue, we adopt the research design introduced by Abdulkadiroğlu, Angrist, Narita, and Pathak (2017, 2021) that builds on the quasi-experimental variation embedded in DA. We explain the strategy briefly in the following and recommend that interested readers consult the original papers for details.

Recall that in NYC, students' applications, priorities, and tie-breakers are the only factors determining assignments. When needed, programs use either *lotteries* or program-specific *non-random tie-breakers* to break ties (see Section 2).

At programs that use *lotteries*, students' assignments are random after controlling for student application and priority (Abdulkadiroğlu, Angrist, Narita, and Pathak, 2017). For other programs that use *non-random tie-breakers*, Abdulkadiroğlu, Angrist, Narita, and Pathak (2021) take a nonparametric regression discontinuity (RD) approach (Hahn, Todd, and Van der Klaauw, 2001) and exploit a subset of assignments that are as good as random. The concern is that non-random tie-breakers might be correlated with students' unobserved abilities or preferences, and thus assignments are no longer random even after controlling for application and priority. However, applicants whose composite scores of priority and tie-breaker are in the small neighborhood around the program's cutoff have a constant risk of clearing the cutoffs of 1/2 (Proposition 1 of Abdulkadiroğlu, Angrist, Narita, and Pathak, 2021).

In practice, we control for the propensity score—the probability of being assigned to treatment schools—rather than all observed cases of student applications and priorities. This is because there are as many unique combinations of applications and priorities as the number of students, and the propensity score reduces the dimension effectively (Rosenbaum and Rubin, 1983). Abdulkadiroğlu, Angrist, Narita, and Pathak (2021) show that DA-generated assignments are independent of any variables unaffected by the treatment after conditioning on the propensity score.[14]

[14]Propensity score denotes the odds of being assigned to a certain type of middle school as a function of student application, priority group, and cutoffs. We can calculate the propensity score for each middle school program for each student. Since DA produces at most one assignment for each student, summing up the propensity score across middle school programs that belong to a certain school type gives the propensity score of being assigned to a middle school of such type. If a student does not apply to middle schools of a certain type, the propensity score is zero. Theorem 1 of Abdulkadiroğlu, Angrist, Narita, and Pathak (2021) provides a compact characterization of such propensity scores using a large market approximation. We provide a simple example of the calculation of the propensity scores in Appendix C.

The treatment effect of interest, the effect of attending a certain type of middle school, is estimated from a two-stage least squares (2SLS) model where the DA *assignment* is used as an instrumental variable for the actual *attendance*.

$$Y_i = \alpha_0 + \beta C_i + \sum_x \alpha_1(x) d_i(x) + g(\mathcal{R}_i) + \delta' Z_i + \eta_i \tag{1}$$

$$C_i = \tilde{\alpha}_0 + \gamma D_i + \sum_x \alpha_2(x) d_i(x) + h(\mathcal{R}_i) + \tau' Z_i + \nu_i \tag{2}$$

Equation (1) is the main equation of interest where β is the treatment effect of interest, and Equation (2) is the respective first-stage regression. Y_i is our outcome of interest describing student i's high school choice behavior or outcomes, C_i is the treatment variable which equals 1 if i attended one of the treatment middle schools and 0 otherwise. D_i is the instrument variable which equals 1 if i was assigned to treatment schools by DA and 0 otherwise. We also include Z_i, the vector of student observable characteristics (ELL, ethnicity, FRL, gender, baseline test scores, and borough of residence) when they were 5th graders i.e., before applying to middle schools. $d_i(x)$ is a dummy variable that equals 1 if i's propensity score equals x and 0 otherwise, and the set of parameters $\alpha_1(x)$ and $\alpha_2(x)$ provides a saturated nonparametric control for all possible values of the propensity score for the DA assignment D_i.[15] $g(\mathcal{R}_i)$ and $h(\mathcal{R}_i)$ are local linear controls for non-random tie-breakers at each program that uses such tie-breakers.[16]

To interpret β as causal, we argue that the exclusion restriction holds. That is, after controlling for propensity scores and non-random tie-breakers, DA assignments D_i are random and do not affect outcomes Y_i other than by affecting the actual attendance C_i.[17] To support this assumption, we provide balance test results in Appendix H.2. The instrumental variable balances the covariates of the students who are assigned to the treatment middle schools

[15] This is possible since the support of the propensity scores is finite. See Abdulkadiroğlu, Angrist, Narita, and Pathak (2021) for more details.

[16] We include a local linear function for each of 104 types of non-random tie-breakers in the data. We also include a set of dummy variables corresponding to each non-random tie-breaker to deal with students who did not apply to a school using that non-random tie-breaker, or students who applied but whose tie-breakers are far from the cutoff following Abdulkadiroğlu, Angrist, Narita, and Pathak (2021). We use the IK bandwidth (Imbens and Kalyanaraman, 2012) separately for each program as suggested by Abdulkadiroğlu, Angrist, Narita, and Pathak (2021).

[17] In principle, controlling for the propensity scores $\{d_i(x)\}$ is enough for the exclusion restriction by Theorem 1 of Abdulkadiroğlu, Angrist, Narita, and Pathak (2021). We further control for student characteristics and non-random tie-breakers to get a more precise estimate of the treatment effect β.

by DA and those who are not, after controlling for the propensity score and non-random tie-breakers among students with non-degenerate risk of being offered (i.e., whose propensity score is in the interval $(0, 1)$ and hence subject to randomization). Based on the balance test result, our preferred specification in the following controls for propensity scores and non-random tie-breakers with non-degenerate risk sample.[18]

3.2 Empirical Results

Table 3 shows our main results. Each panel corresponds to different high school characteristics as the outcome variable.

In Columns (1)-(3), we focus on the average characteristics of the top 5 ranked high school programs.[19] Column (1) presents OLS estimates for comparison. Column (2) presents 2SLS estimates with the full sample, and Column (3) presents our preferred specification—2SLS only with the non-degenerate risk sample. First, we find that OLS overestimates the effects of attending a high-achievement middle school as concerned. For example, in Panel C, the OLS estimate suggests that the average proportion of high-performers in a student's top 5 ranked high school programs increases by 5.19 percentage points when she attends a high-achievement middle school. On the other hand, the 2SLS estimate in Column (2) shows an effect of 3.33 percentage points, and our most preferred estimate in Column (3) is 2.99 percentage points. This contrast confirms the importance of controlling for selection based on unobservables.

Most importantly, our 2SLS estimates show that attending a high-achievement middle school has a causal effect on the characteristics of high school programs a student applies to. In Column (3), we see that the average graduation rate, college enrollment rate, and the proportion of high-performers of the top 5 choices increase by 1.38, 1.76, and 2.99 percentage points on average, respectively.

Next, Columns (4)-(6) illustrate that attending a high-achievement middle school also changes the characteristics of the assigned high school program, not only of the programs students apply to. Attending a high-achievement middle school changes the graduation

[18] Such sample restriction comes with the cost of losing many observations (from N=50,871 to N=8,007). We find that students with non-degenerate risk and those with degenerate risk are quite different: students with non-degenerate risk on average have higher test scores, and more likely to be White. It reconfirms that the 2SLS estimates are local average treatment effect (LATE). Appendix Figure H.2 presents the mean difference between those with non-degenerate offer risk and degenerate offer risk.

[19] Using the average characteristics of the top 1, top 3, or all choices does not significantly change the results.

Table 3: Effect of Attending a High-Achievement MS on HS Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Average of Top 5 Ranked				Assigned	
Model	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Sample	All	All	NDR	All	All	NDR
Panel A: 4yr Graduation Rate (%)						
From High-Achievement MS	1.764***	1.684***	1.379*	3.109***	2.482***	2.422**
	(0.404)	(0.561)	(0.735)	(0.529)	(0.855)	(1.142)
N	44159	44159	7060	41623	41623	6687
R2	0.293	0.318	0.387	0.185	0.202	0.253
$\bar{y}$	83.321	83.321	83.729	78.954	78.954	79.901
Panel B: College Enrollment Rate (%)						
From High-Achievement MS	2.854***	1.716**	1.755*	4.530***	3.025**	3.414**
	(0.516)	(0.780)	(1.011)	(0.669)	(1.191)	(1.566)
N	44158	44158	7060	41546	41546	6679
R2	0.367	0.390	0.459	0.244	0.263	0.310
$\bar{y}$	71.217	71.217	72.197	65.653	65.653	67.204
Panel C: % High-Performing Students						
From High-Achievement MS	5.188***	3.328***	2.986*	6.886***	5.293***	5.292**
	(0.840)	(1.291)	(1.805)	(0.825)	(1.650)	(2.105)
N	44237	44237	7062	42180	42180	6751
R2	0.450	0.473	0.502	0.388	0.406	0.400
$\bar{y}$	39.731	39.731	40.934	33.058	33.058	34.978
Panel D: % White						
From High-Achievement MS	5.080***	2.202***	0.311	5.755***	1.915**	0.301
	(0.750)	(0.729)	(0.655)	(0.793)	(0.819)	(0.832)
N	44237	44237	7062	42180	42180	6751
R2	0.633	0.652	0.717	0.555	0.573	0.621
$\bar{y}$	18.627	18.627	20.334	15.097	15.097	16.761
Panel E: 1(STEM)						
From High-Achievement MS	-0.053***	0.013	0.041	-0.057***	0.022	0.055
	(0.013)	(0.024)	(0.035)	(0.016)	(0.032)	(0.044)
N	44237	44237	7062	42182	42182	6751
R2	0.098	0.126	0.275	0.041	0.059	0.172
$\bar{y}$	0.324	0.324	0.318	0.314	0.314	0.322
First Stage F-stat		146.8	135.2		146.8	135.2

Note: Standard errors clustered at graduating middle school in parentheses. All regressions control for student ethnicity, gender, English Language Learner status, Free/Reduced-price Lunch eligibility, Special Education status, standardized test score in 5th grade, and residential borough in 5th grade. Column (2)-(3) and (5)-(6) control for dummy variables for all possible values of propensity score of being assigned to a high-achievement MS, and local linear control for non-random tie-breakers.
*p<0.1, **p<0.05, ***p<0.01.

rate of the assigned high school program increases by 2.42 percentage points, the college enrollment rate by 3.41 percentage points, and the proportion of high-performing students by 5.29 percentage points (Column (6)). Notably, the magnitude of effects is larger for assigned programs than those a student applies to. This implies that attending a high-achievement middle school not only changes how students value different program characteristics but also how a student is viewed by programs for admission purposes.[20] [21]

4 A Dynamic Model of Middle and High School Choices

We now turn to a dynamic model of middle and high school choices.

The need for a model is twofold. First, students' school assignments are determined as an equilibrium outcome; how *all* students act jointly determines the assignments. While the effect we identified in the previous section is marginal for each *treated* student, any counterfactual policy change will trigger a change in behavior of *all* students, in turn changing the equilibrium. Second, having identified the effects of middle schools on high school choice, we are also interested in exploring how these effects occur. A model is useful to decompose the channels through which middle schools affect high school choice and to quantify each channel's relative importance.

The model is a two-period dynamic model. The first period corresponds to middle school applications and assignments, and the second period corresponds to high school applications and assignments. Based on our empirical findings in Section 3, we incorporate three key features in our model.

First, the model explicitly allows students' tastes for high schools that underlie their applications to depend on the middle school they attend (*application* channel). Students' test scores may change by attending middle schools with different effectiveness, and students may put more/less weight on some high school characteristics depending on their academic

[20] The figures are slightly larger than Corcoran, Jennings, Cohodes, and Sattin-Bajaj (2018) which conducts a field experiment by providing a customized one-page list of proximate high schools with a high graduation rate to students attending high poverty middle schools in NYC.

[21] We supplement the main analysis by additional results in Appendix H.2. We show that the effects are robust to controlling for students' end-of-middle-school test scores and the length of the submitted ROLs. We also explore the heterogeneity in treatment effects by student observable characteristics. Finally, we show that there is no significant effect of attending a high-minority school middle school, which is our second treatment variable.

preparedness (Hastings, Kane, and Staiger, 2005; Abdulkadiroğlu, Agarwal, and Pathak, 2017). Middle schools could also change students' tastes for high schools through other channels than test scores, which we capture as a portmanteau parameter.[22]

Second, how a student is prioritized at each high school program for admissions also depends on the middle school she attended (*priority* channel). First, attending different middle schools may result in different end-of-middle-school test scores, which in turn affects students' admission chances at high school programs that use test scores for admissions. Second, some high schools give eligibility/priority depending on which middle school a student attends.[23]

Third, students may be forward-looking; namely, they may consider those *application* and *priority* channels when they apply to middle schools. More concretely, students form expectations on how they will benefit in the high school choice from attending a particular middle school, which in turn affects how they value different middle school programs.

It is useful to define a few terms before describing our theoretical framework. **Intrinsic priority** is each student's priority at each program that is known ex-ante (e.g., the priority group a student belongs to). Each student with intrinsic priority realizes **ex-post score** at each program which is used by programs to rank students for admissions. For example, in our context, ex-post score is the summation of a student's priority group and the realized lottery draw for the student. A student with a higher score has a higher priority for admissions at each program. Given any matching of students and programs, a program's **ex-post cutoff** is defined as the lowest ex-post score of the admitted students if the seats are filled and $-\infty$ otherwise. Finally, a program is called to be **feasible** to the student if she has a higher ex-post score than the ex-post cutoff, regardless of ranking it.[24]

[22]The *application* channel can include several things beyond what can be captured through the change in test scores. For example, middle schools may directly change students' intrinsic tastes for high schools. Also, attending a better middle school may allow a student to access better information about the quality of high schools, or the high school choice process itself. Peers at different types of middle schools may have differential effects on students' tastes for high schools. Due to data availability, we do not further distinguish among these possibilities but rather include a parameter capturing the 'lump-sum' change in tastes that underlie each student's high school application.

[23]Since schools rank students by pre-announced admission rules, we assume that schools are passive players, as is typical in the literature. This in turn enables one to separately identify the *application* channel from the *priority* channel.

[24]In other words, a feasible program will accept the student should she top-rank the program.

4.1 Theoretical Framework: A Two-Period Model

In the following, denote each student as $i \in \{1, \cdots, I\} = \mathcal{I}$, middle school programs as $m \in \{1, \cdots, J_m\} = \mathcal{M}$, and high school programs as $j \in \{1, \cdots, J_h\} = \mathcal{J}$. We start from period 2 and work backwards.

4.1.1 Period 2: High School Application

High School Application Consider student i who is enrolled in middle school program $m(i)$. Student i has flow utility V_{ij} from enrolling in high school program $j \in \mathcal{J}$

$$V_{ij} = v\left(\tilde{X}_j, Z_i^H, \tilde{d}_{ij}, \gamma_i^H; m(i)\right) + \eta_{ij}$$

where $\tilde{X}_j$ is the vector of observable characteristics of high school program j, Z_i^H is the vector of student observable characteristics when applying to high schools, and $\tilde{d}_{ij}$ is the distance between student i's residence and program j's location. γ_i^H is the vector of student i's unobserved tastes for $\tilde{X}_j$, and η_{ij} is an idiosyncratic preference shock that is iid for each i and j.

Based on the flow utilities and intrinsic priorities, each student submits an ROL, and DA is run with all students' submitted ROLs and ex-post scores to produce high school program assignments (and cutoffs).

Behavioral Assumption In a school choice situation, each student is playing an incomplete information game: each student's assignment is determined not only by her ROL and priorities but also by (i) other students' ROLs and priorities and (ii) the tie-breaking lottery realizations, which are both unknown ex-ante. We assume that in both periods 1 and 2, students submit ROLs such that the resulting assignment outcomes are **ex-post stable** and interpret their choices accordingly.[25] That is, the assigned program of a student is her favorite program

[25] Strictly speaking, we rely on asymptotic stability which implies ex-post stability in a large enough economy (Che, Hahm, and He, 2021). Asymptotic stability (and hence ex-post stability) may be violated when there is a limit on the length of the ROL students can submit and hence the risk of being unassigned is not negligible. In such case, we need to guarantee that there are enough choices ranked to hedge against the risk of being unassigned. In our data, (i) students on average rank 4 high school programs which is much lower than the limit of 12 (recall that only a high school ROL has a length limit) and (ii) the proportion of unassigned students is very small (0.07%). Both indicate that the limit on the length of the ROL and hence the violation of stability are unlikely to be an issue in our context.

among the feasible programs.

Ex-post stability is consistent not only with the implication of the truth-telling assumption but also with students' deviations from truth-telling even in a strategyproof environment (Che, Hahm, and He, 2021).[26] The truth-telling assumption has been traditionally used in the school choice literature (for example, Abdulkadiroğlu, Agarwal, and Pathak, 2017) based on the strategyproofness of DA. However, the truth-telling assumption is not robust to mistakes when such deviations from truth-telling do not affect a student's payoff. For example, a low-performing middle school senior may not choose to apply to her favorite but highly competitive high school program because there is zero chance of admission. This does not entail a payoff loss compared to truth-telling, even if it is one of her most desirable programs. However, the truth-telling assumption would interpret that the student did not like that highly competitive program and hence did not apply to it. On the other hand, ex-post stability will not infer anything about her preference on it since it was not feasible for the student. Therefore, we use ex-post stability as our preferred assumption since is robust to such payoff-irrelevant mistakes. We also estimate the model assuming truth-telling as an additional robustness check in Appendix E.

