Essays in Economics of Education

by

Clémence Idoux

M.Sc., Università Commerciale Luigi Bocconi, 2015

Submitted to the Department of Economics

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2021

Author ....

Department of Economics

May 14, 2021

Certified by ....

Joshua Angrist

Ford Professor of Economics

Thesis Supervisor

Certified by ....

Parag Pathak

Class of 1922 Professor of Economics

Thesis Supervisor

Accepted by ....

Amy Finklestein

John and Jennie S. MacDonald Professor of Economics

Chairman, Department Committee on Graduate Theses

Essays in Economics of Education

by

Clémence Idoux

Submitted to the Department of Economics on May 14, 2021

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Economics

Abstract

This thesis is composed of three essays on the economics of education. The first essay is about the heterogeneity of gains from selective school admission. The question of who benefits from selective school enrollment remains controversial. I show that Boston exam schools have heterogeneous effects on achievement. Impact differences are driven primarily by the quality of an applicant's non-exam-school alternative rather than by student demographic characteristics like race. Admission policies prioritizing students with the weakest schooling alternatives have the potential to increase the impact of exam schools on academic achievement. In particular, simulations of alternative admissions criteria suggests schemes that reserve seats for students with lower-quality neighborhood schools are likely to yield the largest gains.

The second essay is about understanding the impact of selective school admission screens on segregation in New York City schools. 70 years after *Brown v. Board of Education*, US school districts are still economically and racially segregated. School segregation is especially apparent in NYC, the largest US school district. I analyze the impact of two integration plans which reduced the role of screens in admission in two local NYC school districts. I show that abolishing selective admissions reduced both economic and racial segregation. Amending selective admission criteria also elicits substantial behavioral response from applicants. I find evidence that reducing the role of admission screens leads to White and high-income enrollment losses, which decreases the effect of the plans. On the other hand, applicants' changes in application behavior in response to the reforms increased the plans' impact on segregation.

The final essay is about predicting the effect of changes in school admission on students' enrollment. Such predictions are based on estimated student preferences, which in turn are obtained from the ranked order lists they submit. A concern is that an applicant with fixed preferences might submit different lists when faced with different admission criteria. For instance, an applicant could strategically take into account their probability of admission at each school, therefore violating the truthfulness assumption. A solution is to estimate preferences allowing students to strategically choose over all possible lists, but this runs into the curse of dimensionality as the choice space is large. This paper provides a model of applicants' list formation which presumes applicants use a simple heuristic in selecting their lists. In the model, applicants fill their list sequentially, without fully internalizing the dynamic consequences of each choice. Using this simplification, I estimate applicants' preferences, circumventing the dimensionality problem. I leverage an admission reform in NYC to estimate the model. Allowing applicants to deviate from truthfulness affects substantially their estimated preferences.

Acknowledgments

This dissertation would not exist without the guidance and support of my advisors. I am indebted to Joshua Angrist and Parag Pathak for precious feedback in the early stages of my projects, carefully reading through many drafts of my papers, and for inspiring me by example for the past five years. I also want to thank Nikhil Agarwal for reviewing carefully my third chapter and giving many insightful advices.

I am thankful to Daron Acemoglu, David Autor, Amy Filkenstein, and Andrei Shleifer for the many comments and suggestions that made this dissertation better.

I want to thank all the members of the SEII team that have helped me in many ways throughout the years. I am grateful to Eryn Heying and Anna Vallée for dependable administrative support and for helping me draft many presentations and proposals. I also thank Nicolas Jimenez and Vendela Norman for answering many of my data questions and reviewing my drafts carefully. I am particularly grateful to Chetan Patel for exceptional research assistance and for helping me with numerous computational problems.

I thank all my friends, and in particular Peter Badoo, Pauline Bocquet, Aurélien Cir, Florence Gomez, Michele Fornino, Juliette Fournier, Andrew Lilley, Giselle Lilley, Axelle Magnier, Andrea Maneira, Armando Miano, Nicola Rosaia, Elisa Rubbo, Parinitha Sastry, Catherine Seray and Beatriz Vasconcellos for supporting me all this time. For the last two years, I was very lucky to have the best possible officemates, Jonathan Cohen, Garima Sharma, Diana Sverdlin, Charles Serfaty and Martina Uccioli, with whom I shared countless discussions after the applied-micro seminar, lunches in the sun (weather permitting) and funny conversations. Thanks to all of you for having made the PhD experience much more enjoyable and for having managed to keep our office spirit intact despite physical distance. I am especially grateful to my friend Charles with whom I went through each step of the PhD. Thank you for your always good humor, your infinite generosity, your encouraging spirit and for bringing me back so many tasty products from your trips to France. I will miss our time in Cambridge together!

This dissertation is the endpoint of a ten year educational journey around the world, which would not have been possible without the unwavering support of my parents Françoise and Nicolas and of my brothers Martin, Vincent and François. I am also grateful to all the rest of my family for making me feel at home every time I came back.

Last but most important, I want to thank Gianluca Rinaldi, my partner who has supported me in the most difficult years of this journey with his encouraging words and affectionate care. I cannot thank you enough for keeping me grounded and reminding me of the important things in life. I am excited to share with you many other endeavors, intricate debates and happy moments!

Contents

Chapter 1

Who Benefits from Selective School Attendance?[1]

1.1 Introduction

Exam schools are the most sought-after public high schools in the United States. The underrepresentation of minority students at these elite institutions has been at the center of the education policy debate for decades. Recently, exam schools have been under pressure to change their admission criteria in order to increase access for minority and low-income students. School boards are considering a wide range of options, from lotteries to sophisticated place-based admission schemes. Yet, the consequences of these admission systems for achievement have received little attention.

The impact of admission systems on achievement depends on how different students are affected by the schools they attend. Research on exam schools focuses mostly on causal effects for marginal students, that is, students close to admissions

[1]Thanks to the Massachusetts Department of Elementary and Secondary Education and Boston Public Schools for graciously sharing data. The BPS exam school data was provided for this study as part of an ongoing research project studying exam schools at MIT. The views in this paper are those of the author and do not necessarily reflect the official policy or position of Boston Public Schools. I was not commissioned by Boston Public Schools or the Exam School Admissions working group to study any specific policy.

cutoffs (Dobbie and Fryer (2014) and Abdulkadiroglu et al. (2014)). Although mostly small and not statistically significant, these estimates for average marginal students might hide substantial heterogeneity. Applicants' gains from attending exam schools depend both on their personal characteristics and on the quality of non-exam school alternatives. These two sources of differences in potential gains are key to evaluating and comparing the effects of different elite high school admission criteria on overall academic achievement.

This paper estimates the achievement consequences of counterfactual admission criteria accounting for heterogeneity in potential gains. The paper begins with an econometric framework that isolates sources of heterogeneity. This decomposition is relevant to the exam school debate because new admission rules are likely to change the demographics of admitted students as well as the non-exam schools they substitute from. The estimates for Boston exam schools show substantial heterogeneity in exam school gains, driven by differences in the quality of students' non-exam school alternatives. I use these estimates to evaluate admission reforms and to design a scheme that increases overall academic achievement by leveraging differences in expected gains.

I estimate causal exam school effects using the IV methods introduced in Abdulkadiroğlu et al. (2017) and Abdulkadiroglu et al. (2019). I then decompose exam school effects by interacting instruments for exam school enrollment with covariates predicting where applicants would enroll if not offered an exam school seat. I use an applicant's distance to non-exam Boston public schools to predict her non-exam-school alternative. Due to the large number of schools, I split non-exam schools into four groups according to their quality as estimated with an OLS value-added model. This split takes advantage of the fact that the OLS value-added model provides biased but indicative estimates of school quality (Angrist et al. (2017)). Hence, the classification based on OLS value-added is likely to group together schools with similar effects on achievement.

This strategy identifies exam school treatment effect by non-exam-school alternative under the assumption of constant effects within strata (Hull (2015)). In particular, marginal changes in the relative distance to each group of schools should not be correlated with the potential treatment effect of attending an exam school. The data suggests that this assumption is likely to be satisfied, as marginal changes in relative distances are not systematically associated with variation in demographic characteristics or baseline test scores. Moreover, over-identification tests do not reject homogeneous treatment effects along relative distance.

Perhaps surprisingly, my empirical analysis suggests that the least selective of Boston's exam schools (the O'Bryant School) raises achievement the most. In particular, O'Bryant increases 7$^{\text{th}}$ and 8$^{\text{th}}$ grade math test scores by between 0.05 and 0.20 standard deviations for applicants who would have otherwise attended a Boston public school of below-median quality. At the same time, gaining admission at the two most selective Boston exam schools (Boston Latin school or Boston Latin Academy) appears to reduce math test scores with respect to O'Bryant. As a final step to the decomposition, I try to assess the heterogeneity in exam school effects across students with different pre-treatment characteristics but same quality of non-exam-school alternative. This subgroup analysis shows small positive exam school effects for Hispanic and Black applicants and applicants with low baseline math test scores when substituting from certain schools, although these estimates are less precise.

I use these results to compare the achievement effects of counterfactual admission schemes. Two popular alternatives, (i) granting admission to top-ranked middle school applicants and (ii) replacing exam school entrance test (ISEE) scores by state standardized test (MCAS) scores, have little impact on achievement. Similarly, a policy inspired by the temporary admission plan for the 2020-2021 application cycle, which reserves seats to each city zip code based on their school-aged population, does not result in any achievement gains. On the other hand, adopting place-based reserves based on neighborhood schools quality would increase 8$^{\text{th}}$ grade math test score by

0.13 standard deviations for 15% of applicants. Nonetheless, all these alternative admission schemes increase minority students' representation at exam schools.

The rest of the paper is organized as follows. Section **??** presents a framework that distinguishes non-exam-school alternative effects from match effects and discusses the data and the institutional background related to Boston's exam schools. Section **??** details the empirical strategy decomposing exam school effects by outside option. Section **??** presents the exam schools' estimated effects on achievement and the decomposition by non-exam-school alternative. It also discusses heterogeneous exam schools effects for different groups of students. Section **??** simulates the overall achievement gains from adopting several possible admission schemes. Section **??** concludes.

1.2 Background

1.2.1 Going beyond exam school RD estimates

It is crucial to account for treatment effect heterogeneity when evaluating the impact of different admission criteria schemes on academic achievement. Changes in exam school admission criteria affect the population of students that attend those schools and, unless the benefits of attendance are homogeneous, achievement gains estimated using previous cohorts will not apply to the new wave of admits.

In a setting with multiple treatments such as school choice, heterogeneity arises not only from differences in match effects between individuals and treatments but also from differences in outside option across individuals. For instance, Angrist et al. (2019) find that negative exam school effects in Chicago are explained by diversions away from high-performing charters in the Noble Network. Similarly, Chabrier et al. (2016) emphasize the importance of considering the difference in quality of urban and non-urban public schools when comparing the effectiveness of different kinds of charters. In practice, match effects and substitution effects interact, and disentangling them requires additional identification assumptions beyond those typically imposed when estimating elite school treatment effects.

To understand the significance of these sources of heterogeneity, consider a minimal set-up with individuals (students) indexed by i and treatments (schools) indexed by j. Let Y_{ij} denote the potential academic outcome (academic achievement) of student i if she attends school j. The treatment effect of attending school 0 (an exam school) instead of school j for student i can be decomposed as the mean difference in potential outcomes (the substitution effect) and student i specific difference in potential outcomes (the match effect):

$$Y_{i0} - Y_{ij} = \underbrace{E[Y_{i0} - Y_{ij}]}_{\text{substitution effect}} + \underbrace{\epsilon_{i0} - \epsilon_{ij}}_{\text{match effect}} . \tag{1.1}$$

Suppose for simplicity that there is a single student characteristic X_i (e.g. race) that affects potential outcome Y_{ij} differently depending on the school j that the student attends. This corresponds to a situation where students of different races have different benefits from attending each school, but gains are otherwise homogeneous. Formally, this setting can be expressed as $\epsilon_{ij} = E[\epsilon_{ij} \mid X_i = x_i] + \nu_i$. With this simplification, the match effect corresponds to the covariate specific difference in outcome when switching from school 0 to school j:

$$Y_{i0} - Y_{ij} = \underbrace{E[Y_{i0} - Y_{ij}]}_{\text{substitution effect}} + \underbrace{E[\epsilon_{i0} - \epsilon_{ij} \mid X_i = x_i]}_{\text{match effect}} . \tag{1.2}$$

This decomposition may be used to compare students' academic achievement under different exam school admission rules. A change in exam school admission rules affects both the set of schools students are substituting from and the types of students that substitute from these schools. Academic achievement will be larger if students that enroll in exam schools substitute from non-exam-school alternatives in a way that yields large positive substitution effects and large positive match effects.

Hence, comparing the performance of different students assignment to schools in terms of overall academic achievement requires knowing the substitution effect and match effect for each pair of schools and student characteristic. When trying to find the exam school assignment scheme that maximizes academic achievement, substitution effects determine from which school students should be principally reallocated, while match effects indicate which students from within each school should be reallocated. Similarly, the impact of modifying exam schools' admission criteria will depend on the substitution and match effects with respect to exam schools for the students being displaced by the change.

1.2.2 Boston exam schools

Boston is a compelling setting for studying the impact of different elite school admission criteria on achievement. Since Judge W.A. Garrity ordered Boston Latin School, the most selective Boston exam school, to set apart 35% of seats to minority applicants in 1978, exam school admission rules have been a contentious topic. After the unsuccessful attempt of McLaughlin v. Boston Sch. Comm. (1996), Wessmann v. Boston Sch. Comm. (1998) put an end to seat reserves for minority applicants. Admissions thus went back to being solely based on grades and entrance exam results. In October 2020, the Boston School Committee approved a temporary change of admission regime, in part as a response to the challenges created by the Covid-19 pandemic[2]. The plan suspends the entrance exam for a year and sets apart seat reserves for the city's different zip codes. The plan might be adopted permanently if deemed successful. Accurately predicting the achievement effect of a change in admission criteria is thus immediately policy-relevant.

Boston has three elite high schools: Boston Latin School (BLS), Boston Latin Academy (BLA) and the O'Bryant School of Mathematics and Science (OBR). Students can apply for admission either in 7^{th} grade or in 9^{th} grade. Applicants can

[2]The Exam School Admissions Working Group deemed the logistics of administrating the exam school entrance test too complicated

decide to apply to all three schools or to a subset of them and may express their preferences over schools by submitting a rank-order list. Admission at each exam school is based on a school-specific weighted average of middle school GPA and of the Independent School Entrance Examination (ISEE). Each applicant receives a rank at each school she applies to.

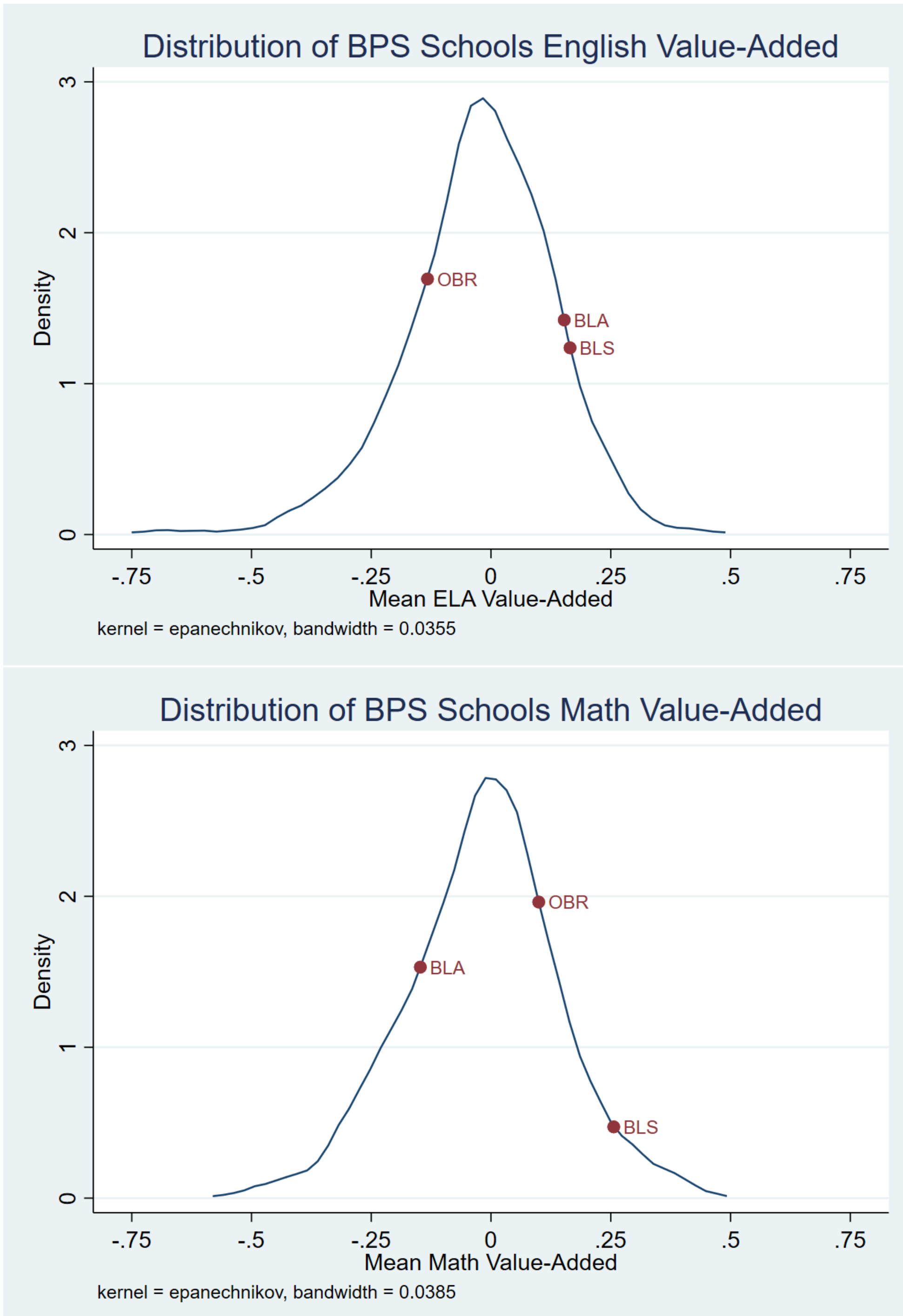
Exam school offers, reconciling applicants' preferences, rankings and schools capacities, are generated using deferred acceptance. This mechanism produces admission cutoffs for each exam school that can be exploited to identify exam school achievement effects. Boston Latin, with an admission rate of 27%, is the most selective of the three schools. It is closely followed by Latin Academy, which on average admits 46 % of its applicants. Finally, O'Bryant admits 56% of its applicants, making it the most easily attainable exam school. While more than 95% of admitted students at Boston Latin and Latin Academy accept their offer, take-up is only 80% at O'Bryant.

Applicants who fail to gain entrance into any of the exam school may enroll in one of Boston public schools. The performance of these schools in raising students' test scores is heterogeneous, as suggested by Figure 1 which displays the distribution of the average estimated value-added of BPS schools [3] on 8^{th} grade MCAS English and math test scores. Schools at the bottom of the estimated value-added distribution increase students' 8th grade test scores by one standard deviation less than schools at the top of the estimated value-added distribution.

Exam schools appear to perform slightly better than the median Boston public school. Boston Latin and O'Bryant appear particularly effective at raising math test scores, while Boston Latin and Latin Academy are among the best schools for increasing English test scores. Although these value-added estimates could be misleading as they are not necessarily unbiased, their dispersion suggests that changing exam school

[3] School value-added is computed following the same model discussed in section 3 for each year between 2004 and 2016. This model controls for student demographics and flexible functions of baseline test scores.

Figure 1.1: English and Math Value-Added of Boston Public Middle Schools



Note: This figure shows the distribution of Boston public middle schools' English and Math Value-added for years 2004-2016. The plots display the mean value added of each school over the period.

assignment could result in achievement gains. Indeed, exam school applicants with non-exam-school alternative in the lower tail of the value-added distribution could benefit from attending an exam school.

1.2.3 Data on Boston students

Students can apply to exam schools either in the spring of their 6^{th} grade for enrollment in 7^{th} grade, or in the spring of their 8^{th} grade for enrollment in 9^{th} grade. Each year, approximately two thousand 6^{th} graders apply to one of these schools. Applicants come from both private sector schools and public sector schools (which includes charters). Applicants who do not receive an offer, or who decline their offer, can choose to enroll in the public school system or in a private school. Admission to most of Boston's regular public schools and charter schools occurs in 6^{th} grade. Admission in 7^{th} grade is thus a peculiarity of exam schools, so most applicants' non-exam-school alternative is the school they are enrolled in when they apply. This detail, combined with the fact that it is the larger application round, motivates my focus on applications to 7^{th} grade.

The analysis sample includes 7^{th} grade applicants from years 2004–2016 for which both baseline and outcome test scores are available.[4] Since I am is interested in studying substitutions from regular public schools, I restrict the analysis sample to students enrolled in a non-charter BPS school in the Boston area prior to application. This excludes students who enroll in a charter school after applying to an exam school, since admissions to Massachusetts charter schools occur either in 5^{th} or 6^{th} grade through lotteries.[5] Hence, the analysis only considers students who either enroll in an exam school or in a non-charter BPS school.

Boston Public Schools (BPS) is the source of the application files for exam schools.

[4] Baseline test scores correspond to MCAS 4^{th} grade test scores. Thus, students who were not enrolled in a Boston Public school in 4^{th} grade are excluded from my analysis.

[5] Among applicants who do not receive an offer from any exam schools, only 3% of non-charter applicants subsequently enroll in a charter school, while 73 % of charter applicants remain in a charter school.

The Massachusetts Department of Elementary and Secondary Education (DESE) provided the enrollment files and MCAS data. Application, enrollment and test scores files are merged using the State unique identifier (SASID). The distance from each sending middle school to each potential school of enrollment is computed using the shortest road distance between the two. Only schools where at least one student enrolled after application during the 2004-2016 period are included in the set of potential non-exam-school alternatives. I include a more detailed description these data sets, how I constructed the distance variable, and the sample restrictions in Appendix A.

1.3 Empirical strategy

1.3.1 Instrument for exam school enrollment

The identification of exam school achievement effects requires an instrument since exam school students are positively selected. I use exam school offer conditional on the probability of receiving an offer to instrument exam school enrollment. Receiving an exam school offer strongly predicts exam school enrollment and is as good as randomly assigned conditional on the propensity score of receiving an offer (Abdulkadiroğlu et al. (2017) and Abdulkadiroglu et al. (2019)). To compute the propensity score of admission at each exam school, I exploit the rank order list of applicants following the methodology in Abdulkadiroglu et al. (2019). Appendix B describes the method in more detail. Although this method is also based on discontinuities at admission cutoffs, it is more general than the approaches used in Abdulkadiroglu et al. (2014) and Dobbie and Fryer (2014) because it combines the variations at each exam school's cutoff.

In practice, since the match yields a single offer, the sum of each exam school's propensity score corresponds to the risk of being assigned at any of the exam schools.The exam school effect on any outcome Y_i can thus be estimated with a just-identified

2SLS procedure that uses an offer from any exam school $D_i = \sum_s D_{is}$ to instrument for enrollment at an exam school E_i, controlling for linear control functions $g_s(\cdot)$ and $h_s(\cdot)$ of the running variables R_{is}, and for the exam school propensity score $\hat{p}_i$. The sample is limited to applicants whose probability of receiving an exam school offer is not equal to 0 or 1. Specifically, I estimate the following regression:

$$Y_i = \beta E_i + \sum_x \alpha_x \mathbb{I}(\hat{p}_i = x) + \sum_s g_s(R_{is}) + \epsilon_i. \qquad \text{(Second stage)}$$

Since there are three exam schools in Boston, the propensity score takes on three different values $x = \{0.895, 0.75, 0.5\}$ for applicants with non-degenerate risk of any exam school offer. For more flexibility, I allow the coefficients associated with each of these values to vary by cohort. The linear control function for each school's running variable is also allowed to vary by cohort. I parameterize these functions as

$$g_s(R_{is}) = \omega_{1s} a_{is} + \kappa_{is} \left[\omega_{2s} + \omega_{3s}(R_{is} - \tau_s) + \omega_{4s}(R_{is} - \tau_s)\mathbb{I}(R_{is} > \tau_s)\right].$$

where a_{is} indicates whether applicant i applied to school s, and $\kappa_{is} = a_{is} \times \mathbb{I}(\tau_s - \delta_s < R_{is} < \tau_s + \delta_s)$ selects applicants in a bandwidth of size δ_s around an admission cutoff τ_s.

