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

Dong Woo Hahm† Minseon Park‡

March 15, 2024

Abstract

We study the dynamic relationship of school choices across different educational stages. Using quasi-random middle school assignments in New York City, we show that middle schools with top quality-review scores cause students to be matched to higher-achievement high schools, in both level and value-added. A decomposition exercise using a sequential model of middle and high school choices shows that such effects of middle schools mainly operate by affecting students' high school applications rather than high school priorities, accounting for nearly 80% of the total effect. By mainly changing students' high school applications, abolishing eligibility restrictions of top quality-review-scored middle schools can not only increase the average quality of attended high schools but also narrow the racial and income disparities in it. Such efficiency and equity gains increase by up to 50% when combined with similar high school admissions reforms.

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

*We deeply appreciate invaluable guidance from Yeon-Koo Che, Pierre-André Chiappori, Chao Fu, Jesse Gregory, John Kennan, Bentley MacLeod, Christopher Taber, Miguel Urquiola, and Seth Zimmerman. We are also grateful to Sandra E. Black, Michael Best, Gautam Gowrisankaran, YingHua He, Seungki Hong, Jean-François Houde, Simon Lee, W. Bentley MacLeod, Bernard Salanié, Jeffrey Smith, Suk Joon Son, Olivier Tercieux, Daniel Waldinger, and Matthew Wiswall for valuable comments and suggestions. Thanks also to the New York City Department of Education, and 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. Any errors are our own.

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

‡Broad Center, Yale School of Management. Email: minseon.park@yale.edu

1 Introduction

Throughout their children's academic journey, households engage with multiple educational decisions that are dynamically interrelated. For example, a child's elementary or middle school may influence which high schools or colleges she aspires to attend and for which she is qualified, and knowing this, parents often factor in these future consequences when making decisions at earlier educational stages. This dynamic linkage in households' decisions underscores the importance of considering how policies targeting one educational stage may also have consequences on other stages. Furthermore, it suggests that policies targeting different schooling stages simultaneously should take potential complementarity or substitutability into account.

We study the dynamic relationship of households' educational decisions and its implications for policy evaluation in the context of school choice. School choice is a choice-based tool for assigning students to schools for which they qualify, involving high-stakes educational decisions for households. Worldwide, numerous school districts implement centralized school choice *across various education levels*.[1] This offers exceptional opportunities for in-depth analysis of households' intertemporal decision-making by generating rich records of households' choices and assignments throughout multiple stages. Nevertheless, prior studies have largely focused on studying school choices *one level at a time*, leading to a lack of empirical evidence and an appropriate framework.

To address this gap, we ask the following research questions:

1. How does a student's earlier school choice influence their subsequent school choices?
2. How can educational policies capitalize on this sequentiality to enhance their effectiveness?

We study these questions in the context of NYC public middle and high school choices. We investigate whether a student's attended middle school, resulting from her middle school choice, affects her high school choice outcomes.

Specifically, we quantify the extent to which this effect occurs by causing students to prioritize different aspects of high schools, such as value-added (Friedman, 1955; Rothstein, 2006; Abdulkadiroğlu, Pathak, Schellenberg, and Walters, 2020), rather than by altering their qualifications for different high schools. Quantifying the relative importance of these two potential channels of dynamic linkage in school choices is crucial for policy implications, which is facilitated by the rich administrative data on households' choices and school admissions rules available in our

[1]Examples include Baltimore, Boston, Denver, Lee (FL), New York City, Newark, New Orleans, Oakland (CA), San Francisco in the US, and Chile, England, Hungary, Paris, Taiwan, Turkey and many more.

context. If student-school match is primarily determined by demand pressure rather than the schools' admissions rules, any admissions reform targeting only high schools (e.g., affirmative action in high schools) might only marginally change student-school match. In contrast, efforts to influence families' demands could prove more effective in achieving policy goals, whether for efficiency or equity. Our second question explores the possibility of an admissions reform *at an earlier stage* as a means to shape demand pressure in the subsequent school assignment processes.

We start by presenting empirical evidence of the impact middle schools have on students' high school applications and assignments. To obtain causal estimates, we leverage the quasi-random assignments induced by tie-breaking rules applied to applicants who hold identical preferences and priorities, following Abdulkadiroğlu, Angrist, Narita, and Pathak (2022).

Our two-stage least squares (2SLS) estimates reveal that attending middle schools recognized as "Well-Developed"—the highest Quality Review Score by the Department of Education and confirmed to have high value-added in Section 3—significantly causes students to *apply to* high schools with better academic performance. Treated students apply to high schools with higher college matriculation rates by 3 percentage points and with higher value-added by 1.4 percentage points. Moreover, these high-performing middle schools boost the likelihood of students' being *assigned to* even higher-caliber high schools. Notably, the middle school effects we identify are more than two times larger than the effects of an information intervention in the same context conducted by Corcoran, Jennings, Cohodes, and Sattin-Bajaj (2022). This suggests that attending a high-quality middle school constitutes a more intensive intervention than providing students with information about proximate high schools' characteristics.

Motivated by the empirical evidence, we turn to a dynamic framework of school choice. The model captures two primary channels through which middle schools influence students' high school choice outcomes. First, it allows families' demand for high schools to depend on the middle schools they attend (the *application* channel). Second, students' admissions chances at high schools also depend on the middle schools they attend (the *priority* channel). Families take these into account when applying to middle schools.

Our model innovates upon the static framework commonly employed in the school choice literature by extending it to a dynamic framework. The main challenge lies in characterizing the continuation value of attending a particular middle school, i.e., the expected payoff of a student's future high school choice that varies across middle schools. To address this challenge, we combine large market matching theory with the dynamic discrete choice framework. We assume the stability of high school assignments—each student is assigned to her most preferred feasible school—which is justified in a large enough economy (Azevedo and Leshno, 2016; Fack, Grenet, and He, 2019; Artemov, Che, and He, 2021; Che, Hahm, and He, 2023). This approach enables us to succinctly characterize the continuation value by interpreting each student's high school assignment and feasible schools (both of which depend on her attended middle school) as her choice and choice set, respectively, as in a standard discrete choice framework. Combined with the distributional assumption on the idiosyncratic preference shocks, this allows us to utilize techniques from the dynamic discrete choice literature.

The main objective of estimation is to separately identify middle school effects through the *application* channel from the serially correlated unobserved heterogeneity in school demand. To achieve this, we exploit the panel structure of our middle and high school application data and the quasi-random variation in middle school assignments. We estimate the model using the expectation-maximization algorithm with a sequential maximization step (Arcidiacono and Jones, 2003) and validate model estimates by comparing simulated middle school effects with our reduced-form estimates.

Using the model, we first show that the effect of middle schools on high school choice outcomes mainly operates through the *application* channel. In our decomposition exercise, we alternatively shut down the *application* and the *priority* channel one at a time. Of the average treatment effects of attending a Well-Developed middle school on the college enrollment rate of assigned high schools, which is 2.6 percentage points, four-fifths can be attributed solely to the *application* channel.

Next, we design and evaluate counterfactual policies considering the dynamic relationship between school choices across different education levels. While any student-school matching is determined by both households' demand and schools' supply of seats, policymakers often find it challenging to influence the demand side. Consequently, most school choice policies have focused on reforming only the supply side.[2] However, recent evidence suggests that the supply-side-only interventions minimally change the student-school matching, largely due to the marked heterogeneity in school demand across students (Oosterbeek, Sóvágó, and van der Klaauw, 2021; Laverde, 2023). Our reduced-form findings and model estimates suggest that *supply-side* reforms of middle school choice, such as changing the admissions rule of middle schools, should influence students' *demand*

[2]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 adopted a similar admissions policy reform (Barry, Ellen, "Boston Overhauls Admissions to Exclusive Exam Schools", *New York Times*, 15 July 2021).

or application for high schools by altering students' middle school assignments.

To quantify this, we evaluate a series of admissions reforms by varying the timing of their implementation. We examine a counterfactual policy in which the city eliminates the geography-based eligibility criteria for Well-Developed middle schools and/or high-college-enrollment-rate high schools, which are predominantly concentrated in a few school districts in the city.

To begin with, the high school admissions reform increases the average college enrollment rate of students' attended high schools by 1%, and reduces racial and income disparities in it by 4.6% and 9.7%, respectively. More importantly, the *middle school reform* results in efficiency and equity gains of a similar magnitude in *high school assignments* to those of the high school admissions reform. Additionally, we find that the two reforms are substitutable; the effects of high school reform increase by up to 50% when implemented in conjunction with a similar middle school admissions reform. These findings suggest that policymakers should consider that intervention on the *supply side* of an earlier school choice induces changes in the *demand side* of subsequent school choice stages. Also, taking into account administrative costs or political burdens, admissions reform at alternative stages can be utilized to design policies targeting specific stages of school choice.

Related Literature The paper is primarily related to three strands of the literature. First, we contribute to the school choice literature by explicitly studying the dynamic relationship between school choice at different educational levels. To this end, we extend the static framework used in the literature to a dynamic framework of school choices at multiple stages. The literature has studied the factors that influence the outcomes of school choice, such as the assignment mechanism (Abdulkadiroğlu, Agarwal, and Pathak, 2017; He, 2017; Agarwal and Somaini, 2018; Calsamiglia, Fu, and Güell, 2020); information provision (Hastings and Weinstein, 2008; Hoxby and Turner, 2015; Ajayi, Friedman, and Lucas, 2017; Luflade, 2017; Corcoran, Jennings, Cohodes, and Sattin-Bajaj, 2022; Kapor, Neilson, and Zimmerman, 2020; Grenet, He, and Kübler, 2022; Ainsworth, Dehejia, Pop-Eleches, and Urquiola, 2023; Campos, 2023); or consideration set (Lee and Son, 2022). A notable exception is Narita (2018), who studies how families' underlying demand changes due to some learning between the main and the supplementary round of NYC's high school choice system. We turn our attention to the schools students attend in earlier educational stages, which directly affect their demand for subsequent schools *and* admissions chances.[3]

[3] In the broader empirical market design literature, several papers have considered dynamics such as on a kidney waitlist (Zhang, 2010; Agarwal, Ashlagi, Somaini, and Waldinger, 2018; Agarwal, Ashlagi, Rees, Somaini, and Waldinger, 2021); public housing (Waldinger, 2021); and dynamic college admissions (Larroucau and Rios, 2020b)

Second, our paper also belongs in the broader literature on the effects of schools on students' future outcomes. Our paper reveals an important mechanism behind the school effects on longer-run outcomes studied in prior studies: the subsequent education choices, which we measure by leveraging rich school application data, are largely determined by earlier education choices. Many researchers have studied the effects on outcomes such as academic performance, including test scores and graduation and college outcomes (Altonji, Elder, and Taber, 2005; Hastings and Weinstein, 2008; Pop-Eleches and Urquiola, 2013; Deming, Hastings, Kane, and Staiger, 2014; Abdulkadiroğlu, Angrist, and Pathak, 2014) and labor market outcomes, such as occupation or wages (Card and Krueger, 1992a,b; Betts, 1995; Clark and Bono, 2016), among many others. Meanwhile, we study the effects of schools on students' future academic choices in a K-12 context.

Lastly, by drawing on the literature that leverages quasi-experimental features built in school assignments, we estimate schools' causal effects on subsequent school choice behavior/assignments, outcomes that are less studied. Previous studies used lotteries in charter school admissions (Hoxby and Rockoff, 2004); the tie-breaking features of centralized assignments (Deming, Hastings, Kane, and Staiger, 2014; Abdulkadiroğlu, Angrist, Narita, and Pathak, 2017, 2022); and test score cutoffs (Pop-Eleches and Urquiola, 2013; Abdulkadiroğlu, Angrist, and Pathak, 2014). Of these, we adopt the methodology of Abdulkadiroğlu, Angrist, Narita, and Pathak (2017, 2022).

2 Institutional Background and Data

2.1 Public School Choice in NYC

NYC is one of the largest school districts worldwide that use centralized school choice to assign students to public schools. School choice starts with 3-year-olds, and students/parents participate in the choice process at subsequent levels: pre-K, kindergarten, elementary, middle, and high school. 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.[4] Since the unit of admission is a program instead of a school, we can

among many others.

[4]Also, 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 have a separate admissions process

consider each program a separate school. In the following, we use the terms "program" and "school" interchangeably when there is no risk of confusion.

Both middle and high school systems use the student-proposing deferred acceptance (DA) algorithm (Gale and Shapley, 1962), which takes students' applications, schools' ranking over students, and the preannounced number of seats as main inputs and produces at most one assignment for each student (see Appendix A for details).

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 can rank up to 12 programs.[5] Schools rank students by preannounced admissions rules, which consist of three layers. First, students are considered only by schools they are eligible for. Second, eligible applicants are classified into a small number of priority groups, such as "students or residents of Manhattan". A program considers all students in the higher priority group prior to any student in lower priority groups. When there is no confusion, we use priority to denote both eligibility and priority groups. Lastly, schools use tie-breaking rules to decide which applicants to admit among those in the same priority group. Some programs actively screen students using *nonrandom tie-breakers*, constructed from the previous year's GPA, statewide standardized test scores, attendance, and punctuality. Other programs break ties using a random *lottery* that is applied to each student across all such programs in the same fashion (a single tie-breaking rule). We leverage the quasi-random assignments that result from these tie-breakers to obtain the effects of middle schools on high school application and assignment in Section 3. Middle and high school programs are classified into subgroups depending on details of their admissions method, which is explained in Appendix A.

Throughout the paper, we focus on the effect of attending "Well-Developed" middle schools, the highest grading from the NYC DOE Quality Review of schools, on high school choice. The Quality Review evaluates how well schools are organized to support student learning and teacher practice. This is the most recent form of a school accountability system that was introduced in 2007. Experienced educators conduct a 2-day school visit to observe classrooms and speak

using a test called the Specialized High Schools Admissions Test (SHSAT). Similarly, we exclude public charter schools because they have separate admissions processes outside of the centralized school choice system.

[5] 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.

with school leaders, teachers, students, and parents/caregivers to give feedback and ratings across 30 sub-indicators of the NYC School Quality Rubric. Schools are grouped into four categories according to the average of schools' ratings across all sub-indicators: Well-Developed, Proficient, Developing, and Underdeveloped. We consider end-of-high school academic outcomes as the high school characteristics of our special interest. Such outcomes include the graduation rate within 4 years, college enrollment rate, PSAT scores, and SAT scores.

Well-Developed status and end-of-high school academic outcomes are good proxies for school quality. For example, Table 1 demonstrates that the Well-Developed status of middle schools is highly correlated with students' baseline characteristics such as mean test scores. Moreover, we find that Well-Developed middle schools enhance students' standardized math test scores by 0.18 standard deviation (Table 3), when we we estimate value-added leveraging quasi-random assignments (see Section 3 for details on the empirical strategy).

More importantly, those characteristics were the main facts included in the middle and high school directories, respectively. A school directory is a widely used handbook that provides parents, students, and teachers with a wide variety of information on schools to help them navigate the admissions process. Although the Quality Review process continues to be used today, the scores were not included in the 2017-18 high school directory.

2.2 NYC School Choice Data

Our analysis sample consists of 47,952 students who participated in the middle school (MS) application process in the academic year 2014-15 and then participated in the high school (HS) application process in the academic year 2017-18.[6] For each student, we have panel data on their middle school and high school applications/assignments. We focus on the applications and assignments of the main round of the admission process. Appendix B provides more details on data sources and sample restrictions.

Table 1 shows that students attending Well-Developed middle schools differ in observable characteristics from students attending other middle schools, which suggests that students sort into schools. For example, the average baseline mean student test scores among all middle schools is -0.11, while it is 0.58 among Well-Developed middle schools. More importantly, the last row of Panel B shows that while 20% of students attending an average high school graduated from a Well-Developed middle school, 30% of students attending an average "high-college rate" high school

[6] Students in our analysis sample have characteristics similar to the whole sample of middle school applicants (Appendix Table B.1).

Table 1: Summary Statistics of Middle and High School Program Characteristics

Variables	(1) Mean	(2) Std	(3) Mean	(4) Std
Panel A: Middle School Program	All Schools (N=670)		Well-Developed (N=150)	
Baseline Mean Test Score	-0.11	1.10	0.58	1.21
% White	14.17	20.88	23.28	27.51
% Black/Hispanic	70.92	30.51	57.18	35.56
% Free/Reduced-price Lunch	76.09	19.06	68.00	23.01
Cohort Size	98.30	90.67	105.10	92.32
Broadly eligible?	0.06	0.23	0.09	0.29
Use Nonrandom Tie-breaker?	0.40	0.49	0.48	0.50
Panel B: High School Program	All Schools (N=745)		High-College Rate (N=248)	
College Enrollment Rate (%)	58.38	17.13	78.06	7.93
4yr Graduation Rate (%)	73.19	15.98	88.72	7.27
Baseline Mean Test Score	-0.20	0.42	0.13	0.41
% White	10.41	15.56	15.28	18.10
% Black/Hispanic	76.58	23.40	66.24	26.93
% Free/Reduced-price Lunch	80.13	15.33	72.72	18.05
Cohort Size	83.04	82.66	110.50	105.30
Broadly eligible?	0.85	0.36	0.86	0.35
Use Nonrandom Tie-breaker?	0.38	0.49	0.59	0.49
Priority to Feeder Schools?	0.14	0.34	0.21	0.41
% From Well-Developed MS	0.20	0.19	0.30	0.24

Note: A high school is considered to have a "high-college rate" if the college enrollment rate of the 2016-17 graduating cohort is greater than the 66th percentile among all high schools. Schools are labeled as "broadly eligible" if they are open to all students from the borough or the city. Other schools are open only to students from the district or attendance zone. The test score is a mean of ELA (English Language Arts) and math test scores. We standardized ELA and math scores, respectively, within each cohort, to have a mean of 0 and a standard deviation of 1. For middle schools, we take the average test scores of 6th-grade students attending the school. For high schools, we take the average of the 8th-grade test scores of 9th-grade students attending the school, since students take NYS tests from 3rd and 8th grades.

did so, in which a high school is labeled as "high-college rate" if the 2016-17 graduating cohort's college enrollment rate is greater than the 66th percentile among all schools. This pattern suggests two possibilities. First, students may have consistent tastes for middle and high school program characteristics.[7] Second, which middle school a student attends may change how she applies and is assigned to high schools. Our aim in the following sections is to explore and distinguish these possibilities.