Ex-post stability plays a significant role in simplifying a rather complicated game situation. In particular, we can focus on outcomes rather than strategies. That is, it enables us to interpret the school choice data such that for each student, her assigned program gives the maximum utility among the programs that were feasible for the student without knowing the exact strategy the student employed to submit ROL. Without ex-post stability, one needs to fully solve the game of incomplete information that each student is facing by enumerating all possible ROLs and finding the optimal strategy profile among them, which would make the estimation of the model extremely heavy in terms of computation. Ex-post stability essentially enables us to interpret the data using a conditional multinomial choice model, where a student's *choice* is the assigned program, and the *choice set* is the ex-post feasible set.[27] Furthermore, it helps us simplify the continuation value of a given middle school to what is known as the 'Emax' term in the dynamic discrete choice literature, as will be seen

[26] Such deviations are often regarded as *mistakes* in the literature. See Hassidim, Romm, and Shorrer (2016); Li (2017); Artemov, Che, and He (2021) for examples of such mistakes in real world and lab experiment settings.

[27] The exogeneity of choice set is satisfied by assuming a large market i.e., the market is large enough that each student cannot affect the cutoffs.

in the description of the first period.

4.1.2 Period 1: Middle School Application

Forward-Looking Behavior Each student is forward-looking. In the first period, each student takes into account that enrolling in a particular middle school program may affect her payoffs in the second period. Hence, we need to model how she forms expectations on the 'continuation value' of each middle school program.

The key concept is ex-post stability. Due to ex-post stability, the ex-ante uncertainties that determine the ex-post scores and cutoffs (in our context, the tie-breaking lottery draws) are the sufficient statistics of the uncertainties present in the economy that affect students' payoffs at their assigned programs. To see this, imagine that a draw of lottery tie-breakers is realized and assigned to each student. DA is then run with the resulting ex-post scores and submitted ROLs, creating ex-post cutoffs of high school programs. Ex-post stability implies that each student is assigned to her favorite high school program among the ex-post feasible high school programs, and hence, knowing the lottery realization is sufficient to know each student's payoff at the assignment.

To this end, let ω denote the uncertainty that determines the ex-post scores and cutoffs in the second period (high school application) with some known distribution $H(\omega)$ where ω is unknown ex-ante. Across different realizations of ω, the high school flow utility V_{ij} is invariant, but the feasibility of a high school program varies, and thus ω affects the expected payoff from high school choice. Let $O_i(Z_i^H, m; \omega)$ denote student i's ex-post feasible set of high school programs given the realization of the uncertainty ω. To capture the aforementioned *priority* channel, $O_i(Z_i^H, m; \omega)$ is explicitly a function of Z_i^H (which may depend on m) and the middle school attendance m.[28]

[28] Recall that the priority channel includes two possible effects of a given middle school. First, the change of test scores which can influence a student's standings at programs that actively screen applicants based on test scores, and second, the change of eligibility or priority group. The former is captured by Z_i^H, and the latter is captured by the additional inclusion of m in the notation.

Middle School Application Now we are ready to describe the first period. Each student i submits ROLs on middle school programs satisfying ex-post stability, based on the utilities

$$U_{im} = \underbrace{u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) + \epsilon_{im}}_{\text{Flow utility of attending } m} + \underbrace{\delta\, E_{\gamma_i^H, \omega, \eta_i, Z_i^H}\left[\max_{j \in O_i(Z_i^H, m; \omega)} V_{ij} \middle| Z_i^M, \gamma_i^M, \epsilon_i, m\right]}_{\text{Continuation value of attending } m} \tag{3}$$

when student i enrolls in middle school program m. X_m is the vector of observable characteristics of middle school program m, Z_i^M is the vector of student observable characteristics when they apply to middle schools, and d_{im} is the distance between student i's residence and program m's location. γ_i^M is the vector of student i's unobserved tastes for X_m, and ϵ_{im} is an idiosyncratic preference shock that is iid for each i and m. δ describes how much each student values the future relative to the current flow payoff, which we later estimate together with other parameters.

Note that U_{im} includes the continuation value of attending m in addition to the flow utility of attending m. By ex-post stability, given ω, student i who attended m will be assigned to the high school program that gives her the maximum utility among those in the ex-post feasible set $O_i(Z_i^H, m; \omega)$. Hence, the continuation value of attending m is the conditional expectation of $\max_{j \in O_i(Z_i^H, m; \omega)} V_{ij}$, where the expectation is with respect to the state variables in the second period (including ω) that are unknown to the student in the first period, and conditional on the state variables known in the first period as well as the middle school program m. Appendix D.3 provides assumptions on the unobservables and explains how those assumptions help simplify the expression of the continuation value.

Based on the utilities and intrinsic priorities, each student submits an ROL, and DA is run with all students' submitted ROLs and ex-post scores to produce middle school program assignments (and cutoffs).

Table 4 summarizes what is known to student i in each period.

4.2 Estimation

Parameterization: Preferences We parameterize the payoff functions using a random coefficient model.

Table 4: Information Available to a Student in Each Period

	Unobserved Taste on School Char. γ_i^M	γ_i^H	Idiosyncratic Preference Shock ϵ_{im}	η_{ij}	Program Characteristics $X_m, \tilde{X}_j$	Student's own Characteristics Z_i^M, Z_i^H	Uncertainty in High School Choice ω
1st Period (Middle School Application)	✓		✓		✓	✓	
2nd Period (High School Application)	✓	✓	✓	✓	✓	✓	

Note: We assume high school program characteristics are exogenous, fixed, and are known to students in the first period. This is supported by the fact that school characteristics are stable over the years. Also, we assume a student has perfect foresight on what Z_i^H she will have by attending each m . Appendix D.2 provides details on how we estimate each middle school's production function of Z_i^H using a value-added model.

First, the flow utilities in each period are

$$\begin{aligned} u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) &= \tilde{u}\left(X_m, Z_i^M, \gamma_i^M\right) - \lambda^M d_{im} \\ &= X_m' \beta_i^M - \lambda^M d_{im} \\ v\left(\tilde{X}_j, Z_i^H, \tilde{d}_{ij}, \gamma_i^H; m(i)\right) &= \tilde{v}\left(\tilde{X}_j, Z_i^H, \gamma_i^H; m(i)\right) - \lambda^H \tilde{d}_{ij} \\ &= \tilde{X}_j' \beta_i^H - \lambda^H \tilde{d}_{ij} \end{aligned}$$

where λ^M and λ^H capture the disutility of traveling, and β_i^M, β_i^H allow students' tastes for program observable characteristics to be heterogeneous across i. We normalize the location of the utilities by setting $\tilde{u}(\cdot) = \tilde{v}(\cdot) = 0$ if all of their arguments are equal to zero. Additionally, we assume that $(\gamma_i^M, \epsilon_{im}) \perp d_{im} \Big| X_m, Z_i^M$ and $(\gamma_i^H, \eta_{ij}) \perp \tilde{d}_{ij} \Big| \tilde{X}_j, Z_i^H, m(i)$ which together with the additive separability of $d_{im}, \tilde{d}_{ij}$ provide nonparametric identification of the utilities $\tilde{u}$ and $\tilde{v}$ (Agarwal and Somaini, 2018).[29]

Let the dimension of X_m, $\tilde{X}_j$ and consequently that of β_i^M, β_i^H be L. For the l-th program

[29]We assume there are common values of outside options 0_m and 0_h to be estimated for middle and high school choices, respectively. That is, $U_{i0_m} = \vartheta_{mi} + \epsilon_{i0_m}$ and $V_{i0_h} = \vartheta_{hi} + \eta_{i0_h}$ where $\epsilon_{i0_m}, \eta_{i0_h}$ both follow EVT1.

characteristic, we parametrize the random coefficients as:

$$\beta_{i,l}^{M} = \beta_{0,l}^{M} + Z_i^{M\prime}\beta_{Z,l}^{M} + \gamma_{i,l}^{M}$$
$$\beta_{i,l}^{H} = \beta_{0,l}^{H} + Z_i^{H\prime}(m(i))\beta_{Z,l}^{H} + \underbrace{\sum_{\tau=1}^{T}\rho_{\tau,l}1\left(\tau(m(i)) = \tau\right)}_{\text{Middle school type effect}} + \gamma_{i,l}^{H}$$

for each $l = 1, 2, \cdots, L$. $\beta_{0,l}^{M}$, $\beta_{0,l}^{H}$ capture the common valuation of all students on the l-th program characteristic in each period. The interaction terms $Z_i^{M\prime}\beta_{Z,l}^{M}$ and $Z_i^{H\prime}\beta_{Z,l}^{H}$ allow individual tastes to depend on individual observable characteristics Z_i^M and Z_i^H, respectively.

Student i's taste over high school characteristics, $\beta_{i,l}^{H}$, is a function of the student's middle school $m(i)$. The student's test score evolves differently depending on $m(i)$, which is captured by $Z_i^H(m(i))$. More importantly, $\sum_{\tau=1}^{T}\rho_{\tau,l}1(\tau(m(i)) = \tau)$ is what we call the *middle school type effect*, where $\tau(m(i))$ is the type of i's attended middle school $m(i)$. It allows students who attend middle schools with some type $\tau = 1, \cdots, T$ to have a different mean valuation of high school program characteristics. $\rho_{\tau,l}$ plays a similar role as the treatment effect β in Equation (1) when the outcome variables are the characteristics of the programs students applied to.[30]

$\gamma_i^M = (\gamma_{i,1}^M, \cdots, \gamma_{i,L}^M)$ and $\gamma_i^H = (\gamma_{i,1}^H, \cdots, \gamma_{i,L}^H)$ capture students' unobservable tastes for middle and high school program characteristics. They are serially correlated, which generates a source of sorting across two periods. We assume:

$$\gamma_i^H = \underbrace{diag(\rho_0)\gamma_i^M}_{\text{serial correlation}} + \xi_i. \tag{4}$$

ξ_i captures the innovation to the unobservable tastes that is only realized in the second period. We assume that $\gamma_i^M \overset{iid}{\sim} \mathcal{N}(0, \Sigma_\gamma)$, $\xi_i \overset{iid}{\sim} \mathcal{N}(0, \Sigma_\xi)$ and they are mutually independent, and we allow Σ_γ and Σ_ξ to be fully flexible.[31]

[30]By using ex-post stability in our model, we implicitly assume that students are aware of not only all the options and their attributes, but also the distribution of admission probabilities at each program. However in real life, students' preferences on programs and also the middle school type effects may operate through information frictions (Luflade, 2017; Neilson, Allende, and Gallego, 2019; Son, 2020). As a result, we follow Allende (2019) and do not interpret our parameter estimates as deep structural preferences but as weights students put on attributes which govern students' application behaviors. Since it is unlikely that these weights will change under the counterfactual scenarios we consider in Section 4.5, the model can be used to predict behaviors. On the other hand, we do not focus on welfare analysis for the reason explained above.

[31]We impose a restriction that $diag(\rho_0)$ is a diagonal matrix. It means that the unobservable taste for one

Finally, we assume that the idiosyncratic preferences ϵ_{im} and η_{ij} both follow Extreme Value Type-I (EVT1) distribution. Together with the assumption on the unobservables, it implies that the continuation value expression can be further simplified to a convenient form (see Appendix D.3).

Source of Identification Our primary identification concern is to distinguish the causal effect of the type of middle school on tastes for high schools ($\{\rho_\tau\}_\tau$) from students' unobservable tastes for those (γ_i^H). The data shows a large correlation between the high school characteristics a student applies and is assigned to and the middle school characteristics she attends (see Table 1). A large part of this relationship can be explained by students' observable characteristics that are constant over time. However, even conditional on observable characteristics, there still is a positive correlation (see Table 3). This could be attributable to either the consistency of the individual student's unobserved tastes over time (i.e., γ_i^M and γ_i^H) or the treatment effect of attending a particular type of middle school ($\{\rho_\tau\}_\tau$).

The key to distinguishing between these sources of explanation comes from the panel structure of the data. That is, we observe each student's middle and high school ROLs. First, the correlation between the unobservable tastes across periods ρ_0 is identified by the degree to which the same student's middle and high school applications look similar after controlling for her observable characteristics.

Next, ρ_τ is identified by how similar the high school applications are across students attending middle schools of the same type. Notably, we implicitly rely on the quasi-random variation in school assignments generated by the tie-breaking rule. The quasi-random assignments generate variation in what type of middle school a student attends beyond her middle school application and intrinsic priorities. Without the quasi-randomness generated by the tie-breaking, observably similar students' attending different middle schools would be all attributable to the difference in γ_i^M once we assume nonparametric identification of the unobserved taste γ_i^M.[32] Thanks to the quasi-random assignments in addition to the distributional assumption on the unobserved tastes, we have variations in which type of

middle school characteristic (e.g., the proportion of high-performers in middle schools) has an impact on the unobservable taste for the corresponding high school characteristics (e.g., the proportion of high-performers in high schools), but not on others (e.g., the proportion of White students in high schools). But the arbitrary correlation among γ_i^M allows $\gamma_i^{M,l}$ to be correlated with $\gamma_i^{H,l'}$ for $l' \neq l$.

[32] Note that priority is also determined based on students' observable characteristics.

middle school a student attends beyond what can be explained by students' observable characteristics and unobserved tastes. Finally, the remaining variation in the application explained by neither within-student consistency nor across-student (of the same middle school) correlation is captured in Σ_ξ.

Estimation For tractability, we focus on students and schools of Staten Island (SI), which is one of the five boroughs of NYC. SI can be effectively treated as an independent school district in NYC since commuting outside of SI is very costly for students.[33] In 2017-18, only 1.8% of the SI middle school students ever ranked a high school program outside of SI. There are 2,626 SI students applying to 20 middle school programs (14 schools) and 47 high school programs (10 schools) in our estimation sample.[34]

For the program characteristics, we use three variables: the proportion of high-performers (current 6th graders (resp., 9th) whose average score of statewide ELA and math exams belongs to the top 1/3), the proportion of White students in current 6th (resp., 9th) grade, and if the program focuses on STEM-related fields. For the student characteristics, we use ethnicity dummy variables (Asian, Black, Hispanic), Free/Reduced-price Lunch (FRL) status, English Language Learner (ELL) status, and the average of most recent math and ELA standardized test scores (normalized to mean 0 and std 1). Finally, we include two types of middle schools, high-achievement middle schools and high-minority middle schools, defined in Section 2.[35]

We estimate via maximum simulated likelihood estimation (MSLE) using sparse grids quadrature (Heiss and Winschel, 2008). Appendix D provides more details on the procedure we use to estimate our model.

[33]One can travel from SI to other boroughs in NYC only via the Staten Island Ferry or the Verrazzano-Narrow Bridge, the only ground transportation route to Brooklyn. See Figure H.4 for the map of NYC school districts.

[34]SI is on average a richer borough with more White and slightly higher performing students compared to the rest of NYC. The proportion of subsidized lunch status was about 54% (72% citywide), the proportion of White students was about 56% (17% citywide), and the average statewide Math exam score was 315 (311 citywide) in academic year 2014-2015. Hence, we do not intend to extrapolate our findings to other boroughs of NYC.

[35]The school types are redefined using only SI schools.

4.3 Results

4.3.1 Model Estimates

Table 5 provides the model estimates of our main specification. The model estimates have mainly three implications.[36]

First and most importantly, we reconfirm that middle schools affect how students value different high school characteristics, as shown by the estimate of the middle school type effect ρ_τ being significantly different from zero. All else equal, attending a high-achievement middle school makes a student willing to travel 0.11 miles more for a 10 pp increase in the proportion of high-performers (resp., 0.31 miles more for a 10 pp increase in the proportion of White students). On the other hand, attending a high-minority middle school makes a student willing to travel 0.17 miles more for a 10 pp increase in the proportion of high-performers (resp., 0.28 miles less for a 10 pp increase in the proportion of White students).[37] Notably, these estimates imply that middle school segregation potentially has a reinforcing effect on high school segregation through the change in students' tastes for high schools. The more high-achieving students attend high-achievement middle schools, the more they will apply to high-achievement high schools than their low-achieving peers. Similarly, attending middle schools with many students of the same race strengthens racial homophily.[38]

Second, the positive (and statistically significant) estimate of δ shows that students prefer middle school programs with higher continuation values. That is, they are forward-looking and value middle school programs that give higher expected utility in the high school choice process. Students are willing to travel 0.81 miles more for one standard deviation increase in

[36]We report the willingness to travel by dividing the coefficient of interest by the coefficient on distance. The average commuting distance to each assigned high school in the data is 2.3 miles.

[37]The estimates of the effect of attending high-minority middle school provide a potential explanation for the nearly null effect we find in Table H.10. In reality, the proportion of high-performing students and the proportion of White students are positively correlated ($r = 0.62$ among SI high school programs), making the effects of high-minority middle schools on the taste for high schools cancel out each other. This results in nearly null treatment effects of high-minority middle schools since we do not consider each program's characteristics simultaneously in the 2SLS in Section 3.

[38]To provide evidence of this argument, we estimate a similar model in which ρ_τ is allowed to be heterogeneous depending on student's own ethnic group. We find that attending a high-minority middle school makes White/Asian (resp., Black/Hispanic) students will to travel 0.34 miles less (resp., 0.22 miles less) for a 10 pp increase in the proportion of White students. It implies that (i) Black/Hispanic students who attend high-minority middle schools with many Black/Hispanic students prefer Black/Hispanic students more, and (ii) White/Asian students who attend *low* minority middle schools with *less* Black/Hispanic students prefer White/Asian more, which we interpret as that middle schools strengthen racial homophily.

the continuation value.