The corresponding first stage is

$$E_i = \gamma D_i + \sum_x \delta_x \mathbb{I}(\hat{p}_i = x) + \sum_s h_s(R_{is}) + \nu_i. \qquad \text{(First stage)}$$

This approach can also identify the treatment effect of each of the exam schools separately. The multi-school specification instruments the dummy for enrollment at each exam school (E_{1i}, E_{2i} and E_{3i}) by the corresponding school offer and controls for the values of each exam school's propensity score. In this case, the sample includes

Table 3. Covariate Balance

	Any Exam Offer		Distance to groups			School Specific Offers		
	Offer gap	p-value	Joint F test	p-value	N	Joint F test	p-value	N
	(1)	(2)	(3)	(4)	(5)	(6)	(8)	(9)
Female	-0.031	0.376	1.144	0.331	5,978	1.242	0.294	8,144
Hispanic	0.025	0.391	1.483	0.219	5,978	0.554	0.646	8,144
Black	0.014	0.693	9.447	0.000	5,978	0.669	0.571	8,144
White	0.006	0.779	0.783	0.504	5,978	0.681	0.564	8,144
Asian	-0.036	0.199	1.171	0.321	5,978	0.665	0.574	8,144
Free/Reduced lunch	-0.006	0.847	0.356	0.785	5,978	1.242	0.294	8,144
Native	0.055	0.134	1.992	0.115	5,978	2.238	0.083	8,144
MCAS English 4th	0.060	0.317	1.603	0.188	5,978	0.984	0.400	8,144
MCAS English 6th	0.029	0.541	0.294	0.830	4,897	0.501	0.682	6,772
MCAS Math 4th	0.020	0.729	1.577	0.194	5,978	0.358	0.784	8,144
MCAS Math 6th	-0.046	0.382	0.241	0.868	5,941	0.335	0.800	8,095

Notes: This table reports estimates of offer and distance to groups effects on covariates for 2004-2016 applicants. Columns 1 and 2 report any-exam offer effects; columns 3-4 show the joint F-statistics of distances to the four groups of schools, columns 5-6 show the joint F-statistics of school-specific offers. Sample sizes in column 5 count the number of observatons with non degenate risk of any-exam offer, sample sizes in column 9 count the number of observatons in the bandwidth of at least one of the schools. Sample sizes are smaller for 6th grade MCAS tests since they were first introduced in 2004 and 2006. Models include school-by-year linear running variable controls and either exam school offer risk or school specific offers risk controls, as specified in the text. Standard errors clustered at the school of origin and application year are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

applicants in the bandwidth of any of the three exam schools.

Table 3 presents encouraging evidence of covariate balance by offer status. In particular, receiving an offer from any of the exam schools is not correlated with higher baseline test scores, which bolsters confidence in the validity of the instrument.

1.3.2 Estimation of exam school treatment effect heterogeneity

Econometric framework

The 2SLS estimators of exam school effects for both all applicants and subgroups of applicants identify a mix of match effects and substitution effects, and are thus not suitable for performing counterfactual analyses. [6] The problem with these estimators is that they fail to address substitution effects, since they ignore the fact that students

[6]See Appendix C for details on what LATE identifies.

are substituting away from different alternatives. [7]

Substitution effects are hard to pin down as they correspond to unobserved choices that need to be inferred. Behaghel et al. (2013), Kline and Walters (2016), Blackwell (2017), Lee and Salanié (2018) and Mountjoy (2019) consider non-parametric identification of multiple treatment channels and multiple treatment-specific instruments[8]. This paper follows most closely the method outlined in Hull (2015), where treatment alternatives are unobserved but vary along observable dimensions. In this case, interacting the instrument with covariates predicting individuals' outside option is an intuitive way of identifying substitution effects. Nonetheless, this approach only allows for identification of both substitution and match effects under additional assumptions. In particular, Hull (2015) establishes that covariate's interaction identifies treatment effect by non-exam-school alternative only under the assumption of constant treatment effect within strata.

Formally, let S_i denote the school in which student i enrolls. S_i takes values from 1 to J for non-exam schools and value 0 for the exam school. Assume there exists a set of $J-1$ covariates $\{W_{ik}\}_{k=2}^{J}$ satisfying Assumption 1. Then Proposition 1 states that interacting these covariates with the instrument for school 0 identifies the treatment effect by non-exam-school alternative .

Assumption 1

1. ***Relevance:*** *$Pr[S_i = j \mid W_i = w_i] \neq Pr[S_i = j \mid W_i = w_i']$ if $w_i \neq w_i'$ $\forall j$ where W_i is the vector of covariates $\{W_{ik}\}_{k=2}^{J}$.*

2. ***Partial unordered monotonicity:*** *For any $w_j < w_j'$ and holding fixed w_{ik} $\forall k \neq j$, $Pr[S_i = j \mid W_{ij} = w_j, \{W_{ik}\}_{k\neq j}] \geq Pr[S_i = j \mid W_{ij} = w_j', \{W_{ik}\}_{k\neq j}]$ and $Pr[S_i = k \mid W_{ij} = w_j, \{W_{ik}\}_{k\neq j}] \leq Pr[S_i = k \mid W_{ij} = w_j', \{W_{ik}\}_{k\neq j}]$ $\forall i$*

[7] Heckman and Urzúa (2010) underlines that IV cannot identify treatment effects for different margins of choice without additional structural assumptions.

[8] Kirkebøen et al. (2017) show how an IV strategy identifies counterfactual specific LATEs when preferred treatment alternatives are directly measured.

3. ***Constant treatment effect within covariate:*** *$E[Y_{i0} - Y_{ij} \mid W_i = w_i] = E[Y_{i0} - Y_{ij} \mid W_i \neq w_i]$ for all j and vector w*

The first condition of Assumption 1 is a first stage condition: the set of covariates must predict individuals' non-exam-school alternative . The second condition generalizes the standard monotonicity assumption from the binary case: each covariate shift renders each treatment either weakly more attractive for all individuals or weakly less attractive for all individuals. This rules out the possibility of compliers flowing in and out of each outside option in response to a shift in the corresponding covariate. The third condition entails that the treatment effect of attending one school instead of another is independent of the covariates predicting the non-exam-school alternative . In the previous framework, it implies that $E[X_i \mid W_{ik}] = E[X_i]\ \forall k$ since X_i is the only determinant of a school's potential effect. This assumption guarantees that variations along interacted covariates influence outcomes only through changes in the outside option.

Proposition 1 (Identification of treatment effect by non-exam-school alternative) *Suppose there exists a valid instrument Z_i for enrollment in school 0 , and a vector of $J-1$ covariates $\{W_{ik}\}_{k=2}^{J}$ that satisfies Assumption 2.*

1. *Conditioning on $\{W_{ik}\}_{k=2}^{J}$ and Z_i, the interaction of Z_i and $(1, \{W_{ik}\}_{k=2}^{J})$ identifies*
   $E[\bar{Y}_0 - \bar{Y}_j \mid D_{i1} > D_{i0}],\ \forall j = 1, ..., J.$

2. *For any covariate X_i, conditioning on $\{W_{ik}\}_{k=2}^{J}$ and X_i, the interaction of Z_i, $(1, \{W_{ik}\}_{k=2}^{J})$ and X_i identifies $E[\bar{Y}_0 - \bar{Y}_j \mid X_i, D_{i1} > D_{i0}], \forall j = 1, ..., J.$*

Proposition 1 establishes that the interaction of covariates predicting each student's outside option with an instrument for enrolling in an exam school identifies exam schools' treatment effects with respect to each non-exam-school alternative, as long as the covariates satisfy the constant treatment within covariate assumption.

This proposition is along the lines of Hull (2015): it extends his result to more than two outside options, and its proof in Appendix D considers a continuous interacted covariate.

Imposing a constant treatment effect within the interacted covariates is less demanding than assuming a constant treatment effect in general, or assuming that interacted covariates are exogenous. A constant treatment effect would rule out any match effect between students and schools since it entails that any two students should benefit equally or lose equally from attending an exam school instead of another school. Exogeneity would limit the set of valid covariates since it requires covariates to not be correlated with potential outcomes (not only treatment effects). A constant treatment effect within interacted covariates, however, only rules out heterogeneity in treatment effect along dimensions that vary substantially with the covariates predicting outside options. It allows heterogeneous match effects between schools and different types of students as long as these types are distributed equally along values of the covariates used to predict non-exam-school alternative.

Testing for a constant treatment effect within the interacted covariates is nonetheless challenging. Over-identification tests of homogeneous treatment effects are typically conducted across covariates, since these tests compare the treatment effect induced by different covariates, not different values of the same covariate. A conclusive over-identification test would thus require all interacted covariates to induce variations of a similar type (e.g. age variations, race variations or geographic variations).

Estimation of exam school effect by non-exam-school alternative

I identify heterogeneity in exam school achievement effect by non-exam-school alternative by interacting distance to schools with the exam school instrument. A long-standing literature has used distances to predict students' school of enrollment (Card (1995), Neal (1997), Booker et al. (2011), Walters (2018) and Mountjoy (2019)).

Nonetheless, contrary to the approach taken in these papers, I do not use relative distances as instruments for school enrollment, but instead as covariates to be interacted with an instrument for enrollment. Hence, my strategy does not require exogeneity of relative distances but rather homogeneity of treatment effect along relative distances.

Considering the large number of potential non-exam-school alternatives, I need to group schools.[9] The decomposition of treatment effect aims at identifying which schools perform worse than exam schools so that students substituting from those would gain from attending an exam school instead. Hence, it is appropriate to group together schools of similar quality .[10] Sorting schools based on their estimated OLS value-added (VA) is likely to result in groups of schools with similar effects on achievement as the bias in the VA model controlling for observables and past achievement is small (Chetty et al. (2014) and Angrist et al. (2017)). Moreover, Angrist et al. (2017) argues that, bias notwithstanding, policy decisions in Boston middle schools based on conventional VA models are likely to generate substantial achievement gains.

Schools are sorted depending on their estimated value-added from a "lagged score" OLS VA model. The model includes indicators for sex, race, subsidized lunch eligibility, special education status (SPED), English-language learner status (ELL) and school year, along with cubic functions of all the baseline math and ELA test scores available.[11] For each application year, I estimate the model on the two previous years' sample of BPS schools enrolling more than 25 students, which captures the value-added of each school at the time of application. As each school's estimated Value-Added is year-specific, a school may be classified in different groups across years.

[9]Moreover, it is not possible to identify the treatment effect of each specific school using distances, as any other point is uniquely defined by its distance from 3 non-collinear reference points on a plane.

[10]Using this method, one may explore classification of schools according to other attributes believed to be relevant. Whichever characteristic is used to categorize schools, the 2SLS procedure should produce unbiased estimates of exam school treatment effects with respect to each group of schools.

[11]MCAS exams for English and math were progressively introduced for each grade in the 2000s. By 2006, BPS students were tested in every grade between 3rd grade and 8$^{\text{th}}$ grade, providing a rich set of past test scores.

Choosing the number of groups represents a trade-off between informativeness, identification and precision. Table A2 offers a comparison of results for different sample splitting procedures. While constant treatment seems to be satisfied to the same extent, estimates become quite imprecise when using more than five groups. The most precise and informative estimates are achieved with four groups, since exam schools appear to perform as well as the median non-exam-school alternative. Schools are thus sorted according to the quartiles of the estimated VA distribution for each year.

As described in Table 1, group 1 schools have the lowest estimated Math VA with an average of -0.22, while group 4 schools have the largest estimated math VA with an average of 0.19. Exam schools have an average estimated math VA of 0.05, which makes them similar to a school in the bottom of group 3. According to these estimates, Boston Latin is the most effective school with a math VA of 0.20, O'Bryant is second at 0.11 and Latin academy lies behind at -0.15. Math and English VA appear to be correlated: schools with low estimated math VA also tend to have low estimated 7^{th} grade English VA. Nevertheless, Latin Academy has the largest English VA of 0.16, while Boston Latin is second at 0.13 and O'Bryant third at -0.13. Surprisingly, Table 1 does not reveal substantial heterogeneity between groups in terms of ethnic composition, share of English learners, SPED, native speakers or baseline achievement. If anything, group 1 schools appear to concentrate a larger share of Asian students than schools in other groups.

To disentangle counterfactual-specific treatment effects, I estimate the following specification where $\{S_{ik}\}_{k=1}^{4}$ are dummies indicating enrollment at group k school:

$$Y_i = \beta_0 + \sum_{k=1}^{4} \beta_k S_{ik} + \epsilon_i.$$

This specification is similar to a value-added model with exam schools as the reference group, so that β_k gives the effect of attending a school of group k with respect to an exam school. This model is evaluated on the sample of applicants that

Table 1. Characteristics of Groups of Schools

	Exam schools (1)	Group 1 (2)	Group 2 (3)	Group 3 (4)	Group 4 (5)
Female	0.55	0.46	0.46	0.47	0.47
Hispanic	0.18	0.32	0.36	0.41	0.39
Black	0.24	0.34	0.40	0.40	0.32
White	0.31	0.08	0.06	0.08	0.10
Asian	0.23	0.22	0.14	0.08	0.13
Other ethnicity	0.04	0.05	0.03	0.04	0.06
Free/reduced lunch	0.47	0.81	0.82	0.82	0.80
Native	0.66	0.50	0.54	0.56	0.57
SPED	0.01	0.19	0.20	0.24	0.22
English learners	0.03	0.22	0.23	0.20	0.20
MCAS Egnlish 4th	1.18	-0.16	-0.27	-0.24	-0.17
MCAS English 6th	1.18	-0.20	-0.30	-0.26	-0.08
MCAS Math 4th	1.23	-0.12	-0.26	-0.23	-0.18
MCAS Math 6th	1.39	-0.18	-0.34	-0.20	-0.12
Mean math VA	0.05	-0.22	-0.07	0.00	0.19
Mean english VA	0.09	-0.23	-0.10	-0.10	0.08
Number of students at risk	4920	861	762	867	734
Number of schools	3	7	7	7	7

Notes: This table presents the characteristics of each group based on students enrolled between 2004 and 2016. Schools in each group are weighted by the number of students at risk with 8th grade Math MCAS enrolling in the school.

have non-degenerate risk, i.e., applicants that are only marginally offered a seat at any of the three exam schools.

To construct instruments for enrollment in each group, I interact receiving an exam school offer with distances from applicants' middle schools to junior high schools. These distances plausibly predict applicants' non-exam-school alternative, since applicants that do not receive an exam school offer are likely to either stay in the school they were enrolled in at the time of application or move to a school in the same neighborhood. Specifically, I interact the distance d_{ik} of each applicant's middle school to the closest school of each group k with a dummy for receiving an exam offer D_i. Thus, enrollment at a school of group $k \in \{1, ..., 4\}$ is instrumented as

$$S_{ik} = \gamma_k D_i + \sum_{j=1}^{K} \lambda_{kj}(d_{ij} \times D_i) + \sum_{j=1}^{K} \phi_{kj} d_{ij} + \sum_{x} \delta(x)_k \mathbb{I}(\hat{p}_i = x) + \sum_{s} h_{sk}(R_{is}) + \eta_{ik}.$$

Table 2 presents the estimated first stages for each group of schools. Relative distance is a good predictor of school enrollment. The first stage estimates appear to be strong, with F-statistics between 64 and 75. The coefficient on the distance to the closest school of each group is systematically negative, meaning that students prefer either staying at the school they attend at the time of application or moving to a school close to it. Moreover, the interaction between the exam school offer and the minimum distance to each group is positive for each group, meaning that an exam school offer shifts people from a school only if it is close enough to constitute one of the potential non-exam-school alternatives. Finally, distance effects appear to be very small and generally insignificant for exam school enrollment, which is consistent with the fact that exam school enrollment depends mostly on receiving an offer.

Test of the constant treatment effect assumption

Proposition 2 states that identifying the treatment effect by non-exam-school alternative requires that the set of covariates interacted with the instrument satisfies the assumption of constant treatment effect within covariate. Since the model controls for the distance to the closest school of each group, the identifying variation comes from differences in relative distances to each group's closest school. Hence, the constant treatment effect assumption requires that marginal changes in relative distance to the closest school of each group are not correlated with changes in the potential treatment effect of attending an exam school. Otherwise, variations in relative distance would not exclusively impact the outcome through changes in the non-exam-school alternative, and would thus not identify the pure substitution effect.

To make this assumption clearer, consider an example with three groups of schools and two students, A and B. These students are equidistant from group 1 and 2 schools, but A lives closer to a group 3 school. The constant treatment effect assumption assumes that A does not gain more or less than B from attending an exam school instead of a group 3 school. This seems plausible since A and B only differ in their relative position to a school of group 3.

Table 2. First Stage Estimates for Fallback School Decomposition

	Group 1 school	Group 2 school	Group 3 school	Group 4 school
	(1)	(2)	(3)	(4)
Exam offer	-0.172***	-0.210***	-0.233***	-0.174***
	(0.041)	(0.042)	(0.046)	(0.043)
Exam offer X distance to group 1	0.113***	-0.036***	-0.042***	-0.038***
	(0.008)	(0.007)	(0.006)	(0.007)
Exam offer X distance to group 2	-0.038***	0.106***	-0.028***	-0.035***
	(0.006)	(0.006)	(0.005)	(0.006)
Exam offer X distance to group 3	-0.041***	-0.047***	0.110***	-0.022***
	(0.006)	(0.007)	(0.008)	(0.007)
Exam offer X distance to group 4	-0.035***	-0.032***	-0.025***	0.089***
	(0.005)	(0.006)	(0.006)	(0.007)
Distance to group 1	-0.128***	0.039***	0.046***	0.042***
	(0.008)	(0.007)	(0.007)	(0.007)
Distance to group 2	0.045***	-0.115***	0.031***	0.039***
	(0.006)	(0.007)	(0.006)	(0.006)
Distance to group 3	0.047***	0.052***	-0.121***	0.023***
	(0.007)	(0.008)	(0.009)	(0.008)
Distance to group 4	0.039***	0.035***	0.027***	-0.100***
	(0.006)	(0.006)	(0.006)	(0.008)
F-statistic	64.1	72.0	**76.1**	**74.5**
N	5,978	5,978	5,978	5,978

Notes: This table reports first stage estimates for enrollment in different types of schools for 2004-2016 exam applicants with non-degenerate any-exam offer risk. The instrumental variables are receiving an exam school offer and its interactions with distances (in kms) to each group's closest school, controlling for each group's distance (in kms). The sample is limited to applicants with MCAS 8th grade math score. All models control for exam schools offers risk, school-by-year linear running variables and the set of variables listed in Table1, except MCAS 6th grade scores. Standard errors clustered at the school of origin and application year are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. .

As a check for the constant treatment effect assumption, Table 3 assesses whether characteristics that could influence treatment effects vary along relative distances to the closest school of each group. All the coefficients are relatively small and only a few are significant. While the probability of being a Black student appears to change along relative distances, all past test scores do not vary significantly. This mitigates the concern that my results are affected by heterogeneity in treatment effect along relative distances.

Finally, the constant treatment assumption can be tested through an over-identification test when distances are used as interacted covariates. Indeed, the model is over-identified since the exam school enrollment dummy is omitted from the model, but the exam school offer dummy and its interactions with distances from all groups are used as instruments. Including one additional distance in the set of instruments allows for an over-identification test for a homogeneous treatment effect along changes in relative distances to schools. This test is particularly powerful since it compares the estimated treatment effect values by varying the geographic definition of complier groups. It is thus a test of homogeneity along variations in the characteristic used to identify the outside option.

1.4 Empirical results

1.4.1 Overall exam school effect

Panel A of Table 4 reports the LATE of enrolling in any of the exam schools on academic achievement. Despite the differences in sample and in method, I find that exam schools have a null or negative effect on test scores, consistent with Abdulkadiroglu et al. (2014) and Dobbie and Fryer (2014). Overall, exam school enrollment appears to reduce an applicant's English score by -0.107 SD in 7^{th} grade and by -0.058 SD in 8^{th} grade, while it is associated with insignificant changes in math achievement of -0.052 SD in 7^{th} grade and 0.029 SD in 8^{th} grade.

Table 4. 2SLS Estimates of Exam School Effects

	MCAS English		MCAS Math	
	7th grade	8th grade	7th grade	8th grade
	(1)	(3)	(5)	(7)
	Panel A: Overall Exam school effect			
Enrollment in exam school	-0.107**	-0.058	-0.052	0.029
	(0.049)	(0.051)	(0.066)	(0.070)
First stage coefficient	0.79	0.79	0.78	0.79
	Panel B: Exam school effect by non-exam alternative school			
Enrollment in group 1 school	0.022	0.013	-0.050	-0.193**
	(0.062)	(0.063)	(0.086)	(0.086)
Enrollment in group 2 school	-0.046	0.005	-0.144*	-0.124
	(0.068)	(0.066)	(0.080)	(0.092)
Enrollment in group 3 school	0.197***	0.093*	0.120	0.047
	(0.062)	(0.056)	(0.077)	(0.086)
Enrollment in group 4 school	0.264***	0.125*	0.319***	0.185**
	(0.063)	(0.067)	(0.090)	(0.092)
Overid p-value	0.426	0.968	0.065	0.221
Implied LATE	-0.108	-0.058	-0.057	0.026
N	5,953	5,978	5,429	5,978

Notes: This table reports 2SLS estimates of school enrollment effects for 2004-2016 exam applicants with non-degenerate any-exam offer risk. The endogenous variable is enrollment either enrollment in 7th grade at any exam school in panel A or at one of the group of schools in Panel B. Panel B coefficients correspond to the effect of enrollment at one of the groups of non-exam alternative schools with respect to enrollment at any exam school. All models control for exam school offer risk, school-by-year linear running variables and the set of variables listed in Table1, except MCAS 6th grade scores. The implied LATE is computed by taking the negative of the sum of panel B coefficients weighted by the share of compliers going to each group of non-exam alternative schools. Standard errors clustered at the school of origin and application year are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. .

The estimated exam school effect is mainly driven by O'Bryant, the least selective exam school. Indeed, figure 2 shows that among applicants with non-degenerate exam school risk, 92 % of offered compliers enroll at O'Bryant and 8% enroll at Boston Latin. These figures are consistent with the fact that 80% of non-offered compliers to a Boston Latin offer enroll either at Latin Academy or O'Bryant, whereas 98% of non-offered compliers to a Latin Academy offer enroll at O'Bryant. Only compliers to an O'Bryant offer enroll a traditional public school more than 40% of the time when not admitted at O'Bryant.

These enrollment patterns among exam school offer compliers are a consequence of the application pattern. Most applicants rank Boston Latin first, then Latin Academy and finally O'Bryant [12]. As a result, most marginal applicants to Boston Latin or Latin Academy clear the admission cutoff at either Latin Academy or O'Bryant. As such, these applicants do not have non-degenerate risk of being seated at any exam school and do not identify the effect of attending an exam school rather than a traditional public school.

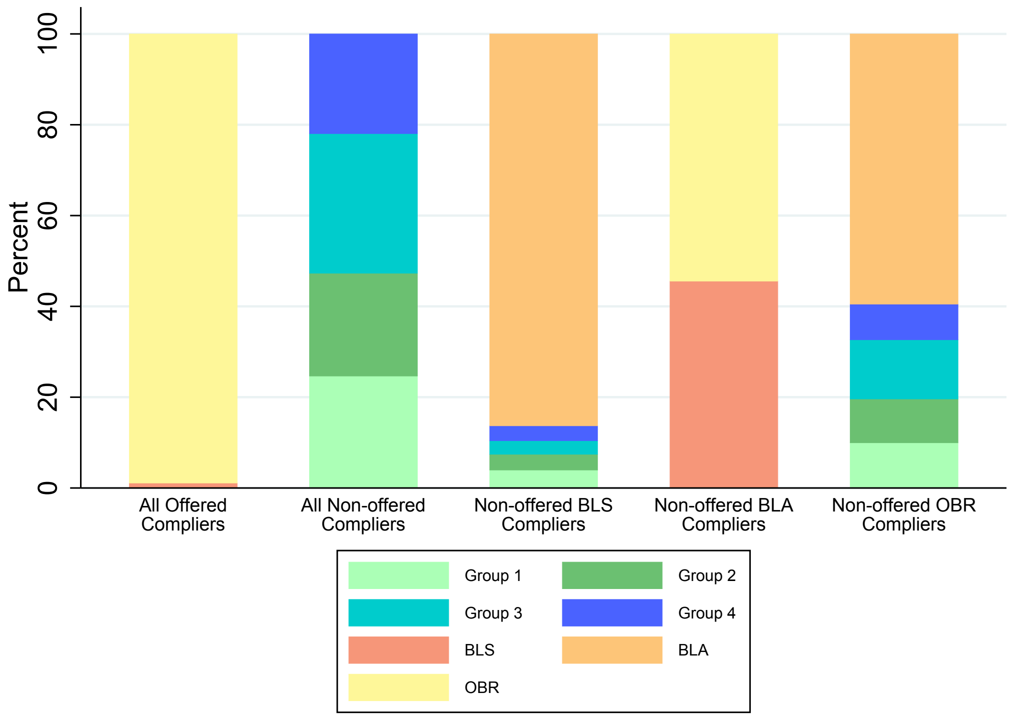
Hence, any comparison between exam schools and traditional public schools will be a comparison between O'Bryant and the traditional sector. Similarly, the decomposition of exam school effect by outside option will give estimates of the effects for the different groups of schools with respect to O'Bryant.