Many high schools employ selective admissions criteria that potentially depend on students' middle schools, which again suggests that students' middle schools play an important role in determining their high school priority standings. Fifty-nine percent of high-college rate high schools

[7]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 school. We control for the borough of residence in Section 3 and the distance to each school in Section 4.

adopt nonrandom tie-breakers relative to the average of 38% across all high schools. Since one important factor in nonrandom tie-breakers is students' middle school academic performance, nonrandom tie-breakers potentially depend on the middle school students attended. Notably, 21% of high-college rate high schools also explicitly gave admissions priorities to students who graduated from certain middle schools, while 14% of all high schools did so.

For interested readers, we present average school characteristics by rank on students' ROLs of middle and high schools in Appendix F.1.

3 Effects of Middle Schools on High School Choice

In this section, we aim to estimate the causal effects of middle school attendance on high school applications and assignments. We first describe our empirical strategy, which uses the quasi-experimental feature built in the centralized assignment system, and present our empirical findings.

3.1 Empirical Strategy

Our main identification concern is that students may sort into different middle schools based on factors unobserved by the researcher, which could simultaneously affect how students choose high schools and where they are assigned. For example, a student who prefers a high-quality middle school more than her peers of the same observable characteristics will likely also prefer a high-quality high school. To deal with this selection issue, we adopt the research design introduced by Abdulkadiroğlu, Angrist, Narita, and Pathak (2017, 2022), which builds on the quasi-experimental variation embedded in DA. We briefly explain the empirical strategy below and recommend that interested readers consult the original papers for details.

Recall that students' applications, priorities, and tie-breakers—either *lotteries* or program-specific *nonrandom tie-breakers*—are the only factors that determine assignment (see Section 2). In programs that use *lotteries*, students' assignments are random after controlling for student application and priority (Abdulkadiroğlu, Angrist, Narita, and Pathak, 2017). For programs that use *nonrandom tie-breakers*, the concern is that nonrandom 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. To deal with this, Abdulkadiroğlu, Angrist, Narita, and Pathak (2022) take a nonparametric regression discontinuity approach and exploit a subset of assignments that are as good as random. 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, 2022), and hence their assignments are as good as random.

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 (2022) formally show that DA-generated assignments are independent of any variables unaffected by the treatment after conditioning on the propensity score.[8]

The treatment effect of interest, the effect of attending a certain type of middle school, is estimated from a 2SLS model in which DA *assignment* is used as an instrumental variable for 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 our main equation, where β is the treatment effect of interest, and Equation (2) is the respective first-stage regression. Y_i is our outcome of interest and describes student i's high school choice behavior or outcomes, and C_i is the treatment variable, which equals 1 if i attended any Well-Developed middle school and 0 otherwise. D_i is the instrumental variable which equals 1 if i was assigned to any Well-Developed middle school 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 $\alpha_2(x)$ provide a saturated nonparametric control for all possible values of the propensity score for the DA assignment D_i.[9] $g(\mathcal{R}_i)$ and $h(\mathcal{R}_i)$ are local linear controls for

[8]It is important to note that propensity scores in this context denote the odds of being assigned to a certain type of middle school as a function of student application, priority group, and cutoffs, which have an analytic formula. We can calculate the propensity score for each middle school program for each student without relying on estimating it by imposing a parametric assumption. Since DA produces at most one assignment for each student, summing the propensity scores across middle school programs that belong to a certain group yields the propensity score of being assigned to such a group of middle schools. If a student does not apply to middle schools of a certain group, the propensity score is zero. Theorem 1 of Abdulkadiroğlu, Angrist, Narita, and Pathak (2022) 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.

[9]This is possible since the support of the propensity scores is finite. See Abdulkadiroğlu, Angrist, Narita, and

nonrandom tie-breakers at each program that uses such tie-breakers.[10] η_i and ν_i are error terms in the main and the first-stage regressions, respectively, and we report robust standard errors.[11]

To interpret β as causal, we argue that the exclusion restriction holds. That is, after controlling for propensity scores and nonrandom tie-breakers, DA assignments D_i are random and do not affect outcomes Y_i other than by affecting the actual attendance C_i.[12] To support this assumption, we provide balance test results in Appendix F.2. The instrumental variable balances the covariates of students who are assigned to the treatment middle schools by DA and those who are not, after controlling for the propensity score and nonrandom tie-breakers among students with non-degenerate risk of being offered admission (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 uses non-degenerate risk samples, controlling for propensity scores and nonrandom tie-breakers.[13]

In the following, we focus on the effect of Well-Developed middle school attendance, although the quasi-experiment is in principle at each middle school level. That is, we consider $C_i = \mathbb{1}(\text{attended any Well-Developed middle school})$ and $D_i = \mathbb{1}(\text{assigned to any Well-Developed middle school})$. This is not only because Well-Developed status is a salient and good proxy for middle school quality (Section 2), but also because many middle schools have only a handful of applicants. Appendix Table F.6 shows that the effects of attending an alternative middle school type, i.e., high-score middle schools whose average of 6th-grade students' standardized test scores is above the 66th percentile among all schools are very similar to those from our main specification.

Pathak (2022) for more details.

[10]We include a local linear function for each of 531 types of nonrandom tie-breakers in the data. We also include a set of dummy variables that corresponds to each nonrandom tie-breaker to deal with students who did not apply to a school using that nonrandom tie-breaker, or students who applied but whose tie-breakers are far from the cutoff following Abdulkadiroğlu, Angrist, Narita, and Pathak (2022). We use the IK bandwidth (Imbens and Kalyanaraman, 2012) separately for each program, as suggested by Abdulkadiroğlu, Angrist, Narita, and Pathak (2022).

[11]According to Abadie, Athey, Imbens, and Wooldridge (2023)'s framework, we observe the population of our sample and our treatment probability is determined at individual level. Abdulkadiroğlu, Angrist, Narita, and Pathak (2022) also report robust standard errors for the same reason.

[12]In principle, controlling the propensity scores $\{d_i(x)\}$ is enough for the exclusion restriction by Theorem 1 of Abdulkadiroğlu, Angrist, Narita, and Pathak (2022). We further control for student characteristics and nonrandom tie-breakers to get a more precise estimate of the treatment effect β.

[13]Such sample restriction comes with the cost of losing many observations. 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 are more likely to be White. This reconfirms that the 2SLS estimates are a local average treatment effect (LATE). Appendix Figure F.5 presents the mean difference between those with non-degenerate offer risk and degenerate offer risk.

3.2 Empirical Results

Table 2 shows that attending a Well-Developed middle school causes students to apply to and be assigned to high schools with better academic outcomes.[14][15] Each panel in the table corresponds to different high school characteristics as the outcome variable. In columns (1)-(3), we focus on the characteristics of the top-ranked high school program.[16] Column (1) presents OLS estimates, column (2) presents 2SLS estimates with the full sample, and column (3) presents our preferred specification—2SLS with the non-degenerate risk sample.

Most importantly, we find that attending a Well-Developed middle school causes students to apply to high school programs in a way that puts more weight on end-of-high school academic outcomes (panels A and B) than on student body composition (panels C and D). Students attending a Well-Developed middle school list high schools with 3.09 percentage points higher college enrollment rates as their top choices (column (3) of panel A). However, there is no such effect on the proportion of Whites in top-ranked high schools (column (3) of panel D). Panel B confirms that students from Well-Developed middle schools apply to high schools with higher value-added on college enrollment, not only the level of college enrollment rate. Appendix Table F.10 shows results with other dimensions of high school characteristics (e.g., graduation rate, % Asian), which confirms that the main results are not driven by the choice of high school characteristics.[17] [18] [19]

Next, columns (4)-(6) illustrate that attending a Well-Developed middle school also changes the assignment outcome, not only the application behavior. Attending a Well-Developed middle

[14]In Table F.6, we explore the effects of differently-defined treatment—i.e.,attending high-score middle schools, whose average of 6th-grade students' baseline test scores is above the 66th percentile across schools. The main story remains the same.

[15]We present Well-Developed middle school attendance effects on the length of the application list and the rank of the assigned school in the Table F.8.

[16]We present Well-Developed middle school attendance effects on top-three ranked and top-five ranked high school programs in Table F.7.

[17]We consider a constant-effects value-added model that controls for students' lagged test scores (Deming, Hastings, Kane, and Staiger, 2014; Chetty, Friedman, and Rockoff, 2014a,b). Details are provided in Appendix D.2.

[18]The panel D shows that OLS overestimates the effects of attending a Well-Developed middle school as one might be concerned. While students from a Well-Developed middle school apply to high schools with more White students (column (1) of panel D), 2SLS estimates show that this is not the effect of Well-Developed middle school attendance (column (3) of the same panel). Rather, it is because students who would apply to those high schools have already sorted into Well-Developed middle schools.

[19]The estimates in columns (2) and (3) differ due to changes in estimates of other covariates. The coefficients of interest for the full sample (in column (2)) vary by whether we control for other covariates, while those with the NDR sample (in column (3)) remain stable (Table F.11). This is because covariates differ between treated and untreated students in the full sample even after controlling for the full set of propensity score dummies Appendix F.2. This confirms the importance of the common support assumption, and in turn, our choice of column (3) as the most preferred specification.

Table 2: Well-Developed MS Attendance Effects on HS Application and Assignment

	(1)	(2)	(3)	(4)	(5)	(6)
Characteristics of		Top-ranked HS			Assigned HS	
Model	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Sample	All	All	NDR	All	All	NDR
Panel A: College Enrollment Rate (%)						
From Well-Developed MS	1.120***	1.820***	2.961***	2.118***	3.851***	4.628***
	(0.139)	(0.508)	(0.616)	(0.156)	(0.560)	(0.678)
N	44489	44489	6320	42055	42055	5975
R2	0.275	0.310	0.441	0.259	0.301	0.419
$\bar{y}$	73.724	73.724	76.096	66.325	66.325	69.040
Panel B: Value-added on College Enrollment Rate (%)						
From Well-Developed MS	-0.218*	0.762*	1.378***	0.262**	1.959***	2.481***
	(0.088)	(0.325)	(0.396)	(0.101)	(0.375)	(0.451)
N	44401	44401	6311	42126	42126	5985
R2	0.049	0.091	0.246	0.040	0.088	0.249
$\bar{y}$	2.748	2.748	3.381	1.164	1.164	2.335
Panel C: % High-Baseline-Score 9th-Graders						
From Well-Developed MS	2.034***	2.606**	2.480*	2.480***	3.146***	2.568**
	(0.245)	(0.858)	(1.045)	(0.210)	(0.745)	(0.847)
N	45081	45081	6393	43019	43019	6097
R2	0.342	0.392	0.547	0.363	0.420	0.546
$\bar{y}$	43.914	43.914	46.963	33.040	33.040	35.797
Panel D: % White						
From Well-Developed MS	2.137***	1.731**	-0.242	3.119***	1.778***	-0.321
	(0.163)	(0.569)	(0.614)	(0.152)	(0.520)	(0.532)
N	45081	45081	6393	43019	43019	6097
R2	0.444	0.507	0.603	0.523	0.581	0.657
$\bar{y}$	19.129	19.129	20.657	14.833	14.833	16.394
First-stage F-stat		411.74	287.42		411.74	287.42

Note: Each panel presents Well-Developed MS attendance effects on different characteristics of high schools that students first-ranked (columns (1)-(3)) or are assigned to (columns (4)-(6)). To construct the outcome in Panel C, we define students to be high-baseline-score if their standardized NYS test score is above the 66th percentile. In columns (3) and (6), we restrict the sample to students with non-degenerate risk of being offered (i.e., whose propensity score is in the interval $(0, 1)$ and hence subject to randomization) from any Well-Developed middle school. Robust standard errors are 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. Columns (2)-(3) and (5)-(6) control for dummy variables for all possible values of propensity score of being assigned to a Well-Developed MS, and local linear control for nonrandom tie-breakers. +p<0.1, *p<0.05, **p<0.01, ***p<0.001.

school changes the level and the value-added of college enrollment rate of assigned high schools, by 4.60 and 2.48 percentage points, respectively (column (6)). Notably, when compared to column (3), the effect is larger on the characteristics of assigned programs than on that of top-ranked programs. Since a student's school assignment is determined by not only her application but also priority standings at each school, this implies that attending a Well-Developed middle school changes not only how students value different programs but also how a student is ranked by programs in admissions. We also present more direct evidence that students' priority standings at high schools on their application lists improve by attending a Well-Developed middle school in the Table F.9.

Accordingly, our model in Section 4 includes a channel through which middle schools affect students' priority standings in high school admissions.

One potential channel for the effects in Table 2 is that Well-Developed middle schools increase students' test scores. For example, attending Well-Developed middle schools increases students' standardized math test scores by 0.18 standard deviation (Table 3). This not only puts students in a better position for admissions[20] but may also affect their applications, as has been widely documented in the literature that students put different weights on school characteristics based on their own test scores (e.g., Abdulkadiroğlu, Agarwal, and Pathak, 2017).

Table 3: Well-Developed MS Attendance Effects on Test Score

	(1)	(2)	(3)
Model	OLS	2SLS	2SLS
Sample	All	All	NDR
Panel A: 8th Grade Z-Math Score			
From Well-Developed MS	0.136***	0.199***	0.182***
	(0.031)	(0.033)	(0.042)
N	32841	32841	4573
R2	0.573	0.592	0.686
$\bar{y}$	0.040	0.040	0.141
Panel B: 8th Grade Z-ELA Score			
From Well-Developed MS	0.071***	0.069**	0.037
	(0.018)	(0.025)	(0.030)
N	43353	43353	6124
R2	0.610	0.623	0.685
$\bar{y}$	0.096	0.096	0.214
First-stage F-stat		411.74	287.42

Note: Robust standard errors are in parentheses. In column (3), we restrict the sample to students with non-degenerate risk of being offered (i.e., whose propensity score is in the interval $(0, 1)$ and hence subject to randomization) from any Well-Developed middle school. 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. Columns (2)-(3) control for saturated dummy variables for all possible values of propensity score of being assigned to a Well-Developed MS, and local linear controls for nonrandom tie-breakers. +p<0.1, *p<0.05, **p<0.01, ***p<0.001.

However, appendix Table F.12 shows that additionally controlling for end-of-middle-school test scores barely changes the main treatment effect of attending Well-Developed middle schools. We take this as suggestive evidence that middle schools affect students' high school application behavior through other channels beyond the evolution of test scores. Notably, since end-of-middle-school test score is an endogenous variable determined after middle school assignments, this is a "bad control"

[20]Table F.9 shows that Well-Developed middle schools cause students' priority scores for a high school to be higher. Well-Developed middle schools change priority scores not only based on test scores, but also on punctuality, GPA, and middle-school-based priority groups.

(Angrist and Pischke, 2009). Thus, we turn to a structural model to decompose the channels of middle school effects in which the model includes a separate channel through which middle schools affect students' high school application beyond the evolution of test scores.

To put our findings in context, we compare our estimates with those of Corcoran, Jennings, Cohodes, and Sattin-Bajaj (2022). They conducted a school-level randomized trial in which students attending treated middle schools were provided with information on proximate high schools, such as whether the graduation rate of the school is below the city median (75%). While treated students did not list high schools with higher graduation rates in their top 3 choices, they avoided high schools with low graduation rates (<70%) and their within-application variability in graduation rates decreased. Consequently, treated students were matched to high schools with 1.5 pp higher graduation rates. Our estimates of the Well-Developed middle school attendance effect are larger; students attending a Well-Developed middle school apply to high schools with a 1.8 pp higher graduation rate ($p<0.01$) and are assigned to schools with 3.4 pp higher graduation rate ($p<0.01$) (Appendix Table F.10).

We attribute this to the fact that our treatment, attending a Well-Developed middle school, can be viewed as a more intensive intervention. For example, middle schools may directly change students' intrinsic tastes for high schools, not only provide students with certain information about high schools. Students attending Well-Developed middle schools may acquire more information on the high school choice process itself, beyond the characteristics of high schools nearby. Middle schools also shape students' academic outcomes differently, which affects how students weigh various aspects of high schools as well as how students are ranked by high schools.

4 A Dynamic Model of Middle and High School Choices

We now turn to a two-period dynamic model of middle and high school choices. The first period corresponds to middle school applications and assignments, and the second period to high school applications and assignments.

Our findings in Section 3 can be summarized in two points. First, a student's attended middle school affects her high school application, beyond what can be explained through the change in end-of-middle-school test scores. Second, the treatment effect is larger in magnitude for assignments than for applications, which suggests that there is an additional role middle schools play through the change in students' priority standings for each high school. We incorporate these findings in our model using the following three key features.

First, the model explicitly allows students' demand 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 preparedness (Hastings, Kane, and Staiger, 2005; Abdulkadiroğlu, Agarwal, and Pathak, 2017). Furthermore, our model allows for the possibility that middle schools could also change students' high school applications through channels other than test scores.

Second, how a student is prioritized by each high school program for admissions also depends on the middle school she attended (*priority* channel). Attending different middle schools may result in different end-of-middle-school test scores, which in turn affects students' admissions chances at high school programs that use test scores for admissions. In addition, some high schools give eligibility/priority depending on which middle school a student attends (Table 1).[21]

Third, students may be forward-looking: 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 process from attending a particular middle school, which in turn affects how they value different middle school programs. Note that while it is not standard to model changes in demand over time, we view our model as capturing the shift in decision-making authority from parents to students between middle and high school choice, i.e., knowing that their children will be in charge of the decision in 3-years, parents apply to middle schools while considering how their children's demand will be shaped by middle schools. Although it goes beyond the scope of this paper to fully model the within-household decision, previous studies have documented that the interaction between parent and child plays an important role in educational choices (Giustinelli, 2016).