Third, unobservable tastes for the program characteristics are serially correlated, which implies that students select into middle and high schools based on unobservable tastes to the researcher. The estimates imply that 28.59% of the variation in the unobservable taste over the proportion of White students in high schools is explained by the consistency of the unobserved taste over the same characteristics in middle schools (18.45% for the proportion of high-performers).

4.3.2 Goodness of Fit

We evaluate how well the model fits the observed data by comparing measures calculated using the data to those calculated using the simulations based on model estimates. In Table 6, we calculate the average characteristics of assigned programs for each type of student and the average characteristics of assigned students for each type of school.

Table 6: Goodness of Fit

Panel A. Average Characteristics of Assigned Programs by Student Type

	Middle Schools				High Schools			
	High-Performers (%)		Black/Hispanic (%)		High-Performers (%)		Black/Hispanic (%)	
	Data	Model	Data	Model	Data	Model	Data	Model
Asian	36.3	37.5	39.8	39.4	32.7	32.6	42.7	42.8
Black	26.9	31.6	65.8	59.6	24.6	24.3	62.4	62.9
Hispanic	31.4	33.7	52.5	49.4	28.6	27.2	52.3	55.4
White	45.2	44.7	23.7	25.2	38.6	37.2	30.0	33.4
English Language Learner	26.5	30.2	58.9	53.4	24.2	23.7	63.0	62.2
Free/Reduced-price Lunch	34.9	36.6	44.9	43.1	30.6	29.6	47.5	49.8

Panel B. Average Characteristics of Assigned Students by School Type

	Middle Schools				High Schools			
	High-Achievement		High-Minority		High-Achievement		High-Minority	
	Data	Model	Data	Model	Data	Model	Data	Model
Asian (%)	9.3	8.9	8.8	8.8	8.7	9.3	7.0	8.0
Black (%)	3.7	3.3	24.8	23.8	4.3	3.5	29.9	22.8
Hispanic (%)	12.3	11.7	40.9	39.4	17.9	14.9	42.0	40.4
White (%)	73.9	75.2	24.6	27.0	68.2	71.4	19.7	27.6
English Language Learner (%)	1.7	1.3	9.9	8.6	2.7	2.6	11.5	9.3
Free/Reduced-price Lunch (%)	41.1	39.6	76.5	73.5	46.3	43.7	78.1	74.1
5th Grade Math Score	322.6	322.6	304.2	307.5	320.0	322.3	301.9	303.7
From High-Achievement MS (%)					56.8	60.6	10.4	9.4
From High-Minority MS (%)					10.5	10.4	62.2	48.5

Note: For model based simulations, we report the average result from 5,000 DA simulations based on the model estimates (100 draws of unobservables × 50 draws of lotteries). The definitions of 'high-achievement' and 'high-minority' are as described in Section 2.2. The scale of 5th grade math score is from 125 to 402.

We find the measures based on model simulations well match those based on the observed data, and hence, our dynamic model can be credibly used to predict the impacts of

Table 5: Preference Estimates

	Middle Schools		High Schools	
	est	se	est	se
Panel A: Preference Estimates				
Proportion of High-Performer				
Main Effect	4.944	(1.144)	0.795	(0.272)
Asian	-1.267	(1.947)	0.827	(0.390)
Black	6.820	(1.961)	-0.199	(0.462)
Hispanic	1.781	(1.288)	-0.275	(0.330)
Free/Reduced-price Lunch	-0.881	(1.130)	-0.922	(0.271)
English Language Learner	-1.804	(2.309)	0.342	(1.177)
5th Grade Test Score	1.088	(0.581)	1.652	(0.141)
Proportion of White				
Main Effect	3.056	(0.875)	4.931	(0.343)
Asian	0.976	(1.588)	-2.011	(0.599)
Black	-6.444	(1.721)	-1.520	(0.613)
Hispanic	-1.666	(1.047)	-1.060	(0.421)
Free/Reduced-price Lunch	-0.565	(0.886)	0.162	(0.346)
English Language Learner	0.752	(1.954)	-0.24	(1.202)
5th Grade Test Score	-0.951	(0.468)	0.341	(0.126)
1(STEM)				
Main Effect	0.281	(0.198)	-0.676	(0.123)
Asian	0.157	(0.324)	-0.174	(0.200)
Black	-0.420	(0.269)	0.090	(0.196)
Hispanic	0.121	(0.213)	0.083	(0.144)
Free/Reduced-price Lunch	-0.122	(0.198)	0.257	(0.126)
English Language Learner	0.062	(0.345)	1.005	(0.326)
5th Grade Test Score	-0.159	(0.096)	0.003	(0.044)
Panel B: Middle School Type Effects				
Type 1 (High-Achievement MS)				
Proportion of High-Performer			0.546	(0.276)
Proportion of White			1.600	(0.318)
1(STEM)			-0.322	(0.137)
Type 2 (High-Minority MS)				
Proportion of High-Performer			0.875	(0.301)
Proportion of White			-1.447	(0.378)
1(STEM)			0.198	(0.136)
Panel C: Unobservable Tastes				
ρ_0			0.074	(0.044)
			0.429	(0.127)
			-0.035	(0.118)
(1,1) of Σ_γ	18.461	(10.853)		
(1,2)	-17.930	(9.653)		
(1,3)	-0.186	(1.626)		
(2,2)	23.168	(10.222)		
(2,3)	2.765	(2.018)		
(3,3)	1.163	(0.697)		
(1,1) of Σ_ξ			0.447	(0.316)
(1,2)			-2.184	(0.950)
(1,3)			0.411	(0.163)
(2,2)			10.670	(2.877)
(2,3)			-2.006	(0.512)
(3,3)			0.377	(0.193)
Panel D: Other Parameters				
Outside option	2.698	(0.367)	-0.371	(0.175)
Distance	0.655	(0.038)	0.509	(0.018)
Discount Factor	0.877	(0.064)		

Note: We report the preference estimates of the main model described in Section 4. School characteristics 'Proportion of High-Performer' and 'Proportion of White' are between 0 and 1, and '1(STEM)' is an indicator variable. In Panel A, Main Effect is the common taste (β_0^M, β_0^H), and we also include interactions of each school characteristics with Asian, Black, Hispanic, Free/Reduced-price Lunch (FRL) status, English Language Learner (ELL) status, 5th Grade Test Score in the following rows (β_Z^M, β_Z^H). Robust standard errors are reported in parentheses.

counterfactual policies in Section 4.5.

First, in Panel A, we find that the average characteristics of the assigned schools for each type of student are very similar across data and model simulations for both middle and high schools. For example, in the data, Asian students are assigned to middle schools with the proportion of high-performers equal to 36.3% and the proportion of Black/Hispanic students equal to 39.8% on average. Using the model estimates, we predict such students are on average assigned to middle schools with 37.5% in terms of the proportion of high-performers and 39.4% in terms of the proportion of Black/Hispanic students. Second, in Panel B, the distributions of student observable characteristics at each type of school are also very similar across data and model simulations. For example, our model almost perfectly predicts the racial composition of each type of middle and high school. Importantly, in the last two rows of Panel B, our model predicts the transition from each type of middle school to each type of high school reasonably well.

4.4 Decomposition of Effects of Middle Schools

Recall that the model allows two channels of middle school effects on high school assignments: the *application* channel and the *priority* channel. To see the relative importance of the two, we perform the following illustrative exercise: what happens if we exogenously make a student currently attending a 'bad' middle school attend a 'good' middle school?

To this end, we randomly select students and counterfactually assign them to a 'bad' middle 'school B' with the lowest average test score in SI as a benchmark. Next, for each student, we counterfactually change their middle school enrollment to a 'good' middle 'school G' with the highest average test score in SI, *one student at a time*. We simulate their high school assignments using the model estimates in the following alternative scenarios.

1. **Full**: both *application* and *priority* channels are active.

2. **Application**: shut down the *priority* channel. That is, we do not allow a student's priorities at each high school to change depending on the middle school she attends.

3. **Priority**: shut down the *application* channel. That is, we do not allow a student's tastes for high school programs to change depending on the middle school she attends.

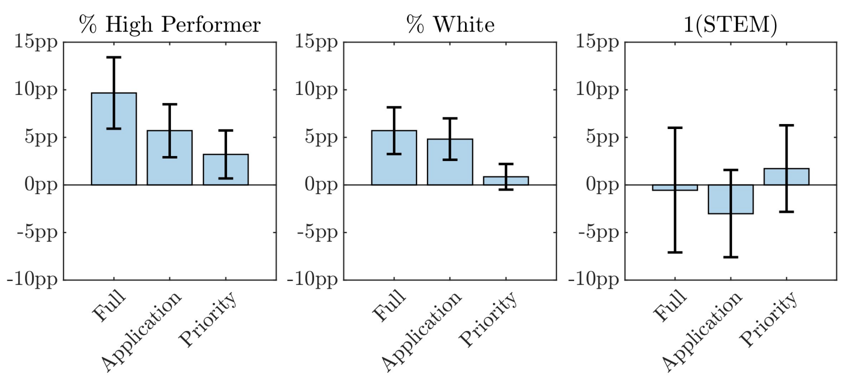
We keep track of how the students' high school assignments change compared to when they attend middle school B in each scenario. We first evaluate the effect in **Full** (the total effect of exogenously changing middle schools) and then investigate to what extent that effect can be explained by the *application* channel (**Application**) or by the *priority* channel (**Priority**). This procedure treats each student essentially as a 'price-taker' who takes the current equilibrium as given and considers how her high school assignment will change when *only* her middle school changes. Also, randomly selecting students and exogenously assigning them to a benchmark school enable us to be free of sorting of students into middle schools based on unobservables. Note that in these regards, the measures we report have an interpretation as the average treatment effect (ATE) of changing middle schools.

Table 7: Alternative Assignment to Middle Schools

	Middle School Types		Average Test Score
	High-Achievement?	High-Minority?	
Middle School B (Lowest-Performing)		✓	602.01
Middle School G (Highest-Performing)	✓	✓	611.42

Note: Average test scores are the average of 8th grade statewide test scores of current seniors (scale: 500 to 650).

Figure 1: Decomposition of Effects of Middle Schools on High School Assignments



Note: We report the decomposition of middle school effects on high school assignments using the model estimates in Table 5. We use a randomly selected subsample of student (10% of the entire sample), and counterfactually assign them to middle school B. Then we calculate the average change in the characteristics of the assigned high school program when they are counterfactually assigned to middle school G. 100 sets of unobservable variables $(\gamma_i^M, \xi_i, \epsilon_{im}, \eta_{ij})$ are drawn and for each set, 50 sets of tie-breaking lotteries are drawn and DA are run, resulting in 5,000 simulated assignments. The bar graphs and 95% confidence intervals are plotted using the average (across unobservables and lottery draws) of mean and standard deviation across students. The corresponding numbers are reported in Table H.11

Figure 1 reports the results. We find that the *application* channel is quantitatively more important than the *priority* channel. For example, when a student attends a 'good' middle school G instead of a ‘bad’ middle school B, we find that about 59% of the total change in the assigned high school’s proportion of high-performers can be explained by the *application* channel. In comparison, the *priority* channel can only explain about 33%. Regarding the proportion of White students, the *application* channel explains about 84%, and the *priority* channel explains about 15% of the total effect.

The decomposition exercise reconfirms that middle schools play an important role in high school choice outcomes and that the effect mainly occurs by affecting how students apply to high schools. Importantly, together with the preference estimates, our analysis shows that first, students’ middle schools affect their tastes for high schools in a way that strengthens segregation (Table 5), and second, these taste changes have an actual impact on students’ high school assignments through the change in their applications (Figure 1), which may affect high school segregation. However, at the same time, the result also motivates the possibility of addressing high school segregation by changing students’ middle school assignments and hence their high school applications, which is explored in the next section.

4.5 Counterfactual Analysis

4.5.1 Segregation in NYC Public High Schools

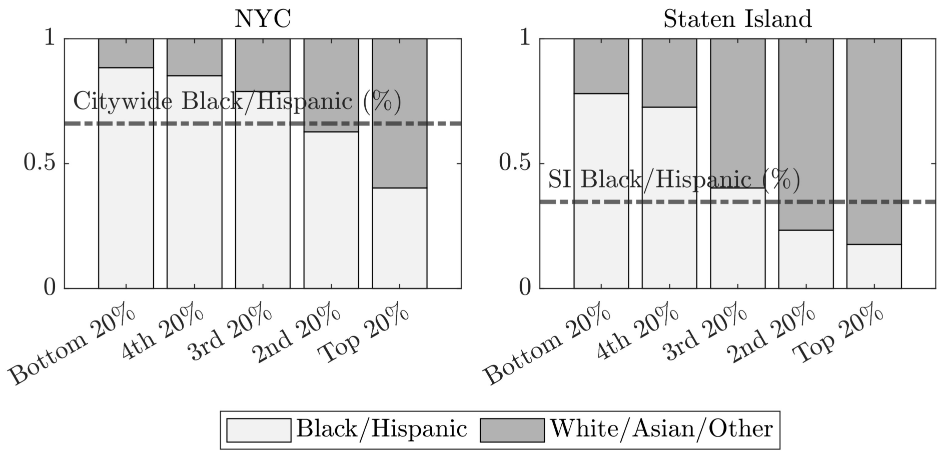
NYC high schools are intensely segregated. Figure 2 plots the racial composition of high schools by quintiles of the average performance of enrolled students. It shows that Black and Hispanic students are underrepresented at high-performing high schools, while they are overrepresented in low-performing high school programs.

The NYC government has long acknowledged this problem. Most recently, partially due to the cancellation of statewide exams due to COVID-19 during 2020 and onwards, the city announced changes to the public school system to deal with racial segregation.

> Mayor Bill de Blasio announced on Friday major changes to the way hundreds of New York City’s selective middle and high schools admit their students. [...] Black and Latino students are significantly underrepresented in selective middle and high schools. [...] The city will eliminate all admissions screening for the schools for at least one year [...] New York will also eliminate a policy that allowed some high schools to give students who live nearby first dibs at spots.
>
> — The New York Times[39]

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

Figure 2: High School Racial Composition by Performance Level



Note: We plot the racial composition of high school programs by the quintiles of average performance level of students. For a given program, the proportion of 9th graders in AY 2017-18 whose average of 8th grade statewide Math and ELA scores fall in the first tercile are calculated, and then the high school programs are classified into each quintile of it. The left panel uses the entire NYC high schools, and the right panel uses the SI high schools. The overall proportion of Black/Hispanic students in each population is plotted in the gray dotted line.

Motivated by NYC's policies, we perform the following counterfactual analysis. Namely, we combine both policies—to eliminate (i) screening based on test scores and (ii) any form of geography-based priority rules, including zoned programs.[40] We evaluate the following three alternative interventions:

1. **MS**: we only get rid of priority rules of middle schools.

2. **HS**: we only get rid of priority rules of high schools.

3. **MSHS:** we get rid of priority rules of both middle and high schools.

In each scenario, we solve the new equilibrium using the model estimates and compare how students' high school assignments change compared to the status quo (**Current**).

Given the importance of middle schools in how students apply and are assigned to high schools, we have two conjectures. First, under **MS**, high school assignments will change through the *application* and *priority* channels induced by the change in middle school

[40]In SI, this amounts to removing priority rules altogether so that schools admit students solely based on lotteries. Hence, assignments are entirely decided by how students apply to schools. This choice of policy intervention highlights the role of how students submit their choices, which is closely related to the main findings of this paper that middle schools mainly affect how students submit choices in subsequent school choices.

assignments. Second, the desegregation effects on high schools will be larger under **MSHS** than under **HS** since **HS** reforms only the 'supply' side (i.e., how high schools select students) while **MSHS** reforms not only the supply side but also the 'demand' side (i.e., how students apply to high schools).[41]

4.5.2 Desegregating Effects of Counterfactual Policies

We evaluate the impacts of counterfactual policy changes along two dimensions: the characteristics of co-assigned peers for minority students and overall segregation measures. We assume that the school characteristics are fixed as under the status quo (**Current**), which gives us the interpretation of the predictions as the short-term impacts.

Effects on Minority Students' Co-assigned Peers Figure 3 plots the relative difference in the average characteristics of co-assigned peers of Black/Hispanic students to those of White/Asian students.[42] We focus on the effects of middle school admission reforms on desegregating high schools.

First of all, notice that middle school-only admissions reform can desegregate middle schools effectively (Panel A). For example, on average, Black/Hispanic students are assigned to middle school programs with 30.1 pp higher in the proportion of Black/Hispanic students than White/Asian students, and middle school-only intervention (**MS**) can undo this gap by 12.7 pp.

Most importantly, this desegregation of middle schools in **MS** leads to desegregating high schools through the *application* and *priority* channels, confirming our conjectures.