1.4.2 Decomposition of exam school effect

While the estimates in Table 4 panel A confirm that enrolling at O'Bryant does not increase academic achievement on average, panel B shows substantial heterogeneity in the O'Bryant treatment effect by non-exam-school alternative. Specifically, applicants with non-exam-school alternatives in the two first quartiles of estimated VA benefit from attending O'Bryant. O'Bryant enrollment increases the 7^{th} and 8^{th} grade math

[12] 61% of applicants in my sample submit these exact preferences, an additional 14 % of applicants invert the order of Boston Latin and Latin Academy, and only 8 % of applicants rank O'Bryant first.

Figure 1.2: Enrollment Destinies by Exam School



Note: This figure shows the enrollment destinies of exam school compliers when offered and not offered an exam school seat. Enrollment compliers are applicants who attend an exam school when offered a seat but not otherwise. The 1^{st} bar plots exam destinies for applicants when accepted in any of the three exam schools. The 2^{nd} bar plots non-exam destinies for applicants rejected from all exam schools. The 3^{rd} bar plots destinies for rejected Boston Latin applicants, the 4^{th} bar for rejected Latin Academy applicants and the 5^{th} bar for rejected O'Bryant applicants. Destinies are estimated as in Abdulkadiroglu et al. (2014). Enrollment rates are measured in the fall following exam school application.

test scores of these applicants by between 0.050 SD and 0.193 SD, while leaving their English test scores unchanged.[13] Nonetheless, only the 8$^{\text{th}}$ grade coefficient for group 1 and 7$^{\text{th}}$ grade coefficient for group 2 are statistically different from zero.

On the other hand, applicants whose non-exam-school alternatives belong to the two highest quartiles of VA achieve worse English and math test outcomes when they enroll at O'Bryant. For instance, a group 4 school increases 8$^{\text{th}}$ grade English MCAS scores by 0.125 SD and 8$^{\text{th}}$ grade math MCAS scores by 0.185 SD with respect to O'Bryant.

Overall, these estimates correspond to lower O'Bryant achievement effects than those implied by the estimated math and English VA presented in Figure 1. Likewise, schools in groups 1 and 2 perform similarly according to the 2SLS estimates even though their estimated average math and English VA are different. This suggests either the existence of some bias in the value-added estimation or some heterogeneity in VA for marginal exam school applicants.

The over-identification test provides reassuring evidence that the assumption of a constant treatment along relative distances is likely to be satisfied. The p-values for the test reported in the bottom of panel B do not reject the null of a homogeneous effect, except marginally for 7$^{\text{th}}$ grade math MCAS. Moreover, the implied LATE, computed by weighting each coefficient in panel B by the estimated share of compliers for the corresponding group of non-exam-school alternatives, is similar to the actual LATE reported in panel A.[14] A stark difference between the two figures would have suggested that the decomposition was picking up some variation not linked to the heterogeneity in non-exam-school alternative.

As a further step to the decomposition, I try to assess the heterogeneity in average

[13] The estimates in this panel correspond to the effect of enrolling in a school from each group instead of an exam school. Hence, a negative coefficient indicates that applicants would benefit from attending O'Bryant instead of a school of the corresponding group.

[14] The share of compliers with a non-exam-school alternative in each group is estimated by regressing a dummy for enrollment in a given group on a dummy for not enrolling at an exam school instrumented by not receiving an exam offer. The second bar of Figure 2 summarizes these results.

effect across compliers with different pre-treatment characteristics. In particular, I explore potential match effects for applicants with a high baseline score, and for Black and Hispanic applicants. Interacting the distance and offer instruments with dummies for each group recovers match effects within each group of non-exam-school alternatives. Nonetheless, the increase in the number of endogenous variables comes at the cost of decreased precision, making the results hard to interpret.

Appendix tables A4 and A5 report the results for the decomposition by baseline math score and minority status. Although the estimates are noisy, match effects appear to be smaller than substitution effects. Moreover, match effects are not consistent across non-exam-school alternative groups and test scores. Only minority applicants with a group 1 non-exam-school alternative appear not to gain from enrolling in an exam school while their non-minority peers do benefit. This suggests that, once accounting for differences in outside options, there is no systematic heterogeneity in treatment effect across students.

1.4.3 Difference in treatment effect of each exam school

As underlined in the previous section, my estimation strategy can only pin down the effect of each exam school with respect to the outside option of applicants with non-degenerate offer risk. Given the enrollment destinies presented in figure 2, it follows that students marginally offered Boston Latin may be used to estimate the effect of attending Boston Latin rather than Latin Academy; while students marginally seated at Latin Academy identify the effect of attending Latin Academy rather than Boston Latin or O'Bryant.

Table 5 presents the results of these comparisons by adding to the model enrollment dummies for each individual exam school. These dummies are instrumented by exam school-specific offers, which are interacted with distances to groups of traditional public schools in order to account for potential differences in substitution patterns [15].

[15] These differences are unlikely to be critical as only a small share of non-offered Boston Latin

Table 5. 2SLS Estimates of Individual Exam School Effects

	MCAS English		MCAS Math	
	7th grade	8th grade	7th grade	8th grade
	(1)	(3)	(5)	(7)
Enrollment in BLS	-0.047	-0.013	-0.124*	-0.330***
	(0.061)	(0.061)	(0.071)	(0.071)
Enrollment in BLA	0.155***	-0.032	-0.111*	-0.355***
	(0.049)	(0.049)	(0.058)	(0.059)
Enrollment in group 1 school	0.008	0.019	-0.067	-0.158*
	(0.061)	(0.061)	(0.083)	(0.084)
Enrollment in group 2 school	-0.050	0.028	-0.107	-0.050
	(0.064)	(0.064)	(0.075)	(0.088)
Enrollment in group 3 school	0.130**	0.066	0.102	0.039
	(0.058)	(0.053)	(0.073)	(0.080)
Enrollment in group 4 school	0.209***	0.150**	0.299***	0.189**
	(0.064)	(0.066)	(0.086)	(0.088)
Overid p-value	0.310	0.127	0.189	0.002
N	8,129	8,144	7,493	8,144

Notes: This table reports 2SLS estimates of school enrollment effects for 2004-2016 exam applicants with non-degenerate exam risk at any of the exam schools. The endogenous variable is enrollment in 7th grade at each specific exam school or at one of the group of schools. Enrollment at O'Bryant is omitted from the model so each coefficient corresponds to the treatment effect with respect to O'Bryant. All models control for exam schools offers risk, school-by-year linear running variables and the set of variables listed in Table1, except MCAS 6th grade scores. Standard errors clustered at the school of origin and application year are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. .

The dummy for enrolling at O'Bryant is excluded from the specification, so that each coefficient corresponds to the effect of attending that school rather than O'Bryant. As expected, including individual exam school enrollment dummies does not substantially affect the coefficients on the non-exam-school alternatives, since these were already implicitly with respect to O'Bryant in the any exam school specification.

On the other hand, Latin Academy appears to perform worse than O'Bryant overall. Attending Latin Academy as opposed to O'Bryant reduces MCAS math test scores by 0.11 SD in 7^{th} grade and by 0.36 SD in 8^{th} grade. The difference in 8^{th} grade MCAS English test score is not statistically different from zero, while Latin Academy appears to increase 7^{th} grade MCAS English test score by 0.16 SD. Moreover, the

compliers enroll at a traditional public school, and they tend to enroll equally in schools of the four VA groups.

table shows no statistically significant gains of enrolling at Boston Latin instead of Latin Academy. Contrary to what applicants seem to believe, it is more beneficial to enroll at O'Bryant than at Latin Academy, and gaining access to Boston Latin from Latin Academy does not appear to make a substantial difference.[16]

1.5 Counterfactual admission criteria

The previous analysis uncovered the effects of Boston's three exam schools on educational attainment for applicants close to each school's admission cutoff. In this section, I use these estimates to evaluate what change in exam school admission criteria would lead to the highest increase in overall achievement. Estimates for marginally seated applicants are particularly relevant for this analysis, as alternative admission criteria primarily change which applicants get admitted at the margin. While there seems to be little scope for match effects, the decomposition by non-exam-school alternative uncovered significant heterogeneity in achievement gains from enrolling at an exam school depending on the quality of an applicant's outside option. Changing exam school admission criteria to leverage this heterogeneity in treatment effect by non-exam-school alternative could thus increase overall achievement.

The optimal admission rule would give priority to applicants whose non-exam-school alternative is of low quality. Since non-exam-school alternatives are not observable, this policy is not implementable in practice. Nonetheless, I can estimate the optimal assignment and use it to benchmark the effects of feasible changes in admission criteria. As for feasible policies, I simulate four alternative admission schemes: granting admission to top-ranked middle school applicants, replacing ISEE scores with MCAS scores, zip code reserves similar to the BPS plan for the 2020-2021 application cycle, and a place-based priority system.

[16] Appendix table A6 explores potential heterogeneity in gains for minority students and applicants with baseline math test scores above the median. Overall, there appears to be no substantial additional gains from getting into a more selective exam school for students in these two groups.

First, I consider the effect of implementing a "Texas top10" style rule. Specifically, I simulate the exam school selection process when granting admission to applicants whose GPA places them in the top 5 or 10% of their middle school.[17] Second, I explore the impact of replacing ISEE scores with MCAS scores in the composite score used to rank applicants. This policy was suggested by Goodman and Rucinski (2018) as a feasible alternative to race-based priorities. Third, I simulate an admission system that reserves seats for each city zip code depending on the size of their school-aged population.[18] This simulation is inspired by the plan adopted by BPS for the 2020-2021 application cycle. Although I try to predict the impact of such a plan on student achievement relative to other admissions policies, I do not intend for this study to serve as an evaluation of Boston's admissions plan. I attempt to reproduce the changes outlined in public documents by the Exam School Admissions working group. However, since I was not commissioned by BPS to study its policy, I do not know the details of the plan's implementation.[19] Finally, I implement a place-based priority system in which applicants, whose closest school belongs to the two lowest quartiles of estimated VA, have priority for 25%, 50% or 75% of reserved seats.[20]

The non-exam-school alternative for each applicant is not directly observable. To predict applicants' non-exam-school alternative, I use a linear model with distances to each group of schools as regressors.[21] Panel A of Table 6 explores the accuracy

[17] To simulate this change in admission criteria, I assume top middle school students' admission behavior is unaffected. That is, all the top students that currently apply would also apply under the new policy, and top students that currently do not apply would still not apply under the new policy.

[18] For each school, available seats are divided between open seats and reserved seats. Open seats represent 20% of seats and are filled first. All applicants have the same priority for open seats. Reserved seats are allocated to each Boston zip code according to the size of its school-aged population. Applicants get priority to the reserved seats allocated to their zip code of residence. For each category of seats, applicants are sorted by priority and composite scores.

[19] My implementation of the admissions plan differs from publicly stated policy in that open seats do not only consider the first choice of each applicant and that I use composite scores to rank applicants and not only GPA.

[20] The address of the school of enrollment at the time of application is used as a proxy for student's address. Reserved seats are filled after open seats, which should favor applicants that qualify for the reserve (Dur et al. (2020)).

[21] Any negative predicted probability of enrolling in a group is set to zero. Predicted probabilities

Table 6. Consequences of Alternative Admission Schemes

Panel A: Accuracy of model predicting applicants' non-exam alternative school

	tribution of fallback schools for non-offered applica			
	Group1	Group2	Group3	Group4
Actual distribution	26.4%	25.2%	26.7%	21.7%
Predicted distribution	26.6%	25.8%	26.3%	21.3%

Panel B: Comparison of counterfactual admission schemes to the actual admission scheme

		Characteristics of offered applicants					Gains in MCAS Math 8th grade	
		Distribution of fallback schools					Net % of	
		Group1	Group2	Group3	Group4	% minority	applicants affected	Average gain
Current admission scheme		27%	26%	27%	21%	42%		
Optimal admission scheme : priority depending on non-exam alternative school		40%	33%	17%	10%	52%	22%	0.14
Automatic admission for top middle school students	top 5%	27%	26%	27%	21%	44%	0%	0.01
	top 10%	26%	26%	27%	21%	48%	0%	0.00
Change in composite score used to ranked applicants	MCAS instead of ISEE	27%	26%	27%	21%	44%	0%	0.03
Replication of plan for 2021 admissions		27%	25%	27%	21%	47%	0%	-0.15
Share of seats reserved to students whose closest school has less than median VA	25 % reserve	28%	27%	25%	20%	42%	3%	0.13
	50 % reserve	32%	30%	21%	18%	43%	9%	0.13
	75 % reserve	34%	32%	18%	15%	48%	15%	0.13

Notes: This table reports the distribution of non-exam alternative schools under different admission schemes for 2004-2016 exam applicants from a BPS school with baseline test scores. Panel A explores the accuracy of the model predicting applicants' non-exam alternative school by comparing the distribution of alternative schools predicted by the model to the actual distribution for non-offered applicants. Panel B compares the distribution of non-exam alternative schools for applicans offered an exam school under five different admission schemes to the distribution of alternative schools for applicants offered an exam school under the actual admission scheme. The table also displays the average gain in test scores and the net share of applicants gaining under each counterfactual admission scheme.

of the prediction model by comparing the actual distribution of non-exam-school alternatives to the distribution predicted by the model, for applicants that do not receive an exam school offer. All predicted shares are less than one percentage point apart from the actual shares.

I use the model to predict the distribution of non-exam-school alternatives for admitted applicants under the different admission schemes, and to estimate the assignment under an optimal admission regime. The optimal assignment is obtained by maximizing the total gain in MCAS math 8th grade test scores. This policy corresponds to maximizing the shares of admitted applicants from groups 1 and 2[22]. Panel B of Table 6 compares the performance of the three different admission schemes to

are then normalized by setting their sum to one for each applicant.

[22]The performance of the different groups of non-exam-school alternatives was gauged with respect to O'Bryant. Nonetheless, as different demographic groups do not seem to benefit more or less from gaining access to Boston Latin or Latin Academy, non-exam-school alternative estimates specify the relative gain or loss from diverting a student from a given school, even when the student enrolls at Boston Latin or Latin Academy. Hence, these estimates may be used to compute the gain from changing the distribution of non-exam-school alternatives for applicants admitted at any of the three exam schools.

the optimal assignment and the actual assignment.

From the table, it appears that only the place-based priority system is likely to improve applicants' overall academic performance. Indeed, while the place-based priority system increases the share of offered applicants whose non-exam-school alternative belongs to the first two quartiles of VA, granting seats to top middle school applicants, replacing ISEE scores with MCAS scores or implementing zip code reserves similar to BPS's temporary plan have almost no effect on the distribution of non-exam-school alternatives among admitted applicants. As a result, the place-based scheme comes the closest to the optimal assignment. Reserving 75% seats to students with a low VA closest alternative increases 8th grade math test score by 0.13 SD for 15% of applicants. For comparison, the optimal assignment results in a similar increase of 0.14 SD for 22% of applicants. On the other hand, granting admission to the top students from each middle school, replacing the ISEE with MCAS or implementing zip code reserves similar to the BPS plan for 2021 admissions affects the test scores of at most 1% of applicants, and the average gain varies between -0.15 SD and 0.03 SD.

The relative performance of each admission criteria can be explained by the correlation between the criteria used and the quality of applicants' non-exam-school alternatives. Considering Table 1, the average baseline MCAS score for students does not vary systematically with the estimated quality of schools. Thus, students with MCAS scores that are higher than their ISEE scores are not more likely to have a low VA non-exam-school alternative. Similarly, each VA group contains the same number of schools and is not concentrated geographically. Selecting applicants from each school of origin or from specific zip codes is thus not likely to affect the distribution of non-exam-school alternatives. On the other hand, an applicant's neighboring school is expected to be her outside option, so giving priority to applicants with low VA neighboring schools is quite effective.

One of the main arguments for reforming exam school admission is that Black and Hispanic students tend to be underrepresented at these elite institutions. Interestingly, all of the alternative admission criteria increase minority students' representation at exam schools, although none targets these students directly.

However, the magnitude of the increase is not correlated with the effectiveness of the admission scheme. Indeed, the admission scheme resulting in the largest minority share (48%) is the priority for the top 10% of students from each middle school, which has no effect on academic achievement. Similarly, implementing zip code reserves similar to BPS's temporary plan also increases minority share at exam schools to 47 %. Finally, the 75% reserve for students with a low-quality outside option increases the share of Blacks and Hispanics among admitted applicants to 48%, that is, by 6 percentage points compared to the actual assignment.

This discrepancy can be explained by the fact that Black and Hispanic students do not appear to have systematically worse outside options (as shown in Table 1). Thus, only some minority applicants would actually benefit from attending an exam school instead of their non-exam-school alternative. In general, Table 1 suggests that no observable characteristic is strongly correlated with the quality of a student's non-exam-school alternative. Hence, targeting based on observable characteristics is unlikely to improve academic achievement.

Granting admission to top-ranked middle school applicants and replacing ISEE scores with MCAS scores does not leverage the heterogeneity in the exam school treatment effect. Thus, implementing these changes would probably not result in substantial changes in overall achievement. On the other hand, directly targeting applicants based on their neighborhood school leverages most of the relevant heterogeneity and results in significant improvements.

1.6 Conclusion

Existing regression discontinuity estimates of exam school effects find zero gain from attending these schools. Nonetheless, I show that these estimates aggregate important heterogeneity in treatment effects and are thus not suitable for performing counterfactual analyses. In particular, applicants whose non-exam-school alternative has an estimated math VA in the bottom two quartiles of Boston public schools benefit from attending an exam school.

It follows that changing the admission scheme for Boston exam schools could increase overall achievement. This improvement comes from identifying and targeting applicants who benefit from attending exam schools. These students are not characterized by specific demographics but rather by low-quality non-exam-school alternatives. This distinction highlights the necessity of separately identifying heterogeneity in substitution effects from match effects when attempting to compare different allocation of students to schools.

More generally, in any setting with multiple treatments, an accurate counterfactual analysis should always involve a full decomposition of the treatment effect. Although identifying these different sources of heterogeneity is challenging, my analysis shows how to leverage applicants' locations. This strategy is particularly relevant since it allows for a compelling test of the identifying assumptions, and my approach may be applicable in other settings where treatment is related to spatial position.

Appendix

Appendix A: data

Boston Public Schools (BPS) is the source of the application files for exam schools. The State of Massachusetts (DESE) provided the enrollment files and MCAS data. Application, enrollment and test scores files are merged using the State unique identifier (SAS-ID). In this section, I describe these data sets, the construction of the distance variable and the additional sample restrictions.

Application data

The exam school application file contains a record for each student consisting of an application id number, state ID (SAS-ID) number, name, gender, race, date of birth, application year, grade of application, preferences over three exam schools, and the composite score for admission. Each record also includes the school where the student receives an offer (if any). This data set covers students with application years from 1995-2017 and applicants for entrance in 7^{th}, 9^{th} and 10^{th} grade. The analysis sample only includes 7^{th} grade applicants from 2004-2016 for which both baseline and outcome test scores are available. Students enroll in the fall of the same year. I exclude duplicate observations and applicants who were missing the application id number from the analysis.

The admission rank cutoff for each year and each school is computed using the rank of the last admitted students, after excluding admitted students with incorrect ranks. Incorrectly ranked admitted students are admitted at an exam school although their rank is much higher than the rank of the last but one admitted at the same school. The admission rank cutoff is used to construct a simulated offer variable for each school, which is preferred to the actual offer variable throughout the analysis. Despite these identified irregularities, the match replicates 99 % on average.

Enrollment data

The Massachusetts enrollment file spans school years 2001-2002 through 2017-2018. Each record contains both a start of the school-year (October) snapshot and at end of the school-year snapshot for each student enrolled in Boston Public Schools, with unique student identifier (the SAS-ID), the student's grade and school, and demographic information. The variables of interest in the enrollment file are grade, year, sex, race, low-income status, special education status (SPED), and native speaker status. The school each student was enrolled prior to application is obtained from the end of the year enrollment files while the October enrollment files are used to determine the school each student enrolls after application. Enrollment data is only available for students enrolled in Boston public schools, excluding students enrolled in private schools before or after application.

Test scores data

The MCAS test scores file spans school years from 2002 to 2018. It includes scores for two subjects: English, and Math, the grade (4^{th}, 6^{th}, 7^{th} and 8^{th} grade) and the year in which the test was taken. I standardize scores among Boston test-takers by year and grade. For each multiple times test takers, the last test score for each grade is considered. 4^{th} grade Math and English MCAS test scores constitute the baseline scores whereas 7^{th} and 8^{th} grade English and 7^{th} and 8^{th} grade Math MCAS test scores are the main outcomes of interest. 6^{th} grade Math and English scores are available only after 2004 and 2006 respectively while 7^{th} grade Math MCAS was introduced only in 2006.

Distance to schools

Only schools where at least one student enrolled after application during the 2004-2016 period are included in the set of potential schools of enrollment. The distance between each sending middle school of the dataset to each potential school of enrollment is computed using the shortest road distance between the two points. It does not take into account differences in traffic or in speed limit across roads.

Sample restrictions

The analysis sample is restricted to students enrolled in a non-charter BPS school prior to application. This excludes the possibility for students to be enrolled in charter after application since admissions to Massachusetts charter schools occur either in 5^{th} or 6^{th} grade through lotteries. Hence, the analysis only consider students that may either enroll in an exam school or in a non-charter BPS school. Students with no baseline test scores, i.e. who were not enrolled in BPS during 4^{th} grade, are also excluded from the analysis.

Description of Value-added sample and model

The Value added models for math and English are estimated on the sample of all non-charter BPS schools. Schools with fewer than 25 students enrolled in 7^{th} grade are excluded from the sample. The math value added model uses Math MCAS test scores from 8^{th} grade but considers the school where students were enrolled in 7^{th} grade. The English value added model uses 7^{th} grade MCAS English test scores. Both model includs indicators for sex, race, subsidized lunch eligibility, special education status, English-language learner status, school year, along with cubic functions of all the baseline Math and ELA test scores available (3rd, 4^{th}, 5^{th} and 6^{th} grades). For each year, the value added of each school is computed using the data from two years prior.

Appendix B: Computation of the propensity scores

Following the methodology in Abdulkadiroglu et al. (2019), I exploit the rank order list of applicants to compute the propensity score of admission at each exam school. Specifically, an applicant whose score is close to a school's admission cutoff has a local risk of admission at this school of one half. Considering the applicant's ranked order list, it is then possible to compute the overall probability of admission at each school by taking into account the applicant's probability of being admitted to a more preferred school.

More formally, let $s = 0, 1, 2, 3$ index exam schools in set S, where $s = 0$ denotes any outside option. $\theta_i = (\succ_i)$ denote applicant i's type, where $\succ_i$ is the applicant's ranking of schools. The school specific rank used for admission is denoted by R_{is}; this is the school specific RD running variable. R_i is the vector of rankings at each school for applicant i. Each applicant gets an offer at school s if and only if its ranking is below the school specific cutoff τ_s, i.e. iff $R_{is} \leq \tau_s$.