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. The effect we identified in the previous section is marginal for each *treated* student. We aim to examine the effects of counterfactual admissions policies and how they interact across schooling levels, which will trigger a change in the behavior of *all* students, and in turn, change 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 for decomposing the channels through which middle

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

schools affect high school choice and quantifying each channel's relative importance.

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 exante (e.g., the priority group a student belongs to). Each student with intrinsic priority realizes an **expost score** at each program that is used by programs to rank students for admissions. For example, in our context, the expost 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 **expost cutoff** is defined as the lowest expost score of the admitted students if the seats are filled and $-\infty$ otherwise. Finally, a program is **feasible** for the student if she has a higher expost score than the expost cutoff, regardless of whether or where in the list she ranks it.[22]

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 $m \in \{1, \cdots, M\} = \mathcal{M}$, and high school programs as $j \in \{1, \cdots, J\} = \mathcal{J}$. We start from period 2 and work backward.

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 perceived utility V_{ij} from enrolling in high school program $j \in \mathcal{J}$,

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

where $\tilde{X}_j$ is the vector of observable characteristics of high school program j, $Z_i^H(m(i))$ 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 $\tilde{\xi}_j$ captures the unobserved vertical quality of j. η_{ij} is an idiosyncratic preference shock that is iid for each i and j. As noted earlier, V_{ij} may depend on the student's middle school $m(i)$ through the change in observable characteristics $Z_i^H(m(i))$ as well as other channels.

Student i has perceived utility V_{i0_h} from the outside option 0_h. The outside option includes private schools, homeschooling, public charter schools, etc.

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

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

Behavioral Assumption In a school choice situation, each student is playing an incomplete information game: Each student's assignment is uncertain from a student's point of view because it is determined not only by her ROL and priorities but also by (i) other students' ROLs and priorities and (ii) tie-breaking lottery realizations, which are both unknown exante. Hence, we need an assumption on the equilibrium behavior of students to interpret the school choice data we observe. We assume that in both periods 1 and 2, students submit ROLs that satisfy **expost stability** in each realization of the uncertainty (Che, Hahm, and He, 2023).[23] Namely, the assigned program of a student is her favorite program among the expost feasible programs.[24] Expost stability is consistent not only with the implication of the truth-telling assumption based on the strategyproofness of DA (e.g., Abdulkadiroğlu, Agarwal, and Pathak, 2017) but also with students' deviations from truth-telling even in a strategyproof environment.[25] Therefore, we use expost stability as our preferred assumption since it imposes a weaker assumption on student behavior.

Expost 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 her ROL. Without expost stability, we need to fully solve the game of incomplete information where each student is facing by enumerating all possible ROLs and finding the optimal strategy profile among them, which would render estimation of the model difficult in terms of computation. Expost stability essentially enables us to interpret the data using a conditional multinomial choice model, in which a student's *choice* is the assigned program, and the *choice set* is the expost feasible set.[26] Furthermore, it helps us simplify the continuation value of a given middle

[23]Che, Hahm, and He (2023) show that a robust equilibrium—a weakening of the weakly dominant strategies equilibrium that allows students' deviations from truthful reporting—satisfies asymptotic stability. This means that as the size of the economy grows, the proportion of students who are assigned their favorite feasible school given each realization of the uncertainty (e.g., tie-breaking lottery) they face converges in probability to 1. Expost stability is implied by asymptotic stability in a large market.

[24]Asymptotic stability (and hence expost 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 a 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 7.6 high school programs, which is 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 small (6.5%). 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.

[25]Such deviations are often regarded as strategic *mistakes* in the literature. See Larroucau and Rios (2020a); Artemov, Che, and He (2021); Hassidim, Romm, and Shorrer (2021); Shorrer and Sóvágó (2023) for examples of such mistakes in real-world and lab experiment settings.

[26]The exogeneity of the choice set is satisfied by assuming a large market—i.e., the market is large enough that each

school to what is known as the "Emax" term in the dynamic discrete choice literature, as will be seen 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 second-period payoffs. Hence, we need to model how she forms expectations on the "continuation value" of each middle school program.

The key concept is expost stability. Due to expost stability, the exante uncertainties that determine the expost scores and cutoffs (in our context, the tie-breaking lottery draws) are 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 expost scores and submitted ROLs, creating expost cutoffs of high school programs. Expost stability implies that each student is assigned to her favorite high school program among the expost feasible high school programs, and hence knowing the lottery realizations is sufficient to know each student's payoff at the assignment.

To this end, let ω denote the uncertainty that determines the expost scores and cutoffs in the second period (high school application) with some known distribution $H(\omega)$ where ω is unknown exante. 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 expost feasible set of high school programs given 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.[27]

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

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

student cannot affect the cutoffs.

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

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 ξ_m captures the unobserved vertical quality of m. 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 calibrate to a specific number. There is also a middle school outside option 0_m, whose utility is denoted as U_{i0_m}. Similar to high school choice, the outside option includes private schools, homeschooling, public charter schools, etc.

Note that U_{im} includes the continuation value of attending m in addition to the flow utility of attending m. By expost 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 expost 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. Table 4 summarizes what is known to student i in each period.[28] Appendix D.3 provides assumptions on the unobservables and explains how those assumptions help simplify the expression of the continuation value when combined with the expost stability of high school match outcome. This simplification is the key trick that enables us to estimate the multistage school choice model.

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

4.1.3 Equilibrium

We assume a large market and define an equilibrium using the uniqueness of stable matching in a large market (Azevedo and Leshno, 2016; Che, Hahm, and He, 2023).[29] An equilibrium is a tuple

[28]We assume high school program characteristics are exogenous and fixed which 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 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.

[29]The demand and supply framework using stable matching was pioneered by Azevedo and Leshno (2016). Che, Hahm, and He (2023) consider a weaker equilibrium concept called robust equilibrium and establish the uniqueness of stable matching in a continuum economy, as well as the asymptotic stability result, which is our key assumption of student behavior.

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 $X_m, \tilde{X}_j, \xi_m, \tilde{\xi}_h$	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 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 obtain by attending each m.

of cutoffs $\{c_j^H\}$, $\{c_m^M\}$ and a tuple of decisions $\{m_i\}$, $\{j_i\}$ that clears the market:

1. Given cutoffs, $S_m = D_m(\{c_j^H\}, \{c_m^M\})$, where
$$D_m(\{c_j^H\}, \{c_m^M\}) = \int_i 1(m_i = argmax_{m' \in O_i^M(\{c_m^M\};\rho^M)} U_{im'}(\{c_j^H\}))di. \tag{4}$$

2. Given cutoffs and $\{m_i\}$ that satisfies the above, $S_j = D_j(\{c_j^H\}, \{c_m^M\})$, where
$$D_j(\{c_j^H\}, \{c_m^M\}) = \int_i 1(j_i = argmax_{j' \in O_i^H(\{c_j^H\}, m_i;\rho^H)} V_{ij'}(m_i)di, \tag{5}$$

where S_m and S_j are the capacities of middle school program m and high school program j, respectively, O_i^M and O_i^H are student i's feasible sets depending on the cutoffs $\{c_j^H\}$, $\{c_m^M\}$ and the lottery draws ρ^M, ρ^H. Note that O_i^H and $V_{ij'}$ are functions of the middle school attended m_i and that $U_{im'}$ is a function of $\{c_j^H\}$ through its dependence on the expected utility from high school choice. Using the large market assumption, we can calculate demand as the sum of the probability of school j being the most preferred feasible one across all students. For example, the aggregate demand for high school j is
$$D_j = \Sigma_i \mathbb{1}(j \in O_i(\{c_{j''}^H\}) \int_\rho \frac{exp(v_{ij}(\cdot))}{\Sigma_{j' \in O_i\{c_{j''}^H\}} exp(v_{ij'}(\cdot))} d\rho. \tag{6}$$
$\{c_{j''}^H\}$ is high school cutoffs. The demand for middle schools is defined analogously.

4.2 Estimation

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

First, the flow utilities in each period are

$$u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) = \tilde{u}\left(X_m, Z_i^M, \gamma_i^M\right) - \lambda^M d_{im} \tag{7}$$

$$= X_m' \beta_i^M - \lambda^M d_{im} \tag{8}$$

$$v\left(\tilde{X}_j, Z_i^H, \tilde{d}_{ij}, \gamma_i^H; m(i)\right) = \tilde{v}\left(\tilde{X}_j, Z_i^H, \gamma_i^H; m(i)\right) - \lambda^H \tilde{d}_{ij} \tag{9}$$

$$= \tilde{X}_j' \beta_i^H + \alpha_{\tau(m(i))} - \lambda^H \tilde{d}_{ij}, \tag{10}$$

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. $\alpha_{\tau(m(i))}$ allows a student who attended a middle school of a certain type to assign a different overall attractiveness value to schools in the choice system, where $\tau(m(i))$ is the type of i's middle school attended $m(i)$. 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. Also, we assume that $(\gamma_i^M, \epsilon_{im}) \perp d_{im} \big| X_m, Z_i^M, \xi_m$ and $(\gamma_i^H, \eta_{ij}) \perp \tilde{d}_{ij} \big| \tilde{X}_j, Z_i^H, \tilde{\xi}_j, m(i)$, 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).

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 characteristic, we parameterize the random coefficients as

$$\beta_{i,l}^M = Z_i^{M\prime} \beta_{Z,l}^M + \gamma_{i,l}^M$$

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

for each $l = 1, 2, \cdots, L$. 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*. 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 role similar to the treatment effect β in Equation (1) when the outcome variables are the characteristics of the programs students applied to. We discuss in more detail what channels this "lumpsum" parameter could capture in Section 4.3.

$\gamma_i^M = (\gamma_{i,1}^M, \cdots, \gamma_{i,L}^M)$ 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 γ_i^M and γ_i^H are jointly distributed according to some discrete distribution G. We do not impose any structure on G, and hence γ_i^M and γ_i^H can be arbitrarily correlated.[30]

Next, we assume the unobserved vertical qualities ξ_m and $\tilde{\xi}_j$ depend on the programs' borough, admissions method, and program focus area. That is,

$$\begin{aligned}\xi_m &= \xi_{1,b(m)} + \xi_{2,a(m)} + \xi_{3,f(m)}\\ \tilde{\xi}_j &= \tilde{\xi}_{1,b(j)} + \tilde{\xi}_{2,a(j)} + \tilde{\xi}_{3,f(j)},\end{aligned}$$

where $b(\cdot), a(\cdot), f(\cdot)$ denote a program's borough, admissions method, and focus area, respectively.

We further 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, this implies that the continuation value expression can be further simplified to a convenient form (see Appendix D.3).

Finally, we model the utility from the outside options as follows.

$$\begin{aligned}U_{i0_m} &= \vartheta_{mi} + \epsilon_{i0_m} && (11)\\ &= \alpha_0^M + Z_i^{M\prime}\alpha_Z^M + \epsilon_{i0_m} && (12)\\ V_{i0_h} &= \vartheta_{hi} + \eta_{i0_h} && (13)\\ &= \alpha_0^H + Z_i^{H\prime}\alpha_Z^H + \eta_{i0_h} && (14)\end{aligned}$$

where $\epsilon_{i0_m}, \eta_{i0_h}$ both follow EVT1. Thus, we allow the attractiveness of the outside options to vary by students' observed heterogeneity. In practice, to keep the computation manageable, we set Z_i as ethnicity. To be clear, α_τ in Equation (10) captures the fact that middle schools can change the propensity to opt out of students' high school assignments.[31]

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 (γ_i^H). The data show a large correlation between the high school characteristics a student applies

[30] In practice, we estimate four types in total, with two types for γ_i^M and γ_i^H, respectively. The main concern regarding having a small number of types is that we may not sufficiently control for the unobserved heterogeneity in school demand by doing so. However, our model estimates fit the data well (Section 4.4.2). In particular, our estimates generate a Well-Developed middle school attendance effect of a magnitude similar to that from the reduced-form analysis. In addition, in Appendix E, we show that average treatment effect estimates from the model with nine types—and thus three types each for γ_i^M and γ_i^H—are very similar to those from our main specification.

[31] For the purpose of normalization, we can have such a term only in either Equation (10) or Equation (14). Note that our model assumes away from non-compliance to another school within the system while incorporating non-compliance by choosing outside options.

and is assigned to and the characteristics of the middle school 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 as seen in our reduced-form analysis in Section 3 (see Table 2). 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 explanations 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 is identified by the degree to which the same student's middle and high school applications are similar after controlling for her observable characteristics.

Next, the identification of ρ_τ relies on the quasi-random variation in school assignments generated by the tie-breaking rule. ρ_τ is identified by how similar the high school applications are across students attending middle schools of the same type, and 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 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] The quasi-random assignments together with the distributional assumption on the unobserved tastes generate variations in which type of middle school a student attends beyond what can be explained by students' observable characteristics and unobserved tastes. Instead of explicitly targeting the reduced-form estimates in the model estimation, we use the reduced-form estimates to validate our model by comparing them with the Well-Developed middle school attendance effects simulated from our model (Section 4.4.2).

Meanwhile, the marginal distribution of unobserved tastes (γ_i^M and γ_i^H) is identified by the variation in the choice sets across students, together with the distributional assumption (logit error) on the idiosyncratic preference shocks, ε and η. The choice sets (=feasible set) are assumed to be exogenously given to individual students given the assumption that the market is large. Then to what extent choices of observably similar students with different choice sets deviate from the independence of irrelevant alternatives assumption identifies the unobserved heterogeneity (Train, 2009).[33]

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

[33]While fully leveraging ROL would have enabled us to nonparametrically identify the marginal distribution of unobserved heterogeneity, we have had to impose the truth-telling assumption.

We set $\delta = 0.75$. An alternative version with $\delta = 0.9$ yields a similar model prediction. Identifying the discounting factor is known to be hard (Rust, 1987), and similarly, in our case, we lack variation that changes the continuation value independent of middle schools' observed and unobserved characteristics. Varying the value of δ mainly changes the coefficient of middle school characteristics, β_i^M, but the model prediction—the average characteristics of assigned middle schools by students' observable characteristics—remains the same, as shown in Appendix E.3. Given the limitation that we do not identify δ, we consider counterfactual scenarios in which high school admissions reform is announced after students choose middle schools in Section 5.2.[34]

Estimation Our estimation sample consists of 30,968 students who participated in the 2014-15 middle school application process, and the 2017-18 high school application process 3 years later. For middle school program characteristics, we include a dummy variable of being Well-Developed according to the Quality Review, the proportion of high-performers, the proportion of White/Asian students, and the proportion of non-Free/Reduced-price Lunch (FRL) eligible students. We construct the last three variables from the characteristics of the 6th graders in 2014-15, the previous application cohort. For high school program characteristics, we include the college enrollment rate of the 2016-17 graduation cohort, the proportion of high-performers, White/Asian students, and non-Free/Reduced-price Lunch (FRL) eligible students among the previous application cohort. For the student characteristics, we use ethnicity dummy variables (Black or Hispanic), FRL status, and the average of most recent math and ELA standardized test scores (normalized to mean 0 and std 1).

We aim to jointly estimate all stages of the model to address the serial correlation in unobserved middle school demand and high school demand (γ_i^M, γ_i^H). To circumvent the computational burden of full information maximum likelihood, we employ the expectation-maximization algorithm with a sequential maximization step proposed by Arcidiacono and Jones (2003). In summary, the idea is to (1) reformulate the full information likelihood function into additive separable terms, each of which represents the likelihood of each stage; (2) update estimates of each stage; and (3) iterate the procedure until convergence. Appendix D provides more details on the procedure we use to estimate our model.

[34]In Online Appendix E.1, we also estimate a static model without the dynamic components of our main model. In particular, δ is set equal to 0 so that students are myopic. Since the static model is a nested model of the full dynamic model, we can perform a likelihood ratio (LR) test, which strongly rejects the static model in favor of our main dynamic model ($p < 0.001$).

4.3 Discussion

Middle School Effect Parameter ρ_τ We allow middle schools to affect students' high school applications through channels other than test scores, which we capture as a lumpsum parameter ρ_τ. Specifically, ρ_τ captures the changes in the weight students assign to high school characteristics as functions of the middle schools attended. Potential channels include changes in preference, information friction on school characteristics and admissions probability, and limited attention, each of which has been studied in the literature (Kapor, Neilson, and Zimmerman, 2020; Arteaga, Kapor, Neilson, and Zimmerman, 2022; Lee and Son, 2022; Campos, 2023). Further distinguishing among these possibilities goes beyond the scope of this paper, since we are interested in the effect of middle schools as a package, and we conduct a set of decomposition and counterfactual exercises that change the allocation of middle school seats rather than specific aspects of middle schools. One potential channel we expect to be of little importance is that students might apply to high schools in order to attend the same high school as their middle school peers. Evidence from surveys and regression analysis shows that this is not a main priority for students when they apply to high schools (Mark, Corcoran, and Jennings, 2021).

Value-added Estimation We also allow middle schools to affect their high school applications through the change in baseline test scores. Appendix D.2 provides details on how we estimate each middle school's production function of Z_i^H using a value-added model. We estimate the value-added via OLS, due to the fact that we are underpowered to leverage lottery variation for many middle schools because the number of applicants with admissions probability strictly between 0 and 1 (non-degenerate risk sample) for each school is often small. However, we find that the OLS estimates and the lottery estimates for a subset of schools are reasonably correlated (=0.56 (p<0.001) for math score; details are in Appendix D.2). We also assume a student has perfect foresight on what Z_i^H she will have as a result of attending each m.

Expost Stability By using expost stability in our model, we implicitly assume that students are aware not only of all the options and their attributes but also the distribution of admissions probabilities at each program. However, in real life, students' preferences on programs and also middle school type effects may operate through information frictions (Luflade, 2017; Neilson, Allende, and Gallego, 2019; Lee and Son, 2022). As a result, we follow Allende (2019) and do not interpret our parameter estimates as deep structural preferences but as weights students assign on school attributes. Since it is unlikely that these weights will change under the counterfactual scenarios we consider in Section 5.2, the model can be used to predict behaviors. On the other hand, we do not focus on welfare analysis for the reason explained above.