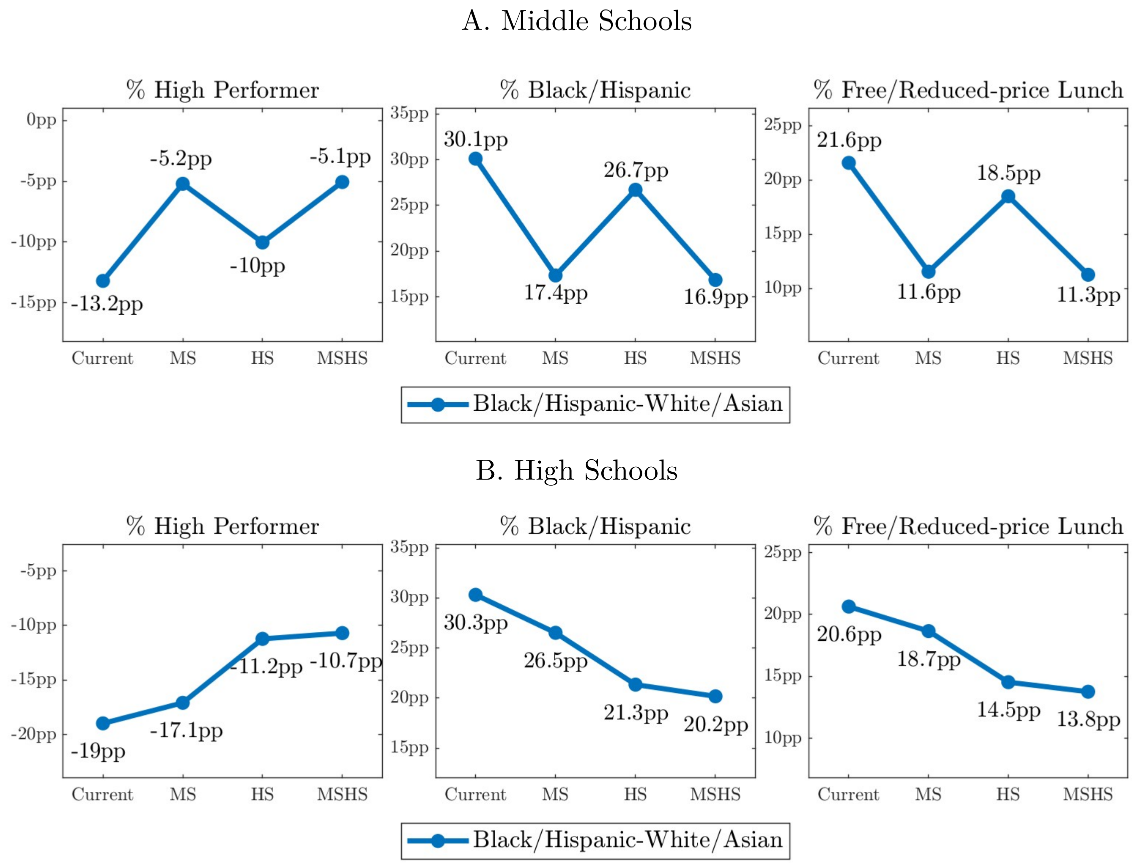
First, we find that intervening only at the middle school level (**MS**) alone can reduce the racial gap in the average characteristics of co-assigned peers in high schools (Panel B). For example, on average, Black/Hispanic students are assigned to high school programs with 30.3 pp higher in the proportion of Black/Hispanic students than White/Asian students, and middle school-only intervention (**MS**) can undo this gap by 3.8 pp, which amounts to nearly 42% of what high school-only intervention (**HS**) can achieve.

Second, the effects of combining both interventions at the middle school level and the high school level (**MSHS**) are larger for high school assignments than only intervening at

[41]One would also expect that due to the forward-looking behavior of students in our model, middle school assignments may change even under **HS**, which is what we find.

[42]See Figure H.5 for the effects on the average characteristics of assigned schools.

Figure 3: Racial Gap in Co-assigned Peers in Staten Island



Note: The graph plots the gap of the characteristics of co-assigned peers between Black/Hispanic versus White/Asian students in each counterfactual scenario. 100 sets of unobservable variables $(\gamma_i^M, \xi_i, \epsilon_{im}, \eta_{ij})$ are drawn and for each set, 50 sets of tie-breaking lotteries are drawn and DA are run, resulting in 5,000 simulated assignments for each counterfactual scenario where the draws are fixed across scenarios. The mean across unobservable and lottery draws are reported for **MS**, **HS** and **MSHS**. We use the observed assignment results in the data for **Current**.

the high school level (**HS**) (Panel B).[43]

Effects on Aggregate Segregation Measures We calculate two measures in order to summarize the aggregate segregation pattern.[44] First, we calculate a measure of racial segregation known as the Theil's H index in Panel A of Table 8. The index calculates a measure of the evenness of ethnic groups across programs based on multigroup entropy scores. It varies between 0 and 1, where 0 means maximum integration and 1 means maximum segregation.

Second, in Panel B of Table 8, we calculate the sorting indices for three student characteristics: 1(Black/Hispanic), baseline standardized test scores, and the neighborhood median income. Each sorting index is between 0 and 1, and is defined as the ratio of the between-program variance of each student characteristic to its total variance (Yang and Jargowsky, 2006; He, Sinha, and Sun, 2021). That is, it measures the fraction of variance of a variable that between-program differences can explain. Hence, 0 means maximum integration and 1 means maximum segregation. Sorting by 1(Black/Hispanic) provides a measure of segregation by race, sorting by test scores of students provides a proxy of sorting by student ability, and sorting by median census tract income provides a proxy of sorting by income.
We find similar patterns as in the effects on minority students' assignments.

Policy Implication Our counterfactual analysis emphasizes the importance of considering the dynamics of school choice in addressing segregation. While most existing policies for desegregation focus on reforming the *supply side*, i.e., how schools select students, it is crucial to consider how we can influence the *demand side* i.e., how students apply to schools. We found in Sections 3 and 4 that students' high school assignments are largely affected by which middle schools they attend, mainly by changing their applications to high schools. Also, the counterfactual analysis showed that intervening in middle schools alone can help desegregate not only middle schools but also high schools. In addition, conditional on intervening at the high school level, there is still room to further desegregate high schools by additionally intervening at the middle school level. Taken together, our findings imply that large school districts can design a more effective school desegregation policy by leveraging that intervention

[43] However the marginal gain of **MS**→**MSHS** (**HS**→**MSHS**) is smaller than that of **Current**→**HS** (**Current**→**MS**), suggesting a possible substitutability of those **MS** and **HS**.

[44] See Appendix G for more details on the description of measures.

Table 8: Aggregate Segregation Measures in Staten Island

		(1) Current	(2) MS	(3) HS	(4) MSHS
Panel A: Racial Segregation Measure					
Theil's H Index	Middle schools	0.216	0.106	0.189	0.102
	High schools	0.207	0.191	0.148	0.135
Panel B: Sorting Indices					
Sorting by Race	Middle schools	0.299	0.173	0.266	0.168
	High schools	0.301	0.263	0.212	0.201
Sorting by Ability	Middle schools	0.162	0.040	0.075	0.037
	High schools	0.357	0.309	0.119	0.117
Sorting by Income	Middle schools	0.456	0.237	0.432	0.230
	High schools	0.346	0.300	0.262	0.242

Note: The table calculates the aggregate segregation measures of schools in each counterfactual scenario. Panel A calculates the Theil's H index, and Panel B calculates the sorting indices by race, ability and income. 100 sets of unobservable variables $(\gamma_i^M, \xi_i, \epsilon_{im}, \eta_{ij})$ are drawn and for each set, 50 sets of tie-breaking lotteries are drawn and DA are run, resulting in 5,000 simulated assignments for each counterfactual scenario where the draws are fixed across scenarios. The mean across unobservable and lottery draws are reported for **MS**, **HS** and **MSHS**. We use the observed assignment results in the data for **Current**.

on the *supply side* of an earlier school choice induces changes in the *demand side* of the subsequent school choice stages.

5 Conclusion

This paper provides a novel, evidence-based dynamic framework of school choices. We show that a student's middle and high school choices are closely related to each other by using student-level panel school choice data from NYC. First, we leverage the quasi-random middle school assignments generated by the tie-breaking feature in DA to provide empirical evidence of middle schools' causal effects on high school applications and assignments. Next, based on the empirical findings, we develop and estimate a dynamic framework of middle and high school choices. We show that the effects of middle schools on high school choice mainly operate by changing students' applications to high schools rather than changing high schools' ranking over students. Finally, we provide a new perspective on understanding and addressing segregation across public schools using the dynamic framework. Segregation patterns in middle and high schools are closely related, and hence the policy intervention for desegregating high schools should begin early enough, and reforming middle school admissions may be one such tool.

Our findings suggest two avenues for future research. First, having confirmed the dynamic relationship between middle school choice and high school choice, we may further move on to directly test for the dynamic complementarity of those two human capital investments (Cunha and Heckman, 2007; Heckman, 2007). While a credible quasi-randomization at two points for a given individual is hard to find, the fact that students are exposed to centralized school choice multiple times opens an avenue for a suitable research design to test for dynamic complementarity. Second, given the importance of the dynamic relationship of school choices, we should consider it in designing assignment mechanisms. For example, one may explore ways to design a student assignment mechanism that considers the dynamic relationship of school choices to achieve more equitable outcomes. We leave these for future research.

References

Abdulkadiroğlu, A., N. Agarwal, and P. A. Pathak (2017): "The Welfare Effects of Coordinated Assignment: Evidence from the NYC HS Match," American Economic Review, 107(12), 3635–89. 6, 18, 20, 66

Abdulkadiroğlu, A., J. D. Angrist, Y. Narita, and P. A. Pathak (2017): "Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation," Econometrica, 85(5), 1373–1432. 6, 7, 13, 48

——— (2021): "Breaking Ties: Regression Discontinuity Design Meets Market Design," Econometrica, Forthcoming. 3, 6, 7, 13, 14, 48, 49

Abdulkadiroğlu, A., Y.-K. Che, and Y. Yasuda (2015): "Expanding "Choice" in School Choice," American Economic Journal: Microeconomics, 7(1), 1–42. 6

Abdulkadiroğlu, A., P. A. Pathak, and A. E. Roth (2005): "The New York City High School Match," American Economic Review Papers and Proceedings, 95, 364–367. 7

——— (2009): "Strategy-Proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match," American Economic Review, 99(5), 1954–1978. 7

Abdulkadiroğlu, A., and T. Sonmez (2003): "School Choice: A Mechanism Design Approach," American Economic Review, 93, 729–747. 45

Abdulkadiroğlu, A., J. Angrist, and P. Pathak (2014): "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools," Econometrica, 82(1), 137–196. 5, 6

Agarwal, N., I. Ashlagi, M. A. Rees, P. Somaini, and D. Waldinger (2021): "Equilibrium allocations under alternative waitlist designs: Evidence from deceased donor kidneys," Econometrica, 89(1), 37–76. 6

Agarwal, N., I. Ashlagi, P. Somaini, and D. Waldinger (2018): "Dynamic Incentives in Wait List Mechanisms," in AEA Papers and Proceedings, vol. 108, pp. 341–47. 6

Agarwal, N., and P. Somaini (2018): "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism," Econometrica, 86(2), 391–444. 6, 23

Ajayi, K. F., W. H. Friedman, and A. M. Lucas (2017): "The Importance of Information Targeting for School Choice," American Economic Review, 107(5), 638–643. 6

Allende, C. (2019): "Competition under social interactions and the design of education policies," Job Market Paper. 24

Artemov, G., Y.-K. Che, and Y. He (2021): “Strategic ’Mistakes’: Implications for Market Design Research,” Working Paper. 20

Azevedo, E. M., and J. D. Leshno (2016): “A Supply and Demand Framework for Two-Sided Matching Markets,” Journal of Political Economy, 124, 1235–1268. 48

Betts, J. R. (1995): “Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth,” The Review of Economics and Statistics, 77(2), 231–250. 6

——— (2011): “Chapter 7 - The Economics of Tracking in Education,” vol. 3, pp. 341–381. Elsevier. 4

Calsamiglia, C., C. Fu, and M. Güell (2020): “Structural estimation of a model of school choices: The boston mechanism versus its alternatives,” Journal of Political Economy, 128(2), 642–680. 6, 48

Calsamiglia, C., G. Haeringer, and F. Klijn (2010): “Constrained School Choice: An Experimental Study,” American Economic Review, 100(4), 1860–1874. 8

Card, D., and A. B. Krueger (1992a): “Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States,” Journal of Political Economy, 100(1), 1–40. 6

Card, D., and A. B. Krueger (1992b): “School Quality and Black-White Relative Earnings: A Direct Assessment,” The Quarterly Journal of Economics, 107(1), 151–200. 6

Che, Y.-K., D. W. Hahm, and Y. He (2021): “Leveraging Uncertainties to Infer Preferences: Robust Analysis of School Choice,” Working Paper. 19, 20

Che, Y.-K., and O. Tercieux (2019): “Efficiency and Stability in Large Matching Markets,” Journal of Political Economy, 127(5), 2301–2342. 6

Chen, Y., and Y. He (2021a): “Information acquisition and provision in school choice: a theoretical investigation,” Economic Theory. 6

Chen, Y., and Y. He (2021b): “Information acquisition and provision in school choice: an experimental study,” Journal of Economic Theory, Forthcoming. 6

Clark, D., and E. D. Bono (2016): “The Long-Run Effects of Attending an Elite School: Evidence from the United Kingdom,” American Economic Journal: Applied Economics, 8(1), 150–176. 6

Corcoran, S. P., J. Jennings, S. Cohodes, and C. Sattin-Bajaj (2018): “Leveling the Playing Field for High School Choice: Results from a Field Experiment of Informational Interventions,” Discussion paper. 6, 17

Corcoran, S. P., and H. M. Levin (2011): “School Choice and Competition in the New York City Schools,” in Education Reform in New York City: Ambitious Change in the Nation’s Most Complex School System, ed. by J. O’Day, C. Bitter, and L. M. Gomez. Harvard Education Press. 46

Cullen, J. B., B. A. Jacob, and S. Levitt (2006): “The Effect of School Choice on Participants: Evidence from Randomized Lotteries,” Econometrica, 74(5), 1191–1230. 6

Cunha, F., and J. Heckman (2007): “The Technology of Skill Formation,” American Economic Review, 97(2), 31–47. 38

Dehejia, R. H., and S. Wahba (2002): “Propensity score-matching methods for nonexperimental causal studies,” Review of Economics and statistics, 84(1), 151–161. 48

Deming, D. J., J. S. Hastings, T. J. Kane, and D. O. Staiger (2014): “School Choice, School Quality, and Postsecondary Attainment,” American Economic Review, 104(3), 991–1013. 5, 6

Dobbie, W., and R. G. Fryer (2014): “The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools,” American Economic Journal: Applied Economics, 6(3), 58–75. 5, 6

Ellison, G., and P. A. Pathak (2021): “The Efficiency of Race-Neutral Alternatives to Race-Based Affirmative Action: Evidence from Chicago’s Exam Schools,” American Economic Review, 111(3), 943–975. 5

Fack, G., J. Grenet, and Y. He (2019): “Beyond Truth-Telling: Preference Estimation with Centralized School Choice and College Admissions,” American Economics Review, 109(4), 1486–1529. 48, 55

Gale, D., and L. S. Shapley (1962): “College Admissions and the Stability of Marriage,” American Mathematical Monthly, 69, 9–15. 2, 7, 45

Grenet, J., Y. H. He, and D. Kübler (2021): “Decentralizing Centralized MatchingMarkets: Implications from Early Offers in University Admissions,” Working Paper. 6

Haeringer, G., and F. Klijn (2009): “Constrained school choice,” Journal of Economic Theory, 144(5), 1921–1947. 8

Hahn, J., P. Todd, and W. Van der Klaauw (2001): “Identification and estimation of treatment effects with a regression-discontinuity design,” Econometrica, 69(1), 201–209. 13

Hassidim, A., A. Romm, and R. I. Shorrer (2016): “” Strategic” Behavior in a Strategy-Proof Environment,” in Proceedings of the 2016 ACM Conference on Economics and Computation, pp. 763–764. 20

HASTINGS, J. S., T. KANE, AND D. STAIGER (2005): "Parental Preferences and School Competition: Evidence from a Public School Choice Program," Discussion paper. 18, 66

HASTINGS, J. S., AND J. M. WEINSTEIN (2008): "Information, School Choice, and Academic Achievement: Evidence from Two Experiments," The Quarterly Journal of Economics, 123(4), 1373–1414. 5, 6

HE, Y. (2017): "Gaming the Boston School Choice Mechanism in Beijing," . 6

HE, Y., S. SINHA, AND X. SUN (2021): "Identification and Estimation in Many-to-one Two-sided Matching without Transfers," arXiv preprint arXiv:2104.02009. 36

HECKMAN, J. J. (2007): "The economics, technology, and neuroscience of human capability formation," Proceedings of the National Academy of Sciences, 104(33), 13250–13255. 38

HEISS, F., AND V. WINSCHEL (2008): "Likelihood approximation by numerical integration on sparse grids," Journal of Econometrics, 144(1), 62–80. 26, 53

HOEKSTRA, M. (2009): "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach," Review of Economics and Statistics, 91(4), 717–724. 6

HOXBY, C. M., AND J. E. ROCKOFF (2004): "The Impact of Charter Schools on Student Achievement," Working Paper. 6

HOXBY, C. M., AND S. TURNER (2015): "What High-Achieving Low-Income Students Know About College," American Economic Review, 105(5), 514–517. 6