Considering applicants' rank order list, let B_{θ_s} be the set of schools type θ prefers to s

$$B_{\theta_s} = \{s' \in S | s' \succ_\theta s\}$$

Given a bandwidth δ, an applicant's type θ and ranks vector R_i, define the risk of being seated at school s as

$$\Psi_s(\theta, R, \delta) = \begin{cases} 0 & \text{if } R_s < \tau_s - \delta \text{ or } R_b > \tau_b + \delta \text{ for some } b \in B_{\theta_s} \\ 0.5^{m_s(\theta,R)} & \text{if } R_s > \tau_s + \delta \text{ and } R_b \leq \tau_b + \delta \text{ for all } b \in B_{\theta_s} \\ 0.5^{1+m_s(\theta,R)} & \text{if } \tau_s - \delta \leq R_s \leq \tau_s + \delta \text{ and } R_b \leq \tau_b + \delta \text{ for all } b \in B_{\theta_s} \end{cases}$$

where $m_s(\theta, T) = |\{b : b \in B_{\theta_s} \quad \text{and} \quad \tau_b - \delta \leq R_b \leq \tau_b + \delta\}|$

Under the Assumption that the distribution of each running variable is continuous at the admission cutoff for each applicants' type, Theorem (1) of Abdulkadiroglu et al. (2019) shows that

$$\lim_{\delta \to 0} \mathbf{E}[D_i(s)|\theta_i = \theta, R_i = R, W_i = W] = \Psi_s(\theta, R, \delta)$$

Where $D_i(s)$ is an indicator for receiving an offer from school s and W_i is a vector of observed and unobserved characteristics of student i. Controlling for the propensity score, offers from school s are locally as good as randomly assigned and can thus be used as instruments for enrollment.

Empirically, the propensity score can be computed as the sample equivalent of the theoretical local propensity score described above using the information contained in the rank order list of students and admission cutoffs. Theorem (2) of Abdulkadiroglu et al. (2019) establishes uniform convergence of the empirical propensity score in an asymptotic sequence that increases market size with a shrinking bandwidth. This justifies conditioning on the empirical propensity score to eliminate OVB in school effect estimates.

When computing the scores, I separately estimate bandwidths for each school and cohort according to Imbens and Kalyanaraman (2012). As an intermediate step, I also estimate the bandwidths for each outcome variable; I keep the smallest bandwidth for each school and cohort.

Appendix C: What does 2SLS identify?

The previous framework showed that question regarding counterfactual assignments cannot be answered without knowing both substitution and match effects. In this section, I show that these effects are hard to identify separately, and that LATE and group specific estimators identify a weighted average of both effects.

I am interested in the treatment effect of enrolling in an exam school, which can indexed as school 0. Let D_i be a dummy equal to 1 when student i enrolls in school 0. The school 0 treatment effect is given by β from the regression of Y_i on D_i,

$$Y_i = \alpha + \beta D_i + \nu_i. \tag{1.3}$$

The decision to enroll in a specific school is typically endogenous and correlated with students' characteristics. Hence, the large literature interested in schools' treatment effects has leveraged instrumental variables based on specific admission rules (lotteries, admission tests, etc.) and student characteristics (distance). Suppose there exists a valid binary instrument Z for enrollment in school 0, and let D_{i1} and D_{i0} denote the potential values of D_i when $Z_i = 1$ and $Z_i = 0$ respectively. Since Z is a valid instrument, it satisfies the following assumptions.

Assumption 2 1. ***Instrument relevance:*** $E[D_{i0}] \neq E[D_{i1}]$

2. ***Random assignment and exclusion:*** *Z_i is independent of $(D_{i1}, D_{i0}, Y_i(D_i, Z_i))$ and $Y_i(D_i, 0) = Y_i(D_i, 1) = Y_i(D_i)$*

3. ***Monotonicity:*** *$D_{i1} \geq D_{i0}$ $\forall i$ and $D_{i1} > D_{i0}$ for some i*

Proposition 1 establishes that the LATE obtained from the 2SLS regression is a weighted average of pairwise comparisons, with weights given by the distribution of outside options for the compliers. The ω_j weight captures the share of compliers that have outside option j, and the $\omega_{x|j}$ weight captures the share of compliers with outside option j that have characteristics x.

Proposition 2 (2SLS identification)
Suppose there exists an instrument Z_i for D_i that satisfies Assumption 1. The 2SLS

regression using Z_i as an instrument identifies

$$\beta_{LATE} = \underbrace{E[Y_{i0}] - \sum_j \omega_j E[Y_{ij}]}_{\textit{substitution effects}} + \underbrace{\sum_x \left(\omega_{x|0} E[\epsilon_{i0} \mid X_i = x] - \sum_j \omega_j \omega_{x|j} E[\epsilon_{ij} \mid X_i = x]\right)}_{\textit{match effects}}. \tag{1.4}$$

Letting $S_i \in \{1, ..., J\}$ denote the non-exam alternative option of student i, the weights are defined as

$$\omega_j = Pr[S_i = j \mid D_{i1} > D_{i0}], \text{ and}$$
$$\omega_{x|j} = Pr[X_i = x \mid D_{i1} > D_{i0}, S_i = j].$$

Proof. Z satisfies Assumption 1 thus β_{LATE} identifies the average treatment effect for compliers:

$$\begin{aligned}\beta_{LATE} &= E[Y_{i1} - Y_{i0} | D_{i1} > D_{i0}] \\ &= E[Y_{i1} | D_{i1} > D_{i0}] - E[Y_{i0} | D_{i1} > D_{i0}]\end{aligned}$$

where $Y_{i1} = Y(D_i = 1)$ and $Y_{i0} = Y(D_i = 0)$.

Using the framework notation:

$$
\begin{aligned}
E[Y_{i1}|D_{i1} > D_{i0}] &= E[\bar{Y}_1 + \bar{Y}_i + E[\epsilon_{i1}|X_i]|D_{i1} > D_{i0}] \\
&= \bar{Y}_1 + E[\bar{Y}_i|D_{i1} > D_{i0}] + \sum_x \omega_{x|1} E[\epsilon_{i1}|X_i] \\
E[Y_{i0}|D_{i1} > D_{i0}] &= E[\bar{Y}_j + \bar{Y}_i + E[\epsilon_{ij}|X_i]|D_{i1} > D_{i0}, D_i = 0] \\
&= \sum_j \omega_j E[\bar{Y}_j + \bar{Y}_i + E[\epsilon_{ij}|X_i]|D_{i1} > D_{i0}] \\
&= \sum_j \omega_j \{\bar{Y}_j + E[\bar{Y}_i|D_{i1} > D_{i0}] + \sum_x \omega_{x|j} E[\epsilon_{ij}|X_i]\} \\
&= \sum_j \omega_j \bar{Y}_j + E[\bar{Y}_i|D_{i1} > D_{i0}] + \sum_x \sum_j \omega_j \omega_{x|j} E[\epsilon_{ij}|X_i]\}
\end{aligned}
$$

Replacing and rearranging these expressions in the formula for $\beta_L ATE$, one obtains the desired decomposition:

$$
\beta_{LATE} = \underbrace{\bar{Y}_1 - \sum_j \omega_j \bar{Y}_j}_{\text{substitution effects}} + \underbrace{\sum_x (\omega_{x|1} E[\epsilon_{i1}|X_i] - \sum_j \omega_j \omega_{x|j} E[\epsilon_{ij}|X_i])}_{\text{match effects}}
$$

□

Corollary 1 points out that 2SLS estimators for different subgroups do not capture pure match effects if compliers with different covariate values substitute differently from non-exam-school alternatives. It is thus necessary to compute match effects controlling for applicants' non-exam-school alternatives. In other words, one need fist to decompose treatment effect by non-exam-school alternatives before attempting to capture match effects.

Corollary 1 (Subgroup 2SLS identification)

Suppose there exists an instrument Z_i for D_i that satisfies Assumption 1. Using Z_i

as an instrument on the subgroup of observations with $X_i = x$ identifies

$$\beta_{LATE|X_i=x} = \underbrace{E[Y_{i0}] - \sum_j \omega_{j|x} E[Y_{ij}]}_{substitution\ effects} + \underbrace{E[\epsilon_{i0} \mid X_i = x] - \sum_j \omega_{j|x} E[\epsilon_{ij} \mid X_i = x]}_{match\ effects}), \tag{1.5}$$

where

$$\omega_{j|x} = Pr[S_i = j \mid D_{i1} > D_{i0}, X_i = x].$$

Appendix D: Proofs

Corollary 1 (Match effect identification)
Suppose there exists an instrument Z_i for D_i which satisfies Assumption 1. Using Z_i as an instrument on the subgroup of observations with $X_i = x$ identifies

$$\beta_{LATE|X_i=x} = \underbrace{\bar{Y}_1 - \sum_j \omega_{j|x} \bar{Y}_j}_{substitution\ effects} + \underbrace{E[\epsilon_{i1}|X_i = x] - \sum_j \omega_{j|x} E[\epsilon_{ij}|X_i = x]}_{match\ effects}) \tag{1.6}$$

where

$$\omega_{j|x} = Pr[S_i = j|D_{i1} > D_{i0}, X_i = x]$$

Proof. This follows directly from Proposition 1 by replacing ω_j by $\omega_j|x$ and dropping the $\sum_x$ since there is only one value of X_i in the subsample.

□

Proposition 1 (Identification of treatement effect by outside option)
If there exists a valid and exogeneous instrument Z_i for enrollment in school 0 and a vector of $J-1$ covariates $\{W_{ik}\}_{k=2}^J$ which satisfies Assumptions 1 and 2 then

- *Conditioning on $\{W_{ik}\}_{k=2}^J$, the interaction of Z_i and $\{W_{ik}\}_{k=2}^J$ identifies*

$E[\bar{Y}_1 - \bar{Y}_j | D_{i1} > D_{i0}] \ \forall j = 2, ..., J$

- *For any covariate X_i, conditioning on $\{W_{ik}\}_{k=2}^J$ and X_i, the interaction of Z_i, $\{W_{ik}\}_2^J$ and X_i identifies $E[\bar{Y}_1 - \bar{Y}_j | X_i, D_{i1} > D_{i0}] \forall j = 2, ..., J$*

Proof. To identify treatment effect of attending school 0 (exam school) by non-exam-school alternative , one would like to estimate

$$Y_i = \alpha + \beta S_{i1} + \sum_{k=2}^{J} \beta_j S_{ij} + u_i$$

The exam school is excluded from the estimation equation. It thus constitutes the school of reference to which school 1 and the other schools are compared, i.e. β corresponds to the treatment gain from attending school 1 instead of the exam school.

For simplicity, let's first consider a case with only two non-exam-school alternatives. Using Frish-Waugh-Lovell theorem, one can reduce every problem with a problem with two non-exam-school alternatives by partialling out the other covariates.

Define $Z_{i2} = W_{i2} \times Z_i \ \forall i$ and assume w.l.o.g. that W_2 is a continuous variable. Denote by $\tilde{Z}_2$, Z_2 partialled out from W_2. Note that since Z is randomly assigned by Assumption 1 $\tilde{Z}_i = Z_i \ \forall i$.

The first stage equations give

$$E[S_{i1} | Z, \tilde{Z}_{i2}] = \alpha_1 + \alpha_1^1 Z_i + \alpha_0^2 \tilde{Z}_{i2}$$
$$E[S_{i2} | Z, \tilde{Z}_{i2}] = \alpha_2 + \alpha_2^1 Z_i + \alpha_2^2 \tilde{Z}_{i2}$$

Plugging these into the second stage, we obtain the following reduced form

$$E[Y_i|Z_i,\tilde{Z}_{i2}] = \alpha + \beta(\alpha_1 + \alpha_1^1 Z_i + \alpha_1^2 \tilde{Z}_{i2}) + \beta_2(\alpha_2 + \alpha_2^1 Z_i + \alpha_2^2 \tilde{Z}_{i2})$$
$$E[Y_i|Z_i,\tilde{Z}_{i2}] = \alpha + \beta\alpha_1 + \beta_2\alpha_2 + \underbrace{(\beta\alpha_1^1 + \beta_2\alpha_2^1)}_{\alpha_y^1} Z_i + \underbrace{(\beta\alpha_1^2 + \beta_2\alpha_2^2)}_{\alpha_y^2} \tilde{Z}_{i2}$$

$E[u_i|Z_i\tilde{Z}_{i2}] = 0$ by Assumption 1.

By definition, 2sls estimator equals reduced form estimates times the inverse of the first stage estimates:

$$\begin{pmatrix} \beta \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \alpha_1^1 & \alpha_2^1 \\ \alpha_1^2 & \alpha_2^2 \end{pmatrix}^{-1} \begin{pmatrix} \alpha_y^1 \\ \alpha_y^2 \end{pmatrix}$$

Solving the system for β and β_2:

$$\beta = \frac{\alpha_2^2\alpha_y^1 - \alpha_2^1\alpha_y^2}{\alpha_1^1\alpha_2^2 - \alpha_2^1\alpha_1^2}$$
$$\beta_2 = \frac{\alpha_1^1\alpha_y^2 - \alpha_1^2\alpha_y^1}{\alpha_1^1\alpha_2^2 - \alpha_2^1\alpha_1^2}$$

Consider a local evaluation point w_2, using the fact that $D_i + S_{i1} + S_{i2} = 1$ and the monotonicity conditions from Assumptions 1 and 2:

$$
\begin{aligned}
\alpha_y^0 &= E[Y_i|Z_i = 1, \tilde{Z}_{i2}] - E[Y_i|Z_i = 0, \tilde{Z}_{i2}] \\
\alpha_y^0 &= E[Y_{i0} + (Y_{i1} - Y_{i0})S_{i1} + (Y_{i2} - Y_{i0})S_{i2}|Z = 1, \tilde{Z}_{i2}] \\
&\quad - E[Y_{i0} + (Y_{i1} - Y_{i0})S_{i1} + (Y_{i2} - Y_{i0})S_{i2}|Z = 0, \tilde{Z}_{i2}] \\
\alpha_y^0 &= E[Y_{i1} - Y_{i0}|0 \leftarrow 1](E[S_{i1}|Z_i = 1, \tilde{Z}_{i2}] - E[S_{i1}|Z_i = 0, \tilde{Z}_{i2}]) \\
&\quad + E[Y_{i2} - Y_{i0}|0 \leftarrow 2](E[S_{i2}|Z_i = 1, \tilde{Z}_{i2}] - E[S_{i2}|Z_i = 0, \tilde{Z}_{i2}]) \\
\alpha_y^0 &= E[Y_{i1} - Y_{i0}|0 \leftarrow 1]\alpha_1^1 + E[Y_{i2} - Y_{i0}|0 \leftarrow 2]\alpha_2^1
\end{aligned}
$$

where $0 \leftarrow 1$ denotes instrument Z compliers moving from treatment 1 to treatment 0 at point (w_2), i.e. from school 1 to the exam school at point (w_2). Similarly, where $0 \leftarrow 2$ denotes instrument Z compliers moving from school 2 to the exam school at point (w_2).

One can derive a similar expression for α_y^2:

$$
\begin{aligned}
\alpha_y^2 &= \frac{\partial E[Y_i|Z_i, \tilde{Z}_{i2}]}{\partial \tilde{Z}_{i2}} \\
\alpha_y^2 &= E[Y_{i1} - Y_{i0}|0 \leftarrow 1(w_2')]\alpha_1^2 + E[Y_{i2} - Y_{i0}|0 \leftarrow 2(w_2')]\alpha_2^2
\end{aligned}
$$

where $0 \leftarrow 2(w_2')$ denotes instrument Z compliers moving from school 2 to the exam school at point $w_2' \downarrow w_2$ but not at point w_2, i.e.

$$
E[Y_{i2} - Y_{i1}|0 \leftarrow 2(w_2')] = \lim_{w_2' \downarrow w_2} E[Y_{i0} - Y_{i1}|S_{2i}(Z_1 = 0, W_{2i} = w_2) = 0, S_{2i}(Z_1 = 0, W_{2i} = w_2') = 1]
$$

These expressions indicate that the multi-treatment estimate aggregates treatment effect from different groups of compliers. Nonetheless, the constant treatment effect within covariate condition from Assumption 2 implies that for any (w_2, w_2')

$$E[Y_{i1} - Y_{i0}|0 \leftarrow 1(w_2)] = E[Y_{i1} - Y_{i0}|0 \leftarrow 1(w_2')]$$
$$E[Y_{i2} - Y_{i0}|0 \leftarrow 2(w_2)] = E[Y_{i2} - Y_{i0}|0 \leftarrow 2(w_2')]$$

Thus, plugging in α_y^0 and α_y^2 on the expressions for β :

$$\begin{aligned}
\beta &= \frac{\alpha_2^2\alpha_1^1 E[Y_{i1} - Y_{i0}|0 \leftarrow 1] + \alpha_2^2\alpha_2^1 E[Y_{i2} - Y_{i0}|0 \leftarrow 2]}{\alpha_1^1\alpha_2^2 - \alpha_2^1\alpha_1^2} \\
&\quad - \frac{\alpha_2^1\alpha_1^2 E[Y_{i1} - Y_{i0}|0 \leftarrow 1] + \alpha_2^1\alpha_2^2 E[Y_{i2} - Y_{i0}|0 \leftarrow 2]}{\alpha_1^1\alpha_2^2 - \alpha_2^1\alpha_1^2} \\
\beta &= \frac{(\alpha_2^1\alpha_1^2 - \alpha_2^2\alpha_1^1)E[Y_{i1} - Y_{i0}|0 \leftarrow 1]}{\alpha_1^1\alpha_2^2 - \alpha_2^1\alpha_1^2} \\
\beta &= E[Y_{i1} - Y_{i0}|0 \leftarrow 1]
\end{aligned}$$

and similarly for β_2

$$\begin{aligned}
\beta_2 &= -\frac{\alpha_1^2\alpha_1^1 E[Y_{i1} - Y_{i0}|0 \leftarrow 1] + \alpha_1^2\alpha_2^1 E[Y_{i2} - Y_{i0}|0 \leftarrow 2]}{\alpha_1^1\alpha_2^2 - \alpha_2^1\alpha_1^2} \\
&\quad + \frac{\alpha_1^1\alpha_1^2 E[Y_{i1} - Y_{i0}|0 \leftarrow 1] + \alpha_1^1\alpha_2^2 E[Y_{i2} - Y_{i0}|0 \leftarrow 2]}{\alpha_1^1\alpha_2^2 - \alpha_2^1\alpha_1^2} \\
\beta_2 &= \frac{(\alpha_2^1\alpha_1^2 - \alpha_2^2\alpha_1^1)E[Y_{i2} - Y_{i0}|0 \leftarrow 1]}{\alpha_1^1\alpha_2^2 - \alpha_2^1\alpha_1^2} \\
\beta_2 &= E[Y_{i2} - Y_{i0}|0 \leftarrow 2]
\end{aligned}$$

$\square$

Appendix E: Appendix Tables

Table A1. Differential Attrition by Distance to Each Group of Schools

	MCAS English		MCAS Math	
	7th grade	8th grade	7th grade	8th grade
	(1)	(2)	(3)	(4)
Exam offer	0.004	0.033**	0.003	0.033**
	(0.012)	(0.016)	(0.012)	(0.016)
Exam offer X distance to group 1	0.001	-0.001	0.001	-0.001
	(0.001)	(0.002)	(0.001)	(0.002)
Exam offer X distance to group 2	0.000	-0.000	0.000	-0.000
	(0.001)	(0.002)	(0.001)	(0.002)
Exam offer X distance to group 3	-0.002	-0.003	-0.002	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)
Exam offer X distance to group 4	0.002	-0.003*	0.001	-0.003*
	(0.002)	(0.002)	(0.002)	(0.002)
N	6,547	6,547	6,547	6,547

Notes: This table reports the effects of offer receipt by distance to each group's closest school on follow-up data availability for 2004-16 applicants. Robust standard errors presented in parentheses. The model controls for exam schools offers risk, school-by-year linear running variables and the set of variables listed in Table1, except MCAS 6th grade scores. Standard errors clustered at the school of origin and application year are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. .

Table A2. 2SLS Estimates of Exam Schools Effect by Non-Exam Alternative School for Different Group Splitting

	MCAS 7th grade			MCAS 8th grade		
	3 groups	4 groups	5 groups	3 groups	4 groups	5 groups
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Second stage for English MCAS					
Enrollment in group 1 school	0.040	0.022	-0.002	0.011	0.013	0.011
	(0.056)	(0.062)	(0.072)	(0.059)	(0.063)	(0.073)
Enrollment in group 2 school	0.062	-0.046	0.018	0.040	0.005	0.003
	(0.058)	(0.068)	(0.073)	(0.053)	(0.066)	(0.070)
Enrollment in group 3 school	0.256***	0.197***	0.069	0.137**	0.093*	0.078
	(0.059)	(0.062)	(0.066)	(0.061)	(0.056)	(0.060)
Enrollment in group 4 school		0.264***	0.284***		0.125*	0.124*
		(0.063)	(0.065)		(0.067)	(0.065)
Enrollment in group 5 school			0.180**			0.068
			(0.075)			(0.075)
overid p-value	0.248	0.426	0.414	0.863	0.968	0.894
N	5,953	5,953	5,953	5,978	5,978	5,978
	Panel B: Second stage for Math MCAS					
Enrollment in group 1 school	-0.076	-0.050	-0.085	-0.206**	-0.193**	-0.208**
	(0.078)	(0.086)	(0.097)	(0.080)	(0.086)	(0.099)
Enrollment in group 2 school	0.005	-0.144*	-0.159*	-0.039	-0.124	-0.216**
	(0.069)	(0.080)	(0.093)	(0.078)	(0.092)	(0.091)
Enrollment in group 3 school	0.276***	0.120	0.053	0.189**	0.047	0.048
	(0.082)	(0.077)	(0.080)	(0.086)	(0.086)	(0.090)
Enrollment in group 4 school		0.319***	0.183**		0.185**	0.102
		(0.090)	(0.088)		(0.092)	(0.092)
Enrollment in group 5 school			0.306***			0.127
			(0.107)			(0.109)
overid p-value	0.071	0.065	0.154	0.347	0.221	0.389
N	5,429	5,429	5,429	5,978	5,978	5,978

Notes: This table reports 2SLS estimates by non-exam alternative school group for different splitting of schools . The 2SLS estimates for MCAS english are reported in Panel A; the 2SLS estimates for MCAS math are displayed in Panel B. All models control for exam schools offers risk, school-by-year linear running variables and the set of variables listed in Table1, except MCAS 6th grade scores. Standard errors clustered at the school of origin and application year are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A3. OLS Estimates of Exam Schools Effect by Non-Exam Alternative School

	MCAS English		MCAS Math	
	7th grade	8th grade	7th grade	8th grade
	(1)	(2)	(3)	(4)
	Panel A: OLS for any exam school enrollment			
Enrollment in group 1 school	-0.258***	-0.178***	-0.415***	-0.306***
	(0.022)	(0.022)	(0.027)	(0.028)
Enrollment in group 2 school	-0.231***	-0.145***	-0.373***	-0.190***
	(0.023)	(0.024)	(0.028)	(0.031)
Enrollment in group 3 school	-0.180***	-0.112***	-0.223***	-0.140***
	(0.022)	(0.021)	(0.026)	(0.028)
Enrollment in group 4 school	-0.005	0.017	-0.080***	0.037
	(0.023)	(0.023)	(0.031)	(0.029)
N	6,027	6,056	5,499	6,056
	Panel B: OLS for school specific enrollment			
Enrollment at BLS	0.226***	0.177***	0.306***	0.107***
	(0.020)	(0.021)	(0.024)	(0.025)
Enrollment at BLA	0.236***	0.063***	0.044**	-0.206***
	(0.019)	(0.019)	(0.023)	(0.023)
Enrollment in group 1 school	-0.150***	-0.154***	-0.414***	-0.443***
	(0.024)	(0.024)	(0.030)	(0.031)
Enrollment in group 2 school	-0.123***	-0.118***	-0.373***	-0.324***
	(0.024)	(0.026)	(0.031)	(0.033)
Enrollment in group 3 school	-0.070***	-0.083***	-0.225***	-0.276***
	(0.024)	(0.023)	(0.029)	(0.030)
Enrollment in group 4 school	0.106***	0.040	-0.079**	-0.098***
	(0.025)	(0.025)	(0.033)	(0.031)
N	8,129	8,144	7,493	8,144

Notes: This table reports OLS estimates of the effect of different types of schools' enrollment for exam school applicants applying 2004-2016. The OLS estimates for enrollment in any of the exam school are reported in Panel A; the school specific OLS estimates are displayed in Panel B. All models control for the set of variables listed in Table1, except MCAS 6th grade scores. Standard errors clustered at the school of origin and application year are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. .