Residential Choice We assume that households' residential locations, and thus distances to schools, are uncorrelated with unobserved heterogeneity conditional on their observable characteristics. Recent papers using a more unified framework of a family's residential and school choice have explored this possibility (Agostinelli, Luflade, and Martellini, 2021; Park and Hahm, 2023). The most relevant concern for the purpose of this paper is that households' moving decisions might also be affected by which middle school students attend. If this is the case, the application channel would be affected by the additional force of changing locations. Table F.14 shows that Well-Developed middle schools do not shift households' propensity to move across boroughs, school districts, or Census tracts.[35]

Preference for Peers While the analysis of peer effects in both value-added and school demand is interesting, it is not the main focus of our model. Rather, we use the previous cohort's composition as school characteristics following the literature (Abdulkadiroğlu, Agarwal, and Pathak, 2017; Calsamiglia, Fu, and Güell, 2020).[36] Endogenizing peer composition in students' school demands particularly complicates the counterfactual analysis, since it involves the multiple equilibria issue. While this assumption does not cause concern in interpreting our middle school effects and decomposition results, predictions for the counterfactual policy might be of concern. Accordingly, we interpret our counterfactual exercise as partial equilibrium changes although we endogenously solve new admissions cutoffs.

4.4 Results

4.4.1 Model Estimates

Table 5 provides the selected model estimates where the full set of estimates are reported in Appendix Table F.13. 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 Well-Developed middle school causes a student to be willing to travel 0.14 miles more to attend a high school with one standard deviation increase (=14.4 pp) in the college matriculation rate, which amounts to a 6.4% increase

[35]Meanwhile, overall mobility is quite high; 23% of households changed their residential Census tracts in between middle and high school application processes. Thus, we measure the distances to middle and high schools from students' residential Census tracts in their 5th and 8th grades, respectively.

[36]Allende (2019); Idoux (2022) are few exceptions.

from the average commuting distance of high school students.[37] Meanwhile, students who attend a Well-Developed middle school consider outside options, either public charter schools or non-public schools, to be more attractive. ($\alpha_\tau = -0.939$).[38] Also, as expected, non-minority students have higher values of the outside options in both middle and high school choices (e.g., private schools, homeschooling, public charter schools.)

Table 5: Selected Demand Estimates

	Middle Schools		High Schools	
	est	se	est	se
Panel A: Middle School Type Effects				
Attending Well-Developed Middle School				
Proportion of White/Asian/Other			0.441	(0.082)
Proportion of non-FRPL			0.458	(0.149)
Proportion of High-Performers			1.151	(0.145)
Proportion of College Enrollment Rate			0.471	(0.113)
Panel B: Other Parameters				
α_τ			-0.939	(0.071)
Outside option (White/Asian/Other)	5.291	(0.123)	6.016	(0.095)
Outside option (Black/Hispanic)	3.378	(0.084)	4.049	(0.091)
Distance	-0.655	(0.023)	-0.556	(0.004)

Note: We report a selected subset of preference estimates in Appendix Table F.13. School characteristics "Proportion of White/Asian/Other," "Proportion of non-FRPL," "Proportion of High-Performers," "College Enrollment Rate" are between 0 and 1. Standard errors acquired from 50 bootstrap samples are reported in parentheses.

The coefficients that represent the characteristics of middle and high schools, along with observed heterogeneity, align with anticipated patterns. Specifically, there is pronounced demand among students for middle schools categorized as Well-Developed and for high schools that exhibit higher college matriculation rates. There is also a marked preference for schools that boast a student population with baseline test scores above the 66th percentile. Notably, this tendency is more pronounced among applicants who themselves have higher baseline test scores.

[37] We report the willingness to travel by dividing the coefficient of interest by the coefficient of distance. The average commuting distance to each assigned high school in the data is 2.24 miles.

[38] One might think that this can be attributed to Well-Developed middle school graduates' being more often assigned to specialized exam high schools, rather than public charter schools or non-public schools. There are 8 specialized high schools in the city (e.g., Stuyvesant and Bronx High School of Science) excluding Fiorello H. LaGuardia High School of Music & Art and Performing Arts, and they admit students based on the Specialized High Schools Admissions Test (SHSAT) score. These specialized high schools have their own centralized school assignment system, separate from the regular high schools that comprise our estimation sample. By running Equation (1), we find no evidence that Well-Developed middle schools cause students to take the SHSAT more or get an offer from specialized high schools more. Results are presented in Table F.15

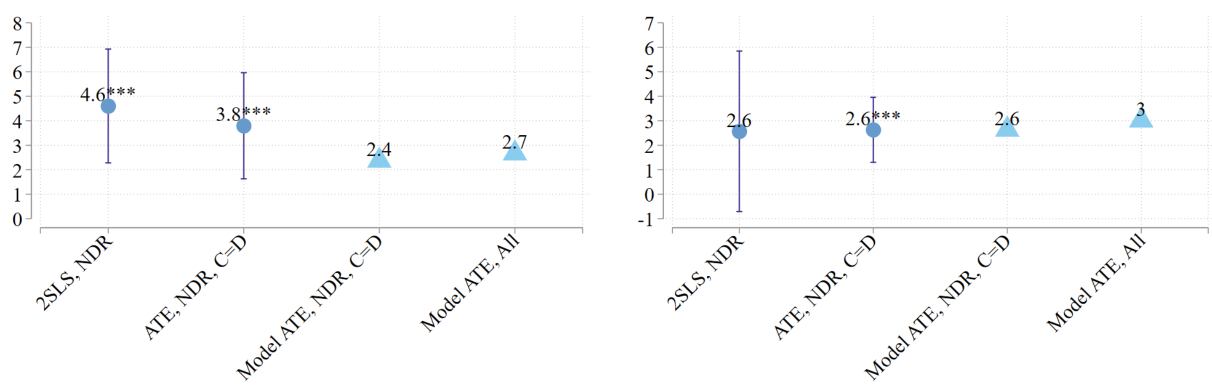
4.4.2 Model Validation

We evaluate how well the model fits the observed data by comparing measures calculated using the data with those calculated using the simulations based on our model estimates along two dimensions: the treatment effects of middle school attendance on high school assignments and the average characteristics of assigned programs.[39] We find that 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 counterfactual policies in Section 5.2.

Treatment Effects of Middle School Attendance on High School Assignments The effects of attending Well-Developed middle schools based on our model simulation closely match the causal estimates from quasi-random assignment using data. The leftmost numbers (**2SLS, NDR**) in each subplot of Figure 1 are the 2SLS estimates from Table 2 panel A, column (6), which is essentially the local average treatment effect (LATE) among compliers—i.e., students who attended Well-Developed middle schools because they were assigned to those.

Figure 1: Effects of Middle Schools on High School Assignments: Data vs. Model

(a) % College of High School (pp) (b) % High-performer of High School (pp)



Note: The figure plots middle school effects on high school assignments from the data and the simulation. Figure (a) plots the effect on the college enrollment rate of assigned high schools. Figure (b) plots the effect on the proportion of high-performers of assigned high schools. Students are labeled to be high-performing if their standardized test score is above the 66th percentile of their cohort in the system. In each subfigure, we plot 2SLS estimates, average treatment effects (ATE) on the subsample, simulated average treatment effect (ATE) on the subsample, and simulated average treatment effect (ATE) of all students. We explain in the text which students are included in the subsample and how we calculate the ATE of this subsample from the data. We present 95% confidence intervals of treatment effects calculated from the data. For the model simulation, we assign each student to her most preferred Well-Developed middle school and non-Well-Developed middle school, respectively, and compare the assigned high schools' characteristics in each scenario. We simulate over 10,000 high school lotteries for each scenario.

In each subplot, the second number (**ATE, NDR, C=D**) is the average treatment effect (ATE) for a subset of students for whom we can calculate it as the inverse probability weighting (IPW) estimate.

[39]We do not calculate the effects on high school applications, since our model does not pin down a unique strategy (i.e., application) of a student as long as it satisfies expost stability.

The subset includes students whose Well-Developed middle school admissions probability is strictly between 0 and 1, and who enrolled in a Well-Developed middle school after being assigned. (see Appendix D.5 for more detail).

Then, for the same subset of students, we simulate the ATE using our model estimates. For each student, we counterfactually assign one student at a time to her most preferred Well-Developed middle school and non-Well-Developed middle school, respectively, and compare the high school assignment results. We calculate the average difference across students and label it as "**Model ATE, NDR, C=D**".[40] The third column in Figure 1 plots the model ATEs. Both for college enrollment rate and proportion of high-performers, the model ATEs are within the confidence interval of ATEs from the data. This confirms that our model fits the data well, even though we do not directly target the ATEs from the data in the model estimation.

Furthermore, we can calculate the ATEs for the whole sample using the model estimates because we model the unobserved heterogeneity and its serial correlation. The fourth column illustrates ATEs for the whole sample (**Model ATE, All**) and we find that they are in magnitude similar to the ATEs for the subset. In the following Section 5.1, we decompose the source of this ATE for the whole sample.

Average Characteristics of Assigned Programs In Table 6, we also calculate the average characteristics of assigned programs, for the entire sample and by students' observable characteristics. We find that the average characteristics of the assigned schools for each type of student are very similar between the data and model simulations for both middle and high schools.

5 Sources of Middle School Effects and Policy Analysis

Using model estimates, we conduct a decomposition exercise to quantify the relative importance of the *application* channel and the *priority* channel, through which the middle school effects on high school assignments operate in our model. This is essentially a partial equilibrium exercise in which we switch the middle school assignment of one student at a time, holding other students fixed.

Next, we conduct a series of counterfactual policy analyses in which we compare the equilibrium effects of interventions with different timings. We consider that a policy reform would induce changes in many students' middle and high school applications/assignments, and thus affect middle

[40] Essentially, this procedure treats each student 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. Exogenously assigning a student to a benchmark school enables us to be free of students' sorting into middle schools based on unobservables. These facilitate the interpretation of the average difference we calculate as ATE.

Table 6: Goodness of Fit

	All		By Racial Group				By FRL Status			
			Black/Hispanic		White/Asian/Other		FRL		Non-FRL	
	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
Panel A: Middle School Characteristics										
Proportion of White/Asian/Other	0.33	0.32	0.19	0.19	0.61	0.60	0.28	0.28	0.48	0.45
Proportion of Non-FRL	0.24	0.23	0.19	0.19	0.34	0.33	0.21	0.20	0.35	0.33
Proportion of High Performers	0.30	0.29	0.23	0.23	0.44	0.44	0.27	0.27	0.40	0.39
$\mathbb{1}$(Grade A)	0.25	0.25	0.21	0.20	0.35	0.35	0.24	0.24	0.30	0.29
Panel B: High School Characteristics										
Proportion of White/Asian/Other	0.30	0.26	0.20	0.18	0.51	0.45	0.27	0.24	0.42	0.32
Proportion of Non-FRL	0.21	0.18	0.17	0.16	0.28	0.23	0.19	0.17	0.28	0.22
Proportion of High Performers	0.27	0.25	0.23	0.22	0.36	0.34	0.26	0.24	0.33	0.29
College Enrollment Rate	0.62	0.57	0.59	0.55	0.67	0.62	0.61	0.57	0.66	0.58

Note: For model-based simulations, we report the average result from multiple simulations using 100 draws of middle school lotteries and 10,000 high school lotteries. For a given draw of the lottery, we assign a student to her most preferred feasible middle school and high school, respectively. Students are labeled as high-performing if their standardized test score is above the 66th percentile of their cohort in the system.

and high school equilibrium cutoffs. Therefore, we recalculate the equilibrium using our model estimates. Meanwhile, we assume that students apply to middle and high schools taking school characteristics as fixed. Thus, we interpret our predictions as short-run effects.

5.1 Decomposition of Effects of Middle Schools

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 assign students to their most preferred eligible non-Well-Developed middle schools and change their assignment to their most preferred eligible Well-Developed middle school, one student at a time. Then we simulate their high school assignments in each case in the following alternative scenarios.

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

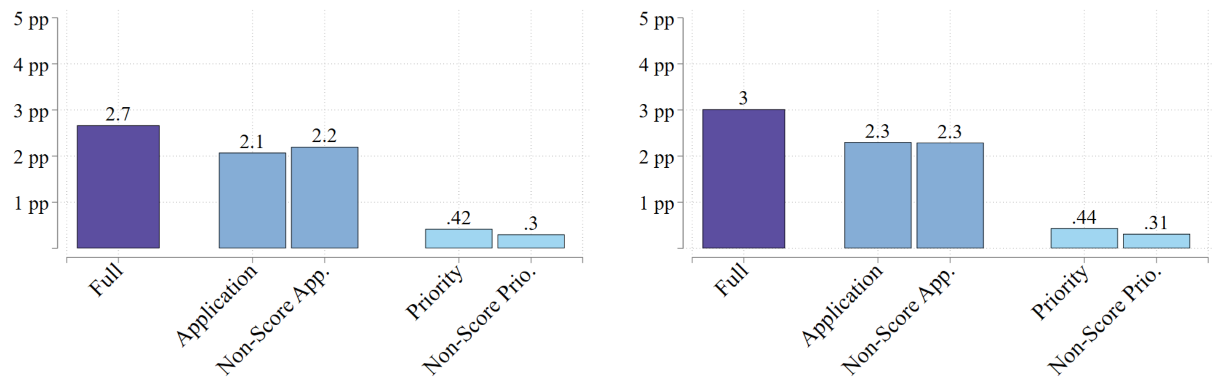
2. **Application**: shut down the *priority* channel. That is, we do not allow a student's priorities at each high school to change depending on the middle school she attends.

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 track how students' high school assignments change compared with when they attend a non-Well-Developed middle school in each scenario. We first evaluate the effect in **Full** (the total effect of exogenously changing middle schools). Note that this corresponds to **Model ATE, All** in Figure 1. We then investigate to what extent that effect can be explained by the *application* channel

(**Application**) or by the *priority* channel (**Priority**). Similar to calculating the model ATEs in Section 4.4.2, this procedure treats each student as a "price-taker" who takes the current equilibrium as given, and hence gives us the interpretation as a partial equilibrium exercise.

Figure 2: Decomposition of Effects of Middle Schools on High School Assignments
(a) % College of High School (pp) (b) % High-performer of High School (pp)



Note: We report the decomposition of middle school effects on high school assignments using the model estimates in Table F.13. We assign each student to her most preferred Well-Developed and non-Well-Developed middle school, respectively, and simulate her high school assignment in each case. 10,000 high school lotteries are drawn and we plot the average across simulations.

Figure 2 reports the results. We find that the *application* channel is quantitatively more important than the *priority* channel. In Figure 2(a), we find that about 80% of the full ATE effect on the college enrollment rate of assigned high schools can be explained by the *application* channel. In comparison, the *priority* channel only explains about 15%. In the third and the fifth bars, we further shut down changes in the end-of-middle-school test scores for the *application* and *priority* scenario, respectively. Overall, the results validate that the modest impact of test scores observed in our reduced-form section (Table F.12) also extends to the ATE. The third bar illustrates that the influence stemming from the *application* channel primarily originates from the non-score middle school effect ρ_τ. Likewise, approximately 70% (0.3/0.42) of the influence attributed to the *priority channel* is ascribed to the non-test score aspect—i.e., high schools granting direct admissions priority to graduates from specific middle schools. The overall results are very similar when we use other high school characteristics, such as the proportion of high-performers as illustrated in Figure 2(b).

5.2 Policy Analysis

The decomposition results suggest that the effects of middle schools primarily manifest through the *application* channel, which implies that reforms in middle school admissions could influence high school matching outcomes by altering how students apply to high schools. This possibility is examined in this current section.

Outcomes of Interest We examine the effects of potential policy changes on three key metrics: the average college enrollment rate of the high schools to which students are assigned, the disparity in this enrollment rate linked to students' free or reduced-price lunch (FRL) status, and the disparity in the same measure between Black/Hispanic and White/Asian/Other students. The latter two metrics are chosen due to the pronounced disparities in high school quality attended by students based on ethnicity and socioeconomic status in this context (see Table 6 for more details).

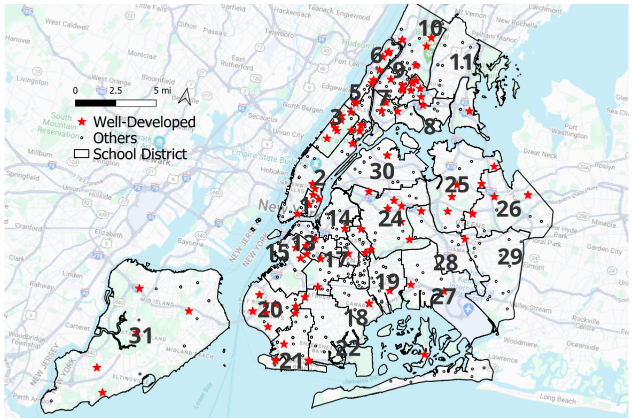
Importantly, these gaps may originate as early as middle school placements. Our data show that only 21% of Black and Hispanic students enroll in a Well-Developed middle school, compared with 35% of their White and Asian peers. Likewise, 24% of FRL students attend a Well-Developed middle school, in contrast to 30% of non-FRL students. These statistics highlight the importance of understanding how such trends might evolve under different policy interventions.

Counterfactual Policies We examine a counterfactual policy in which the city abolishes the eligibility criteria for selected middle and high schools—specifically, Well-Developed middle schools and high schools with college enrollment rates above the 66th percentile. Under this policy, students who would have been ineligible in the baseline scenario are considered the lowest priority group for these targeted schools. Within our study period, 9% of targeted middle schools and 85% of targeted high schools were accessible to students from either the borough or the city as a whole (Table 1), with the remainder open exclusively to those within certain school districts or attendance zones.[41]

The fact that even academically distinguished schools are not operating at full capacity suggests that this counterfactual policy could lead to gains in the average quality of the schools students attend. During the period we studied, Well-Developed middle schools had 4.5 vacant seats on average, and accounted for a total of 530 vacant seats across such schools. These vacancies may persist under the current system due to the eligibility restriction, even though there are students who, if eligible, would likely apply and gain admission to these Well-Developed schools. Crucially, our findings in previous sections suggest that, when more students attend Well-Developed middle schools, they in turn apply to high schools with higher college matriculation rates. This higher demand for such high schools can be reflected in student placements, given that these high schools also had on average 4.5 empty seats, totaling 3,097 seats citywide.