ICELAND, J. (2004): "The multigroup entropy index (also known as Theil's H or the information theory index)," US Census Bureau. Retrieved July, 31, 2006. 60

IDOUX, C. (2022): "Integrating New York City Schools: The Role of Admission Criteria and Family Preferences," Working Paper. 5

IMBENS, G., AND K. KALYANARAMAN (2012): "Optimal bandwidth choice for the regression discontinuity estimator," The Review of economic studies, 79(3), 933–959. 14

JACKSON, K. C. (2010): "Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago," The Economic Journal, 120(549), 1399–1429. 5

LARROUCAU, T., AND I. RIOS (2020): "Dynamic college admissions and the determinants of students' college retention," . 6

LAVERDE, M. (2022): "Distance to Schools and Equal Access in School Choice Systems," . 5

LI, S. (2017): "Obviously Strategy-Proof Mechanisms," American Economic Review, 107(11), 3257–3287. 20

Luflade, M. (2017): "The value of information in centralized school choice systems," Working Paper. 6, 24

Mark, N. D., S. P. Corcoran, and J. Jennings (2021): "Choosing Alone? Peer Similarity in High School Choices," Working Paper, (396). 6

Neilson, C., C. Allende, and F. Gallego (2019): "Approximating the Equilibrium Effects of Informed School Choice," Discussion paper. 24

Pop-Eleches, C., and M. Urquiola (2013): "Going to a Better School: Effects and Behavioral Responses," American Economic Review, 103(4), 1289–1324. 5, 6

Rosenbaum, P. R., and D. B. Rubin (1983): "The central role of the propensity score in observational studies for causal effects," Biometrika, 70(1), 41–55. 13

Smith, J. A., and P. E. Todd (2005): "Does matching overcome LaLonde's critique of nonexperimental estimators?," Journal of econometrics, 125(1-2), 305–353. 48

Son, S. J. (2020): "Distributional impacts of centralized school choice," Discussion paper. 24

Waldinger, D. (2021): "Targeting in-kind transfers through market design: A revealed preference analysis of public housing allocation," American Economic Review, 111(8), 2660–96. 6

Yang, R., and P. A. Jargowsky (2006): "Suburban development and economic segregation in the 1990s," Journal of Urban Affairs, 28(3), 253–273. 36

Zhang, J. (2010): "The sound of silence: Observational learning in the US kidney market," Marketing Science, 29(2), 315–335. 6

Appendix to

A Dynamic Framework of School Choice: Effects of Middle Schools on High School Choice

Dong Woo Hahm Minseon Park

July 2022

List of Appendices

A Details of NYC School Choice Process

A.1 Student-Proposing Deferred Acceptance Algorithm

In detail, DA works as follows (Gale and Shapley, 1962; Abdulkadiroğlu and Sonmez, 2003):

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

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

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

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

A.2 NYC School Admission Methods

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

High school programs use similar admission methods as middle schools—Unscreened, Limited Unscreened, Screened, Screened: Language, Screened: Language & Academics, Zoned, Audition, Educational Option, and Continuing 8th Graders. Audition programs are similar to Talent Test middle school programs, and Educational Option is a mixture of

Unscreened and Screened.[A-1] Continuing 8th Graders programs are open only to continuing eighth graders in the same school. Other admission methods are similar to middle school choice.

A.3 The Timeline of Admission Process

The timeline of the admission process is as follows (Corcoran and Levin (2011), Directory of NYC Public High Schools). By December, students are required to submit their ROLs. By March, DA algorithms are run which determine students' assignments. Students who accept their offer finalize, and if a student rejects an offer, then she goes to the next round. This describes the main round of the entire system. A majority of students finalize in the main round (about 85% each year). Students who are not assigned in the main round or rejected the assignment go to the Supplementary round, which is similarly organized to the main round and includes school-programs that did not fill up their capacities in the main round, or programs that are newly opened. Finally, there is an administrative round in which students who are not assigned a school even after the second round are administratively assigned to a school.

B Data and Sample Restriction

B.1 Data Sources

The main data used is the administrative data acquired from the New York City Department of Education, focusing on the 8th grade cohort in the academic year 2017-2018. This cohort applied to middle schools in the academic year 2014-15, and to high schools in the academic year 2017-18.

There are four sets of data used to construct information on students. First, high school application (HSAP) data includes information on each round of the application process (ROL, rank, priority, eligibility, assignment, etc.) related to high school application and standardized test scores information. Second, middle school application (MSAP) data includes similar variables as HSAP but for middle school applications. Third, yearly June biographic data includes more comprehensive biographic data of the students, including ethnicity, gender,

[A-1]Educational Option programs have the purpose of serving students at diverse academic performance levels. These programs divide students into high (highest 16%), middle (68%) and low ELA (lowest 16%) levels. 50% of the seats in each group are filled using school-specific criteria like a screened program and the other 50% are filled randomly similarly as an unscreened program. (NYC DOE Introduction to High School Admissions) disability status, as well as information on attendance and punctuality. Lastly, Zoned DBN data includes information about students' residence (census tract level)[A-2] and which elementary, middle, and high schools the students are zoned to. We merge all data sets using a unique student ID.

School information is constructed using the 2014-15 NYC Middle School Directory and 2017-18 NYC High School Directory that are published every year before the application process starts. The School Directory includes each program's previous year's capacity and the number of students who applied in the previous year, admission criteria (eligibility and priority), accountability data such as progress report, graduation rate, college enrollment rate, and types of language classes provided, etc. Other variables about current 6th graders in middle schools and 9th graders in high schools, such as the composition of ethnicity or the proportion of high-performing students are constructed using the previous year's student-level data.

B.2 Sample Restriction

We start with 72,318 observations in the middle school application data. Out of 72,318, 67,153 students participated in the main round of the middle school application. We drop students with missing demographic characteristics or invalid standardized test scores, and are left with 62,972 students. Among the remaining students, 54,012 students participated in high school application after three years.[A-3] We present summary statistics and balance test results of these 54,012 students in Section 2.[A-4] For new middle and high schools, school characteristics are missing. After excluding students who went to a new middle school and whose high school ranked ordered list is filled only with new high schools, we have 44,237 students. The estimates in Table 3 are based on this sample.

[A-2] In the current data set, the finest level of geographic information of a student is census-tract level. The distance between students and schools is calculated as follows. For each census tract in NYC, we use the latitude and longitude coordinates of the centroid from corresponding year's US Census gazetteer file. School's coordinates are calculated using their exact street addresses with Google API. Next we calculate the distance between the coordinates of the exact school location and students' census tract of residence centorid based on the Haversine formula.

[A-3] Those who participated in the middle school choice but not participated in the high school choice do not appear in the data afterwards. Examples might include drop-outs, those who attend private or charter high schools, and those who moved out of NYC. These are more likely to be low-performers, subsidized lunch status, or Black students.

[A-4] 801 students applied only to new middle schools on which there is no characteristics of the previous cohort. We present summary statistics and balance test results on middle school application behavior for the rest (n=53,211).

C An Example of Calculating Propensity Scores

The following example illustrates how to calculate the propensity scores (=admission probabilities) following Abdulkadiroğlu, Angrist, Narita, and Pathak (2017, 2021).[A-5]

Consider student i who submits a rank-ordered list A-B-C where A is her most preferred option and C is her least preferred option. Priority score used for admissions is a sum of priority group and a tie-breaker, where priority group lexicographically dominates tie-breakers. That is, student i's score at program j is

$$\text{score}_{ij} = \underbrace{\text{PG}_{ij}}_{\text{priority group}\in\mathbb{N}} + \underbrace{\text{TB}_{ij}}_{\text{tie-breaker}\in[0,1]}$$

where i has higher priority than i' at j if and only if $\text{score}_{ij} > \text{score}_{i'j}$. Programs A and B share a random tie-breaker $\text{TB}_{iA} = \text{TB}_{iB} \overset{iid}{\sim} U[0,1]$, and programs C uses a non-random tie-breaker $\text{TB}_{iC} \sim F_i$, where F_i is unknown and potentially depends on the student and has a support on [0,1]. A cutoff of program j is given by the minimum of scores of admitted students at j if all seats are filled, and $-\infty$ if some seats are left unfilled. Let us assume a large market (Azevedo and Leshno, 2016; Fack, Grenet, and He, 2019; Calsamiglia, Fu, and Güell, 2020) and denote each program's degenerate large market cutoff by cutoff_j. Student i is admitted to program j if $\text{score}_{ij} \geq \text{cutoff}_j$ and at the same time rejected from all programs ranked above j.

Table C.1: Example of Propensity Score

Programs	A	B	C
PG_{ij}	1	1	2
Cutoff	2.2	1.4	2.6
Admission Prob.	0	1×0.6	$1 \times 0.4 \times (1 - F_i(0.6))$
Local Admission Prob.	0	1×0.6	$1 \times 0.4 \times \mathbf{0.5}$

Table C.1 illustrates how to calculate the propensity score for student i in this example. Student i has no chance of being admitted to program A, since no realization of the tie-breaker is large enough to clear the cutoff of program A. Next, the probability of being assigned to program B is the probability of being rejected from program A (=1) times the probability of

[A-5]Note that the propensity score in this context denotes the exact probability of being treated, and involves no prediction of the odds by estimating a logit or a probit model, which is typically found in papers with propensity score matching (for example, Dehejia and Wahba, 2002; Smith and Todd, 2005).

getting accepted to program B. The cutoff of B is 1.4, so i can be assigned to program B as long as her lottery number is greater than 0.4, which happens with a probability of 0.6. Hence, student i's admission probability at program B is $1 \times 0.6 = 0.6$. Next, i gets assigned to program C if she is rejected from all previous options (which happens with probability 1×0.4) and then clears the cutoff of program C. While it is impossible to get the exact probability of clearing the cutoff, $1 - F_i(0.6)$, Theorem 1 of Abdulkadiroğlu, Angrist, Narita, and Pathak (2021) suggest that i clears the cutoff with half chance if i's tie-breaker TB_{iC} is close enough to the cutoff. In that case, the local admission probability is given by $1 \times 0.4 \times 0.5$.

D Additional Procedures for Computation

D.1 Constructing Priority Scores and Simulating Uncertainties

Each student's priority scores are necessary to use ex-post stability. First, to calculate the continuation value, we need to simulate the set of feasible schools in each realization of ex-post cutoffs by running DA algorithm multiple times, which takes students' priorities when attending different middle school programs. Second, to interpret data as a conditional multinomial logit model, we need to construct the feasible set of programs for each student, regardless of if she ranked them or not.

In NYC, priority scores consist mainly of three ingredients: eligibility, priority group, and priority ranks at programs involving screening. First, eligibility and priority group are determined in a deterministic manner, based on the pre-announced rule in NYC Middle School Directory and NYC High School Directory published every year before public school applications.

Next, when it comes to priority ranks, while the data set includes the priority rank of applicants to each program, there is no information on the ranks of those who did not apply to that particular program. In addition, the exact formula that each program uses is not available. Therefore, we estimate the priority ranks for Screened, Screened: Language, Screened: Language & Academics, and the screened part of Education Option programs. To this end, we assume there exists a program-specific latent variable as a function of various student characteristics, which determines the rank of students at each program. Specifically, let w_{ij} be the latent variable of i at an actively ranking program j as a function of student characteristics Z_i. We assume:

$$w_{ij} = \beta_j Z_i + e_{ij}$$

and

$$i \text{ is ranked higher than } i' \text{ if and only if } w_{ij} > w_{i'j}$$

where Z_i includes standardized statewide Math and ELA exam scores; Math, Social Sciences, English, Science GPA; and days absent and days late. We assume e_{ij} is iid as EVT1. From the data, we gather all possible pairs of applicants to program j, and maximize the following likelihood:

$$\sum_{i>i',i,i'\in\mathcal{I}_j} \log\left(\frac{\exp(w_{ij})1\{i \text{ is ranked higher than } i'\} + \exp(w_{i'j})1\{i' \text{ is ranked higher than } i\}}{\exp(w_{ij}) + \exp(w_{i'j})}\right)$$

where $\mathcal{I}_j$ is the set of applicants to program j which is observed in the data. Using the estimates $\hat{\beta}_j$, we predict $\hat{w}_{ij} = \hat{\beta}_j Z_i$ for all i and reconstruct the priority ranks based on $\hat{w}_{ij}$.

Finally, we describe how to simulate the uncertainties in the economy, ω. We draw 40,000 sets of lotteries and run DA 40,000 times. To additionally account for the uncertainty in other students' types, we repeat the procedure by bootstrapping 200−1 times from the data and creating multiple economies. We use the resulting empirical distribution as the distribution of ω.

D.2 Evolution of Test Scores

Some of the student characteristics are time-invariant while others change especially as functions of middle school a student attends. In particular, a student's test scores may change depending on the middle school she attends, because different middle schools may have different effectiveness. We estimate each middle school's 'production function'[A-6] using a value-added model.

Specifically, let $y^H_{i,m}$ be the potential end-of-middle-school test score when student i attends middle school m. We assume 'selection on observables':

$$E\left[y^H_{i,m} \mid Z^M_i, m\right] = \alpha_m + Z^{M\prime}_i \beta_m, \quad m \in \mathcal{M}$$

and estimate via OLS of $y^H_{i,m(i)}$ on school indicators interacted with Z^M_i where $m(i)$ is the actual middle school attendance in the data and $y^H_{i,m(i)}$ is the observed y^H_i in the data. We include baseline test scores, sex and ethnicity dummy variables, English Language Learner

[A-6]To ensure enough sample size, we estimate the value-added of each middle school instead of middle school program.

status, disability status, and free/reduced-price lunch status in Z_i^M.

Table D.2: Mean and Standard Deviation of VA Coefficients Across Schools

	Math		ELA	
	est	se	est	se
Baseline Test Score	0.346	0.035	0.331	0.033
	(0.060)	(0.015)	(0.040)	(0.013)
Female	1.591	1.650	3.077	1.517
	(1.425)	(0.412)	(2.327)	(0.352)
Asian	6.002	3.993	6.029	3.402
	(4.892)	(2.108)	(4.617)	(1.547)
Black	-2.422	4.542	-2.502	4.642
	(6.194)	(2.527)	(3.826)	(3.216)
Hispanic	-2.309	2.708	-0.738	2.472
	(3.945)	(1.260)	(3.391)	(1.008)
English Language Learner	-2.862	5.691	1.239	6.066
	(7.230)	(2.669)	(6.273)	(3.045)
Student with Disability	-6.885	2.345	-5.571	2.212
	(3.192)	(0.690)	(2.122)	(0.663)
Free/Reduced-price Lunch	-1.380	2.264	-1.501	2.013
	(2.124)	(1.190)	(1.974)	(0.863)

Table D.2 reports the mean and standard deviations of the coefficients $\hat{\alpha}_m, \hat{\beta}_m$ and their standard errors across middle schools. First, students with higher baseline test scores tend to have higher test scores, reflecting their higher academic ability. Second, there exist significant variation across middle schools as well as heterogeneity based on student observable characteristics.

D.3 Assumptions on Unobservables and Computation of Continuation Value

Collect the first stage state variables and middle school option m in:

$$\Psi_{1i} = (Z_i^M, \gamma_i^M, \epsilon_i, m)$$

where $\epsilon_i = (\epsilon_{i1}, \cdots, \epsilon_{iJ_m})$. Ψ_{1i} contains variables conditional on which student i takes the expectation of the second period payoff in calculating the continuation value of m. Note that m is not the actual student's middle school attendance but is an exogenous option given to the student in the middle school choice.

We assume the following relationships on the unobservables.

$$\eta_{ij} \perp \epsilon_{im} \Big| \gamma_i^M, \xi_i, \quad \forall i, j, m \tag{D.1}$$

$$\xi_i, \eta_i, \Psi_{1i} \text{ are mutually independent}, \quad \forall i \tag{D.2}$$

$$\omega \perp (\xi_i, \eta_{ij}) \Big| \Psi_{1i} \text{ and } \omega \perp \Psi_{1i}, \quad \forall i, j, m \tag{D.3}$$

The first assumption states conditional on the first period unobserved taste γ_i^M and the innovation to it ξ_i, the idiosyncratic preferences in each period, ϵ_{im} and η_{ij}, are independent for all i, j, m. The second assumption states that the innovation to the unobserved tastes in the second period, the second period idiosyncratic preferences, and the first period state variables are mutually independent.[A-7]

Finally, the third assumption states that ω, the uncertainty determining high school feasibility, is independent of the unobservable tastes for high school programs and the idiosyncratic preferences in the second period, conditional on the state variables in the first period and middle school attendance m. In addition, ω is independent of the state variables in the first period and middle school attendance m. This assumption is valid as long as the economy is large enough so that each student acts like a 'price-taker' and cannot affect the cutoffs of high schools.