Table A4. Exam Schools Effect for Applicants Below and Above the Median Baseline Math Score

	MCAS English		MCAS Math	
	7th grade	8th grade	7th grade	8th grade
	(1)	(2)	(3)	(4)
Panel A: Second stage for any exam school enrollment				
Enrollment exam school	-0.090*	-0.027	-0.013	0.058
	(0.052)	(0.051)	(0.068)	(0.072)
Enrollment exam school X above median	-0.045	-0.081**	-0.106**	-0.075*
	(0.035)	(0.035)	(0.047)	(0.046)
First Stage F-statistic	1433.0	1431.5	1327.7	1431.5
Panel B: Second stage for any exam school enrollment by non-exam alternative school				
Enrollment in group 1	-0.012	-0.029	-0.116	-0.221**
	(0.069)	(0.068)	(0.091)	(0.090)
Enrollment in group 1 X above median	0.078	0.073	0.148	0.028
	(0.088)	(0.087)	(0.104)	(0.102)
Enrollment in group 2	-0.069	-0.010	-0.173*	-0.146
	(0.080)	(0.071)	(0.092)	(0.099)
Enrollment in group 2 X above median	0.109	0.120	0.128	0.052
	(0.094)	(0.084)	(0.097)	(0.107)
Enrollment in group 3	0.191***	0.061	0.090	0.030
	(0.069)	(0.061)	(0.089)	(0.098)
Enrollment in group 3 X above median	0.004	0.079	0.057	0.014
	(0.064)	(0.069)	(0.085)	(0.093)
Enrollment in group 4	0.246***	0.077	0.281***	0.105
	(0.076)	(0.071)	(0.098)	(0.099)
Enrollment in group 4 X above median	0.012	0.094	0.160	0.324**
	(0.091)	(0.104)	(0.145)	(0.137)
First Stage F-statistic	181.3	176.2	158.8	176.2
overid p-value	0.657	0.384	0.095	0.150
N	5,953	5,978	5,429	5,978

Notes: This table reports 2SLS estimates of the effect of different types of schools' enrollment for above and below median applicants applying 2004-2016. The sample is restricted to applicans with non degenerate any-exam offer risk. Above median applicants correspond to applicants with 4th grade Math MCAS test score above the median score of students with non degenerate any-exam school risk. Instruments and distance controls are interacted with above the median status. The set of instruments and controls is otherwise as described in Tables 2 and 5 . Standard errors clustered at the school of origin and application year are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. .

Table A5. Exam Schools Effect for Minority and Non-Minority Applicants

	MCAS English		MCAS Math	
	7th grade	8th grade	7th grade	8th grade
	(1)	(2)	(3)	(4)
Panel A: Second stage for any exam school enrollment				
Enrollment in exam school	-0.069	-0.017	-0.015	0.040
	(0.055)	(0.057)	(0.071)	(0.080)
Enrollment in exam school X minority	-0.070	-0.075*	-0.066	-0.020
	(0.044)	(0.039)	(0.050)	(0.061)
First Stage F-statistic	1,200	1,187	1,053	1,186
Panel B: Second stage for any exam school enrollment by non-exam alternative school				
Enrollment in Group 1	-0.078	-0.023	-0.202**	-0.288***
	(0.083)	(0.082)	(0.084)	(0.106)
Enrollment in Group 1 X minority	0.186**	0.037	0.257***	0.164
	(0.089)	(0.081)	(0.096)	(0.119)
Enrollment in Group 2	-0.031	-0.029	-0.046	-0.041
	(0.088)	(0.091)	(0.110)	(0.122)
Enrollment in Group 2 X minority	-0.035	0.051	-0.174	-0.151
	(0.081)	(0.090)	(0.117)	(0.118)
Enrollment in Group 3	0.191**	0.010	0.029	0.008
	(0.083)	(0.071)	(0.093)	(0.104)
Enrollment in Group 3 X minority	0.005	0.153**	0.133	0.060
	(0.080)	(0.068)	(0.082)	(0.102)
Enrollment in Group 4	0.224***	0.089	0.378***	0.234*
	(0.083)	(0.080)	(0.107)	(0.122)
Enrollment in Group 4 X minority	0.090	0.056	-0.072	-0.079
	(0.097)	(0.098)	(0.115)	(0.129)
First Stage F-statistic	130.5	125.3	128.7	125.3
overid p-value	0.190	0.016	0.065	0.474
N	5,953	5,978	5,429	5,978

Notes: This table reports 2SLS estimates of the effect of different types of schools' enrollment for minority and non-minority applicants applying 2004-2016. The sample is restricted to applicans with non degenerate any-exam offer risk. Minority applicants refer to applicants who are either Black or Hispanic. Instruments and distance controls are interacted with minority status. The set of instruments and controls is otherwise as described in Tables 2 and 5 . Standard errors clustered at the school of origin and application year are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. .

Table A5. Exam Schools Effect for Minority and Non-Minority Applicants

	MCAS English		MCAS Math	
	7th grade	8th grade	7th grade	8th grade
	(1)	(2)	(3)	(4)
Panel A: Second stage for any exam school enrollment				
Enrollment in exam school	-0.069	-0.017	-0.015	0.040
	(0.055)	(0.057)	(0.071)	(0.080)
Enrollment in exam school X minority	-0.070	-0.075*	-0.066	-0.020
	(0.044)	(0.039)	(0.050)	(0.061)
First Stage F-statistic	1,200	1,187	1,053	1,186
Panel B: Second stage for any exam school enrollment by non-exam alternative school				
Enrollment in Group 1	-0.078	-0.023	-0.202**	-0.288***
	(0.083)	(0.082)	(0.084)	(0.106)
Enrollment in Group 1 X minority	0.186**	0.037	0.257***	0.164
	(0.089)	(0.081)	(0.096)	(0.119)
Enrollment in Group 2	-0.031	-0.029	-0.046	-0.041
	(0.088)	(0.091)	(0.110)	(0.122)
Enrollment in Group 2 X minority	-0.035	0.051	-0.174	-0.151
	(0.081)	(0.090)	(0.117)	(0.118)
Enrollment in Group 3	0.191**	0.010	0.029	0.008
	(0.083)	(0.071)	(0.093)	(0.104)
Enrollment in Group 3 X minority	0.005	0.153**	0.133	0.060
	(0.080)	(0.068)	(0.082)	(0.102)
Enrollment in Group 4	0.224***	0.089	0.378***	0.234*
	(0.083)	(0.080)	(0.107)	(0.122)
Enrollment in Group 4 X minority	0.090	0.056	-0.072	-0.079
	(0.097)	(0.098)	(0.115)	(0.129)
First Stage F-statistic	130.5	125.3	128.7	125.3
overid p-value	0.190	0.016	0.065	0.474
N	5,953	5,978	5,429	5,978

Notes: This table reports 2SLS estimates of the effect of different types of schools' enrollment for minority and non-minority applicants applying 2004-2016. The sample is restricted to applicans with non degenerate any-exam offer risk. Minority applicants refer to applicants who are either Black or Hispanic. Instruments and distance controls are interacted with minority status. The set of instruments and controls is otherwise as described in Tables 2 and 5 . Standard errors clustered at the school of origin and application year are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. .

Chapter 2

Selective admissions and school segregation [1]

2.1 Introduction

In its landmark 1954 *Brown v. Board of Education* ruling, the Supreme Court asserted that racial segregation in public schools was "inherently unequal" and mandated that U.S. schools should integrate "with all deliberate speed" (*Brown v. Board of Educ. II*, 1955). In the years following *Brown*, however, legal and societal momentum to desegregate schools slowed. Redlining, housing discrimination, and white flight all contributed to *de facto* school segregation. Subsequent Supreme Court rulings compounded these challenges by allowing racial isolation as long as it is not explicitly mandated (*Milliken v. Bradley*, 1974), reducing court oversight of desegregation efforts (*Board of Education v. Dowell*, 1991), and limiting the use of race in admissions decisions (*Parents Involved v. Seattle School District No. 1*, 2007). Because of these legal and societal forces, many American schools are still divided by both race, ethnicity, and class (Lutz (2011), Reardon et al. (2012)). For instance, more than 75%

[1] Thanks to the NYC Department of Education for graciously sharing data. The views in this paper are those of the author and do not necessarily reflect the official policy or position of the NYC Department of Education.

of enrolled students are both low income and Black or Hispanic in 17% of American schools (GAO 2016).

Meanwhile, substantial and persistent achievement gaps exist. In 2000, students from low-income (10th percentile) families scored almost 1.25 standard deviations lower on standardized reading tests, and Black students scored about 0.75 standard deviations lower than their white peers by the same measure (Reardon (2013)). While these achievement gaps might not be solely a consequence of school segregation, the Black-White test score gap is higher in more segregated cities (Card and Rothstein (2007), Vigdor and Ludwig (2007)). Moreover, the gap increased in school districts that suspended race-based admissions (Billings and Hoekstra (2019), Cook (2018)). On the other hand, studies on the effects of integration policies suggest that integration may have both academic and psychological benefits to students (Angrist and Lang (2004), Guryan (2004), Johnson (2011), Johnson (2019), Zebrowitz et al. (2008)).

Such stark racial and socioeconomic gaps are especially apparent in large, urban school districts such as New York City. A 2012 analysis revealed that more than half of NYC's public schools enrolled student populations consisting of 90 percent or more Black and Hispanic students (Fessenden (2012)). The city has experimented with numerous proposals to diversify its schools (Shapiro (2019), Johnson (2019)). Most recently, Mayor de Blasio campaigned on the promise to make the schools "reflect the city better" (Harris (2018)).

School segregation persists in NYC despite its centralized school choice process which should afford considerable opportunity for social and racial integration. Indeed, school choice emerged in large part from a desire to stem the return to segregated neighborhood schools by decoupling school environment from residential sorting. Each year, more than 70 000 5^{th} grade students and 79 000 8^{th} grade students can choose among most of 479 NYC Middle schools and 418 NYC high schools. In practice, students' choice is somewhat limited by schools' selection criteria that tend to favor high performing students and nearby applicants.

Since centralized assignment systems reconcile students' preferences with schools' preferences, segregation is an equilibrium outcome of demand and supply side factors. School choice may result in a segregated school system either because of students' preferences (demand side) or because of schools' selection criteria (supply side). On the demand side, students may prefer enrolling in schools with similar classmates (Bjerre-Nielsen and Gandil (2020)) or close to their home (Laverde (2020)). In the latter case, school segregation will arise as a consequence of residential segregation. On the supply side, schools may implement admission criteria based on academic achievement or residence that result in homogeneous student body. Both mechanisms are likely to be important in NYC, which is one of the most residentially segregated city (Reardon et al. (2008)) in the United States and where roughly a third of Middle and High schools screen applicants on academic achievement.

The contemporary public debate over segregation in NYC often focuses on supply side factors and, in particular, on the practice of screened admissions. Selective enrollment schools are often accused to perpetuate racial and economic segregation, by allowing white and upper income families to bypass mostly-minority and low-income public schools (Hu and Harris (2018)).[2] Nonetheless, the importance of screened admission, as opposed to families' preferences or residential sorting, is not clear. Conversely, selective schools could help retain White and high income students into the public system. Indeed, screened schools came to prominence in the 1970s, when the City was facing an exodus of mostly White and Asian middle class families from the public school system.

In this paper, I study the contribution of screened admissions to the observed pattern of segregation in NYC middle schools by analyzing admission reforms in two local NYC school districts. In 2019, two of the 32 NYC school districts launched district-wide integration plans which reduced the role of screens in middle school admissions.

[2] This viewpoint is reflected in the New York Times' widely-followed 2020 podcast, Nice White Parents.

The Brooklyn Northwest district (district 15) eliminated traditional screening criteria and set aside 52 percent of the seats in each school for students who are low-income, English language learners, or homeless. At the same time, the Manhattan Upper West Side district (district 3) kept screened admission but set aside 25 percent of seats in each school for students who come from low-income families and earned low grades in elementary school.

I take advantage of student-level application and enrollment administrative data to disentangle the channels through which the two integration plans affect school segregation. This comprehensive data allows me to analyze how applicants adapt their application and enrollment behavior to changes in selective admissions, and how applicants' behavioral responses amplify or diminish the plans' effects on school segregation. In particular, I focus on changes in enrollment outside the public school sector and in the preferences applicants submit to the central assignment system.

I find that Northwest Brooklyn's integration plan substantially decreased economic and racial segregation at the district's schools. Applicants residing in the district attended middle schools that were 24 % less economically segregated and 16 % less racially segregated. On the other hand, Manhattan Upper West Side's integration plan was less successful and only decreased economic segregation by 9 %. The more substantial impact of Northwest Brooklyn's plan is consistent with the more far-reaching nature of its plan. On the other hand, the larger decline in economic segregation than in racial segregation is consistent with the fact that both integration plans targeted directly low-income students.

Despite their differences in final impact, both integration plans elicited important behavioral responses from applicants. During the application stage, families responded to the integration plans by changing the ranked order list of schools they submit to the central assignment system. During the enrollment phase, White applicants and higher-income applicants were much more likely to turn down their match offer and enroll outside the public school system after the implementation of the plans.

Overall, White student and higher-income student enrollment losses halved the impact of the integration plans effects on racial and economical segregation in both districts. On the other hand, the changes in ranked order lists reinforced the plans' effects on diversity which resulted in twice more diverse match offers thanks to applicants' behavioral response during the application period. Hence, applicant behavioral response at the enrollment and assignment phases essentially cancelled out.

To understand how applicants would respond to alternative reforms of school admissions, I try to identify the specific changes entailed by the integration plans that caused applicants' responses. I show that changes in match offers' take-up rates can be fully explained by changes in achievement of potential peers. White and higher-income applicants were more likely to turn down their match offer because their assigned school would have, on average, lower achieving potential peers after the implementation of the integration plans. Changes in application behaviors are harder to explain without explicitly modelling applicants' formation of ranked order list. Nonetheless, developing and estimating a model of applicants' choice is beyond the scope of the present paper.

The rest of the paper is organized as follows. Section **??** describes the NYC school assignment system and the data. Section 2.3 details the measure of segregation used throughout the paper and describes NYC residential and school segregation. Section 2.4 discusses the integration plans and explains the channels through which the integration plans may affect school economic and racial segregation. Section 2.5 presents the effects of the integration plans on school diversity. This section also details applicants' response to the plans. Section 2.6 concludes.

2.2 Background

2.2.1 The NYC middle school match

The NYC public school system counts 450 middle schools which enroll 70 000 new 6^{th} graders each fall. In the preceding winter, rising 5^{th} graders submit applications to NYC public middle schools through a centralized admission system run by the NYC Department of Education (DOE). Applicants apply to academic programs and are asked to rank them by order of preference. Subsequently, academic programs also submit a ranking of all their applicants. A school may operate more than one program. In the spring, the centralized admission system combines the information and makes a single school offer to each applicant.

Since 2003, the NYC DOE uses Deferred Acceptance (DA) to reconcile student and school preferences. Applicants that are unassigned at the end of DA are manually placed in programs with extra seats based on geographic proximity and expressed interests. About 92 % of students that completed their application[3] are matched in the main round. Finally, 12 % of students appeal their final offer. In this case, they submit a new rank ordered list of up to three programs (which may include programs they had applied for) and receive a new tiebreaker which is used to run a second round of DA. Nonetheless, less than 10% of those receive a new offer.

Applicants report their preferences to the mechanism through a ranked-order list, which was limited to 12 choices in 2017. To support families in the application process, The NYC DOE provides both a physical admission guide and access to a personalized website. Besides practical guidance on the application process, the guides include an information page about each school applicants may be eligible for. A school's information page includes a brief statement of its mission; a list of offered programs, courses and extracurricular activities; the performance of its students on standardized tests; admission priorities and selection criteria for each of its programs; the number

[3] The NYC DOE includes in the match any public school students enrolled in 5^{th} grade even if they did not log-in to the application website and/or did not rank any school.

of applicants per seat and the priority of last admitted applicant. The DOE also issues annual school reports that list basic demographics, teacher characteristics, and detailed statistics about performance levels and environment.

Each academic program has specific eligibility and admission criteria. Based on the factors used to rank applicants, programs may be classified in two broad categories: screened and unscreened. Both screened and unscreened programs may rank groups of students based on program-specific priorities, related to applicants' residential zones, school at the time of application and, rarely, attendance to an information session. However, only screened programs may also rank individual students based on prior grades, standardized test scores, talent tests scores and attendance at the previous school. Ties between applicants with the same rank are broken using a unique tiebreaker.

Contrary to NYC high schools, NYC middle schools are intended to serve students living in their neighborhood. Most middle schools only consider applicants from specific residential areas or give priority to local students. In 2018, 83% of programs had zone or district eligibility requirements, 14% were borough-wide programs and only the remaining 3% were city-wide programs. In addition, 23 % of borough-wide or city-wide programs gave priority to applicants residing or attending school in specific districts. Because of these rules, 85% of students attend a middle school in their district.

2.2.2 Data

The data for this paper is obtained from the DOE administrative information system. It covers all students enrolled in the New York City public school system. These data include the application and match data for NYC middle schools for enrollment years 2015-2016 to 2019-2020. The application and match files contain information on applicant's choices, applicants' priorities and rankings at the programs they applied to, applicants' main round offer, manual offer if applicable, and final offer. All applicants receive a final offer. The data also contains information about the disability status of the applicant as students with disabilities are matched to specific seats. I am able to replicate the main round offers received by 93% of applicants in the 2015-2016 match and by 99% of applicants in the most recent matches.

The application data can be matched through a unique identifier to data on school enrollment, students' demographics, standardized test scores, and residential location. The DOE collects school enrollment data in June of each year. Besides the grade and school enrolled[4], the data also contains information about the ethnicity, the poverty status which proxies for free or reduced price lunch (FRPL) status, and English language learner (ELL) status of each NYC student. The test score files include the results to NY State ELA and math standardized tests administered in grades 3 to 8. A performance level of 1, 2 ,3 , or 4 is associated to each scaled score. Students that score above 3, which corresponds roughly to a performance in the 60^{th} percentile, are considered high performers. Finally, the DOE provided students' census tract and zip code of residence. The distance to schools is computed as the flying air distance between the centroid of a student's census tract and each school.

2.3 Segregation in NYC

Although residential segregation appears to have decreased since its peak in the 1970's (Cutler et al. (1999)), American cities remain substantially segregated by race (Platt Boustan (2012)) and New York City is not an exception. Reardon et al. (2008) describes racial segregation in the 40 largest U.S metropolitan areas at different geographic scales. NYC is amongst the five most segregated metropolises when considering small neighborhoods of 500m radius. Nonetheless, the city fares better when segregation is measured at 4000m scale. In other words, the segregation patterns in

[4] The school enrolled corresponds to the last school a student was enrolled in during the academic year. Students that leave the NYC public school system mid-year have as school of enrollment the latest NYC public school they attended.

NYC are due more to small-scale variation in racial composition than to large-scale variation.

School choice may ameliorate the segregation inherent to neighborhood schools by allowing students to attend more distant schools. The policy has more leverage when segregation is at smaller scale level, as in NYC, since students do not need to travel long distances to mix with different peers.

The scope for within-district choice to increase integration may be limited by the share of non-minority students remaining in the district. Desegregation plans in the 1970's and 1980's were associated with significant white enrollment losses that offset part of the plans effect (Reber (2005)). Nonetheless, the NYC 6^{th} grade population was still one third white or Asian in 2018. This suggests substantial scope for school choice to affect racial mixing in NYC middle schools.

2.3.1 Measure of Segregation

I use an isolation index to measure segregation. The index corresponds to the probability that a member of one group meets a member of the same group within a given geographical unit. Let X_i be a dummy corresponding to an applicant's characteristic (e.g. being White) and S_i denote the geographic unit to which an applicant belongs (e.g. a neighborhood or a school). The isolation index for Whites is the expected share of Whites at the neighborhood or school level weighted by the distribution of whites across neighborhoods or schools.

$$I_X = E[E[X|S]|X] \tag{2.1}$$

To assess the level of integration, the isolation index has to be compared to the marginal distribution of the groups in the population studied. Indeed, if groups are identically distributed across locations, the index is equal to the marginal probability of each group in the population.[5] On the other hand, if groups are perfectly segregated, the index takes a value of 1. Therefore, I often standardize the isolation index by the group marginal probability in the population of interest. The standardized isolation index can be interpreted as the percentage deviation from perfect integration.

$$SI_X = \frac{E[E[X|S]|X] - E[X]}{1 - E[X]} \tag{2.2}$$

2.3.2 NYC Residential and School Segregation

Table 2.3.1: NYC Segregation in 2018

		Isolation index for different geographical units			
	Marginal dist.	Census	School	Possible MS assignment	
	NYC	tract	district	Real match	No screen match
	(1)	(2)	(3)	(4)	(5)
Black	0.22	0.54	0.40	0.48	0.47
Hispanic	0.41	0.59	0.52	0.57	0.56
Black + Hispanic	0.64	0.80	0.73	0.78	0.77
Asian	0.18	0.47	0.32	0.40	0.39
White	0.16	0.48	0.28	0.40	0.38
FRPL	0.75	0.81	0.77	0.80	0.80
N	78,723	78,723	78,723	78,723	78,723

Note: This table reports isolation indexes for different demographic groups and different geographic units in NYC in 2018. The isolation index corresponds to the probability that a student's peer in a given geographical unit is of the same group as hers. Columns 1, 2 and 3 report the probability that this happens at the scale of the city, the census tract and school district, and columns 3 and 4 at the scale of the middle school. The sample is restricted to 6^{th} grade applicants offered or enrolled in match schools who have non-missing demographic information. Columns (4) and (5) consider the school the student is offered in the match when available (92% of cases). The school offered in column (5) results from a simulation of the match where schools can rank students using only tie-breakers.

[5] At the population level, the index simplifies to the marginal distribution.

Table 2.3.1 documents the extent of residential and school segregation of middle school applicants in 2018. As the second column shows, residential segregation at the census tract level is substantial. For instance, although roughly 22 % of NYC's middle schools applicants are Black, more than half of a Black applicant's census tract neighbors are Black. The pattern is even more pronounced for White and Asian applicants: half of an Asian or White applicant's neighbors are of her same race when only approximately 15 % are so at the city level. The segregation is also economical as FRPL applicants also tend to be slightly more represented in some census tracts.

Residential segregation, as measured by the isolation index, falls when computed at the school district level. For example, columns 2-3 show that the share of neighboring students of the same-race for White or Asian students is down from around 50% in the same census tract to roughly 30% in the same school district. Since most students attend a middle school in their district of residence, the school district isolation indexes constitute a more relevant benchmark to evaluate school segregation than city-level proportions.

Schools appear to facilitate integration as same-race exposure is smaller in schools than at the census tract level. Nonetheless, schools are still segregated, even when compared to school districts. Column 4 of Table 2.3.1 shows that around 40% of a White or Asian 6^{th} grader's classmates are White or Asian,[6] which is halfway between the census tract and district levels' values. The same pattern holds for every race as well as for FRPL status. Overall, the NYC centralized seat assignment results in middle schools which are less diverse than school districts. In other words, the match fails to achieve the potential level of integration that eligibility and priority rules would allow.

[6] Classmates are defined as match participants who receive the same school offer or non match participants that enroll in the school. Hence, the measure does not account for the effect of applicants' take up of match offers on school segregation.

2.3.3 Why does school choice fail to integrate?

"School choice" may result in a segregated school system either because of students' preferences (demand side) or because of schools' selection criteria (supply side). Students may prefer enrolling in schools with similar classmates or very close to their homes (Laverde (2020)). In the latter case, school segregation arises as a consequence of residential segregation. On the other hand, schools may implement admission criteria that result in a homogeneous student body. In NYC, both channels could be important as 33% of middle schools screen their applicants on academic achievement and behavioral measures.

To isolate the contribution of screening to school segregation, I simulate a counterfactual central school assignment assuming NYC middle schools had not used screens but applicants' preferences had remained unchanged. Specifically, I simulate the mechanism using the ranked order lists submitted to the DOE but ranking applicants at each program uniquely based on their priority and tiebreaker.[7] As column 5 of Table 2.3.1 shows, the simulated assignment does not display much more integration than the actual match. For instance, the probability that a White or Asian 6^{th} grader's classmate is of the same race falls from 0.40 in the real match to 0.38 in the simulated match. According to these estimates, screening appears to account for about only 10 % of the difference between district level and school level segregation measures while the remaining difference is explained by applicants' preferences.

Nonetheless, the validity of the simulation's results relies upon the assumption that application behavior is unaffected by changes in programs' ranking criteria. This assumption is unlikely to hold if applicants incur a small cost when applying to each additional program. Indeed, Fack et al. (2019) underlines that, in the presence of an application cost, DA is not weakly strategy proof as applicants are better off omitting unlikely choices and limiting the length of their list if their first choices

[7] To make sure that all applicants receive an offer, I extend applicants' ranked order list to include all programs to which applicants are eligible. The added programs are ranked by proximity and appended to the end of the submitted rank order lists.

appear likely. Since abolishing screens substantially affects admission probabilities at screened programs, it is plausible that applicants would have modified their ranked order list in response to the change. The simulation is thus insufficient to identify the contribution of both schools' screens and students' preferences to school segregation. The remainder of the paper takes advantage of a natural experiment in two NYC districts to get a more accurate decomposition.