[41] See Appendix A for further information on geographic divisions within the city.

Figure 3: Location of Well-Developed Middle Schools



Note: The figure plots the spatial distribution of all middle schools in 2014-15. Red stars denote Well-Developed middle schools.

Also, the unequal spatial distribution of Well-Developed middle schools suggests that citywide middle school choice could reduce the disparity in school characteristics among students from different socioeconomic backgrounds. As illustrated in Figure 3, some districts, such as Districts 3, 20, or 25, have a higher number of Well-Developed schools, while areas like the northern Bronx and southern Brooklyn and Queens have fewer of these schools. Students from underrepresented districts could gain access to Well-Developed middle schools outside of their districts under the policy, and subsequently their high school match outcomes could change.

Notably, the NYC Department of Education (DOE) plans to implement this policy at the middle school level starting with the 2025-26 application cohort, motivated by these potential gains.[42]

Our subsequent analysis aims to gauge the potential changes in high school matching outcomes driven by this comparatively moderate policy shift, by experimenting with the timing of the policy's implementation. Specifically, we assess the impact of three intervention scenarios with alternative timings:

1. **HS**: Remove eligibility rules for high schools only.

2. **MS**: Remove eligibility rules for middle schools only.

[42] Since this proposal does not alter the priorities of currently eligible students, it has been met with relatively less opposition than more extensive reforms, such as abolishing all admissions rules. Prompted by this, the city introduced a scaled-down version in Bronx middle schools in 2019. To be specific, the city abolished school-district eligibilities and opened all middle schools in the Bronx to all Bronx students (Zingmond, Laura, Bronx Middle School Best Tets, InsideSchools, October 20, 2020).

3. **MSHS:** Remove eligibility rules for both middle and high schools.

For each scenario, we determine a new equilibrium based on our model estimates and compare how students' high school assignments change compared to the status quo (**Current**). In each scenario, for targeted schools using nonrandom tie-breakers, we predict the values of a tie-breaker for the newly eligible students using the methodology outlined in Appendix D.

We proceed under the assumption that any reform of high school admissions is announced only after students have submitted their middle school applications. This is because we do not have a causal estimate of δ, the parameter that quantifies the importance of the continuation value relative to other middle school attributes. Also, we assume that school characteristics remain constant under the current system (**Current**), which means that our predictions should be interpreted as indicative of short-term effects only.

Equilibrium Simulation and Selection For each counterfactual scenario, we recalculate the admissions cutoffs. We search for middle and high school cutoffs such that the demand for each school is weakly smaller than the supply according to our notion of equilibrium in Section 4.1.3. Specifically, we search for middle and high school cutoffs as fixed points of Equation (6) (and of its analogous version for middle schools).

Even though we do not uniquely pin down students' equilibrium strategies, our large market assumption, together with the uniqueness of stable matching in a large market (Azevedo and Leshno, 2016; Che, Hahm, and He, 2023), uniquely predicts the matching outcomes, which are our main interest.

Results In Table 7, we present the main characteristics of middle and high schools that students are assigned to in the baseline scenario in column (1), and the percent change in the outcomes of interest in each policy scenario in columns (2)-(4).

First of all, we find that the middle school-only admissions reform (**MS**, column (3) of table Table 7) changes high school match outcomes almost as effectively as the high school-only admissions reform (**HS**, column (2) of table Table 7). As expected, in scenario **MS**, changes in the middle school match precede, as shown in Panel C. This initiates the change through the *application* and *priority* channels of middle schools to shape students' high school matches. The similarity in the size of the effects between the two scenarios can be attributed to the fact that 91% of the targeted middle schools were only accessible to students from the same district or an even smaller geographical area, whereas this was the case for only 15% of the targeted high schools. This result suggests that even

Table 7: Simulation of Policy Effects

	(1) Current pp	(2) HS	(3) MS	(4) MSHS
		% Change from (1)		
Panel A: % College of High School				
Mean	64.83	+1.66%	+1.03%	+1.61%
Racial Gap	9.91	-4.52%	-4.56%	-7.97%
Income Gap	3.89	-9.45%	-9.68%	-13.57%
Panel B: % High-performer of High School				
Mean	34.44	+1.34%	+0.70%	+1.14%
Racial Gap	23.45	-5.74%	-4.97%	-6.23%
Income Gap	9.85	-8.46%	-3.03%	-4.05%
Panel C: 1(Assigned to a Grade A Middle School)				
Mean	0.26	-	+5.88%	+5.88%
Racial Gap	0.15	-	-3.40%	-3.40%
Income Gap	0.05	-	-14.03%	-14.03%

Note: The table reports changes in the main outcomes of interest in each counterfactual policy analysis. Column (1) presents the main characteristics of middle and high schools that students are assigned to in the baseline scenario, and columns (2)-(4) the percentage change in the outcomes of interest in each policy scenario. The racial gap is measured as the difference in the averages of outcome variables between Black and Hispanic students and others. Similarly, the income gap is measured between Free-or-educed-lunch-eligible students and others.

when high school choice is already saturated, leaving little room for reforms, policymakers can leverage the dynamics in school choice to bring about meaningful changes.

Furthermore, the combined intervention at both the middle school and high school level (**MSHS**) yields larger effects on high school assignments compared with the intervention at the high school level only (**HS**). Specifically, when focusing on the percentage of college enrollment in assigned high schools, the **HS** intervention increases the overall mean by 1.66% and reduces racial and income gaps by 4.52% and 9.45%, respectively. In contrast, the **MSHS** intervention increases the overall mean by 1.61% and reduces the gaps by 7.97% and 13.57%. However, the marginal gain of **MS**→**MSHS** (**HS**→**MSHS**) is smaller than that of **Current**→**HS** (**Current**→**MS**), which suggests a possible substitutability between **MS** and **HS** interventions.

6 Policy Implication and Conclusion

Policy Implication Our counterfactual analysis emphasizes the importance of considering the dynamics of school choice in designing school choice reforms. While many existing policies focus on reforming the admissions criteria—*supply side* reforms— recent studies have shown that students' middle or high school assignments remain unchanged largely because of marked heterogeneity in school demand or location across students (Oosterbeek, Sóvágó, and van der Klaauw, 2021; Laverde, 2023; Idoux, 2022). We suggest that by bringing the timing dimension of such reforms into the main consideration, policymakers can design more effective admissions reform. 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 induce changes in high school assignments as well as middle school assignments. In addition, there is still room to further change the type of high schools students attend by intervening at the middle school level even after intervening at the high school level, although the interventions at these two timings are substitutable in our context.

Taken together, our findings imply that large school districts can design a policy with larger effects by leveraging the fact that intervention on the *supply side* of an earlier school choice induces changes in the *demand side* of subsequent school choice stages.

Concluding Remarks This paper extends a static framework of school choice to a dynamic framework of school choices at multiple stages. We showed that top-rated middle schools cause students to be matched to higher achievement high schools, mainly by changing families' high school application behavior. We argued that this dynamic relationship in school choice behavior should be considered in designing school choice policies since a reform targeting one stage can affect student-school match in other stages.

We conclude by suggesting two avenues for future research. First, having confirmed the dynamic relationship between middle school choice and high school choice, we may further directly test for the dynamic complementarity of human capital investments at different educational stages (Cunha and Heckman, 2007; Heckman, 2007). While a credible quasi-randomization at multiple times for a given individual is hard to find, the fact that students are exposed to centralized school choice multiple times provides a suitable research design. Second, given the importance of the dynamic relationship of school choices, we can 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.

References

Abadie, A., S. Athey, G. W. Imbens, and J. M. Wooldridge (2023): "When should you adjust standard errors for clustering?," The Quarterly Journal of Economics, 138(1), 1–35. 12

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. 5, 15, 17, 19, 28

Abdulkadiroğlu, A., J. D. Angrist, Y. Narita, and P. Pathak (2022): "Breaking ties: Regression discontinuity design meets market design," Econometrica, 90(1), 117–151. 3, 6, 10, 11, 12, 4

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, 10, 4

Abdulkadiroğlu, A., P. A. Pathak, J. Schellenberg, and C. R. Walters (2020): "Do Parents Value School Effectiveness?," American Economic Review, 110(5), 1502–1539. 2

Abdulkadiroğlu, A., and T. Sonmez (2003): "School Choice: A Mechanism Design Approach," American Economic Review, 93, 729–747. 1

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. 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. 5

Agarwal, N., I. Ashlagi, P. Somaini, and D. Waldinger (2018): "Dynamic Incentives in Wait List Mechanisms," AEA Papers and Proceedings, 108, 341–47. 5

Agarwal, N., and P. Somaini (2018): "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism," Econometrica, 86(2), 391–444. 5, 23

Agostinelli, F., M. Luflade, and P. Martellini (2021): "On the Spatial Determinants of Educational Access," HCEO Working Paper. 28

Ainsworth, R., R. Dehejia, C. Pop-Eleches, and M. Urquiola (2023): "Why do households leave school value added on the table? The roles of information and preferences," American Economic Review, 113(4), 1049–1082. 5

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. 5

Allende, C. (2019): "Competition under social interactions and the design of education policies," Working Paper. 27, 28

Altonji, J. G., T. E. Elder, and C. R. Taber (2005): "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools," Journal of Political Economy, 113(1), 151–184. 6

Angrist, J. D., and J.-S. Pischke (2009): Mostly harmless econometrics: An empiricist's companion. Princeton University Press. 16

Arcidiacono, P., and J. B. Jones (2003): "Finite mixture distributions, sequential likelihood and the EM algorithm," Econometrica, 71(3), 933–946. 4, 26, 10

Arteaga, F., A. J. Kapor, C. A. Neilson, and S. D. Zimmerman (2022): "Smart matching platforms and heterogeneous beliefs in centralized school choice," The Quarterly Journal of Economics, 137(3), 1791–1848. 27

Artemov, G., Y.-K. Che, and Y. He (2021): "Strategic 'Mistakes': Implications for Market Design Research," Working Paper. 4, 19

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. 4, 21, 36, 5

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

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. 5, 28

Calsamiglia, C., G. Haeringer, and F. Klijn (2010): "Constrained School Choice: An Experimental Study," American Economic Review, 100(4), 1860–1874. 7

Campos, C. (2023): "Social Interactions and Preferences for Schools: Experimental Evidence from Los Angeles," Working Paper. 5, 27

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 (2023): "Leveraging uncertainties to infer preferences: Robust analysis of school choice," Working Paper. 4, 19, 21, 36

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014a): "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates," American Economic Review, 104(9), 2593–2632. 13

——— (2014b): "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood," American Economic Review, 104(9), 2633–79. 13

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 (2022): "Leveling the Playing Field for High School Choice: Results from a Field Experiment of Informational Interventions," NBER Working Paper Series. 3, 5, 16

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. 2

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. 4

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. 6, 13

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. 4

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. 4, 5

FRIEDMAN, M. (1955): "The role of government in education," Economics and the Public Interest, 2(2), 85–107. 2

GALE, D., AND L. S. SHAPLEY (1962): "College Admissions and the Stability of Marriage," American Mathematical Monthly, 69, 9–15. 7, 1

GIUSTINELLI, P. (2016): "Group decision making with uncertain outcomes: Unpacking child–parent choice of the high school track," International Economic Review, 57(2), 573–602. 17

GRENET, J., Y. HE, AND D. KÜBLER (2022): "Preference discovery in university admissions: The case for dynamic multioffer mechanisms," Journal of Political Economy, 130(6), 1427–1476. 5

HAERINGER, G., AND F. KLIJN (2009): "Constrained school choice," Journal of Economic Theory, 144(5), 1921–1947. 7

HASSIDIM, A., A. ROMM, AND R. I. SHORRER (2021): "The limits of incentives in economic matching procedures," Management Science, 67(2), 951–963. 19

HASTINGS, J. S., T. KANE, AND D. STAIGER (2005): "Parental Preferences and School Competition: Evidence from a Public School Choice Program," NBER Working Paper Series. 17

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," Working Paper. 5

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

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. 5

Idoux, C. (2022): "Integrating New York City Schools: The Role of Admission Criteria and Family Preferences," Working Paper. 28, 37

Imbens, G., and K. Kalyanaraman (2012): "Optimal bandwidth choice for the regression discontinuity estimator," The Review of Economic Studies, 79(3), 933–959. 12

Kapor, A. J., C. A. Neilson, and S. D. Zimmerman (2020): "Heterogeneous beliefs and school choice mechanisms," American Economic Review, 110(5), 1274–1315. 5, 27

Larroucau, T., and I. Rios (2020a): "Do "short-list" students report truthfully? Strategic behavior in the Chilean college admissions problem," Working Paper. 19

——— (2020b): "Dynamic college admissions and the determinants of students' college retention," Working Paper. 5

Laverde, M. (2023): "Distance to Schools and Equal Access in School Choice Systems," Working Paper. 4, 37

Lee, J., and S. J. Son (2022): "Distributional Impacts of Centralized School Choice," Working Paper. 5, 27

Luflade, M. (2017): "The value of information in centralized school choice systems," Working Paper. 5, 27

Mark, N. D., S. P. Corcoran, and J. Jennings (2021): "Choosing Alone? Peer Similarity in High School Choices," Working Paper, (396). 27

Narita, Y. (2018): "Match or mismatch? Learning and inertia in school choice," Working Paper. 5

Neilson, C., C. Allende, and F. Gallego (2019): "Approximating the Equilibrium Effects of Informed School Choice," Working Paper. 27

Oosterbeek, H., S. Sóvágó, and B. van der Klaauw (2021): "Preference heterogeneity and school segregation," Journal of Public Economics, 197, 104400. 4, 37

Park, M., and D. W. Hahm (2023): "Location Choice, Commuting, and School Choice," Working Paper. 28

Pop-Eleches, C., and M. Urquiola (2013): "Going to a Better School: Effects and Behavioral Responses," American Economic Review, 103(4), 1289–1324. 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. 11

Rothstein, J. M. (2006): "Good principals or good peers? Parental valuation of school characteristics, Tiebout equilibrium, and the incentive effects of competition among jurisdictions," American Economic Review, 96(4), 1333–1350. 2

Rust, J. (1987): "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher," Econometrica: Journal of the Econometric Society, pp. 999–1033. 26

Shorrer, R. I., and S. Sóvágó (2023): "Dominated Choices in a Strategically Simple College Admissions Environment," Journal of Political Economy Microeconomics. 19

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. 4

Train, K. E. (2009): Discrete choice methods with simulation. Cambridge University Press. 25

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. 5

Zhang, J. (2010): "The sound of silence: Observational learning in the US kidney market," Marketing Science, 29(2), 315–335. 5

(For Online Publication)

Online Appendix to

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

Dong Woo Hahm   Minseon Park

March 2024

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.

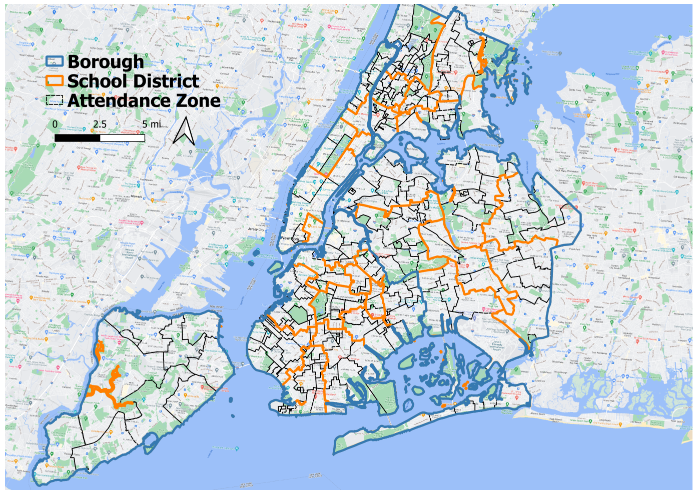
- The algorithm terminates either when there are no new proposals or equally when all rejected students have exhausted their preference lists.

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

A.2 NYC School Admission Methods

Depending on the eligibility criteria, middle schools are classified into three types—district schools, borough schools, and citywide schools. The city is divided into 5 boroughs and 32 community school districts. A student's residence or elementary school decides eligibility at each type of school. In the academic year 2014-2015, 14, 39, and the rest were citywide, borough, and district programs, respectively, among 670 programs. By contrast, the high school choice is fully citywide—all students were eligible for more than 95% of high school programs in NYC in the academic year 2017-18. Figure A.1 illustrates geographic divisions of the city.

Figure A.1: Geographic Division of NYC



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 admission methods similar to 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 8th graders in the same school. Other admissions methods are similar to middle school choice.

A.3 Timeline of The 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 that determine students' assignments. Students who accept their offer finalize, and if a student rejects an offer she goes to the next round. This describes the main

[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, similar to an unscreened program (NYC DOE Introduction to High School Admissions).

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 organized similarly to the main round and includes schools/programs that did not fill 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 are administrative data 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.

Four sets of data are used to construct information on students. First, high school application (HSAP) data include information on each round of the application process (ROL, rank, priority, eligibility, assignment, etc.) related to high school application and standardized test score information. Second, middle school application (MSAP) data includes variables similar but for middle school applications. Third, yearly June biographic data includes more comprehensive biographic data on students, including ethnicity, gender, and disability status, as well as information on attendance and punctuality. Lastly, Zoned DBN data include information on 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, which 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 reports, graduation rate, college enrollment rate, types of language classes offered, etc. Other variables for 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.

[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 the 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 the students' census tract of residence centroid based on the Haversine formula.

B.2 Sample Restriction

We start with 72,318 observations in the middle school application data. Out of 72,318 students, 67,153 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 3 years.[A-3] We present summary statistics and balance test results for 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 rank-ordered list is filled only with new high schools, we have 44,237 students. Estimates in Table 2 are based on this sample.