Given the assumptions, the continuation value of middle school program m in Equation (3)

[A-7] Recall that the second period unobserved taste is $\gamma_i^H = diag(\rho_0)\gamma_i^M + \xi_i$ and hence it is serially correlated with γ_i^M, the first period unobserved taste. By assuming the first two independence conditions, we effectively assume that the correlation in the unobserved tastes is enough to model students' tastes that are consistent over the two periods but not captured by observable characteristics.

can be simplified as:

$$
\begin{aligned}
& E_{\gamma_i^H,\omega,\eta_i,Z_i^H}\left[\max_{j\in O_i(Z_i^H,m;\omega)} V_{ij}\middle|\Psi_{1i}\right] \\
=& E_{\gamma_i^H,\omega,\eta_i}\left[\max_{j\in O_i(Z_i^H,m;\omega)} V_{ij}\middle|\Psi_{1i}\right] && : Z_i^H \text{ is perfectly predictable} \\
=& \int_\omega E_{\gamma_i^H,\eta_i}\left[\max_{j\in O_i(Z_i^H,m;\omega)} V_{ij}\middle|\Psi_{1i}\right] dH(\omega) && : (\text{D.3}) \\
=& \int_\omega \int_{(\xi_i,\eta_i)} \max_{j\in O_i(Z_i^H,m;\omega)} V_{ij} dF(\eta_i,\xi_i|\Psi_{1i}) dH(\omega) \\
=& \int_\omega \int_{\xi_i} \left(\int_{\eta_i} \max_{j\in O_i(Z_i^H,m;\omega)} V_{ij} dF(\eta_i)\right) d\Phi(\xi_i) dH(\omega) && : (\text{D.2}) \\
=& \int_\omega \int_{\xi_i} \left(\mu + \log\left(\sum_{j\in O_i(Z_i^H,m;\omega)} \exp(v_{ij})\right)\right) d\Phi(\xi_i) dH(\omega) && : \eta_{ij} \overset{iid}{\sim} EVT1
\end{aligned}
$$

where $v_{ij} \equiv V_{ij} - \eta_{ij}$, F denotes a cdf function of its argument and $\Phi(\cdot)$ is a cdf of some multivariate normal variable, and μ is the Euler-Mascheroni constant.

In the final expression, the first integral over ω is calculated by using the empirical distribution of ω as described in Appendix D.1. The second integral over ξ_i is calculated using sparse grids quadratures (Heiss and Winschel, 2008). We use students' residence in the first period to calculate the distance to each high school program in calculating the continuation value.

D.4 Likelihood Function

Let student i's assigned middle and high school programs be m_i, j_i and the respective feasible sets be O_i^m, O_i^h. Note that $m_i \in O_i^m$ and $j_i \in O_i^h$. Let u_{im} and v_{ij} denote the part of U_{im} and V_{ij} excluding the idiosyncratic preference terms ϵ_{im} and η_{ij}. Also, denote the parameters to be estimated as

$$\theta = \left(\beta_0^M, \beta_Z^M, \beta_0^H, \beta_Z^H, \rho_0, \{\rho_\tau\}_\tau, \Sigma_\gamma, \Sigma_\xi, \vartheta_m, \vartheta_h, \delta, \lambda^M, \lambda^H\right)$$

Then for student i, conditional on γ_i^M, ξ_i,

$$\begin{aligned} P_i(\theta, \gamma_i^M, \xi_i) &= P(\text{observe } m_i, j_i | \gamma_i^M, \xi_i, \theta) \\ &= P\left(\begin{array}{c} U_{im_i} = \max_{m \in O_i^m} U_{im} \quad\quad \text{and} \\ V_{ij_i} = \max_{j \in O_i^h} V_{ij} \text{ given } m_i \end{array} \middle| \gamma_i^M, \xi_i, \theta \right) \\ &= \frac{\exp(u_{im_i}(\gamma_i^M, \theta))}{\sum_{m \in O_i^m} \exp(u_{im}(\gamma_i^M, \theta))} \frac{\exp(v_{ij_i}(\gamma_i^M, \xi_i, \theta; m_i))}{\sum_{j \in O_i^h} \exp(v_{ij}(\gamma_i^M, \xi_i, \theta; m_i))} \qquad : \text{(D.1)} \end{aligned}$$

where the second equality comes from the ex-post stability, and the third equality comes from the distributional assumptions on the unobservables. Then, since $\gamma_i^M \perp \xi_i$,

$$P_i(\theta) = \int_{\gamma_i^M} \int_{\xi_i} P_i(\theta, \gamma_i^M, \xi_i) d\Phi(\xi_i | \Sigma_\xi) d\Phi(\gamma_i^M | \Sigma_\gamma)$$

where $\Phi(\cdot|\Sigma)$ is the cdf of a multivariate normal with zero mean and covariance matrix Σ, and hence

$$\prod_i P_i(\theta), \text{ or } \sum_i \log P_i(\theta)$$

is the final likelihood function to be maximized.

E Alternative Specifications

E.1 Static Model

Recall the key features of the dynamic model: forward-looking agents, serial correlation of the unobservable tastes, and middle school type effects. To highlight the importance of including those features in the model, we estimate a restricted static model without the dynamic components of the model.

The static model has the same main components as the main model, but with three marked differences. First, we assume students are myopic so that they do not consider the high school application when making middle school choices ($\delta = 0$). Second, we do not allow the unobserved tastes for program characteristics to be serially correlated ($\rho_0 = 0$), and third, middle school type effects are absent ($\rho_\tau = 0, \forall \tau$). Table E.3 reports the preference estimates of the static model and the goodness-of-fit measures are reported in Table F.5.

Importantly, the goodness-of-fit measure of the restricted model is worse than our preferred specification in terms of middle school applications and more or less similar in terms of high school applications. This is as expected as the static model does not consider the forward-looking behavior of students in the middle school application stage and thus does a worse job of fitting the corresponding data. On the other hand, high school application is the *last* stage of the multi-period game and hence unlikely to be affected by whether including a dynamic feature or not.

Since the static model is a nested model of the full dynamic model in which the restriction that ρ_0, δ, and $\rho_\tau, \forall\tau$ are equal to zero is imposed, we can perform a likelihood ratio (LR) test. The result is reported in Panel C. The static model is strongly rejected in favor of our main dynamic model ($p < 0.001$), reconfirming the importance of including the forward-looking behavior of students, middle school effects on tastes, and serial correlation of unobservable tastes in the model.

E.2 Strict Truth-Telling (STT)

We also estimate the model with a different assumption on student behavior. Strict Truth-Telling (STT) assumes that (Fack, Grenet, and He, 2019)

1. Students rank all acceptable programs (i.e., better than the outside option) in their true preference order.
2. All unranked programs are unacceptable to the student. That is, they are worse than the outside option.

Hence, the likelihood used in the estimation for STT is as follows.

Let student i's ROL on middle school programs and high school programs be $R_i^M = (R_{i,1}^M, \cdots, R_{i,|R_i^M|}^M)$ and $R_i^H = (R_{i,1}^H, \cdots, R_{i,|R_i^H|}^H)$ respectively. We will use the notation $\succ_i$ to denote the inferred preferences by STT. Note that for each ranked c, $\{c' : c' \not\succ_i c\}$ includes c itself, programs ranked below c on the ROL, and the programs that are not ranked.

Table E.3: Preference Estimates: Static

	Middle Schools		High Schools	
	est	se	est	se
Panel A: Preference Estimates				
Proportion of High-Performer				
Main Effect	-7.574	(1.273)	0.929	(0.287)
Asian	-0.788	(1.773)	0.870	(0.411)
Black	10.573	(1.915)	-0.169	(0.469)
Hispanic	2.682	(1.246)	-0.356	(0.339)
Free/Reduced-price Lunch	0.804	(1.039)	-1.008	(0.280)
English Language Learner	-0.988	(2.358)	0.523	(1.227)
5th Grade Test Score	0.881	(0.545)	1.995	(0.165)
Proportion of White				
Main Effect	8.805	(1.382)	6.264	(0.401)
Asian	0.621	(1.458)	-1.513	(0.642)
Black	-10.161	(1.828)	-2.347	(0.645)
Hispanic	-2.729	(1.090)	-1.654	(0.447)
Free/Reduced-price Lunch	-1.628	(0.841)	-0.231	(0.371)
English Language Learner	-0.009	(2.030)	-1.288	(1.238)
5th Grade Test Score	-0.597	(0.435)	-0.109	(0.194)
1(STEM)				
Main Effect	0.396	(0.260)	-0.679	(0.124)
Asian	-0.076	(0.273)	-0.144	(0.203)
Black	-0.690	(0.294)	0.078	(0.192)
Hispanic	-0.206	(0.207)	0.089	(0.142)
Free/Reduced-price Lunch	-0.135	(0.176)	0.241	(0.127)
English Language Learner	0.208	(0.298)	0.874	(0.321)
5th Grade Test Score	0.113	(0.081)	-0.030	(0.063)
Panel B: Middle School Type Effects				
Type 1 (High-Achievement MS)				
Proportion of High-Performer			0.159	(0.288)
Proportion of White			0.627	(0.391)
1(STEM)			-0.307	(0.137)
Type 2 (High-Minority MS)				
Proportion of High-Performer			0.936	(0.321)
Proportion of White			-2.770	(0.479)
1(STEM)			0.173	(0.135)
Panel C: Unobservable Tastes				
(1,1) of Variance of Random Taste	40.970	(13.423)	0.730	(0.408)
(1,2)	-37.851	(12.154)	-3.058	(1.052)
(1,3)	-1.330	(2.585)	0.496	(0.164)
(2,2)	35.893	(13.309)	12.813	(2.760)
(2,3)	1.790	(3.262)	-2.078	(0.492)
(3,3)	0.384	(0.721)	0.337	(0.159)
Panel D: Other Parameters				
Outside option	-1.851	(0.198)	-0.112	(0.173)
Distance	0.718	(0.031)	0.496	(0.018)

Note: We report the preference estimates of the static model. School characteristics 'Proportion of High-Performer' and 'Proportion of White' are between 0 and 1, and '1(STEM)' is an indicator variable. In Panel A, Main Effect is the common taste (β_0^M, β_0^H), and we also include interactions of each school characteristics with Asian, Black, Hispanic, Free/Reduced-price Lunch (FRL) status, English Language Learner (ELL) status, 5th Grade Test Score in the following rows (β_Z^M, β_Z^H). Robust standard errors are reported in parentheses.

Then for student i, conditional on γ_i^M, ξ_i,

$$\begin{aligned}
P_i(\theta, \gamma_i^M, \xi_i) =& P(\text{observe } R_i^M, R_i^H | \gamma_i^M, \xi_i, \theta) \\
=& P\left(\begin{matrix} U_{i,R_{i,1}^M} > \cdots > U_{i,R_{i,|R_i^M|}^M} > U_{i0_m} > U_{im'}, \forall m' \in \mathcal{M} \setminus R_i^M \text{ and} \\ V_{i,R_{i,1}^H} > \cdots > V_{i,R_{i,|R_i^H|}^H} > V_{i0_h} > V_{ij'}, \forall j' \in \mathcal{J} \setminus R_i^H \end{matrix} \middle| \gamma_i^M, \xi_i, \theta \right) \\
=& \frac{\exp(u_{i0_m})}{\exp(u_{i0_m}) + \sum_{m' \notin R_i^M} \exp(u_{im'}(\gamma_i^M, \theta))} \prod_{m \in R_i^M} \left(\frac{\exp(u_{im}(\gamma_i^M, \theta))}{\exp(u_{i0_m}) + \sum_{m' \not\succ_i m} \exp(u_{im'}(\gamma_i^M, \theta))} \right) \\
& \times \frac{\exp(v_{i0_h})}{\exp(v_{i0_h}) + \sum_{j' \notin R_i^H} \exp(v_{ij'}(\gamma_i^H, \xi_i, \theta))} \prod_{j \in R_i^H} \left(\frac{\exp(v_{ij}(\gamma_i^H, \xi_i, \theta))}{\exp(v_{i0_h}) + \sum_{j' \not\succ_i j} \exp(v_{ij'}(\gamma_i^H, \xi_i, \theta))} \right)
\end{aligned}$$

Then,

$$P_i(\theta) = \int_{\gamma_i^M} \int_{\xi_i} P_i(\theta, \gamma_i^M, \xi_i) d\Phi(\xi_i | \Sigma_\xi) d\Phi(\gamma_i^M | \Sigma_\gamma)$$

and hence

$$\prod_i P_i(\theta), \text{ or } \sum_i \log P_i(\theta)$$

is the likelihood.

Table E.4 reports the preference estimates based on STT. The main difference is the negative and statistically significant estimate on δ, the valuation on the continuation value. The intuition is as follows. In case of payoff-irrelevant mistakes in which students omit favorable yet infeasible middle school programs from their ROLs, STT interprets that those omitted programs are less preferred than all ranked programs as well as the outside option. However, these competitive programs with high cutoffs will have high continuation value since they are likely to provide higher opportunities for getting into more favorable high school programs. As a result, STT would incorrectly infer that middle schools with high continuation values as unfavorable, resulting in a negative estimate for the discount factor.

The goodness-of-fit measures are reported in Table F.5. STT overall is outperformed by ex-post stability. Especially, the mean predicted fraction of students assigned to the observed assignments is nearly decreased to half, showing that the truth-telling assumption may be problematic even in a strategyproof environment.

Table E.4: Preference Estimates: Strict Truth-Telling (STT)

	Middle Schools		High Schools	
	est	se	est	se
Panel A: Preference Estimates				
Proportion of High-Performer				
Main Effect	4.385	(0.335)	-0.628	(0.141)
Asian	0.87	(0.527)	0.708	(0.209)
Black	1.253	(0.460)	0.601	(0.231)
Hispanic	-0.042	(0.376)	-0.278	(0.164)
Free/Reduced-price Lunch	0.080	(0.324)	-0.217	(0.133)
English Language Learner	-0.218	(0.806)	0.721	(0.481)
5th Grade Test Score	0.959	(0.172)	1.907	(0.066)
Proportion of White				
Main Effect	0.496	(0.190)	3.659	(0.139)
Asian	0.328	(0.354)	-0.693	(0.226)
Black	0.417	(0.327)	-1.102	(0.242)
Hispanic	0.605	(0.256)	-0.239	(0.163)
Free/Reduced-price Lunch	0.170	(0.217)	-0.202	(0.132)
English Language Learner	-0.070	(0.565)	-1.205	(0.455)
5th Grade Test Score	-0.414	(0.114)	-0.904	(0.062)
1(STEM)				
Main Effect	0.630	(0.069)	-0.305	(0.065)
Asian	-0.241	(0.129)	0.201	(0.102)
Black	-0.468	(0.113)	-0.051	(0.102)
Hispanic	-0.200	(0.091)	0.122	(0.074)
Free/Reduced-price Lunch	-0.205	(0.080)	-0.010	(0.064)
English Language Learner	0.175	(0.169)	0.286	(0.175)
5th Grade Test Score	0.074	(0.041)	-0.264	(0.026)
Panel B: Middle School Type Effects				
Type 1 (High-Achievement MS)				
Proportion of High-Performer			0.290	(0.141)
Proportion of White			0.321	(0.134)
1(STEM)			-0.101	(0.067)
Type 2 (High-Minority MS)				
Proportion of High-Performer			1.320	(0.177)
Proportion of White			-0.774	(0.177)
1(STEM)			0.024	(0.076)
Panel C: Unobservable Tastes				
ρ_0			0.457	(0.095)
			0.495	(0.279)
			0.147	(0.097)
(1,1) of Σ_γ	3.307	(0.841)		
(1,2)	-0.930	(0.373)		
(1,3)	-1.237	(0.218)		
(2,2)	0.262	(0.148)		
(2,3)	0.348	(0.108)		
(3,3)	0.463	(0.093)		
(1,1) of Σ_ξ			3.040	(0.306)
(1,2)			-2.538	(0.250)
(1,3)			-0.344	(0.092)
(2,2)			2.312	(0.265)
(2,3)			-0.009	(0.078)
(3,3)			0.491	(0.054)
Panel D: Other Parameters				
Outside option	-0.528	(0.256)	1.974	(0.038)
Distance	0.746	(0.012)	0.486	(0.007)
Discount Factor	-0.65	(0.062)		

Note: We report the preference estimates of the model based on STT. School characteristics 'Proportion of High-Performer' and 'Proportion of White' are between 0 and 1, and '1(STEM)' is an indicator variable. In Panel A, Main Effect is the common taste (β_0^M, β_0^H), and we also include interactions of each school characteristics with Asian, Black, Hispanic, Free/Reduced-price Lunch (FRL) status, English Language Learner (ELL) status, 5th Grade Test Score in the following rows (β_Z^M, β_Z^H). Robust standard errors are reported in parentheses.

F Additional Goodness-of-Fit Measures

We provide additional goodness-of-fit measures along two dimensions: how well it predicts the assignments and how well it predicts students' revealed preferences. Column (1) is our main specification, Column (2) is the static model in Appendix E.1, and Column (3) is the model based on STT in Appendix E.2.

Table F.5: Additional Goodness of Fit Measures

	(1) Dynamic Ex-post Stability		(2) Static Ex-post Stability		(3) Dynamic Strict Truth-Telling	
	MSAP	HSAP	MSAP	HSAP	MSAP	HSAP
Panel A. Simulated versus observed assignment (100 simulated samples)						
Mean predicted fraction of students assigned to observed assignments	0.5709	0.2022	0.5539	0.2018	0.3129	0.1111
	(0.0053)	(0.0049)	(0.0058)	(0.0048)	(0.0050)	(0.0048)
Panel B. Predicted versus observed partial preference order						
Mean predicted probability that a student's partial preference order among the programs in her ROL coincides with the submitted rank order	0.3848	0.1395	0.3215	0.1422	0.3769	0.1261
Panel C. Likelihood Ratio Test						
H_0: $\delta = \rho_0 = \rho_\tau = 0, \forall\tau$			Reject H_0 ($p < 0.001$)			

Note: Panel A calculates the average success rate of predicting the observed assignments in the data using the model estimate. 100 sets of unobservable variables ($\gamma_i^M, \xi_i, \epsilon_{im}, \eta_{ij}$) are drawn and for each set, 50 sets of tie-breaking lotteries are drawn and DA are run, resulting in 5,000 simulated assignments. The average and standard deviation (in parentheses) across the unobservable draws are reported. Panel B calculates the average predicted probability of a student's partial preference order among the programs ranked on her ROL coincides with the submitted rank order.