2.4 NYC integration plans

2.4.1 Description of the integration plans

Since 2017, some of NYC middle schools are participating in a "Diversity in Admission" initiative which aims at increasing diversity within their schools. This school-based initiative encourages schools to give admission priority for part of their seats to disadvantaged students. In practice, participating schools are free to decide the portion of reserved seats and the group of students that receive admission priority for reserved seats. Appendix table A1 describes the specific policies adopted by participating schools since 2017.

As part of this momentum, NYC school Districts 3 and 15 launched in 2019 district-wide "Diversity in Admission" initiatives that changed substantially programs' admission criteria for all district schools. Brooklyn's District 15 eliminated screening criteria at all its middle schools, and started reserving 52 percent of the seats in each school for students who are low-income (eligible for reduced-price or free lunch), English language learners, or homeless. Manhattan's District 3 plan represented a compromise between the traditional screened system and District 15's new policy; students are still screened, but some schools set aside 25 percent of seats for students who come from low-income families, struggle on state tests, or earn low report card grades.[8] The plans affected eleven middle schools in District 15 and sixteen

[8] 10 % of seats are reserved to FRPL-eligible students who score an average below 2 on a composite

middle schools in District 3. Both districts began implementing the changes as students applied for enrollment in the 2019-2020 school year. Although the details of the plans differed across districts, their aims were similar: limit the extent of screening to promote integration by expanding access to most selective schools, especially for low income and minority students.

The adoption of these plans by districts 3 and 15 provide two natural experiments to evaluate the impact of screens on school segregation. District-level plans are most likely to have an impact when the distribution of students in district schools doesn't match the demographics of the district, since they mostly affect the distribution of students living in a district among schools in the same district. Both district 3 and district 15 satisfy this condition. As shown in Table 2.4.1, both districts are more diverse than the city. In particular, White students, who constitute respectively 40 % and 31 % of district 3 and 15's student populations, are twice more represented in the two districts than city-wide. The two districts have also the lowest shares of FRPL students in NYC, at about 42% and 54%.

Yet, prior to the adoption of the plans, schools in both districts were also more racially and economically segregated than schools in other NYC districts. Prior to 2019, Black and Hispanic students' standardized isolation indexes reached values of 0.44 in district 3 and 0.26 in district 15 against 0.03 on average in NYC. Similarly, FRPL students' isolation was equal to 0.47 in district 3 and 0.34 in district 15 compared to 0.08 citywide. Moreover, admission screens were particularly prevalent in both districts before the reforms. Prior to 2019, 57% of district 3 programs and 80% of district 15 programs screened their applicants on grades and behavioral measures, while only 33% of NYC programs did so. Hence, if admission screens play an important role in school segregation, reducing their role in these two districts should result in a substantial increase in school diversity.

Studying district level plans allows one to assess the consequences of screens more

of 4th grade math and ELA scores and 15 % of seats are reserved to FRPL-eligible students who earn an average between 2 and 3 on the same composite score

Table 2.4.1: Characteristics of District 3 and District 15

	NYC (1)	D3 (2)	D15 (3)
Panel A: Characteristics of Schools and Applicants			
# of middle schools	491	20	11
# of programs	686	21	13
% of screened programs	33%	57%	80%
# of applicants	71512	1134	2536
% Asian applicants	0.18	0.07	0.22
% Black applicants	0.23	0.20	0.06
% Hispanic applicants	0.41	0.29	0.38
% White applicants	0.16	0.40	0.31
% FRPL applicants	0.72	0.42	0.54
% ELL applicants	0.13	0.05	0.16
Applicants mean math proficiency	2.8	3.3	3.1
Applicants mean english proficiency	2.7	3.1	2.9
Panel B: Standardized School Isolation index			
Black + Hispanic	0.03	0.44	0.26
White	0.17	0.22	0.19
FRPL	0.08	0.47	0.34

Note: This table presents the characteristics of NYC school districts. Column (1) includes all NYC middle school applicants, while columns (2) and (3) include applicants residing in district 3 and district 15 respectively at the time of application. Panel A describes the population of middle school applicants that enroll in 6^{th} grade in years 2015-2018. The number of schools and programs correspond to the average number of schools and programs in each year, not to the total number of schools and programs that ever existed. Mean math and English proficiency are computed based on the proficiency level obtained in 4^{th} grade state tests. Panel B presents the standardized school isolation index for three groups of applicants (Black and Hispanic, White and FRPL). Standardized school isolation indexes are computed for each NYC district separately by standardizing each group's school isolation index by its share among applicants residing in the district. The average standardized school isolation index for NYC corresponds to the average of district-level standardized school isolation indexes weighted by the share of NYC students of each group living in each district.

accurately than analyzing school-level pilots. It is unlikely for applicants to gain admission at a school out of their district, while it is easy to not apply to a single school. Hence, if reducing screens improves integration, we should observe an increase in integration in both districts as students are constrained to mix within their district or to leave the public school system altogether. In addition, these district-level policies allow one to evaluate how screening affects the retention of more advantaged and White students in the public school system.

2.4.2 How did the integration plans impact school diversity?

The district 3 and 15 plans aimed at increasing school diversity by changing admission criteria. To this aim, both plans lowered admission barriers at selective schools for low income (FRPL) students with low baseline test scores. Specifically, the district 3 plan increases admission odds of FRPL applicants with low baseline test scores at screened schools by reserving them seats. On the other hand, the district 15 plan increases their odds of admission at previously screened schools by cancelling screens and reserving seats to low-income students.

As a consequence, the plans decreased admission probabilities at competitive schools for non-FRPL applicants with high baseline test scores. Indeed, these students had high admission odds prior to the plans because of their test scores, and they are not eligible for the reserves created by the plans. Table 2.4.2 summarizes the effects of the integration plans on admission probabilities for students with different FRPL status and baseline test scores.

Table 2.4.2: Changes in Admission Probabilities Due to the Integration Plans

	FRPL students	non-FRPL students
	District 3	
Low baseline test scores	↑	↓
High baseline test scores	↓	↓
	District 15	
Low baseline test scores	↑	Ambiguous
High baseline test scores	Ambiguous	↓

While both integration plans target directly economic segregation, by increasing the admission odds of some low-income students, they may also affect racial segregation. Black and Hispanic students are more likely to benefit from the changes in admission odds as they are often FRPL-eligible and have lower baseline test scores on average.[9] Hence, approximately 70% of district 3 Black and Hispanic students and 50% of district 15 Black and Hispanic students are eligible for the reserves created under the integration plans. By contrast, only 3% and 15% of White students qualify to the reserve in district 3 and district 15 respectively. It follows that both integration plans could reduce racial segregation as well as economic segregation.

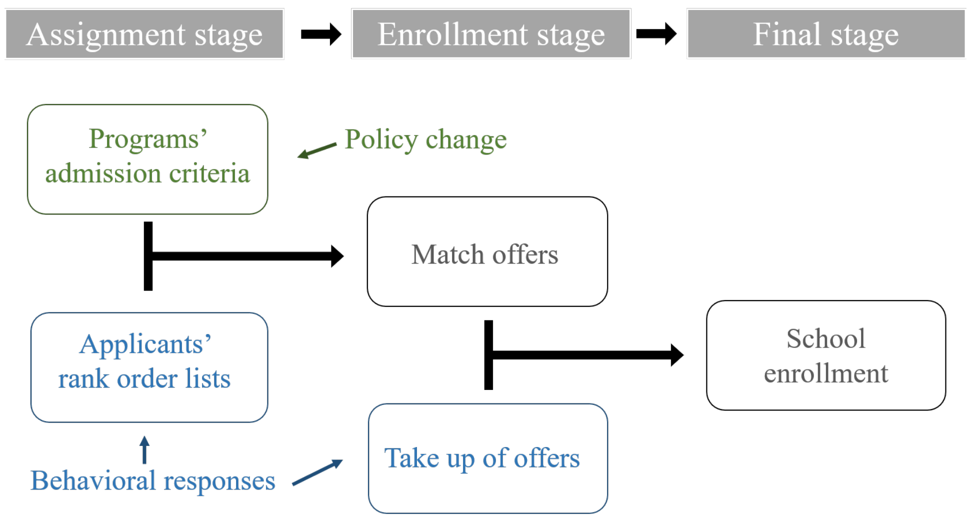
The extent to which both plans increase diversity at the districts' schools depends not only on the plans themselves but also on students' application and enrollment behaviors. Considering the structure of NYC middle school match, a change in admission scheme corresponds to a change in the preferences of schools in the match. As shown in the flow chart of Figure 2.4.1, school preferences affect which students enroll in each school through DA offers. Students' behavior influences the enrollment effect of a change in school preferences by impacting both which DA offers are made and which DA offers get translated into enrollment.

As DA reconciles schools and students' preferences, students' preferences affect the extent to which a change in schools' admission criteria may translate into more diverse offers. Changing admission criteria cannot increase diversity if students only apply to schools attended by similar peers. As such, districts 3 and 15 plans' effect on diversity relies upon the fact that low income and minority applicants apply to schools enrolling mostly high income and non-minority students. On top of pre-existing application patterns, students may modify the ranked order lists they submit to the mechanism in response to the change in admission criteria, which will affect further the offers made by the mechanism.

Students might impact the effect of the plans on school diversity also through their decision of taking-up their match offer. Once they receive their final offer from the

[9] See Appendix table A2 for a description of students' characteristics by race.

Figure 2.4.1: From admission criteria to final enrollment



match, students may decide whether to enroll in their offered school or in a school outside the public school system. Students' take up decisions affect which match offers get translated into enrollment. Hence, applicants' take-up behavior determines whether match offers are more or less diverse than actual enrollment. As such, a change in applicants' take up in response to the diversity plans will affect its final effect on school diversity.

Depending on the nature and extent of applicants' behavioral responses, the effect of a change in admission criteria on school diversity could be amplified or diminished. For instance, if high income applicants assigned to schools enrolling a majority of low income applicants systematically reject their offer and exit the public school system, then the change in admission criteria will not result in an increase in school diversity. On contrary, school diversity could even decrease as fewer high income applicants attend public schools. On the other hand, if low income applicants start listing competitive schools because they anticipate a higher probability of admission or a more welcoming environment, then the effect of the plans on diversity will be reinforced.

2.5 Impact of the integration plans

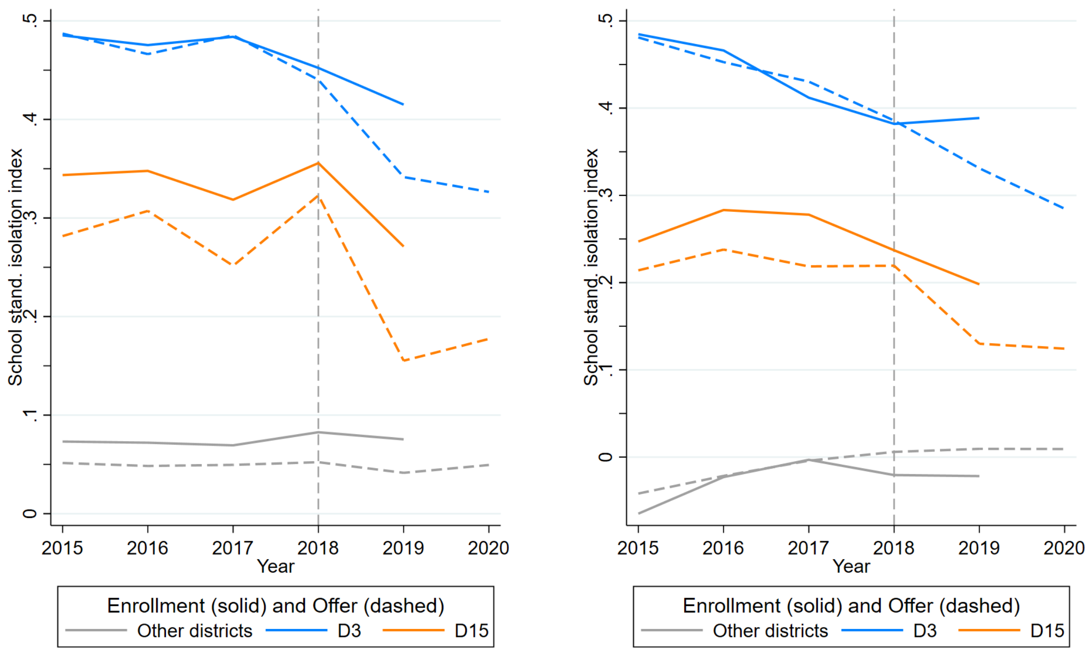
2.5.1 Total effects on school diversity

As a first step in the analysis, I consider the total effect on diversity of both integration plans. In an early evaluation of the plans, Margolis et al. (2020) find large statistically significant effect of the D15 integration plan. According to their analysis, economic segregation and racial segregation in district 15 decreased by 55% and 38% respectively. On the other hand, they find no evidence of a statistically significant impact of the district 3 integration plan.

Nonetheless, the scope of their analysis is limited by their use of school level enrollment data. They measure within-district school segregation as the the mean absolute percentage point difference between the proportion of a particular group of students enrolled in each school and the district. Since their segregation index is with respect to the population of students that enrolled in district schools, it cannot account for potential changes in the demographic composition of students enrolling in the district. In particular, it cannot capture the effect on segregation of students enrolling in schools outside their district in response to the plans. If many white and high-income students exit their district, the segregation indexes for minority and low-income students would be lower because their share among students enrolled in the district rose.

To take into account changes in the demographics of students enrolling in each district, I standardize the school isolation index for each group by the proportion of applicants residing in the district belonging to the group. This standardization with respect to the population of students residing in the district allows me to control for changes in applicants' demographics makeup in each district over time while taking into account changes in students' exit behaviors. Hence, this guarantees that within-district minority students' and low-income students' isolation is larger if more non-minority and high-income students enrolls schools outside their district of residence.

Figure 2.5.1: Evolution of Standardized School Isolation Indexes for FRPL and Minority Students



(a) Free and reduced price lunch students (b) Minority students

Note: These figures plot the evolution of school standardized isolation indexes for district 3, district 15 and other NYC districts between 2015 and 2020. Panel A displays the standardized index for applicants classified as disadvantaged by the DOE, a proxy for FRPL status. Panel B displays the standardized index for Black, Hispanic, Native American and multi-racial applicants. To obtain the plotted values, school isolation indexes are standardized by the share of students belonging to the group considered among applicants residing in the district. The standardized school isolation index for other NYC districts correspond to the weighted average of district level standardized indexes, with weights equal to the shares of NYC students belonging to the group considered residing in each district. Dashed lines give the value of the standardized school isolation index at the offer stage, that is if all students were to enroll in the school they are offered in the match. Solid lines correspond to the value of the index after enrollment. Students that leave the NYC public school system are not considered to compute the school isolation index but are included in the standardization.

Figure 2.5.1 shows the trends in the school standardized isolation indexes for FRPL students (panel a) and ,minority students (panel b) in district 3, district 15 and other NYC districts. Prior to the integration plans, economic segregation in both districts had remained fairly stable, with some year-to-year variation. On the other hand, racial segregation was stable in district 15 prior to the plan implementation but was already on a declining trend in district 3. Thus, a further decline in racial segregation in district 3 could be due to factors unrelated to the integration plans.

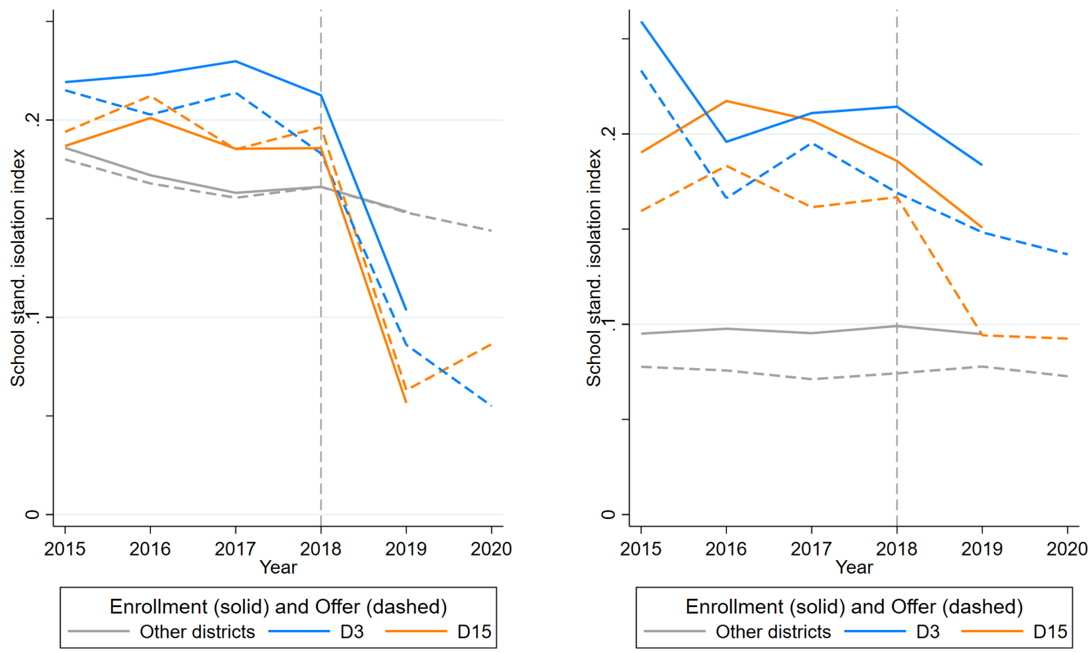
The plans resulted in a decrease in economic segregation in both districts, although the drop was larger in district 15. In 2019, the first year of its integration plan, FRPL isolation at the school of enrollment in District 3 dropped by 9% from the prior year's level of segregation. In the same year, FRPL isolation in District 15 dropped by 24%. The more substantial impact of district 15 plan on economic segregation follows from the more far-reaching nature of its plan, which eliminated screening at all middle schools and reserved 52% of seats for more disadvantaged students.

The integration plans also resulted in a decrease in racial segregation in district 15, albeit to a lesser extent. The Black and Hispanic isolation index decreased by 16% in district 15 between 2018 and 2019. The larger declines in economic segregation than in racial segregation are consistent with the fact that both integration plans targeted directly low-income students.

The decrease in racial segregation in D15 after the plans' implementation occurred mainly as a result of Hispanic and White students attending more integrated schools. As shown in Figure 2.5.2, the school standardized isolation index fell by 72% for White students and by 19% for Hispanic students in D15 after the reforms. On the other hand, there was no noticeable change for Asian and Black students (appendix figure A1).

Although there was no change in minority students' isolation in district 3, isolation decreased by 49% for white students. This apparent inconsistency arises from the fact, which will be explored in detail in the next section, that a larger share of white students exit public school after the implementation of the integration plan. As a result of this selective attrition, White students remaining in the district attended schools that enrolled fewer white students on average. Their school of enrollment had thus on average a share of White students closer to the share of White applicants residing in the district. A similar substantial increase in White students' exit also explains the very sharp drop in isolation for Whites in district 15 after the plan implementation.

Figure 2.5.2: Evolution of Standardized School Isolation Indexes for White and Hispanic Students



(a) White students

(b) Hispanic students

Note: These figures plot the evolution of school standardized isolation indexes for district 3, district 15 and other NYC districts between 2015 and 2020. Panel A displays the standardized index for White applicants. Panel B displays the standardized index for Hispanic applicants. Standardized school isolation indexes are computed as in figure 2.5.1.

From Figures 2.5.1 and 2.5.2, we can see that the economic segregation index declined in District 15 and 3 in 2019-20. However, we also see a fair amount of year-to-year volatility prior to 2019-20, even though no district-wide integration plan was being implemented. Inference on district-level indexes is challenging as each district-year combination corresponds to only one observation. To get a sense of how much "noise" we might expect in the isolation indexes, Figure 2.5.3 plots the trends in standardized school isolation indexes for each NYC district. For each district, segregation indexes are normalized to zero for 2018, the year prior to the implementation of the integration plans. To get a more accurate comparison between districts, the plot for each demographic group includes only districts in which at least 10% and no more than 90% of the student population belongs to the group.[10]

From the figure, it appears that the district 15 decline in economic and racial segregation is most likely attributable to its integration plan. Compared to other NYC districts, district 15 had among the largest drop in isolation in 2019 for all demographic groups. Moreover, the district had amongst the smallest variations for years prior to the plan implementation. Hence, the drop in district 15 isolation indexes is not likely due to chance or pre-existing trends.

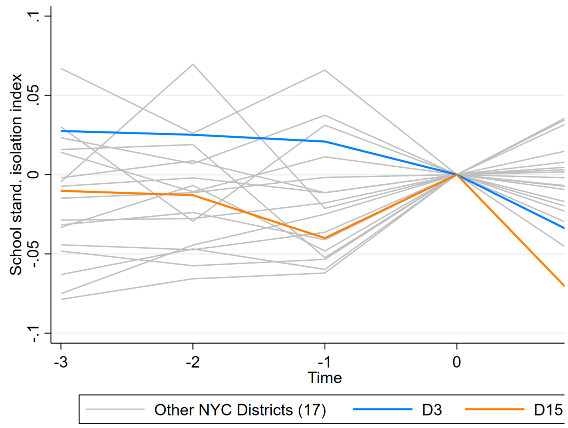
The evidence is less convincing for district 3, where the standardized isolation index for both FRPL and minority students were in a declining trend prior to the implementation of the diversity plan. The pre-existing declining trend is particularly marked for Black and Hispanic students, for which district 3 has amongst the largest values two and three years prior to the intervention. Moreover, the 2019 decline in isolation for minority students and Hispanic students is only modest compared to other districts. While it is plausible that the decrease in economic segregation in district 3 resulted from its integration plan, the decrease in racial segregation appears to be limited and within the bounds of typical year-to-year fluctuations.

Taken as a whole, the analyses presented here provides evidence that District 15's policy decreased economic segregation, as it was designed, and also had an indirect smaller impact on reducing racial segregation. In District 3, we see little evidence of a decrease in economic or racial segregation. While the point estimate for economic segregation is negative, it seems within the bounds of ordinary year-to-year noise.

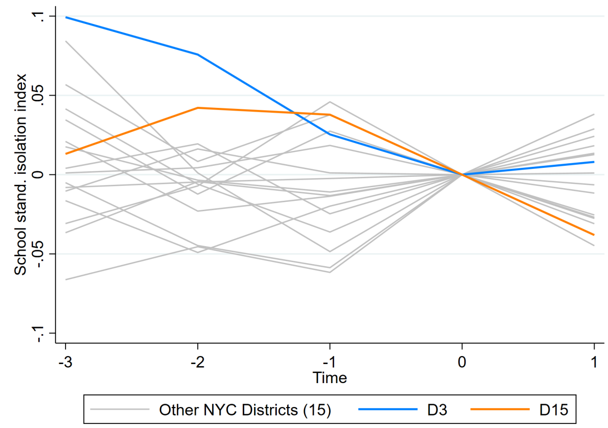
These results are consistent with Margolis et al. (2020) but much smaller in magnitude. The difference in magnitude arises from the fact that my standardization

[10] As it is divided by one minus the share of the group in the population, the normalized isolation index is sensitive to small changes in this share when the share is large. On the other hand, the isolation index depends on the choice of few students when the share is small.