Table B.1 presents summary statistics of baseline student characteristics. Columns (1)-(2) present summary characteristics of all middle school applicants (whole sample, $N = 47,952$), and Columns (3)-(4) present those of middle school applicants net of attrition (main sample, $N = 45,833$). The majority of students are either Black (23%) or Hispanic (41%) and Free/Reduced-price Lunch (FRL) eligible (75%), and 30% of students ranked a Well-Developed middle school as their first choice. While a student lists 0.91 Well-Developed middle schools on average, there is remarkable variation from one student to another, which is captured by the sizable standard deviation.

C An Example of Calculating Propensity Scores

The following example illustrates how to calculate propensity scores (=admission probabilities) following Abdulkadiroğlu, Angrist, Narita, and Pathak (2017, 2022).[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. The priority score used for admissions is a sum of priority group and a tie-breaker, where the priority group lexicographically dominates tie-breakers. That is, student i's score at program j is

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

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 program C uses a nonrandom tie-breaker

[A-3]Those who participated in the middle school choice but did not participate in the high school choice do not appear in the data afterward. 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, for which there are 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).

[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).

Table B.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 Z-score	-0.01	0.991	0.009	0.984
5th Grade Math Z-score	0.004	0.988	0.025	0.981
English Language Learner (ELL)	0.064	0.244	0.061	0.241
Free/Reduced-price Lunch (FRL)	0.757	0.429	0.753	0.431
Asian	0.178	0.383	0.182	0.386
Black	0.238	0.426	0.235	0.424
Hispanic	0.417	0.493	0.414	0.493
White	0.150	0.358	0.153	0.360
Panel B: Middle School Application Behavior				
Ranked Well-Developed MS 1st?	0.305	0.460	0.304	0.460
# of Well-Developed MS Ranked	0.911	1.223	0.912	1.232
N	47,952		45,833	

Note: Summary statistics of student characteristics in 5th grade.

$\text{TB}_{iC} \sim F_i$, 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.2: 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.2 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 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 by 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 (2022) suggests that i clears the cutoff with half a 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 for each middle and high schools are necessary to use expost stability. First, to interpret data as a conditional multinomial logit model, we need to construct the feasible set of programs for each student, regardless of whether she ranked them. Second, to calculate the continuation value, we need to describe how students' priority scores at high schools would change by which middle school they attend.

In NYC, priority scores consist mainly of three ingredients: eligibility, priority group, and priority ranks at programs that involve screening. First, eligibility and priority groups are determined in a deterministic manner, based on the preannounced rule in the 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 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 in 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:

$$w_{ij} = \beta_j Z_i + e_{ij} \text{ and } i \text{ is ranked higher than } i' \text{ if and only if } w_{ij} > w_{i'j}$$

where Z_i includes standardized statewide math and ELA exam scores; math, social sciences, english, and 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 also 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 the 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 consider a constant-effects value-added model that controls for students' lagged test scores. To ensure enough sample size, we estimate the value-added of each middle school instead of the middle school program.

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":

$$y^H_{i,m} = \alpha_m + Z_i^{M\prime}\beta + u_i, \quad m \in \mathcal{M}$$

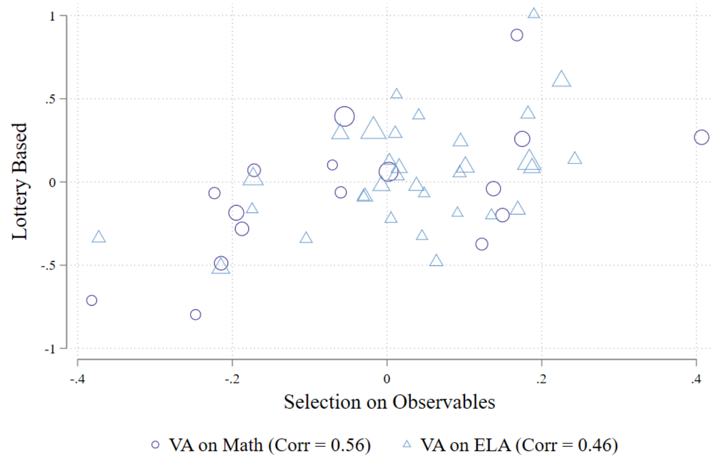
and estimate via OLS of $y^H_{i,m(i)}$ on school indicators where $m(i)$ is the actual middle school attendance in the data and $y^H_{i,m(i)}$ is the observed y^H_i in the data. $Z_i^{M\prime}\beta$ includes baseline test scores, sex and ethnicity dummy variables, English Language Learner status, disability status, and free/reduced-price lunch status in Z_i^M.

In principle, we can leverage the lottery variation built in the DA assignment system in NYC. However, we are underpowered to do so for many middle schools because the number of applicants with admissions probability strictly between 0 and 1 (non-degenerate risk sample) for each school is often small.

Thus, we estimate middle schools' value-added by relying on the standard selection-on-observable assumption. However, we also estimate value-added using the lottery variation for a subset of middle schools whose number of applicants with non-degenerate risk is equal to or greater than 150. For this subset of middle schools, we estimate value-added using the lottery variation following Equation (1), by changing treatment, instrumental variables, and propensity scores accordingly (for example, $C_i = \mathbb{1}(i \text{ attends school} m)$). Figure D.2 shows that the value-added estimates from the OLS model are highly correlated with the value-added estimates that leverage the lottery variation.

We also estimate high schools' value-added on the college enrollment rate using OLS with the selection-on-observable assumption.

Figure D.2: Validation of Value-added Estimates from the OLS Model



Note: The figure plots middle schools' value-added on math and ELA scores for schools whose number of applicants with non-degenerate risk is equal to or greater than 150. The y-axis shows the value-added estimated using lottery variation and the x-axis shows the value-added estimated by relying on the standard selection-on-observable assumption, with students' lagged test scores controlled for.

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 $\epsilon_i = (\epsilon_{i1}, \cdots, \epsilon_{iM})$. 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.

$$\eta_{ij} \perp \epsilon_{im} \big| \gamma_i^M, \gamma_i^H, \ \ \forall i, j, m \tag{D.2}$$

$$\gamma_i^H \perp \eta_i \big| \Psi_{1i}, \ \ \forall i \tag{D.3}$$

$$\eta_i \perp \Psi_{1i}, \ \ \forall i \tag{D.4}$$

$$F(\gamma_i^H \big| \Psi_{1i}) = G(\gamma_i^H \big| \gamma_i^M) \tag{D.5}$$

$$\omega \perp (\gamma_i^H, \eta_{ij}) \big| \Psi_{1i} \text{ and } \omega \perp \Psi_{1i}, \ \ \forall i, j, m \tag{D.6}$$

The first assumption states that conditional on the unobserved tastes γ_i^M and γ_i^H and the idiosyncratic preferences in each period, ϵ_{im} and η_{ij}, are independent for all i, j, m. The second assumption states that conditional on the first-period state variables, the unobserved tastes in the second period and the second-period idiosyncratic preferences are independent. The third assumption states that the second-period idiosyncratic preferences and the first-period state variables are independent.[A-6] The fourth assumption states that the distribution of the unobserved tastes in the second period depends on the first-period state variables only through the first-period unobserved tastes, where F denotes a generic cdf function of its arguments.

Finally, the fifth assumption states that ω, the uncertainty that determines 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) can be simplified as

$$
\begin{aligned}
&E_{\gamma_i^H,\omega,\eta_i,Z_i^H}\left[\max_{j\in O_i(Z_i^H,m;\omega)} V_{ij}\middle|\Psi_{1i}\right] \\
=&E_{\gamma_i^H,\omega,\eta_i}\left[\max_{j\in O_i(Z_i^H,m;\omega)} V_{ij}\middle|\Psi_{1i}\right] && : Z_i^H \text{ is perfectly predictable} \\
=&\int_\omega E_{\gamma_i^H,\eta_i}\left[\max_{j\in O_i(Z_i^H,m;\omega)} V_{ij}\middle|\Psi_{1i}\right] dH(\omega) && : (\text{D.6}) \\
=&\int_\omega \int_{(\gamma_i^H,\eta_i)} \max_{j\in O_i(Z_i^H,m;\omega)} V_{ij} dF(\eta_i,\gamma_i^H|\Psi_{1i}) dH(\omega) \\
=&\int_\omega \int_{\gamma_i^H}\left(\int_{\eta_i} \max_{j\in O_i(Z_i^H,m;\omega)} V_{ij} dF(\eta_i)\right) dG(\gamma_i^H|\gamma_i^M) dH(\omega) && : (\text{D.3}),(\text{D.4}),(\text{D.5}) \\
=&\int_\omega \int_{\gamma_i^H}\left(\mu + \log\left(\sum_{j\in O_i(Z_i^H,m;\omega)} \exp(v_{ij})\right)\right) dG(\gamma_i^H|\gamma_i^M) dH(\omega) && : \eta_{ij} \overset{iid}{\sim} EVT1
\end{aligned}
$$

where $v_{ij} \equiv V_{ij} - \eta_{ij}$ and F denotes a generic cdf function of its argument 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. We use students' residence in the first period to calculate the

[A-6] Recall that the first- and second-period unobserved tastes can be arbitrarily correlated. By the first three assumptions, 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.

distance to each high school program in calculating the continuation value.

D.4 EM Algorithm

First of all, we derive the full likelihood function as follows. 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

$$\theta = \left(\beta_0^M, \beta_Z^M, \beta_0^H, \beta_Z^H, \{\alpha_\tau\}_\tau, \xi_m, \tilde{\xi}_j, \{\gamma^{M,p}\}, \{\gamma^{H,q}\}, \{G(p,q)\}_{p,q}, \{\rho_\tau\}_\tau, \vartheta_m, \vartheta_h, \lambda^M, \lambda^H\right)$$

where, with some abuse of notation, $G(p,q)$ is the probability that γ_i^M and γ_i^H equal the $\gamma^{M,p}$, $\gamma^{H,q}$, the p-th and q-th types, respectively. Then for student i, conditional on γ_i^M, γ_i^H,

$$\begin{aligned}
P_i(\theta, \gamma_i^M, \gamma_i^H) &= P(\text{observe } m_i, j_i | \gamma_i^M, \gamma_i^H, \theta) \\
&= P\left(\begin{array}{cc} U_{im_i} = \max_{m \in O_i^m} U_{im} & \text{and} \\ V_{ij_i} = \max_{j \in O_i^h} V_{ij} \text{ given } m_i & \end{array} \middle| \gamma_i^M, \gamma_i^H, \theta \right) \\
&= \frac{\exp(u_{im_i}(\gamma_i^M, \theta))}{\sum_{m \in O_i^m} \exp(u_{im}(\gamma_i^M, \theta))} \frac{\exp(v_{ij_i}(\gamma_i^H, \theta; m_i))}{\sum_{j \in O_i^h} \exp(v_{ij}(\gamma_i^H, \theta; m_i))} \qquad : (\text{D.2})
\end{aligned}$$

where the second equality comes from the expost stability and the third equality comes from the distributional assumptions on the unobservables. Then,

$$\begin{aligned}
P_i(\theta) &= \int_{(\gamma_i^M, \gamma_i^H)} P_i(\theta, \gamma_i^M, \gamma_i^H) dG(\gamma_i^M, \gamma_i^H) \\
&= \sum_{p,q} G(p,q) P_i(\theta, \gamma^{M,p}, \gamma^{H,q})
\end{aligned}$$

and hence $\prod_i P_i(\theta)$, or $\sum_i \log P_i(\theta)$ is the final likelihood function to be maximized.

To overcome the computational challenge coming from our dynamic model with many parameters, we estimate our model using the expectation-maximization algorithm with a sequential maximization step proposed by Arcidiacono and Jones (2003). First, we reformulate the full likelihood function in the following expectation form.

$$\mathcal{E}(P, \gamma, \theta \mid \hat{P}, \hat{\gamma}, \hat{\theta}) = \Sigma_i \Sigma_p \Sigma_q h\left(p, q \mid \hat{P}, \hat{\gamma}, \hat{\theta}\right) \log P(p,q) \tag{D.7}$$

$$+ \Sigma_i \Sigma_p \Sigma_q h\left(p, q \mid \hat{P}, \hat{\gamma}, \hat{\theta}\right) \log P_i^M(p; \theta^{MS}, \theta^{HS}, \gamma_m, \gamma_h, P) \tag{D.8}$$

$$+ \Sigma_i \Sigma_p \Sigma_q h\left(p, q \mid \hat{P}, \hat{\gamma}, \hat{\theta}\right) \log P_i^H\left(q; \theta^{HS}, \gamma_h\right) \tag{D.9}$$

where $h\left(p, q \mid \hat{P}, \hat{\gamma}, \hat{\theta}\right) = \frac{\hat{P}(p,q)\hat{P}_i^M(p)\hat{P}_i^H(q)}{\Sigma_{p'}\Sigma_{q'}\hat{P}(p',q')\hat{P}_i^M(p')\hat{P}_i^H(q')}$. Then we take the following steps.

Steps

1. Start with an initial guess and calculate the conditional probability h_0
2. Update $\hat{P}_1(p,q) = \frac{1}{I}\Sigma_i h_0\left(p, q \mid \hat{P}_0, \hat{\gamma_0}, \hat{\theta}_0\right)$
3. Update (θ^{HS}, γ_h) by maximizing $f^h \equiv \Sigma_i f_i^h \equiv \Sigma_i \Sigma_p \Sigma_q h_0(p,q) \log P_i^H(q; \theta^{HS}, \gamma_h)$
   - Outer loop for nonlinear parameters
   - Inner loop for α_s^H
     - For $s = 2, \cdots, Jsch_h$,

$$\begin{aligned}
\frac{\partial f_i^h}{\partial \alpha_s^M} &= \sum_p \sum_q h_0(p,q)\left(\frac{\partial P_i^H(q)}{\partial \alpha_s^H} \bigg/ P_i^H(q)\right) \\
&= \sum_p \sum_q h_0(p,q)\left(P_i^H(q)\left[\frac{\partial v_{ij_i}(q,\tau(m_i))}{\partial \alpha_s^H} - \frac{1}{\sum_{j \in B_i^H} \exp(v_{ij}(q))}\sum_{j \in B_i^H} A_j\right] \bigg/ P_i^H(q)\right) \\
&= \sum_p \sum_q h_0(p,q)\left[\frac{\partial v_{ij_i}(q,\tau(m_i))}{\partial \alpha_s^H} - \frac{1}{\sum_{j \in B_i^H} \exp(v_{ij}(q))}\sum_{j \in_i^H} A_j\right] \\
&= \sum_p \sum_q h_0(p,q)\left[1\{s(j_i) = s\} - \frac{\sum_{j \in B_i^H \& s(j)=s} \exp(v_{ij}(q))}{\sum_{j \in B_i^H} \exp(v_{ij}(q))}\right],
\end{aligned}$$

where $A_j = \exp(v_{ij}(q))\frac{\partial v_{ij}(q,\tau(m_i))}{\partial \alpha_s^H}$. Then,

$$\begin{aligned}
\frac{\partial f^h}{\partial \alpha_s^M} = &\sum_{i:s(j_i)=s}\left(\sum_p \sum_q h_0(p,q)\left[1 - \frac{\sum_{j \in B_i^H \& s(j)=s} \exp(v_{ij}(q))}{\sum_{j \in B_i^H} \exp(v_{ij}(q))}\right]\right) \\
&- \sum_{i:s(j_i)\neq s}\left(\sum_p \sum_q h_0(p,q)\frac{\sum_{j \in B_i^H \& s(j)=s} \exp(v_{ij}(q))}{\sum_{j \in B_i^H} \exp(v_{ij}(q))}\right)
\end{aligned}$$

so that the FOC leads to

$$\sum_{i:s(j_i)=s}\sum_p \sum_q h_0(p,q) = \sum_i\left(\sum_p \sum_q h_0(p,q)\frac{\sum_{j \in B_i^H \& s(j)=s} \exp(v_{ij}(q))}{\sum_{j \in B_i^H} \exp(v_{ij}(q))}\right) \tag{D.10}$$

4. Update (θ^{MS}, γ_m) by maximizing $f^m \equiv \Sigma_i f_i^m \equiv \Sigma_i \Sigma_p \Sigma_q h_0(p,q) \log P_i^M(\theta^{MS}, \hat{\theta}_1^{HS}, \gamma_m, \hat{\gamma}_{h,1}, \hat{P}_1)$

- Calculate EU with up-to-date estimates $\hat{P}_1, \hat{\theta}_1^{HS}, \hat{\gamma}_{h,1}$
- Outer loop for nonlinear parameters
- Inner loop for α_s^M: similarly to step 3

5. Iterate until convergence

D.5 Average Treatment Effect for a Subset

Abdulkadiroğlu, Angrist, Narita, and Pathak (2022) shows that the following conditional independent assumption (CIA) holds for students whose Well-Developed middle school admissions probability is strictly between 0 and 1:

$$(Y_{i1}, Y_{i0}) \perp D_i | \{d_i(x)\}, \mathcal{R}_i, \tag{D.11}$$

where Y_{i1} is the outcome of student i when she attends a Well-Developed middle school, Y_{i0} is the outcome when she does not attend any Well-Developed middle school, $d_i(x)$ is a full set of dummies for each value of the propensity score, and $\mathcal{R}_i$ is the local linear control for nonrandom tie-breakers of middle schools. Further restricting the sample to those with $C_i = D_i$,

$$(Y_{i1}, Y_{i0}) \perp C_i | \{d_i(x)\}, \mathcal{R}_i, C_i = D_i, \tag{D.12}$$

where $C_i = \mathbb{1}(\text{Assigned to a Well-Developed middle school})$, and $D_i = \mathbb{1}(\text{Attended a Well-Developed middle school})$. Using this independence assumption, we can obtain the ATE using the IPW for this subset, where we use the probability of being assigned to a Well-Developed middle school as the weight while controlling for $\mathcal{R}_i$. Note that this subset includes not only the compliers but also the always-takers who are assigned to a Well-Developed middle school and the never-takers who are not assigned to any Well-Developed middle school.