First, Panel A of Table F.5 compares each student's predicted assignment to the observed one. We have about 57.1% success rate for middle schools and 20.2% for high schools.[A-8] The higher success rate for middle schools can be explained by the fact that first, the number of programs are much smaller for middle schools, and second, middle schools have much more 'Zoned Guarantee' programs that guarantee admissions to students who are zoned to the school as long as they rank them.

Next in Panel B, we take as given the programs that a student has included in her submitted ROL, and compute the probability of observing this particular preference order among the ranked programs. Given the distributional assumptions on $\gamma_i^M, \xi_i^H, \epsilon_{im}, \eta_{ij}$, we can calculate the probabilities without relying on Monte Carlo simulations. We have 38.5% for

[A-8]We provide two benchmarks. First, the upper bound is calculated using submitted ROLs in the data, without relying on any estimates or the model: 78.0% for MSAP and 61.2% for HSAP. They do not equal 100% due to lottery draws (8 (40%) middle school programs and 29 (62%) high school programs use lottery draws for tie-breaking in our sample). Next, the lower bound is calculated using random prediction. That is, we let students randomly apply to programs and programs randomly select who to admit. On average, we find 5.9% for MSAP and 2.3% for HSAP.

middle schools and 14.0% for high schools. The difference in probabilities between middle schools and high schools is due to the larger number of high school programs and longer high school ROLs.[A-9]

G Segregation Measures

G.1 Theil's H Index

Theil's H Index is also known as the Information Theory Index or the Multigroup Entropy Index. We closely follow the definition used by the United States Census Bureau to describe housing patterns (Iceland, 2004).[A-10]

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

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

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

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

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

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

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

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

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

[A-9] On average, students rank 2 middle school programs (std 1.22), and 4 high school programs (std 2.60).
[A-10] See https://www.census.gov/topics/housing/housing-patterns/about/multi-group-entropy-index.html

G.2 Sorting Index

Sorting index for a given characteristic is defined by the ratio of the between-school variance to the total variance, measuring the fraction of the variance of a given characteristic that can be explained by the between-school differences. Specifically, let y_{ij} be the student i's characteristic of interest who is enrolled in school j. Then, the sorting index for y is simply obtained by the R^2 of the following linear regression:

$$y_{ij} = \alpha_j + e_{ij}$$

It varies between 0 and 1 by definition, and 0 means maximum integration, and 1 means maximum segregation.

H Additional Tables and Figures

H.1 Additional Tables and Figures from Section 2

Average School Characteristics by Rank on Students' ROL Tables H.6 and H.7 summarize the averages of school characteristics by rank on students' ROLs of middle schools and high schools, respectively. There are mainly three patterns. First, students tend to rank distant schools from their homes lower on their ROLs. Notably, the average distance of ranked programs is larger for high school programs than for middle school programs. As mentioned before, this possibly reflects that high school application has a higher degree of citywide school choice. Next, students rank schools with high student achievement higher on their ROLs. Third, students rank schools with a high proportion of subsidized lunch status, Black/Hispanic students lower on their ROLs.

H.2 Additional Tables and Figures from Section 3

Balance Test We present the students' test scores, demographic characteristics, and variables that describe middle school application behavior of the students who are assigned to the treatment middle schools by DA (*offered* students) and those who are not (*non-offered* students).

First, **Raw Difference** shows the sharp raw difference of covariates between the offered and the non-offered students. The offered have higher test scores and are less likely to be FRL, ELL, Black/Hispanic, or need special education, with all statistically significant differences. They also rank more high-achievement middle schools (recall this is our treatment of interest)

Table H.6: Middle School Program Characteristics on ROLs

	1	2	3	4	5	6	7	8	9	10	11	12 or longer
# Students Ranked	52789	40428	33980	24435	13419	7439	4131	2859	2013	1557	961	729
% Students Ranked	97.7	74.9	62.9	45.2	24.8	13.8	7.6	5.3	3.7	2.9	1.8	1.3
Distance (miles)	1.4	1.7	1.8	1.9	1.9	2	2.1	2.3	2.4	2.6	3	3.2
Mean Score (6th grade)	308.2	307.7	305.8	305.1	305.7	305.8	306.2	306.2	305.3	305.8	306.8	304.2
Mean Score (8th grade)	300.7	300.4	299.2	298.5	299.1	299.6	300.7	300.3	299.3	299.2	299.6	296
% Black/Hispanic	63.8	65.9	69.5	69.2	67.3	67.5	66.5	65.6	65.7	68.2	63.8	70.3
% Female	49.9	50.3	50.1	50.1	49.7	49.9	49.9	49.6	49.4	49.6	50.2	49.9
% Free/Reduced-price Lunch	69.4	69.9	71.6	72.2	72.7	73.7	74	73.9	75.1	73.8	71.8	73
6th Grade Size (100s)	1.6	1.3	1.2	1.2	1.1	1.1	1.1	1.1	1.1	1	1.1	0.9

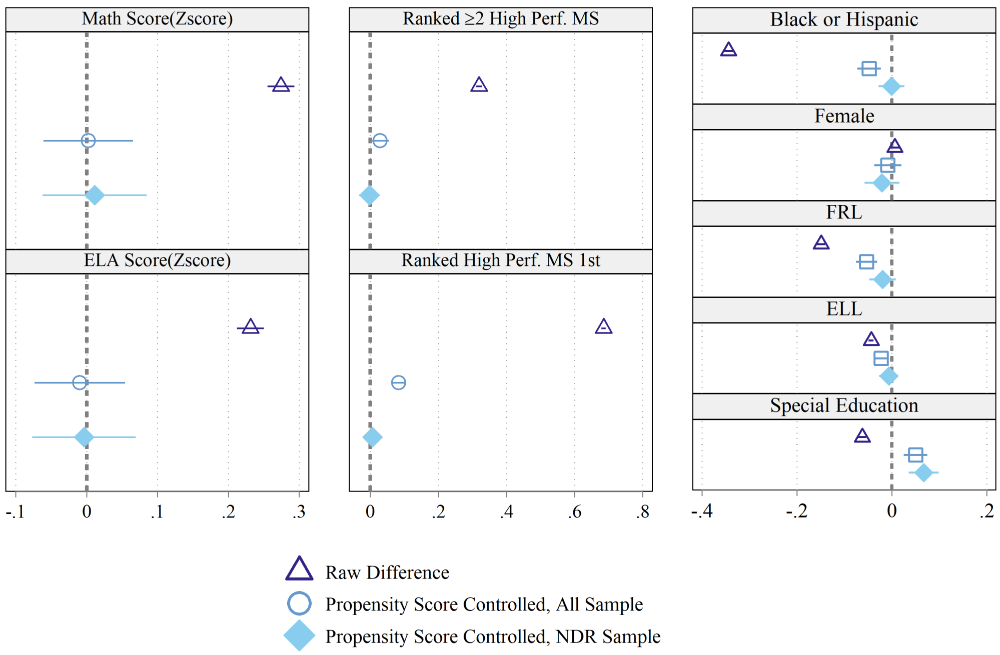
Note: The table calculates the average characteristics of the middle school programs on students' ROLs, by the rank on the ROL (N=54,012). % Black/Hispanic, % Female and % Free/Reduced-price Lunch are calculated using the characteristics of the currently enrolled 6th graders in AY 2014-15. Mean Score (6th grade) and Mean Score (8th grade) are calculated using the average of the statewide standardized Math and ELA exams of currently enrolled 6th graders and 8th graders in AY 2014-15, where the scale is from 110 to 410.

Table H.7: High School Program Characteristics on ROLs

	1	2	3	4	5	6	7	8	9	10	11	12
# Students Ranked	53187	49070	47234	44381	41062	37011	32413	28235	23943	20435	16952	13402
% Students Ranked	98.5	90.9	87.5	82.2	76.0	68.5	60.0	52.3	44.3	37.8	31.4	24.8
Distance (miles)	3.2	3.4	3.5	3.6	3.7	3.8	3.9	3.9	4.0	4.0	4.0	3.9
Mean Score (9th grade)	312.7	310.8	309.1	307.9	307.1	306.0	305.6	304.7	304.1	303.2	302.5	301.2
4yr Grad Rate (%)	85.4	84.1	83.4	82.8	82.5	82.1	82.0	81.6	81.4	80.9	80.2	79.2
Enroll in College (%)	73.8	72.3	71.3	70.7	70.2	69.7	69.6	69.1	68.8	68.1	67.3	66.1
% Black/Hispanic	58.2	59.5	60.7	62.4	63.5	65.0	66.0	67.4	68.5	69.9	70.7	71.6
% Female	53.4	51.9	51.1	50.7	50.4	50.2	50.1	50.0	49.7	49.8	49.7	49.5
% Free/Reduced-price Lunch	69.8	71.2	72.2	73.3	73.8	74.5	74.9	75.5	76.0	76.6	77.4	78.0
9th Grade Size (100s)	1.7	1.5	1.5	1.4	1.4	1.3	1.3	1.3	1.3	1.2	1.2	1.3

Note: The table calculates the average characteristics of the high school programs on students' ROLs, by the rank on the ROL (N=54,012). % Black/Hispanic, % Female and % Free/Reduced-price Lunch are calculated using the characteristics of the currently enrolled 9th graders in AY 2017-18. Mean Score (9th grade) are calculated using the average of the 8th grade statewide standardized Math and ELA exams of currently enrolled 9th graders in AY 2017-18, where the scale is from 130 to 400. 4yr Grad Rate and Enroll in College are calculated using the average of the graduating cohort in AY 2017-18.

Figure H.1: Covariate Balance Test: Offered Students v.s. Non-offered Students



Note: **Raw Difference** shows the t-test results of covariate mean difference between the offered and the non-offered. **Propensity Score Controlled, All Sample** shows the coefficient of the offered when we regress the covariate on the offered dummy variables, the nonparametric controls for propensity score, and the local linear function of non-random tie-breaker, using the entire sample. **Propensity Score Controlled, NDR Sample** is similar to **Propensity Score Controlled, All Sample** but we only include the sample of which propensity score is neither 0 nor 1. We plot the relative difference of each covariate of the offered students to that of the non-offered students, and the unit is standard deviation for the left panel and fraction for the middle and right panels. Markers show the exact estimates, and 95% CIs are presented. Robust standard errors are estimated. N=8,007 for **Propensity Score Controlled, NDR Sample**, and N=50,871 for other estimates.

than the non-offered students and are more likely to list them first on their ROLs, which makes natural sense as such behavior will unambiguously increase the odds of being offered such schools.

Next, we control for the propensity scores and include local linear control of tie-breakers in the following two specifications denoted by **Propensity Score Controlled, All Sample** and **Propensity Score Controlled, NDR Sample** in Figure H.1. Specifically, we run

$$W_i = \alpha_0 + \gamma D_i + \sum_x \alpha_1(x) d_i(x) + h(\mathcal{R}_i) + e_i \tag{H.4}$$

where W_i is the student covariates that we test balance, and D_i, $\{d_i(x)\}_x$ and $h(\mathcal{R}_i)$ are the same as in our main specification Equation (1).

Propensity Score Controlled, All Sample in Figure H.1 presents estimates on γ with students of all possible propensity scores, including 0 and 1. Controlling for the propensity score and non-random tie-breakers effectively balances covariates. Next, **Propensity Score Controlled, NDR Sample** shows the γ only with students with non-degenerate risk of being offered, i.e., subject to randomization. Further restricting the sample to those with non-degenerate risk provides an almost perfect balance between the offered and the non-offered group.

Figure H.2 presents the mean difference between those with non-degenerate offer risk and degenerate (0 or 1) offer risk when the treatment variable is 'attended a high-achievement middle school'. In our data, 2/3 of degenerate risk sample have the propensity score equal to 0, which means they did not apply to any of the high-achievement middle schools or had zero chance of getting in conditional on applying, suggesting they are different from the non-degenerate risk sample. Indeed, we find that students with non-degenerate risk and those with degenerate risk are quite different: students with non-degenerate risk have higher test scores, and less likely to be Black/Hispanic, and obviously ranked many treatment middle schools. It reconfirms that the 2SLS estimates we find in Section 3.2 are local average treatment effect (LATE).

Robustness Check We further investigate if the reduced-form evidence of middle school effects on high school choice we identified is mainly driven by the increase in students' test scores. Moreover, we control for the length of high school application list because students submit lists of different lengths, and the average characteristics of schools change along the rank on the list, as shown in Tables H.6 and H.7.

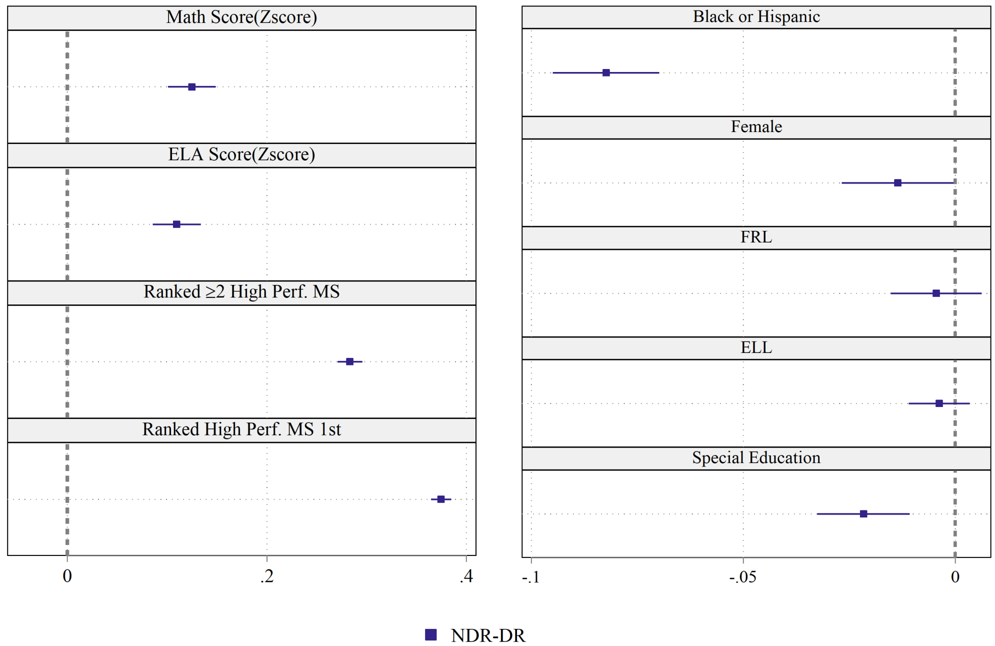
First, Table H.8 uses the same identification strategy as in Section 3 to show that attending

Table H.8: Effect of Attending a High-Achievement MS on Other Outcomes

	(1)	(2)	(3)
Model	OLS	2SLS	2SLS
Sample	All	All	NDR
Panel A: 8th Grade ELA Score (Zscore)			
From High-Achievement MS	0.064***	0.032	0.032
	(0.019)	(0.035)	(0.042)
N	42516	42516	6826
R2	0.610	0.616	0.639
$\bar{y}$	0.090	0.090	0.183
Panel B: 8th Grade Math Score (Zscore)			
From High-Achievement MS	0.173***	0.115**	0.152**
	(0.029)	(0.056)	(0.067)
N	32935	32935	5562
R2	0.582	0.591	0.651
$\bar{y}$	0.051	0.051	0.202
Panel C: Length of Application List			
From High-Achievement MS	-0.718***	-0.716*	-1.161**
	(0.265)	(0.429)	(0.585)
N	44237	44237	7062
R2	0.117	0.171	0.360
$\bar{y}$	7.579	7.579	7.049
Panel D: 1(Assigned to the First Ranked School)			
From High-Achievement MS	0.006	0.062*	0.035
	(0.015)	(0.033)	(0.042)
N	41312	41312	6571
R2	0.037	0.053	0.154
$\bar{y}$	0.472	0.472	0.464

Note: Standard errors clustered at graduating middle school in parentheses. All regressions control for student ethnicity, gender, English Language Learner status, Free/Reduced-price Lunch eligibility, Special Education status, standardized test score in 5th grade, and residential borough in 5th grade. All columns also control for saturated dummy for all possible values of propensity score of being assigned to a high-achievement MS, and local linear controls for non-random tie-breakers.
*p<0.1, **p<0.05, ***p<0.01.

Figure H.2: Covariate Balance Test: Non-degenerate v.s. Degenerate Risk Samples



Note: This table shows the t-test results of covariate mean difference between those with non-degenerate offer risk and those with degenerate offer risk. Markers show the exact estimates, and 95% CIs are presented. Robust standard errors are estimated (N=50,871).

a high-achievement middle school has a causal impact on the increase of students' 8th grade math test scores and shorter high school ROLs.