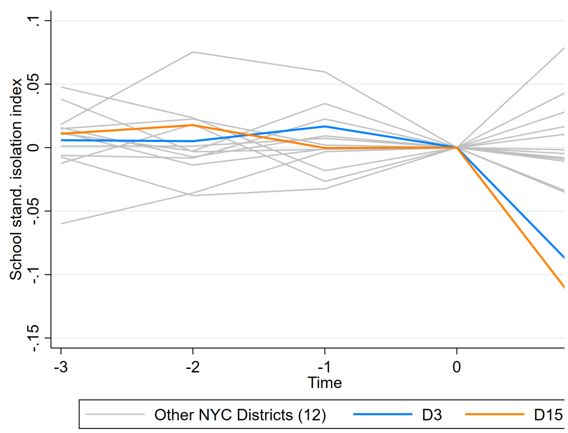
Figure 2.5.3: Comparison of Changes in Standardized School Isolation Indexes in all NYC Districts



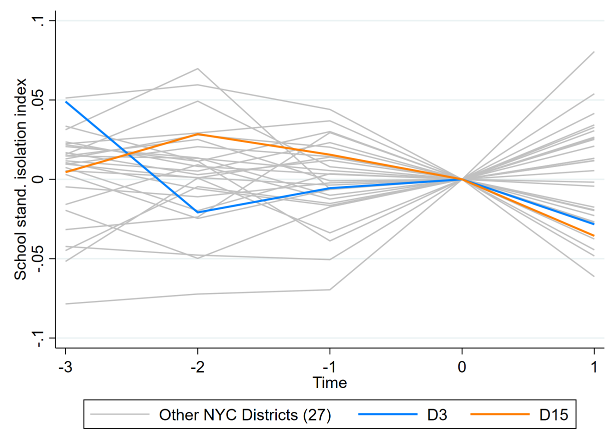
(a) Free and reduced price lunch students



(b) Minority students



(c) White students



(d) Hispanic students

Note: These figures plot changes in standardized school isolation index with respect to 2018 for all NYC districts. Panel A considers the index for FRPL eligible applicants, panel B for Black, Hispanic, Native American and multi-racial applicants, panel C for White applicants, and Panel D for Hispanic applicants. Each plot only includes districts in which at least 10% and no more than 90% of the student population belongs to the group considered. The data is also winsorized by excluding districts that display variations in standardized school isolation index larger than 10 % prior to 2018.

takes into account the effect of exits. Indeed, if one were to consider segregation at the match offers stage as shown in dashed lines in the graphs, the decline in economics and racial segregation would be close to the ones founds in Margolis et al. (2020).

The gap between segregation at the offer stage and enrollment stage widened after the implementation of the plans, pointing to the fact that changes in take up behaviors might have mitigated the effects of the reforms on diversity.

2.5.2 Effects on applicants' exit from the public school system

Reber (2005) documents the substantial decline in White students enrollment following the implementation of court-ordered segregation plans in the 1970's and 1980's. Since these plans, the demographic make-up of urban school districts has changed significantly as Black and White students have been progressively replaced by Hispanic and Asian students. In 2018, only 38% of NYC middle school students were either Black or White. Given this difference in context, it is not obvious that district 3 and district 15 integration plans would lead to similar flights than those observed in the 1970's and 1980's.

Nonetheless, the differences in effects of the integration plans on segregation at the match offer stage and enrollment stage hint to an increase in the share of White students and high-income students enrolling outside of NYC public schools. Indeed, computing isolation indexes based on match offers corresponds to assuming perfect take-up of offers. According to figure 2, changes in offer take-up reduced the decrease in economic segregation by 25 p.p. and the decrease in racial segregation by 15 p.p in both districts.

To estimate the effect of the integration plans on applicants' take up behavior, I implement a differences-in-differences regression that controls for district and year fixed effects. Specifically, I estimate the following regression

$$Y_{itd} = \lambda_t + \delta_d + \beta_1 \mathbb{I}(d = 3) \times \mathbb{I}(t = 2019) + \beta_2 \mathbb{I}(d = 15) \times \mathbb{I}(t = 2019) + \epsilon_{itd} \quad (2.3)$$

where λ_t and δ_d are year and district fixed effects. Y_{itd} is a dummy that takes a value of 1 when applicant i residing in district d enrolls in a school outside the public sector in year t. β_1 and β_2 capture the increase or decrease in the probability that students attend a school outside the traditional public school sector after the implementation of the integration plans.[11]

The differences-in-differences estimates presented in panel A of Table 2.5.1 suggest that both integration plans substantially affected applicants' enrollment decision. In particular, the plans resulted in a large increase in the shares of White students and non-FRPL students enrolling outside the public school system. Both shares went up by almost 7 p.p. in district 3 and by almost 8 p.p. in district 15. On the other hand, the integration plans had limited effects on exit from public school for Black, Hispanic, Asian and FRPL students. The point estimates for most of these students were close to zero and insignificant in both districts. The only exceptions are Asian students in district 3, who were 9 p.p. more likely to exit the public school system, and Black students in district 15, who were 7 p.p. less likely to exit the public school system. Nonetheless, these changes in Asian students and Black students' exit had limited impact on the district overall, as Asian students an Black students represent less than 7% of applicants in district 3 and district 15 respectively.

Can this decrease in offer take-up for White and non-FRPL students be rationalized by these applicants getting less desirable assignments as a result of the integration plans? As emphasized in the previous section, both integration plans entailed a decrease in admission odds at more competitive programs for most White and non-FRPL students. As suggested by panel A of Table 2.5.2, these lower admission probabilities did lead to White applicants and non-FRPL applicants being admitted to less preferred schools after the implementation of the integration plans. Compared to previous years, White applicants and non-FRPL applicants were offered on average a choice ranked 0.7 and 1.4 position lower in their list in district 3 and district 15 respectively.

[11] I consider Charter schools to be outside the traditional public school sector as they do not take part in the match and the NYC DOE only collects partial enrollment data for these schools.

Table 2.5.1: D-in-D Estimates of Changes in Out-of-district Enrollment

	Probability of exiting the public school system						
	All	Black	Hispanic	Asian	White	FRPL	non-FRPL
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
D3 × 2019	0.04***	-0.01	0.04*	0.09**	0.06***	0.00	0.07***
	(0.01)	(0.03)	(0.02)	(0.05)	(0.02)	(0.02)	(0.02)
D15 × 2019	0.03***	-0.07**	0.01	-0.00	0.08***	-0.01	0.08***
	(0.01)	(0.03)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
mean D3 pre-2019	0.10	0.13	0.07	0.12	0.10	0.09	0.11
mean D15 pre-2019	0.12	0.18	0.09	0.07	0.18	0.09	0.17
N	332,491	73,261	136,303	60,515	56,173	239,494	92,997

Note: This table reports differences-in-differences estimates of integration plan effects for 2015-2019 middle school applicants. The endogeneous variable is the interaction of dummies for residing in district 3 and district 15 and applying for admission in 2019. The dependent variable is a dummy equal to one for applicants that do not enroll in a NYC public school at any point of the school year following admission in middle school. The last two rows report the mean of the dependent variable among 2015-2018 applicants. All models control for year and district fixed effects. Robust standard errors on year are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. .

Nonetheless, this change in the rank of the offer obtained might not capture a change in the desirability of the offer obtained if applicants adapted their list in response to the integration plans. Hence, panel B of Table 2.5.2 considers the effect of the plans on a proxy for school desirability: the mean math baseline test score of offered students.[12] This choice of proxy is motivated by Abdulkadiroğlu et al. (2020) which find that peer achievement is the main determinant of NYC high school popularity. Results using potential peers mean baseline achievement are consistent with panel A estimates. The mean peer math test score offered to White applicants and non-FRPL applicants was about 0.25 standard deviation lower compared to previous years after the implementation of the diversity plans.

[12] I use 5^{th} grade math achievement since fewer students have a missing math test score than a missing English test score. The peer mean math baseline test score is computed leaving out each student to avoid simultaneity bias.

Table 2.5.2: D-in-D Estimates of Changes in Desirability of Assigned School

	All (1)	Black (2)	Hispanic (3)	Asian (4)	White (5)	FRPL (6)	non-FRPL (7)
	Panel A: Applicant's rank of match offer						
D3 × 2019	0.49***	0.19	0.23**	1.00***	0.72***	0.13*	0.67***
	(0.05)	(0.12)	(0.10)	(0.22)	(0.08)	(0.08)	(0.08)
D15 × 2019	0.55***	-0.32**	-0.01	0.05	1.41***	-0.11**	1.35***
	(0.05)	(0.13)	(0.06)	(0.11)	(0.11)	(0.05)	(0.10)
mean D3 pre-2019	1.56	1.89	1.69	1.54	1.51	1.81	1.52
mean D15 pre-2019	2.13	2.23	1.90	2.53	2.30	2.10	2.33
	Panel B: Mean math baseline test score of match offer peers						
D3 × 2019	-0.11***	-0.04	-0.01	-0.24***	-0.23***	0.07**	-0.23***
	(0.02)	(0.05)	(0.04)	(0.07)	(0.02)	(0.03)	(0.02)
D15 × 2019	-0.10***	-0.01	0.02**	-0.05***	-0.25***	0.03***	-0.24***
	(0.01)	(0.03)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
mean D3 pre-2019	0.37	-0.16	0.05	0.81	0.77	-0.11	0.72
mean D15 pre-2019	0.31	0.19	0.11	0.36	0.550	0.15	0.51
N	396,958	85,967	163,110	72,553	67,453	286,690	110,268

Note: This table reports differences-in-differences estimates of integration plans effects for 2015-2020 middle school applicants. In all panels, the endogeneous variable is the interaction of dummies for residing in district 3 or district 15 and applying for admission in 2019. Panel A dependent variable indicates the ranking of the school applicants were assigned to in the match. If an applicant is assigned through manual placement to a choice she did not listed, the variables takes the value of the length of her list plus one. Panel B dependent variable is the leave-out mean 5^{th} grade math test score among applicants offered the same school by the match. The last two rows of each panel report the mean of the dependent variable among 2015-2018 applicants. All models control for year and district fixed effects. Robust standard errors on year are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. .

Table 2.5.2 provides suggestive evidence that the changes in student enrollment due to the plans are associated with changes in the desirability of the school obtained through the match. To push it one step further, I investigate whether all the changes in offers take up are mediated through changes in the characteristics of potential peers. Let W_{itd} be a characteristic of the offer received by student i residing in district d

and applying in year t. For instance, W_{itd} is the average baseline math test score of applicants that received the same match offer as student i. If the offer take-up decision depends on W_{itd} and students characteristics X_{itd}, it can be summarized by a multivariate function which can be approximated by an additive linear probability model.

$$E[Y_{itd}|W_{itd}, X_{itd}] = \lambda_t + \gamma_d + \mu W_{itd} + X'_{itd}\pi \tag{2.4}$$

Under this model, district 3 and district 15 plans provide with exogenous variation in W_{itd}. So equation 2.4 can be estimated using the integration plans as instruments for W_{itd}. Let Z_{itd} denote district 3 reform instrument, i.e. the interaction of the 2019 dummy and the dummy for district 3. Excluding district 15 applicants from the sample, the reduced form equation is

$$E[Y_{itd}|Z_{itd}, X_{itd}] = \lambda_t + \gamma_d + \mu E[W_{itd}|Z_{itd}, X_{itd}] + X'_{itd}\pi \tag{2.5}$$

Differencing by Z_{itd} conditional on X_{itd}, we obtain:

$$E[Y_{itd}|Z_{itd} = 1, X_{itd}] - E[Y_{itd}|Z_{itd} = 0, X_{itd}] = \mu(E[W_{itd}|Z_{itd} = 1, X_{itd}] - E[W_{itd}|Z_{itd} = 0, X_{itd}]) \tag{2.6}$$

Equation 2.6 implies that the reduced form effect of instrument Z_{itd} for each X_{itd} and any Z_{itd} is driven by the corresponding covariate-specific change in W_{itd} induced by Z_{itd}. In other words, the covariate-specific reduced form effect of any of the two integration plans on exit should be proportional to the corresponding covariate-specific first stage effect of the plans on potential peers average math test score. The slope of the line linking differences in exit rates due to the integration plans to the corresponding differences in potential peers achievement is the causal effect of potential peers achievement.

The proportionality hypothesis embodied in equation 2.6 generates two sets of testable restrictions. The first entails that IV estimates using district 3 reform and district 15 reform as instruments for potential peers achievement should be equal. The second is a constant-effects assumption which says that IV estimates across covariate-defined subgroups should be equal. Both restrictions should hold if all the effect of the integration plans on public school exit is mediated through a decrease in potential peers achievement.

Table 2.5.3: 2SLS Estimates of Potential Peers Achievement Effect on Exit from Public School

	Plans instruments			Plans instruments with covariate interactions		
	district 3 (1)	district 15 (2)	district 3 and district 15 (3)	district 3 (4)	district 15 (5)	district 3 and district 15 (6)
Potential peer achievement	-0.29** (0.14)	-0.35*** (0.09)	-0.33*** (0.08)	-0.20*** (0.06)	-0.30*** (0.05)	-0.27*** (0.04)
First stage F	4.8	14.1	18.3	11.1	32.3	41.8
Overid p-value			0.75	0.40	0.41	0.44
Overid DF			1	8	8	17
N	298,802	304,942	309,959	298,802	304,942	309,959

Note: This table reports alternative IV estimates of the effects of potential peers achievement on the probability of exiting the public school system for 2015-2019 middle school applicants. Potential peers achievement corresponds to the leave-out mean 5^{th} grade math test score among applicants offered the same school by the match. Estimates in columns (1) to (3) were computed by instrumenting potential peers mean achievement by district 3 and district 15 integration plans. For the estimates in columns (4) to (6), the instrument list includes the two integration plans plus interactions with covariates (English learner, race, interactions of low baseline test scores with FRPL eligibility). The table also reports first-stage F-statistics and overidentification test p-values and degrees of freedom. For each specification, the sample is restricted to applicants with non-missing covariates. Robust standard errors are reported in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%.

As a point of reference, Table 2.5.3 reports in columns (1) and (2) just-identified IV estimates computed using separately district 3 and district 15 integration plans to instrument potential peers achievement. The resulting estimates are similar, although the estimate using district 3 reform is less precise due to the smaller first stage. Being offered a school where peers perform 0.1 standard deviation better on average decreases the probability of exiting the public school system by approximately 3 p.p.

As can be seen in column (3) in Table 2.5.3, 2SLS estimates using both districts integration plans to instrument potential peers achievement are a little more precise and qualitatively similar than the just-identified estimates computed using district 15 reform alone. The overidentication test statistic associated with 2SLS procedure that uses both district plans as instruments gives a formal test of the equality of IV estimates computed using districts' plans as instruments one at a time. The p-value for the overidentification test rejects the hypothesis of a difference in estimates with a value of 0.75.

The overidentification test statistic associated with a 2SLS procedure that uses the integration plans and plans-covariate interactions as instruments implicitly tests equality of IV estimates computed separately for covariate subgroups. This test therefore evaluates the second set of restrictions implied by equation 2.6. In column (4), I first implement this test by instrumenting potential peers achievement with district 3 plan and 8 covariate interactions, in a model that controls for covariate main effects (as well as their interactions with district and year fixed effects). With 8 covariate interactions and an integration plan main effect in the instrument list, the resulting overidentification test has 8 degrees of freedom. The covariates used in the estimation are dummies for English learner status, race and the interactions of low baseline tests scores with Free and reduced price lunch eligibility.

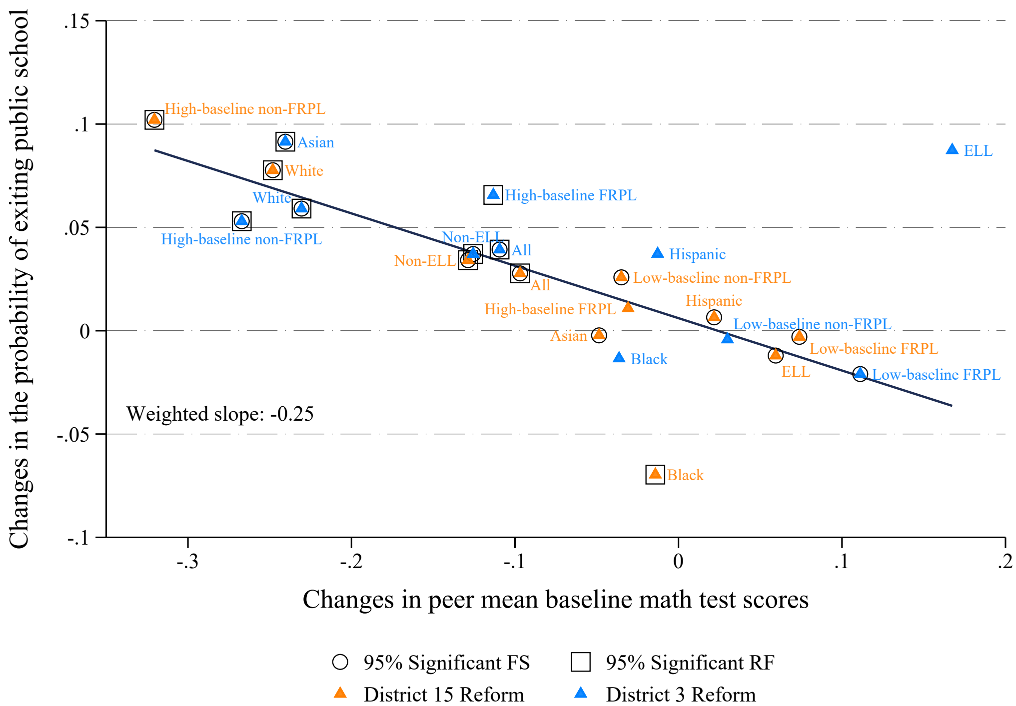
The 2SLS estimate using district 3 plan and plan-covariate interactions as instruments for potential peers achievement is similar to the one using district 3 integration plan alone. Moreover, the overidentification test statistic associated with the estimate provide little evidence of differences in impact across covariate subgroups since its p-value is 0.40. Column (5) of Table 2.5.3 also presents estimates computed using an instrument list that interacts district 15 plan with covariate subgroups. The overidentification p-value again show little evidence of heterogeneous effects.

Finally, column (6) reports results from a 2SLS procedure that combines both integration plans and the associated sets of covariate interactions for a total of 18 instruments. The resulting estimate lie between the corresponding 2SLS estimates computed using covariate interactions with each district plan alone and is otherwise close to the estimates using district plans alone. The overidentification test statistics for this model (with 17 degrees of freedom) offers little evidence against the homogeneity restrictions described by equation 2.6.

This remarkable homogeneity is reflected in Figure 2.5.4, which presents a graphical representation of 2SLS estimates with covariate interactions in the instrument list. Specifically, the figure plots covariate-specific reduced form estimates for public school exit probability against the corresponding first stage estimates for potential peers mean achievement, constructed using both districts' integration plans instruments. Red numbers in the figure label results using the district 3 plan instrument, while black numbers label those using district 15 plan. Each reduced form and first stage estimate in the figure is at one-at-a-time for a variety of (potentially overlapping) covariate cells. For example, points labeled with 7 show first stage and reduced form estimates for White applicants.

Consistent with the direct effects of the integration plans in Table 2.5.1 and panel B of Table 2.5.2, the integration plans generate large positive first stage decreases in potential peers math baseline achievement and a corresponding set of positive reduced form estimates for White and non-FRPL applicants. On the other hand, the first-stage effects for low-baseline FRPL students of the integration plans on potential peers achievement are positive, while the corresponding reduced form estimates are negative. Finally, the points for different subgroups fall roughly on a straight line with a slope equal to about -0.25 . This visual IV estimate indeed appears to rationalize the complete collection of reduced form and first stage estimates plotted in the figure.

Figure 2.5.4: Covariate VIV for the Effects of Potential Peer Achievement on Exit from Public School



Note: This is a visual instrumental variable (VIV) plot of reduced form effects of district 3 and district 15 integration plans on the probability of exiting public school against the corresponding first stage effects of district 3 and district 15 integration plans on potential peer math baseline test scores, separately for a set of 11 covariate-defined groups. district 15 plan effects are plotted in black; district 3 plan effects are plotted in red. Covariate-specific estimates are computed one at a time in the relevant subsamples and labeled from 1-11.The slope of the line through these estimates is -0.25. Fitted lines are not forced to pass through the origin, as implied by the proportionality restriction described in the text.

It thus seems plausible that all the effect of the integration plans on public school exit is mediated through a decrease in potential peer achievement.

Finally, Table 2.5.4 explores whether changes in other peers characteristics which might be correlated with math baseline achievement could instead explain the changes in public school exit probability. Specifically, the table considers the proportion of minority students and the proportion of FRPL students among applicants that receive the same offer, in addition to math achievement.

The 2SLS estimates using both integration plans and their covariate interactions as the instruments list for each potential peers characteristics are reported in columns (1) to (3). The proportion of minority students and FRPL students among potential peers are both associated with higher exit rates. Being offered a school with 10 p.p. more minority students or FRPL students increases the probability of enrolling outside the public sector by approximately 4 p.p. Moreover, both estimates are statistically significant at the 1% level.

Nonetheless, the overidentification p-values for models treating proportion of minority students and proportion of FRPL students as single endogeneous variables reject the equality of IV estimates computed separately for covariate subgroups. Indeed, for both models, they are lower than 0.01. This violates the proportionality condition embedded in equation 2.6 and thus suggests that changes in peer racial and social background cannot fully explain the changes in enrollment outside the public school sector after the implementation of the integration plans.

Finally, columns (4) to (6) report the IV estimates of models including together each pair of potential peer characteristics. Models with two endogenous variables capture pairs of causal effects at the same time. These models, identified by differences in the two integration plans effect on school composition, allow for the possibility that different sorts of causal effects are reinforcing or offsetting.

These two endogeneous variable models suggest that peer achievement is the relevant peer characteristics that affect students' take-up decision. Including potential peers achievement as an additional regressor flips the signs of the estimates for the proportion of minority students and for the proportion of FRPL students. Moreover, the estimate for minority students loses its statistically significance, while the one for FRPL students becomes only significant at the 10% level. On the other hand, the estimate for potential peer achievement is hardly affected and retain its statistical significance. Potential peers math achievement appears thus to be the dimension of peers characteristics that applicants consider when deciding whether to turn down their match offer.

Table 2.5.4: 2 SLS Estimates of Potential Peers Characteristics on Exit from Public School

	Probability of exiting the public school system					
	(1)	(2)	(3)	(4)	(5)	(6)
Potential peers achievement	-0.27***			-0.31***	-0.39***	
	(0.04)			(0.06)	(0.08)	
Potential proportion minority		0.37***		-0.16		-0.51**
		(0.09)		(0.11)		(0.20)
Potential proportion FRPL			0.40***		-0.25*	0.72***
			(0.07)		(0.14)	(0.17)
First stage F	41.8	18.7	30.6	21.4	21.4	15.4
Overid (DF = 17) p-value	0.43	0.00	0.01	0.45	0.60	0.02
N	309,959	309,959	309,959	309,959	309,959	309,959

Note: This table reports IV estimates of the effects of different peers characteristics on the probability of exiting the public school system for 2015-2019 middle school applicants. Potential peers achievement corresponds to the leave-out mean 5^{th} grade math test score among applicants offered the same school by the match. Potential proportion of minority corresponds to the leave-out share of Black, Hispanic, native American and Multi-racial students among applicants offered the same school by the match. Potential proportion of FRPL corresponds to the leave-out share of FRPL students among applicants offered the same school by the match. Estimates computed by instrumenting potential peers characteristics by district 3 and district 15 integration plans plus interactions with covariates (English learner, race, interactions of low baseline test scores with FRPL eligibility). The table also reports first-stage F-statistics and overidentification test p-values. Robust standard errors are reported in parentheses; * significant at 10%; **significant at 5%; *** significant at 1%.

Summing up, District 3 and district 15 integration plans were associated with White and non-FRPL enrollment losses. White and non-FRPL applicants were more likely to turn down their match offer because they were assigned on average to schools with lower achieving potential peers after the integration plans. These changes in offer take-up offset part of the impact of the integration plans on school segregation. According to figure 2, changes in offer take-up reduced the decrease in economic segregation by 25 p.p. and the decrease in racial segregation by 15 p.p in both districts. In other words, White and non-FRPL students' exit from the public system halved the effect of the integration plans.

2.5.3 Effects on applicants' submitted preferences

At the enrollment stage, applicants' behavioral response to the integration plans offset part of their effect on school diversity. Was there also a behavioral response of applicants at the assignment stage? Specifically, did applicants modify the ranked ordered list they submit to the mechanism in response to the two integration plans? And, if they did so, did that amplify or diminish the impact of the integration plans?

To answer these questions, it is important to first understand why applicants might modify their ranked order lists in response to the changes in admission criteria entailed by the district 3 and district 15 reforms. I assume that when submitting their list to the centralized admission system, applicants are trying to obtain the best choice possible according to their preferences. This implies that applicants choose the list that gives them the highest expected utility, as a function of the utility they derive from being assigned to different schools.

Formally, let u_{is} denotes the utility that applicant i derived from being assigned to school s. Schools are indexed in $\{0, ..., S\}$ where school 0 corresponds to being unassigned. Applicant i has probability $a_{is}(\ell_i, \hat{p}_{is})$ to be assigned to school s. This school specific probability of admission depends on the ordering of schools in applicant i ranked order list ℓ_i. Indeed, since DA considers applicants' choices in order, listing a school lower in the list decreases mechanically the probability of assignment as the applicant might obtain a higher ranked choice. Moreover, a school that is not ranked has a null probability of admission. The probability of admission to school s also depends on $\hat{p}_{is}$, applicant's i subjective probability of qualifying at s, which itself depends on subjective beliefs about others applicants ranked order list $\hat{\ell}_{-i}$ and about schools ranking of applicants $\{r_s\}_{s=1}^{S}$. Finally, applicant i might incur some cost of application $C(\ell_i)$.