E Alternative Specifications

E.1 Static Model

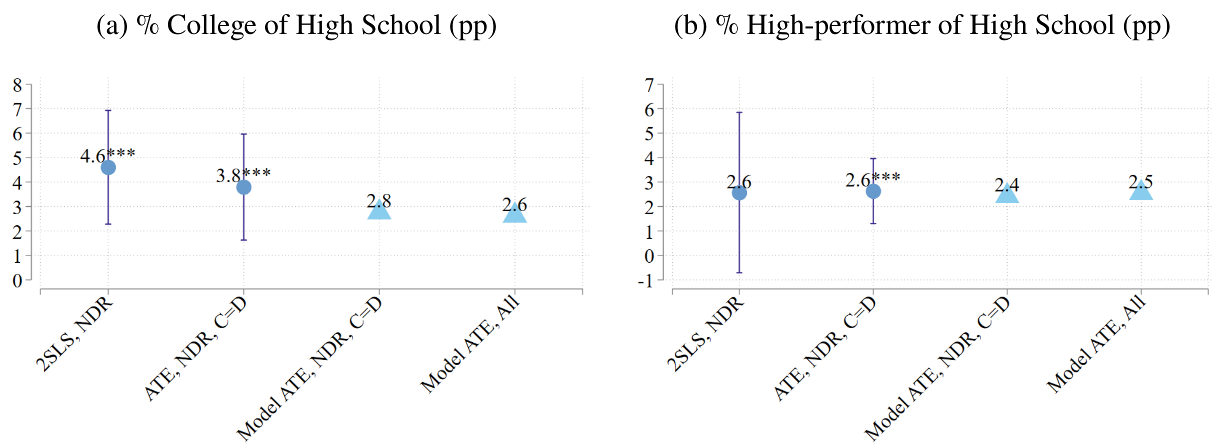
Recall the key features of the dynamic model: forward-looking agents, serial correlation of unobservable tastes, and middle school-type effects. To highlight the importance of including those features in the model, we estimate a restricted static model without the dynamic components of the model. The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

$G(p,q) = G_M(p)G_H(q), \forall p, q$, where G_M, G_H are the marginal distributions of unobserved tastes in middle and high schools, respectively. Third, middle school type effects are absent—i.e., $\rho_\tau = \alpha_\tau = 0, \forall \tau$. Since the static model is a nested model of the full dynamic model in which the above restrictions are imposed, we can perform a likelihood ratio (LR) test. The static model is strongly rejected in favor of our main dynamic model ($p < 0.001$), which reconfirms the importance of modeling the dynamics.

E.2 More Unobserved Types

The benefit of potentially having more unobserved types in the model is that by doing so, we can more flexibly control for the unobserved heterogeneity. However, it comes with a larger computing burden. While our four-type model already fits the treatment effect estimate from the data well (Section 4.4.2), we also estimated a nine-type model as a robustness check. Figure E.3 shows that the treatment effects from the nine-type model are similar to those from the four-type model. This reassures that the four-type model captures the unobserved heterogeneity in our data reasonably well.

Figure E.3: Effects of Middle Schools on High School Assignments: Data vs. Model with 9 Types



Note: The figure plots middle school effects on high school assignments from the data and the simulation. Figure (a) plots the effect on the college enrollment rate of assigned high schools. Figure (b) plots the effect on the proportion of high-performers in assigned high schools. Students are labeled to be high-performing if their standardized test score is above the 66th percentile of their cohort in the system. In each subfigure, we plot 2SLS estimates, average treatment effects (ATE) on the subsample, simulated average treatment effect (ATE) on the subsample, and simulated average treatment effect (ATE) of all students. We explain in the text which students are included in the subsample and how we calculate the ATE of this subsample from the data. We present 95% confidence intervals of treatment effects calculated from the data. For the model simulation, we assign each student to her most preferred Well-Developed middle school and non-Well-Developed middle school, respectively, and compare the assigned high schools' characteristics in each scenario. We simulate over 10,000 high school lotteries for each scenario.

E.3 Alternative Values of the Discount Factor

As discussed in the main text, we set $\delta = 0.75$ due to the lack of variation in the data to identify the discount factor. As a robustness check, we estimated the model with an alternative choice of

$\delta = 0.9$. It leads to model predictions similar to the version with $\delta = 0.75$, as demonstrated by the model fit in terms of the average characteristics of assigned middle schools. See Table E.3.

Table E.3: Goodness of Fit: $\delta = 0.9$

	All		By Racial Group				By FRL Status			
			Black/Hispanic		White/Asian/Other		FRL		Non-FRL	
	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
Middle School Characteristics										
Proportion of White/Asian/Other	0.33	0.32	0.19	0.19	0.61	0.60	0.28	0.28	0.48	0.45
Proportion of Non-FRL	0.24	0.23	0.19	0.19	0.34	0.33	0.21	0.20	0.35	0.33
Proportion of High Performers	0.30	0.30	0.23	0.23	0.44	0.44	0.27	0.27	0.40	0.39
$\mathbb{1}$(Grade A)	0.25	0.25	0.21	0.20	0.35	0.35	0.24	0.24	0.30	0.29
High School Characteristics										
Proportion of White/Asian/Other	0.30	0.29	0.20	0.20	0.51	0.48	0.27	0.27	0.42	0.39
Proportion of Non-FRL	0.21	0.20	0.17	0.17	0.28	0.27	0.19	0.19	0.28	0.26
Proportion of High Performers	0.27	0.26	0.23	0.22	0.36	0.34	0.26	0.25	0.33	0.29
College Enrollment Rate	0.62	0.58	0.59	0.56	0.67	0.63	0.61	0.58	0.66	0.60

Note: For model-based simulations, we report the average result from multiple simulations using 100 draws of middle school lotteries and 10,000 high school lotteries. For a given draw of the lottery, we assign a student to her most preferred feasible middle school and high school, respectively. Students are labeled high-performing if their standardized test score is above the 66th percentile of their cohort in the system.

F Additional Tables and Figures

F.1 Additional Tables and Figures from Section 2

Average School Characteristics by Rank on Students' ROL Tables F.4 and F.5 summarize the averages of school characteristics by rank on students' ROLs of middle schools and high schools, respectively. There are three main patterns. First, students tend to rank schools distant 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 stated 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.

F.2 Additional Tables and Figures from Section 3

Balance Test We present students' test scores, demographic characteristics, and variables that describe the middle school application behavior of students who are assigned to 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 offered and non-offered students. The offered have higher test scores and are less likely to be FRL, ELL, Black/Hispanic, or need special education, all with statistically significant differences. They also rank more high-achievement middle schools (recall this is our treatment of interest) than the non-offered students and are more likely to list them first on their ROLs, which makes sense because such

Table F.4: 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 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 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 for currently enrolled 6th graders and 8th graders in AY 2014-15, where the scale is from 110 to 410.

Table F.5: 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 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 Scores (9th grade) are calculated using the average of the 8th grade statewide standardized math and ELA exams for currently enrolled 9th graders in AY 2017-18, where the scale is from 130 to 400. 4-year Grad Rate and Enroll in College are calculated using the average of the graduating cohort in AY 2017-18.

behavior will unambiguously increase the odds of being offered such schools.

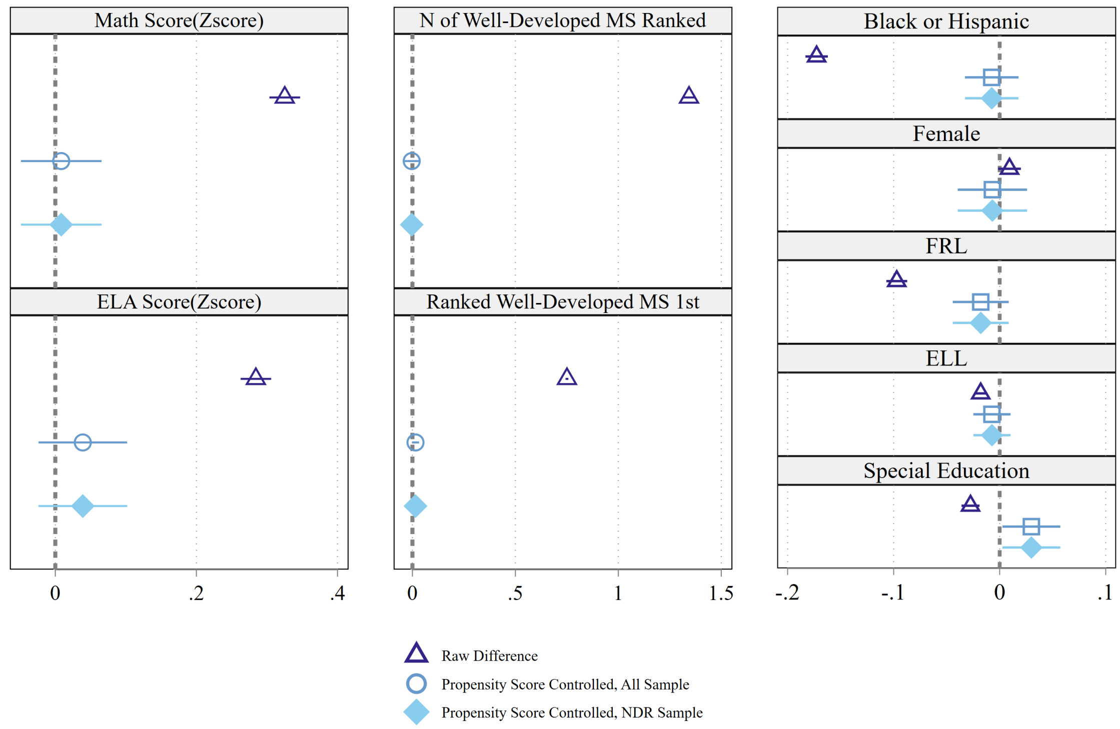
Next, we control for the propensity scores and include local linear control of tie-breakers in the following two specifications denoted by **Propensity Score Controlled, All Sample** and **Propensity Score Controlled, NDR Sample** in Figure F.4. Specifically, we run

$$W_i = \alpha_0 + \gamma D_i + \sum_x \alpha_1(x) d_i(x) + h\left(\mathcal{R}_i\right) + e_i \tag{F.13}$$

where W_i is the student covariates which we test balance on, 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 F.4 presents estimates on γ with students with all possible propensity scores, including 0 and 1. Controlling for the propensity score and nonrandom tie-breakers effectively balances covariates. Next, **Propensity Score Controlled, NDR Sample** shows the γ only for students with non-degenerate risk of being offered—i.e., subject to

Figure F.4: Covariate Balance Test: Offered Students vs. Non-offered Students

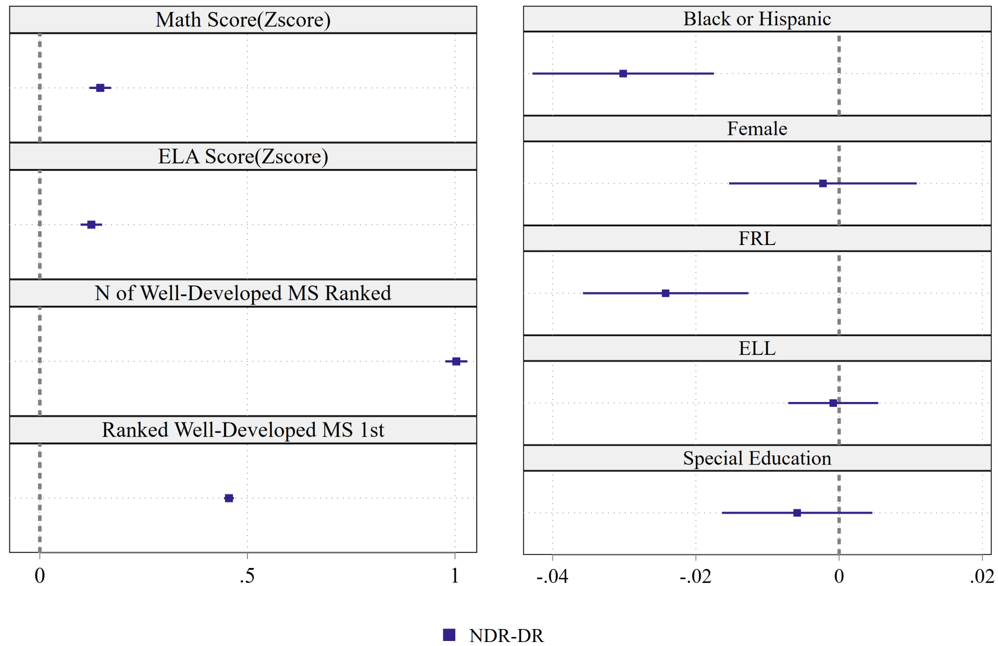


Note: **Raw Difference** shows the t-test results of covariate mean difference between the offered and the non-offered. **Propensity Score Controlled, All Sample** shows the coefficient of the offered when we regress the covariate on the offered dummy variables, the nonparametric controls for propensity score, and the local linear function of nonrandom 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 whose 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 the standard deviation for the left panel and proportion for the middle and right panels. Markers show the exact estimates, and 95% CIs are presented. Robust standard errors are estimated. N=6,610 for **Propensity Score Controlled, NDR Sample**, and N=46,618 for other estimates.

randomization. Further restricting the sample to those with non-degenerate risk provides an almost perfect balance between the offered and the non-offered groups.

Figure F.5 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 the degenerate risk sample have a 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, which suggests that 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, are less likely to be Black/Hispanic, and obviously ranked many treatment middle schools. This reconfirms that the 2SLS estimates we find in Section 3.2 are local average treatment effects (LATE).

Figure F.5: Covariate Balance Test: Non-degenerate vs. 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=46,618).

Additional Specifications of the Main Results Table F.6 presents high-score middle school attendance effects on the characteristics of high schools to which students apply and get assigned. We define a middle school to be high-score if the average of 6th-grade students' baseline test scores is above the 66th percentile across schools. Similar to the effect of Well-Developed middle schools in Table 2, attending middle schools with high baseline student test scores causes students to apply to and get matched to high schools with better academic peformance in both level and in value-added.

Table F.7 presents Well-Developed middle school attendance effects on the mean characteristics of top 3-ranked and top 5-ranked high schools. We find smaller effects, and in particular the mean of college enrollment rates does not seem to change. This is because Well-Developed middle schools motivate students to submit shorter lists (Table F.8), and thus many students from Well-Developed middle schools are dropped from these regressions. Notably, Well-Developed middle school graduates' shorter lists does not come at the expense of the probability of being assigned to any school on their lists. Rather, the probability of being assigned to any school in the main round increases by 2.3 pp, and the probability of being assigned to their first-ranked school increases by 8.6 pp (Table F.8).

The fact that students attending Well-Developed middle schools submit shorter high school

Table F.6: High-Score MS Attendance Effects on HS Application and Assignment

Characteristics of	(1)	(2) Top-ranked HS	(3)	(4)	(5) Assigned HS	(6)
Model	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Sample	All	All	NDR	All	All	NDR
Panel A: College Enrollment Rate (%)						
From High-Score MS	2.634***	2.377***	2.107**	4.133***	4.250***	3.764***
	(0.147)	(0.722)	(0.806)	(0.163)	(0.780)	(0.891)
N	39360	39360	6826	37237	37237	6460
R2	0.277	0.316	0.442	0.262	0.309	0.411
$\bar{y}$	73.382	73.382	76.524	66.008	66.008	69.129
Panel B: Value-added on College Enrollment Rate (%)						
From Well-Developed MS	0.134	0.949*	1.140*	0.792***	1.643**	1.635**
	(0.089)	(0.457)	(0.519)	(0.105)	(0.512)	(0.593)
N	39275	39275	6818	37282	37282	6475
R2	0.053	0.099	0.239	0.044	0.099	0.262
$\bar{y}$	2.659	2.659	3.317	1.089	1.089	2.185
Panel C: % High-Baseline-Score 9th-Graders						
From Well-Developed MS	5.670***	3.824**	2.927*	6.894***	6.240***	5.936***
	(0.242)	(1.181)	(1.342)	(0.201)	(0.981)	(1.068)
N	39844	39844	6900	38017	38017	6592
R2	0.362	0.413	0.505	0.388	0.445	0.508
$\bar{y}$	43.743	43.743	48.036	33.025	33.025	36.423
Panel D: % White						
From Well-Developed MS	5.173***	1.156	-0.026	5.756***	0.824	-0.342
	(0.156)	(0.748)	(0.789)	(0.137)	(0.677)	(0.703)
N	39844	39844	6900	38017	38017	6592
R2	0.475	0.537	0.575	0.556	0.612	0.627
$\bar{y}$	19.303	19.303	20.262	15.096	15.096	15.508
First-stage F-stat		189.44	129.94		189.44	129.94

Note: Each panel presents High-Score MS attendance effects on different characteristics of high schools that students first-ranked (columns (1)-(3)) or are assigned to (columns (4)-(6)). We define a middle school to be high-score if the average of 6th-grade students' baseline test score is above the 66th percentile across schools. To construct the outcome in Panel D, we define students to be high-baseline-score if their standardized NYS test score is above the 66th percentile. In columns (3) and (6), we restrict the sample to students with non-degenerate risk of being offered (i.e., whose propensity score is in the interval $(0, 1)$ and hence subject to randomization) from any Well-Developed middle school. Robust standard errors are 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. Columns (2)-(3) and (5)-(6) control for dummy variables for all possible values of propensity score of being assigned to a High-Score MS and local linear control for nonrandom tie-breakers. +p<0.1, *p<0.05, **p<0.01, ***p<0.001.

application lists yet are more likely to be matched with schools ranked high on their list suggests that their priority standing at the high schools they applied to should improve as a result of Well-Developed middle school attendance. In Table F.9, we use priority score (Equation (C.1)) to confirm this.

Table F.10 shows results with other dimensions of high school characteristics (e.g., graduation rate, % Asian), which confirms that the main results—attending a Well-Developed middle school makes students apply to high school programs in a way that puts more weight on the end-of-high school academic outcomes than students' body composition— are not driven by the choice of high school characteristics.