Nevertheless, Table H.9 shows that the 2SLS estimates are robust to controlling for the end-of-middle-school test score and the length of ROLs. Notably, students with higher end-of-middle-school test scores apply to high schools with better academic performance, the pattern described well in previous studies that estimate school demand (e.g., Hastings, Kane, and Staiger, 2005; Abdulkadiroğlu, Agarwal, and Pathak, 2017). For example, the estimates in Column (2) illustrate that the increase of a student's ELA test score by one standard deviation is associated with 0.9 pp increase in the average graduation rate of high schools on her application list. However, even with the test scores controlled, the main treatment effect of attending a high-achievement middle school is 1.6 pp, which is comparable to the effect in Column (4) of Table 3. This mediation analysis shows that there is still a treatment effect of attending a high-achievement middle school, even controlling for test scores. Motivated by this finding, we include a separate component that captures the effect of middle school beyond its value-added on test scores in the structural model presented in Section 4.

Table H.9: Mediation Analysis

	(1)	(2)	(3)	(4)
Model	2SLS	2SLS	2SLS	2SLS
Dependent Variable	Average of Top 5 Ranked		Assigned	
Sample	NDR	NDR	NDR	NDR
Panel A: 4yr Graduation Rate (%)				
From High-Achievement MS	1.353*	1.565**	2.341**	2.274**
	(0.700)	(0.766)	(1.119)	(1.104)
8th Grade ELA Score	0.909***		1.601***	
	(0.170)		(0.288)	
8th Grade Math Score	0.765***		1.006***	
	(0.206)		(0.326)	
Number of Programs Ranked		0.160**		-0.126
		(0.072)		(0.086)
N	7060	7060	6687	6687
R2	0.398	0.390	0.264	0.253
$\bar{y}$	83.729	83.729	79.901	79.901
Panel B: College Enrollment Rate (%)				
From High-Achievement MS	1.751*	1.846*	3.301**	3.038**
	(0.967)	(1.018)	(1.542)	(1.444)
8th Grade ELA Score	1.314***		2.070***	
	(0.205)		(0.328)	
8th Grade Math Score	0.910***		1.416***	
	(0.231)		(0.374)	
Number of Programs Ranked		0.078		-0.320***
		(0.095)		(0.115)
N	7060	7060	6679	6679
R2	0.471	0.460	0.324	0.314
$\bar{y}$	72.197	72.197	67.204	67.204
Panel C: % High-Performing Students				
From High-Achievement MS	2.913*	3.567*	5.185**	5.232**
	(1.748)	(1.899)	(2.061)	(2.039)
8th Grade ELA Score	2.114***		3.023***	
	(0.351)		(0.409)	
8th Grade Math Score	1.258***		1.315**	
	(0.397)		(0.522)	
Number of Programs Ranked		0.500***		-0.050
		(0.119)		(0.122)
N	7062	7062	6751	6751
R2	0.513	0.510	0.415	0.400
$\bar{y}$	40.934	40.934	34.978	34.978

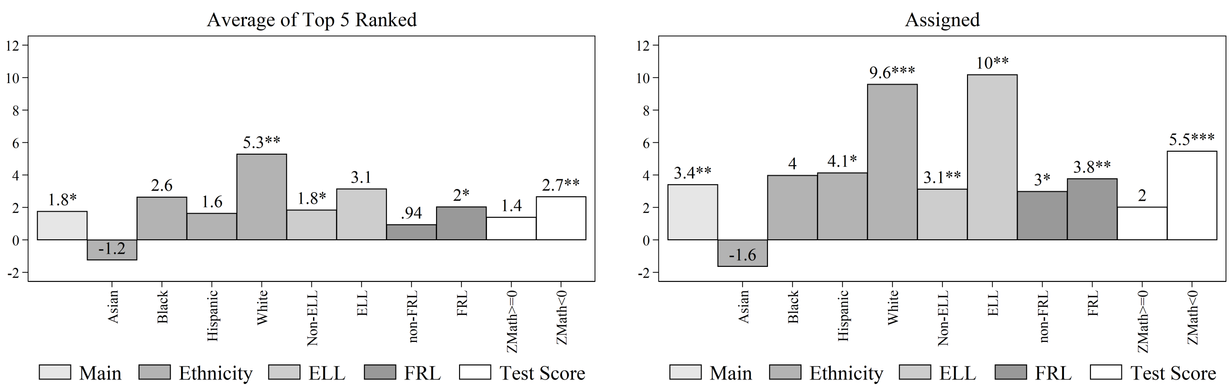
Note: All columns show 2SLS estimates. Standard errors clustered at graduating middle school in parentheses. All regressions control for student ethnicity, gender, English Language Learner status, Free/Reduced-price Lunch eligibility, Special Education status, standardized test score in 5th grade, and residential borough in 5th grade. All columns also control for dummy variables for all possible values of propensity score of being assigned to a high-achievement MS, and local linear controls for non-random tie-breakers.
*p<0.1, **p<0.05, ***p<0.01.

Heterogeneous Treatment Effects Next, we provide results on possible heterogeneous treatment effects. Figure H.3 presents the heterogeneity analysis results by student demographic characteristics. The effect is broad-based, but there is a larger effect among students whose baseline mean of the outcome variable is lower than their peers. For example, while attending a high-achievement middle school increases the college enrollment rate of an ELL student's assigned school by 10 pp, it does by 3.1 pp for a non-ELL student. The baseline level of the matched school's college enrollment rate is 59 and 66 percent among ELL/non-ELL students, respectively. In a similar vein, attending a high-achievement middle school has a larger effect for FRL students than their non-FRL peers and for Black, Hispanic, White students than their Asian peers. These results suggest that attending a high-achievement middle school could *level the field* for different groups of students.

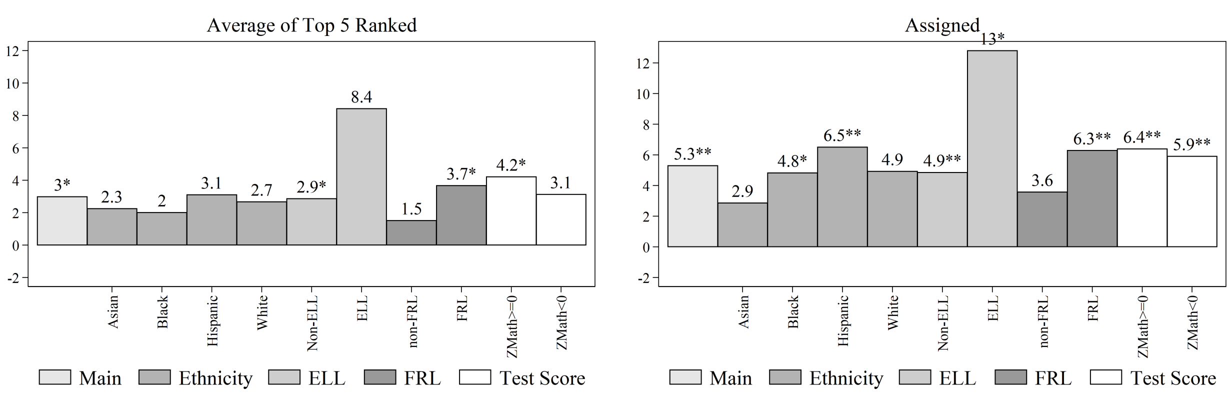
Treatment Effect of Attending High-Minority Middle Schools Next, we explore the treatment effects of attending a middle school with a high proportion of Black/Hispanic students in Table H.10. We do not find significant effects.

Figure H.3: Effect of Middle School on High School Choice By Student Characteristics

A. College Enrollment Rate (%)



B. % High-Performing Students



Note: Standard errors are clustered at graduating middle school. We label a student ‘high-performing’ if the standardized test score is above 66th percentile of the distribution. The estimates are derived by running the 2SLS model (Equation (1)) separately with students of the corresponding characteristics. Baseline student characteristics are controlled, excluding the demographic variable of interest. For instance, regression only with Asian students does not include the set of ethnicity dummy variables, but include ELL status, FRL status, and test score.
*p<0.1, **p<0.05, ***p<0.01

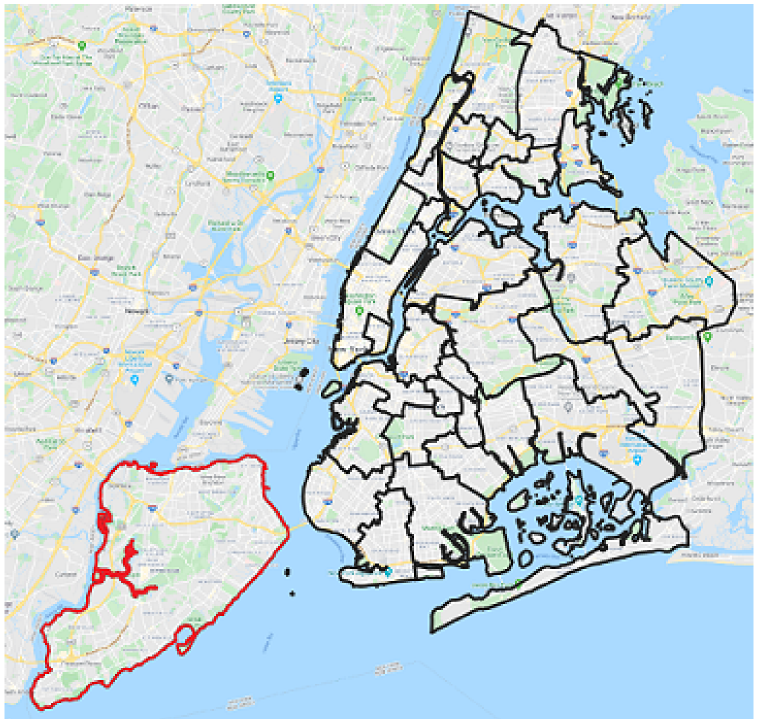
Table H.10: Effects of Attending Highly Minority MS on HS Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Average of Top 5 Ranked			Assigned		
Model	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Sample	All	All	NDR	All	All	NDR
Panel A: 4yr Graduation Rate (%)						
From High-Minority MS	-1.306***	-0.549	-0.225	-1.480***	-0.034	-1.039
	(0.455)	(1.017)	(1.182)	(0.559)	(1.798)	(2.057)
N	46631	46631	3308	43927	43927	3103
R2	0.291	0.308	0.312	0.180	0.192	0.207
$\bar{y}$	83.438	83.438	79.593	79.097	79.097	74.068
Panel B: College Enrollment Rate (%)						
From High-Minority MS	-1.686***	-0.314	0.248	-2.189***	-0.713	-0.794
	(0.553)	(1.267)	(1.459)	(0.661)	(2.052)	(2.383)
N	46630	46630	3307	43843	43843	3091
R2	0.363	0.378	0.358	0.237	0.250	0.260
$\bar{y}$	71.371	71.371	66.679	65.829	65.829	60.183
Panel C: % High-Performing Students						
From High-Minority MS	-4.024***	1.900	3.188	-3.875***	1.534	3.957*
	(0.850)	(1.745)	(2.084)	(0.800)	(2.046)	(2.240)
N	46723	46723	3317	44579	44579	3163
R2	0.441	0.455	0.370	0.376	0.387	0.333
$\bar{y}$	39.839	39.839	28.252	33.146	33.146	21.158
Panel D: % White						
From High-Minority MS	-4.758***	-0.029	0.415	-4.152***	-0.930	-0.056
	(0.560)	(0.669)	(0.597)	(0.525)	(0.748)	(0.651)
N	46723	46723	3317	44579	44579	3163
R2	0.616	0.627	0.367	0.535	0.544	0.288
$\bar{y}$	18.518	18.518	7.045	14.957	14.957	4.242
Panel E: 1(STEM)						
From High-Minority MS	0.034**	0.029	0.045	0.022	-0.036	-0.017
	(0.014)	(0.043)	(0.055)	(0.018)	(0.062)	(0.077)
N	46723	46723	3317	44582	44582	3163
R2	0.089	0.113	0.241	0.037	0.051	0.164
$\bar{y}$	0.325	0.325	0.376	0.315	0.315	0.361

Note: Standard errors clustered at graduating middle school in parentheses. All regressions control for student ethnicity, gender, English Language Learner status, Free/Reduced-price Lunch eligibility, Special Education status, standardized test score in 5th grade, and residential borough in 5th grade. Column (2)-(3) and (5)-(6) control for dummy variables for all possible values of propensity score of being assigned to a high-minority middle school, and local linear controls for non-random tie-breakers.
*p<0.1, **p<0.05, ***p<0.01.

H.3 Additional Tables and Figures from Section 4.2

Figure H.4: Staten Island and NYC Community School Districts



Source: NYC Open Data.
Note: The map shows 32 community school districts (CSD) in NYC. The red bordered is Staten Island, which is CSD 31 and well-separated from the rest of NYC.

H.4 Additional Tables from Section 4.4

Table H.11: Decomposition of Effects of Middle Schools on High School Assignments

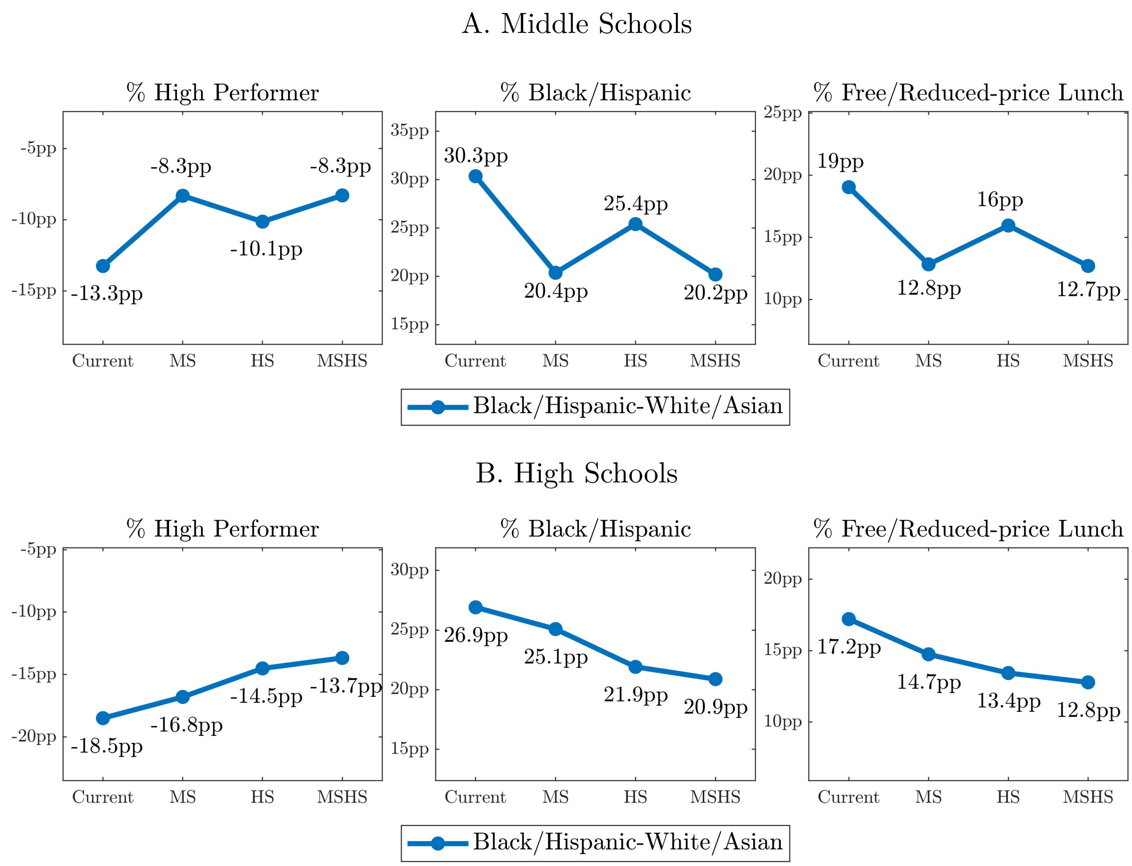
		(1) Full	(2) Application	(3) Priority
Middle School B to G	% High-Performer	9.67	5.70	3.20
		(1.11)	(0.73)	(0.78)
	% White	5.71	4.82	0.85
		(0.76)	(0.65)	(0.45)
	1(STEM)	-0.54	-3.01	1.72
		(2.11)	(1.55)	(1.44)

Note: Average standard deviation across unobservable and lottery draws is reported in parentheses.

H.5 Additional Figures from Section 4.5

Counterfactual Policy Predictions on Characteristics of Assigned Schools We report the results of counterfactual analyses on the racial gap in the characteristics of the assigned schools.

Figure H.5: Racial Gap in Characteristics of Assigned School Programs in Staten Island



Note: The graph plots the gap of the characteristics of assigned school programs between Black/Hispanic versus White/Asian students in each counterfactual scenario. 100 sets of unobservable variables $(\gamma_i^M, \xi_i, \epsilon_{im}, \eta_{ij})$ are drawn and for each set, 50 sets of tie-breaking lotteries are drawn and DA are run, resulting in 5,000 simulated assignments for each counterfactual scenario where the draws are fixed across scenarios. The mean across unobservable and lottery draws are reported for **MS**, **HS** and **MSHS**. We use the observed assignment results in the data for **Current**.