Table F.7: Well-Developed MS Attendance Effects on HS Application

Characteristics of	(1)	(2) Top3-ranked HS	(3)	(4)	(5) Top5-ranked HS	(6)
Model	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Sample	All	All	NDR	All	All	NDR
Panel A: College Enrollment Rate (%)						
From Well-Developed MS	0.758	1.275+	1.307	0.589	1.491*	0.858
	(0.488)	(0.686)	(0.800)	(0.546)	(0.731)	(0.939)
N	39471	39471	5408	34222	34222	4662
R2	0.209	0.243	0.387	0.185	0.223	0.379
$\bar{y}$	71.334	71.334	72.992	70.354	70.354	71.947
Panel B: Value-added on College Enrollment Rate (%)						
From Well-Developed MS	0.155	-0.207	-0.325	-0.046	0.159	0.248
	(0.301)	(0.401)	(0.489)	(0.297)	(0.443)	(0.542)
N	39401	39401	5400	34180	34180	4660
R2	0.040	0.073	0.219	0.034	0.068	0.227
$\bar{y}$	2.741	2.741	3.096	2.795	2.795	3.066
Panel C: % High-Baseline-Score 9th-Graders						
From Well-Developed MS	1.934*	3.153**	2.937**	1.849*	3.431*	2.719*
	(0.863)	(1.029)	(1.005)	(0.841)	(1.378)	(1.381)
N	40041	40041	5471	34799	34799	4734
R2	0.265	0.313	0.445	0.231	0.283	0.441
$\bar{y}$	39.740	39.740	41.890	37.824	37.824	40.402
Panel D: % White						
From Well-Developed MS	1.453+	0.956	1.166	0.989	2.230*	1.645+
	(0.758)	(0.812)	(0.758)	(0.610)	(1.041)	(0.948)
N	40042	40042	5471	34799	34799	4734
R2	0.368	0.425	0.550	0.288	0.349	0.507
$\bar{y}$	17.604	17.604	20.393	15.751	15.751	18.608
First-stage F-stat		411.74	287.42		411.74	287.42

Note: Each panel presents Well-Developed MS attendance effects on different characteristics of high schools that students top 3-ranked (columns (1)-(3)) or top 5-ranked (columns (4)-(6)). To construct the outcome in Panel D, we define students to be high-baseline-score if their standardized NYS test score is above the 66th percentile. In columns (3) and (6), we restrict the sample to students with non-degenerate risk of being offered (i.e., whose propensity score is in the interval $(0, 1)$ and hence subject to randomization) from any Well-Developed middle school. Robust standard errors are 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. Columns (2)-(3) and (5)-(6) control for dummy variables for all possible values of propensity score of being assigned to a Well-Developed MS and local linear control for nonrandom tie-breakers. +p<0.1, *p<0.05, **p<0.01, ***p<0.001.

The estimates in columns (2) and (3) in Table 2 differ due to changes in estimates of other covariates. The coefficients of interest with the full sample (in column (2)) vary by whether we control for other covariates or not, while those with NDR sample (in column (3)) remain stable (Table F.11). This is because covariates differ between treated and untreated student in the full sample even after controlling for the full set of propensity score dummies Appendix F.2. This reassures the importance of common support assumption, and in turn, our choice of column (3) as most preferred specification.

Table F.12 shows that additionally controlling for end-of-middle-school test scores barely changes the main treatment effect of attending Well-Developed middle schools. We take this as a suggestive evidence that middle schools affect students' high school application behavior through other channels

Table F.8: Well-Developed MS Attendance Effects on HS Application and Assignment—Other Outcomes

	(1) Length of List	(2) $\mathbb{1}$(Unassigned)	(3) Rank of Assigned Program	(4) $\mathbb{1}$(Assigned to Top-rank)
From Well-Developed MS	-1.479***	-0.023+	-0.165	0.86***
	(0.181)	(0.012)	(0.107)	(0.241)
N	6413	6413	5988	6413
R2	0.345	0.149	0.210	0.223
$\bar{y}$	7.394	0.066	2.524	44.160

Note: The table presents Well-Developed MS attendance effects on four outcomes. The outcome variable in column (1) is the total number of programs included in students' rank-ordered list, in column (2) is an indicator of whether the student is unassigned to any school on her list, in column (3) is the rank of matched school on her rank-ordered list, and in column (4) is an indicator of whether the student is assigned to her top-ranked school. We restrict the sample to students with non-degenerate risk of being offered (i.e., whose propensity score is in the interval $(0, 1)$ and hence subject to randomization) from any Well-Developed middle school. Robust standard errors are 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 control for dummy variables for all possible values of propensity score of being assigned to a Well-Developed MS and local linear control for nonrandom tie-breakers. +p<0.1, *p<0.05, **p<0.01, ***p<0.001.

Table F.9: Well-Developed MS Attendance Effects on HS Priority Score

	(1)	(2)	(3)	(4)	(5)
	Score at 1^{st} ranked	Score at 2^{nd} ranked		Score at 3^{rd} ranked	
From Well-Developed MS	-0.209***	-0.057*	-0.064*	-0.031	-0.009
	(0.034)	(0.025)	(0.025)	(0.025)	(0.025)
Application FEs	1^{st}-ranked	2^{nd}-ranked	Up to 2^{nd}-ranked	3^{rd}-ranked	Up to 3^{rd}-ranked
N	6413	5688	5688	5474	5474
R2	0.661	0.828	0.858	0.840	0.897
$\bar{y}$	1.318	1.417	1.417	1.459	1.459

Note: The table presents Well-Developed MS attendance effects on priority scores at high schools students applied to. Priority score is constructed as in Equation (C.1); the integer maps onto the priority group and the decimal part maps onto nonrandom tie-breaker. The lower the value, the higher the priority standing. The outcome variable in column (1) is the priority score at students' first-ranked high schools. In columns (2)-(3), it is the priority score at their second-ranked high schools, and the third-ranked high schools in columns (4)-(5). To control for the fact that different high schools use different priority groups and tie-breaking rules, we control for the full set of dummy variables for Nth-ranked high schools (columns (1), (2), (4)) or up to Nth-ranked high schools (columns (3) and (5)). We restrict the sample to students with non-degenerate risk of being offered (i.e., whose propensity score is in the interval $(0, 1)$ and hence subject to randomization) from any Well-Developed middle school. Robust standard errors are 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 control for dummy variables for all possible values of propensity score of being assigned to a Well-Developed MS and local linear control for nonrandom tie-breakers. +p<0.1, *p<0.05, **p<0.01, ***p<0.001.

beyond the evolution of test scores.

F.3 Additional Tables and Figures from Section 4

F.3.1 Demand Estimates

Table F.13 reports demand estimates.

F.3.2 Discussion of the Modeling Choice and Relevant Data Patterns

The Null Grade A Attendance Effect on Moving Table F.14 shows that Well-Developed middle schools did not shift households' propensity to move across boroughs, school districts, or Census tracts. We run Equation (1) with the same set of students as in the main empirical pattern in Table 2.

Table F.10: Well-Developed MS Attendance Effects on HS Application and Assignment

Characteristics of	(1)	(2) Top-ranked HS	(3)	(4)	(5) Assigned HS	(6)
Model	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Sample	All	All	NDR	All	All	NDR
Panel A: Graduation Rate (%)						
From Well-Developed MS	0.648***	1.207**	1.757***	1.480***	2.935***	3.467***
	(0.113)	(0.415)	(0.512)	(0.129)	(0.462)	(0.565)
N	44522	44522	6324	42139	42139	5982
R2	0.209	0.244	0.375	0.215	0.254	0.380
$\bar{y}$	85.443	85.443	86.897	79.675	79.675	81.453
Panel B: Value-added on Graduation Rate (%)						
From Well-Developed MS	-0.448***	0.329	0.507	-0.172*	1.136***	1.467***
	(0.071)	(0.259)	(0.328)	(0.080)	(0.297)	(0.358)
N	44401	44401	6311	42126	42126	5985
R2	0.143	0.184	0.372	0.024	0.071	0.236
$\bar{y}$	0.904	0.904	0.592	1.149	1.149	1.198
Panel C: % White or Asian						
From Well-Developed MS	2.032***	1.449+	0.047	3.230***	1.860**	-0.392
	(0.218)	(0.772)	(0.905)	(0.209)	(0.698)	(0.750)
N	45081	45081	6393	43019	43019	6097
R2	0.496	0.547	0.655	0.519	0.579	0.693
$\bar{y}$	41.282	41.282	43.952	33.308	33.308	35.540
Panel D: % Free-or-Reduced Lunch Eligible						
From Well-Developed MS	-1.112***	-2.133***	-0.427	-1.713***	-2.323***	-0.421
	(0.169)	(0.600)	(0.675)	(0.146)	(0.508)	(0.514)
N	45081	45081	6393	43019	43019	6097
R2	0.373	0.447	0.562	0.430	0.509	0.608
$\bar{y}$	70.100	70.100	68.039	75.696	75.696	73.836
First-stage F-stat		411.74	287.42		411.74	287.42

Note: Each panel presents Well-Developed MS attendance effects on different characteristics of high schools that students first-ranked (columns (1)-(3)) or are assigned to (columns (4)-(6)). To construct the outcome in Panel C, we define students to be high-baseline-score if their standardized NYS test score is above the 66th percentile. In columns (3) and (6), we restrict the sample to students with non-degenerate risk of being offered (i.e., whose propensity score is in the interval $(0, 1)$ and hence subject to randomization) from any Well-Developed middle school. Robust standard errors are 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. Columns (2)-(3) and (5)-(6) control for dummy variables for all possible values of propensity score of being assigned to a Well-Developed MS and local linear control for nonrandom tie-breakers. +p<0.1, *p<0.05, **p<0.01, ***p<0.001.

The Null Grade A Attendance Effect on Specialized High School Match Table F.15 shows that Well-Developed middle schools did not shift students' propensity to take the SHSAT more or to get an offer from specialized high schools more. We run Equation (1) with the same set of students in the main empirical pattern in Table 2. There are 11 specialized high schools in the city, including Stuyvesant and the Bronx High School of Science, and they admit students based on the SHSAT score. These 11 high schools have their own centralized school assignment system, separate from those of district schools, which comprise our estimation sample.

Table F.11: Estimates With and Without Covariates

	(1)	(2)	(3)	(4)
Dependent Variable: College Enrollment Rate of	Top-ranked HS		Assigned HS	
Model	2SLS	2SLS	2SLS	2SLS
Sample	All	NDR	All	NDR
Panel A: With Covariates				
From Well-Developed MS	1.820***	2.961***	3.851***	4.628***
	(0.508)	(0.616)	(0.560)	(0.678)
N	44489	6320	42055	5975
R2	0.310	0.441	0.301	0.419
$\bar{y}$	73.724	76.096	66.325	69.040
Panel B: Without Covariates				
From Well-Developed MS	2.588***	3.040***	4.949***	4.890***
	(0.531)	(0.637)	(0.590)	(0.696)
N	46391	6623	43875	6268
R2	0.145	0.362	0.167	0.352
$\bar{y}$	73.643	76.063	66.271	69.038

Note: The dependent variables are college enrollment rates of top-ranked high schools (columns (1) and (2)) and those of assigned high schools (columns (3) and (4)). In columns (2) and (4), we restrict the sample to students with non-degenerate risk of being offered (i.e., whose propensity score is in the interval $(0, 1)$ and hence subject to randomization) from any Well-Developed middle school. Robust standard errors are in parentheses. Panel A controls 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. Panel B does not include those covariates. All columns control for dummy variables for all possible values of propensity score of being assigned to a Well-Developed MS and local linear control for nonrandom tie-breakers. +p<0.1, *p<0.05, **p<0.01, ***p<0.001.

Table F.12: Controlling For 8th Grade Test Scores

	(1)	(2)	(3)	(4)
Dependent Variable	College Enrollment Rate (%)		VA on College Enrollment Rate (%)	
Characteristics of	Top-ranked	Assigned	Top-ranked	Assigned
From Well-Developed MS	3.572***	4.738***	1.970***	2.369***
	(0.859)	(0.924)	(0.554)	(0.627)
8th Grade ELA Score	1.385***	1.919***	0.019	0.632*
	(0.378)	(0.387)	(0.230)	(0.260)
	(0.354)	(0.369)	(0.228)	(0.252)
8th Grade Math Score	2.163***	1.977***	-0.017	0.592*
	(0.351)	(0.389)	(0.219)	(0.254)
N	4492	4263	4485	4267
R2	0.404	0.425	0.259	0.295
$\bar{y}$	73.617	66.456	3.563	1.897

Note: All columns show 2SLS estimates with NDR sample. Robust standard errors 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 Well-Developed MS and local linear controls for nonrandom tie-breakers. +p<0.1, *p<0.05, **p<0.01, ***p<0.001.

Table F.13: Demand Estimates

	Middle Schools est	Middle Schools se	High Schools est	High Schools se
Panel A: Preference Estimates				
Proportion of White/Asian/Other				
Black/Hispanic	-1.661	(0.125)	-2.477	(0.085)
FRPL	0.191	(0.146)	-0.462	(0.081)
5th Grade Test Score	0.106	(0.156)	-0.023	(0.037)
Proportion of non-FRPL				
Black/Hispanic	0.902	(0.232)	-2.136	(0.171)
FRPL	-1.207	(0.199)	-1.205	(0.214)
5th Grade Test Score	-0.351	(0.086)	-0.571	(0.088)
Proportion of High-Performers				
Black/Hispanic	0.004	(0.175)	0.285	(0.140)
FRPL	0.207	(0.159)	0.376	(0.150)
5th Grade Test Score	0.735	(0.065)	1.182	(0.059)
1(Well-Developed Middle School)				
Black/Hispanic	-0.083	(0.065)		
FRPL	0.182	(0.064)		
5th Grade Test Score	-0.016	(0.027)		
College Enrollment Rate				
Black/Hispanic			-0.327	(0.113)
FRPL			0.686	(0.092)
5th Grade Test Score			-0.662	(0.036)
Panel B: Middle School Type Effects				
Attending Well-Developed Middle School				
Proportion of White/Asian/Other			0.441	(0.082)
Proportion of non-FRPL			0.458	(0.149)
Proportion of High-Performers			1.151	(0.145)
Proportion of College Enrollment Rate			0.471	(0.113)
Panel C: Unobservable Tastes				
Type 1: Proportion of White/Asian/Other	2.552	(0.470)	3.775	(0.105)
Proportion of non-FRPL	2.102	(0.407)	5.047	(0.196)
Proportion of High-Performers	3.125	(0.181)	0.894	(0.166)
Proportion of College Enrollment Rate	0.362	(0.091)	4.145	(0.123)
Type 2: Proportion of White/Asian/Other	-1.207	(38.251)	14.752	(0.801)
Proportion of non-FRPL	0.216	(68.322)	15.032	(0.823)
Proportion of High-Performers	3.027	(27.741)	-12.292	(1.513)
Proportion of College Enrollment Rate	-0.045	(10.120)	-13.341	(0.820)
Panel D: Unobservable Tastes Probabilities				
Pr(MS type=1 & HS type=1)	0.649	(0.110)		
Pr(MS type=1 & HS type=2)	0.049	(0.007)		
Pr(MS type=2 & HS type=1)	0.275	(0.105)		
Pr(MS type=2 & HS type=2)	0.027	(0.010)		
Panel E: Other Parameters				
α_τ			-0.939	(0.071)
Outside option (White/Asian/Other)	5.291	(0.123)	6.016	(0.095)
Outside option (Black/Hispanic)	3.378	(0.084)	4.049	(0.091)
Distance	-0.655	(0.023)	-0.556	(0.004)

Note: We report the preference estimates of the main model described in Section 4. School characteristics "Proportion of White/Asian/Other," "Proportion of non-FRPL," "Proportion of High-Performers," "College Enrollment Rate" are between 0 and 1, and "1(Well-Developed Middle School)" is an indicator variable. Panel A reports coefficients on the interactions of each school characteristics with Black/Hispanic, FRPL status, and 5th Grade Test Score. Standard errors acquired from 50 bootstrap samples are reported in parentheses. We do not report the fixed effects estimates.

Table F.14: Well-Developed MS Attendance Effects on Moving

	(1) $\mathbb{1}$(Across Boroughs)	(2) $\mathbb{1}$(Across Districts)	(3) $\mathbb{1}$(Across Census Tracts)
From Well-Developed MS	0.004	-0.006	-0.017
	(0.009)	(0.016)	(0.021)
N	6371	6371	6371
R2	0.231	0.189	0.153
$\bar{y}$	0.038	0.117	0.235

Note: The table presents Well-Developed MS attendance effects on the propensity to move across boroughs, school districts, and Census tracts. There are 5 boroughs, 32 school districts, and 2,164 Census tracts in the city. We restrict the sample to students with non-degenerate risk of being offered (i.e., whose propensity score is in the interval $(0, 1)$ and hence subject to randomization) from any Well-Developed middle school. Robust Ssandard errors 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 control for dummy variables for all possible values of propensity score of being assigned to a Well-Developed MS and local linear control for nonrandom tie-breakers. +p<0.1, *p<0.05, **p<0.01, ***p<0.001.

Table F.15: Well-Developed MS Attendance Effects on Specialized High School Offer

	(1) $\mathbb{1}$(Took SHSAT)	(2) $\mathbb{1}$(Offer from any Specialized High School)
From Well-Developed MS	0.022	0.003
	(0.019)	(0.010)
N	6413	6413
R2	0.485	0.449
$\bar{y}$	0.397	0.091

Note: The table presents Well-Developed MS attendance effects on the propensity to take the specialized High Schools Admissions Test (SHSAT) or receive an offer from one of the Specialized high schools. There are 11 specialized high schools in NYC, including Stuyvesant and Bronx High School of Science, and students must take the SHSAT to apply to one of those schools. We restrict the sample to students with non-degenerate risk of being offered (i.e., whose propensity score is in the interval $(0, 1)$ and hence subject to randomization) from any Well-Developed middle school. Robust standard errors 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 control for dummy variables for all possible values of propensity score of being assigned to a Well-Developed MS and local linear control for nonrandom tie-breakers. +p<0.1, *p<0.05, **p<0.01, ***p<0.001.