Given this notation, I assume that applicant i chooses her ranked order list ℓ_i to maximize the expected utility from her list minus the list's cost:

$$\max_{\ell_i} U_i(\ell_i) = \sum_{s=0}^{S} u_{is} \times a_{is}(\ell_i, \hat{p}_{is}) - C(\ell_i)$$

Given this maximization problem, Fack et al. (2019) shows that if $C(\ell) = 0$ then DA is strategy proof and all applicants should rank their school in order of preferences. On the other hand, if $C(\ell) \neq 0$ for some list ℓ then applicants will take into account their probabilities of assignment at each school a_{is} when forming their lists.

By changing students' ranking $\{r_s\}_{s=1}^{S}$ at each district school, district 3 and district 15 integration plans affect applicants' ranked order lists through two different channels. First, the plans may change the utility u_{is} each applicant derived from being assigned to each school since the characteristics of students enrolling in each school might change. Second, the plans change the probability of assignment to each school for each applicant a_{is}, since they affect applicants probability of qualifying at each school p_{is} as well as potentially other applicants' ranked order lists ℓ_{-i}. This second channel only matters if applicants take into account probabilities of admission when forming their list, that is if $C(\ell) \neq 0$.

Table 2.4.2 summarized the effects of district 3 and district 15 integration plans on admission probabilities for applicants depending on FRPL status and baseline test scores. Both integration plans increased admission odds at more competitive programs for FRPL applicants with low baseline test scores, while they decreased admission odds for non-FRPL applicants with high baseline test scores. The changes in admission probabilities for the two remaining groups of applicants are more ambiguous. Given the increase in their admission odds, one should expect that FRPL applicants with low baseline test scores are more likely to rank more competitive schools after the implementation of the integration plans.

To provide with some suggestive evidence on the changes in competitiveness of schools ranked by different applicants, Table 2.5.5 reports mean baseline achievement, a proxy for a school desirability, for schools ranked first on their list by different groups of applicants before and after the implementation of the integration plans.[13] Specifically, the mean baseline achievement of each school is computed using the 5^{th} grade math test score of students enrolled in the school for years prior to 2019, so that it captures school competitiveness prior to the implementation of the integration plans. Applicants are sorted in four distinct groups depending on their FRPL status and whether they had a 4^{th} grade math test score below or above the median test score in the district.[14]

The first panel of Table 2.5.5 shows the average mean baseline math test score of the school ranked first by the four different groups of applicants in district 3, district 15 and other districts. In all districts, students with lower baseline test scores tend to rank first less competitive schools than students with the same economic background but with higher baseline test scores. Moreover, students with similar baseline achievement list first more competitive schools if their family have a higher income. Finally, district 3 and district 15 applicants from each group rank first more competitive schools than their counterparts in other districts on average, which is consistent with the fact that students in both districts have higher math and ELA test scores than the average NYC student.

Panel B of Table 2.5.5 reports the change in competitiveness of the school applicants list first in 2019 and 2020 with respect to previous years. Consistent with the increase in admission odds, FRPL applicants with low baseline test scores rank first more competitive schools than in previous years in both district 3 and district 15. On average, the schools they rank first enrolled students with respectively 0.09 and 0.07 larger baseline math test scores. These changes are statistically significant at the 1 percentage level.

[13] I only consider the school ranked first to avoid the selection bias that would arise if the plans also affected the decision of ranking more choices. All applicants rank at least one school and thus the plans do not affect this decision.

[14] 4^{th} grade math test scores are used to define low and high baseline students as admission at most screened school is in part based on 4^{th} scores and math test scores are missing for fewer applicants than English test scores.

Table 2.5.5: Estimates of Competitiveness of Applicants' First Choices

		Mean math score in school ranked first		
		District 3	District 15	Others
		(1)	(2)	(3)
Panel A: Mean effect for 2015-2020 applicants				
FRPL	low baseline	0.00	0.24***	-0.060***
		(0.02)	(0.01)	(0.002)
	High baseline	0.56***	0.69***	0.22***
		(0.03)	(0.01)	(0.00)
Non-FRPL	low baseline	0.41***	0.48***	0.19***
		(0.02)	(0.01)	(0.00)
	High baseline	0.94***	0.87***	0.56***
		(0.01)	(0.01)	(0.00)
Panel B: Additional effect for 2019-2020 applicants				
FRPL	Low baseline	0.09***	0.07***	0.02***
		(0.03)	(0.02)	(0.00)
	High baseline	0.06	0.01	0.04***
		(0.05)	(0.01)	(0.00)
Non-FRPL	Low baseline	0.07*	0.04*	0.05***
		(0.04)	(0.02)	(0.01)
	High baseline	0.01	-0.10***	0.03***
		(0.01)	(0.01)	(0.01)
N		5,995	13,432	349,453

Note: This table reports the competitiveness of the first choice ranked for four groups of applicants. The competitiveness of the first choice corresponds to 5^{th} grade math test score of students enrolled prior to 2019 in the school ranked first. Applicants are sorted in four distinct groups depending on their FRPL status and whether they had a 4^{th} grade math test score below or above the median grade in their district of residence. Panel A reports the mean competitiveness of the first choice for each group of 2015-2019 applicants. Panel B reports the change in competitiveness of the first choice after the implementation of the integration plans in 2019 for each group of applicants. Column (1) and (2) are restricted to applicants residing in district 3 and district 15 respectively. Column (3) includes applicants residing in any other NYC district. Robust standard errors are reported in parentheses; * significant at 10%; **significant at 5%;*** significant at 1%.

On the other hand, the integration plans did not induce other groups of applicants to increase the competitiveness of their first choice. Non-FRPL applicants with low baseline test scores did list slightly more competitive schools after the implementation of the plans, but the increase is only marginally statistically significant. While, in district 15, non-FRPL applicants with high baseline test scores rank first schools that enrolled students with 0.10 smaller baseline math test scores.

For comparison, the competitiveness of applicants' first choices did not change heterogeneously for the different groups of applicants in other NYC districts. All applicants list more competitive schools as first choice in 2019 and 2020 compared to previous years, regardless of their FRPL status and their baseline achievement. Heterogeneous changes in first choice competitiveness were thus specific to the two NYC districts for which admission probabilities also changed heterogeneously for different groups of applicants. These differences in patterns suggest that applicants did adapt their first choices to the changes in admission probabilities entailed by the integration plans.

As additional evidence that applicants changed their ranked order lists in response to the integration plans, I consider the effect of the plans on the length of the ranked order lists and probability of being unassigned at the end of the main round of DA.[15] While students' admission odds were negatively or positively affected by the plans depending on their demographics and baseline achievement, the plans resulted in increased uncertainty in admission for all applicants, especially in district 15. Indeed, prior to the plans, students could evaluate their odds of admission at different programs depending on their test scores. Applicants with higher elementary test scores could suppose they had a good chance to qualify at competitive screened programs that ranked their applicants based on prior academic achievement, while applicants with lower grades could presume that their chances were lower. By replacing screened admissions with lotteries, district 15 made applicants more uncertain about their admission chances as admissions became random and not based on students' background. Similarly, the district 3 plan created seat reserves for lower achieving FRPL students that increased the chances of students that were less likely to qualify

[15] Applicants who remain unassigned after the DA algorithm has been run are manually placed into a school by the DOE.

Table 2.5.6: D-in-D Estimates of Changes in Applicants' Ranked Order Lists

	All	FRPL		non-FRPL	
		Low baseline	High baseline	Low baseline	High baseline
	(1)	(2)	(3)	(4)	(5)
	Panel A: Length of the rank-ordered list submitted				
D3 × 2019	1.01***	0.75***	1.20***	0.63**	1.086***
	(0.09)	(0.20)	(0.25)	(0.27)	(0.12)
D3 × 2020	1.20***	0.47**	0.99***	0.89***	1.61***
	(0.09)	(0.19)	(0.30)	(0.32)	(0.13)
D15 × 2019	2.56***	2.07***	1.94***	2.88***	3.45***
	(0.07)	(0.13)	(0.15)	(0.23)	(0.12)
D15 × 2020	2.84***	1.58***	1.63***	3.92***	4.61***
	(0.07)	(0.12)	(0.14)	(0.22)	(0.12)
mean D3 pre-2019	4.0	4.5	4.6	4.2	3.5
mean D15 pre-2019	4.9	4.3	5.2	4.7	5.3
	Panel B: Probability of being unassigned				
D3 × 2019	0.01*	0.01	-0.02	0.02	0.00
	(0.01)	(0.01)	(0.03)	(0.03)	(0.01)
D3 × 2020	0.02***	0.00	0.01	0.05*	0.01
	(0.01)	(0.01)	(0.03)	(0.03)	(0.01)
D15 × 2019	-0.00	-0.01	-0.02	-0.02**	0.04***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
D15 × 2020	-0.01	0.01	-0.02	-0.03***	0.01
	(0.00)	(0.01)	(0.02)	(0.01)	(0.01)
mean D3 pre-2019	0.06	0.05	0.08	0.06	0.05
mean D15 pre-2019	0.07	0.07	0.10	0.07	0.05
N	397,744	156,576	110,242	37,183	64,870

Note: This table reports differences-in-differences estimates of integration plan effects for 2015-2020 middle school applicants. In all panels, the endogeneous variables are the interaction of dummies for residing in district 3 and district 15 and applying for admission in 2019 and 2020. Panel A dependent variable is a count variable that indicates the length of applicants' ranked order lists. Panel B dependent variable is a dummy that is equal to one for applicants who are unassigned by the algorithm and have to be manually placed. The last two rows of each panel report the mean of the dependent variable among 2015-2018 applicants. All models control for year and district fixed effects. Robust standard errors on year are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. .

and decreased the chance of students that were more likely to qualify. Hence, more students had a more uncertain probability of admission after the district 3 plan as well. In response to the increase in admission uncertainty, one would expect students to list more choices to avoid being unassigned by the mechanism and manually placed to a school they might not like.

Using the same differences-in-differences specification as in equation 2.3, Table 2.5.6 explores the consequences of the integration plans on the length of the list submitted in panel A and the probability of being unassigned in panel B. Results are reported separately for the two cohorts of applicants that applied, after the adoption of the reforms, for admissions in the fall of 2019 and in the fall of 2020 respectively. One should expect some adjustment between the two years as the effects of the plans on the admission probabilities for different applicants became observable. Column (1) reports the effects of the plans for all applicants while columns (2) to (5) report them for the same four distinct groups of applicants considered in Table 2.5.5.

The integration plans resulted in an increase in the number of choices listed by all applicants in both districts. The increase was more substantial for applicants in district 15, consistent with a larger increase in admission uncertainty. On average, district 3 applicants listed 1 additional choice in 2019 and 1.2 additional choices in 2020, while district 15 applicants listed 2.6 additional choices in 2019 and 2.8 additional choices in 2020. All these changes are significant at the 1% level [16]

Despite their longer ranked order lists, applicants were slightly more likely to be unassigned after the implementation of the integration plan in district 3. The share of applicants manually places was 1% and 2% larger in 2019 and 2020 respectively. On the other hand, the increase in ranked order lists adequately mitigated the increase in uncertainty in district 15 as the share of unassigned students remained stable in both years after the reform.

[16] As a robustness check, appendix figures A2 and A3 display event study graphs for the change in ranked order list length after 2019 in all NYC district. District 3 and district 15 expansions of ranked order list length in 2019 and 2020 are clear outliers when compared to other NYC districts.

The heterogeneity in effects of the plans for applicants with different baseline achievement and FRPL status in columns (2) to (5) is consistent with applicants listing more choices to reduce the probability of manual placement. In both districts, non FRPL applicants with high baseline test scores increased the length of their lists the most. These applicants were the ones whose probability of admission were the most affected by both reforms as they were not eligible for reserved seats and had high admission probabilities at more selective screened programs prior to 2019. On the other hand, FRPL applicants with low baseline test scores benefit the most from the reform in terms of increased admission odds and were also the group of applicants that added fewer schools to their lists. Finally, the small and mostly insignificant changes in the probability of manual placement across all columns suggest that each group of students adjusted the length of their ranked order lists in proportion to the change in admission uncertainty induced by the integration plans. The only exception is for district 15 non-FRPL students with low baseline test scores whose probability of being unassigned was marginally smaller in 2019 and 2020 than in previous years.

Taken as a whole, this reduced form analysis provide suggestive evidence that applicants adapted their ranked order list in response to the integration plans. First, it appears that applicants whose admission odds increased because of the changes in admission criteria applied to more competitive schools. Second, applicants included more schools in their lists to compensate for the higher uncertainty in admission. The increase was more pronounced for applicants for which it became harder to predict their assignment after the implementation of the plans. It is plausible that the integration plans affected other dimensions of applicants' ranked order lists in both districts. Nonetheless, this partial evidence still indicates that changes in application behavior might have impacted the effect of the integration plans.

As a final step, I assess the extent to which changes in ranked order lists affected the impact of the integration plans on segregation. This can be evaluated by estimating what would have been the plans' effects on diversity at the match offer stage if applicants did not change their application behavior. I use 2018 data to simulate match offers under district 3 and district 15 plans, but in absence of any change in applicants' ranked order lists. The differences in segregation between the simulated 2018 match and the real 2018 match capture the mechanical effects of the changes in admission criteria. The impact of applicants' behavioral response can be estimated by subtracting these mechanical effects from the observed declines in isolation indexes between 2018 and 2019. These calculations are accurate under the assumption that applicants' submitted preferences would have remained stable between 2018 and 2019 if there had been no reforms.

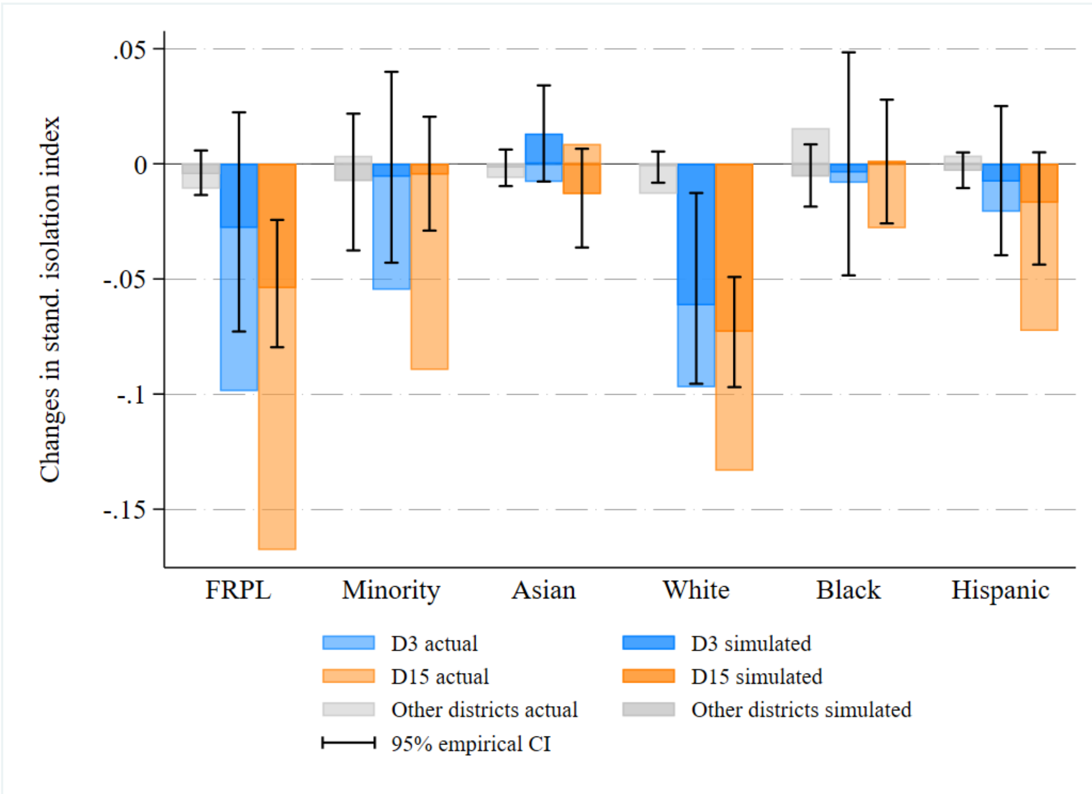
Results from the simulation, as displayed in Figure 2.5.5, establish that changes in students' application behavior explains more than two-thirds of the increase in economic and racial diversity at the match offer stage. The decline in FRPL and minority student standardized isolation indexes would have been more than three times smaller if applicants had not adapt their ranked order lists to the integration plans. Applicants' behavioral response appears to be similarly important for the decrease in White and Hispanic students' isolation.

This difference in the plans impact due to applicants' behavioral response is unlikely to be due to chance. Indeed, the actual effects of both integration plans are not comprised in the 95% confidence interval around the simulated plans effect with no behavioral response. These empirical confidence intervals accounts for the variability in standardized isolation indexes that may arise from differences in applicants' population and tiebreakers sequence. They are generated by simulating 100 times the 2018 match under the two integration plans redrawing each time a sample of applicants and a sequence of tie-breakers.[17]

Summing up, applicants did change their rank order lists in response to the integration plans. This behavioral response at the assignment stage reinforced the effects of the plans on diversity, they made the plans three times more effective. Hence,

[17] Applicants are sampled with replacement from each district independently. SWD and non-SWD applicants are also sampled separately as they participate in two distinct matches.

Figure 2.5.5: Decomposition of Changes in stand. school isolation index



Note: This bar chart plots the changes in standardized school isolation indexes at the match offers stage between 2018 and 2019. Each bar gives the observed change in the index between 2018 and 2019. The shaded part of each bar corresponds to the effect of the integration policies absent behavioral response, i.e. using 2018 applicants' ranked order list. A manual placement round is run after DA in which applicants ranked all available schools by distance and schools rank applicants by tie-breaker. The 95% confidence intervals for the effect of the integration plans absent behavioral response are computed by simulating 100 the 2018 match under integration plans. For each simulation, a new sequence of tiebreakers is drawn and applicants are sampled with replacement, through sampling stratified by SPED indicator and district.

not accounting for applicants' response to changes in admission criteria would be misleading as it might underestimate the impact of these changes.

2.6 Conclusion

The integration plans of District 3 and District 15 aimed at increasing school diversity by lowering barriers to admission at most selective schools to low income students with lower baseline achievement. While the district 15 plan successfully decreased economic and racial segregation at the district's schools, the district 3 plan was less successful. Despite their differences in final impact, both integration plans elicited important behavioral responses from applicants. Applicants responded to the integration plans both by changing the ranked order lists they submit to the mechanism during the assignment stage and by selectively taking-up their match offer during the enrollment stage.

The change in take-up of match offers in response to the integration plans mitigated their impact on diversity. Indeed, White applicants and non-FRPL applicants were much more likely to turn down their match offer and enroll outside the public school system after the implementation of the plans. As a result of these white student and high income students enrollment losses, the integration plans' effects on racial and economic segregation were halved in both districts.

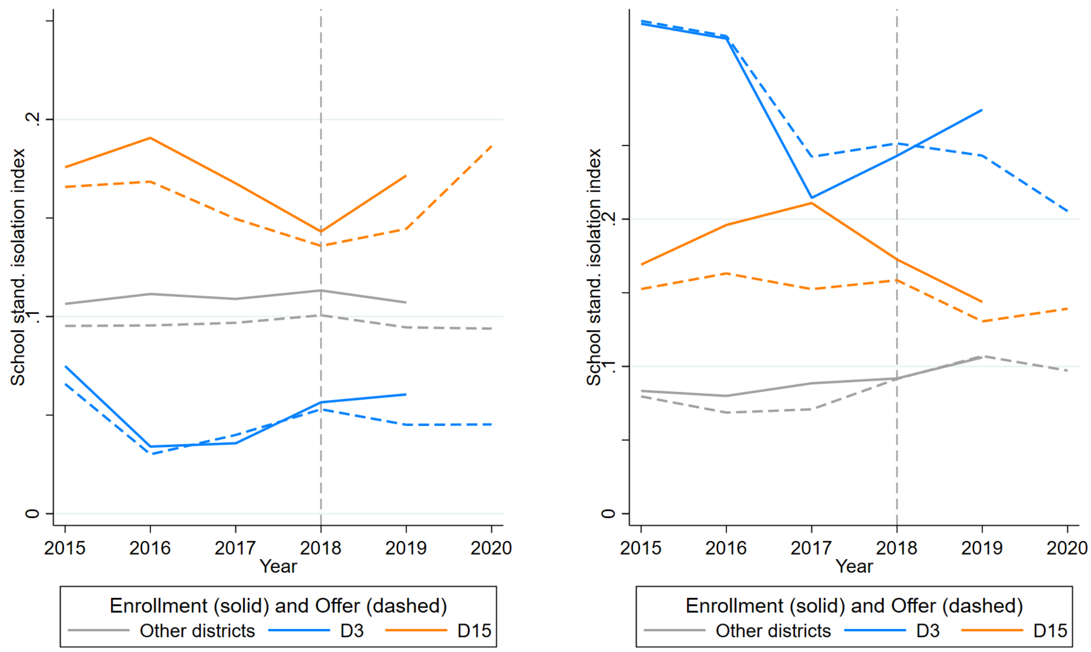
These changes in match offers' take-up rates can be explained by the changes in achievement of potential peers. White and non-FRPL applicants were more likely to turn down their match offer because they were assigned on average to schools with lower achieving potential peers after the integration plans' implementation. This decrease in potential peers achievement resulted from the direct effect of the plans on peers composition as well as the decline in admission probabilities at most selective schools for White and non-FRPL students.

On the other hand, the change in applicants' ranked order lists in response to the integration plans reinforced their effects on diversity. The changes in admission criteria in both districts resulted in twice as much diverse match offers thanks to applicants' behavioral response at the assignment stage. Given the multi-dimensionality of ranked order lists, it is not possible to understand which implications of the integration plans matter the most for applicants' response without modelling the formation of a ranked order list. In particular, a model of applicants' formation of ranked order lists would allow to understand the extent to which applicants' responded to the change in admission probabilities and to the anticipated change in school demographics composition.

Appendix

Appendix 2.A Additional Figures

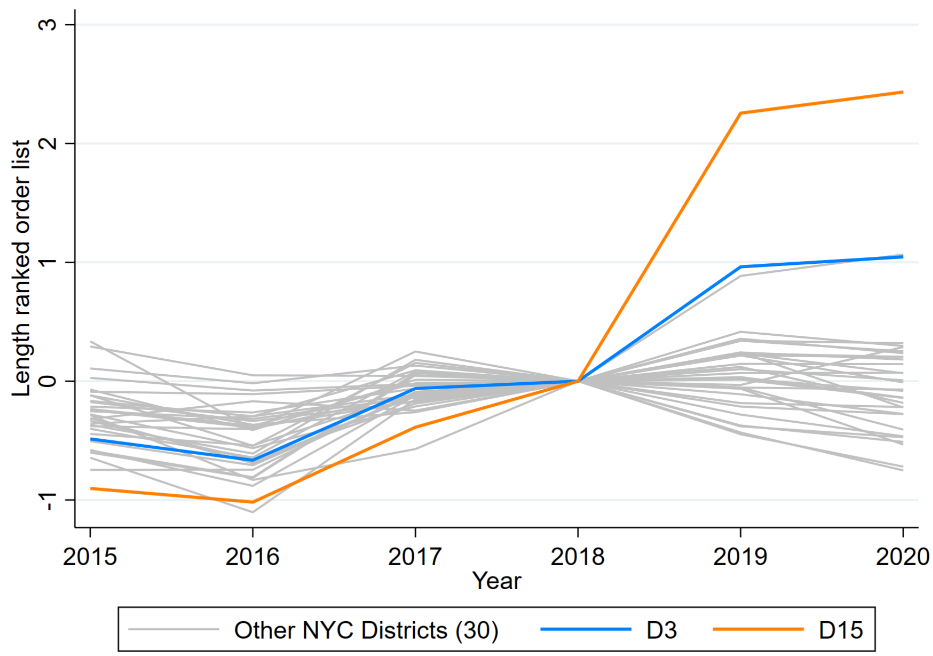
Figure A1: Evolution of Standardized School Isolation Indexes for Asian and Black Students



(a) Asian students (b) Black students

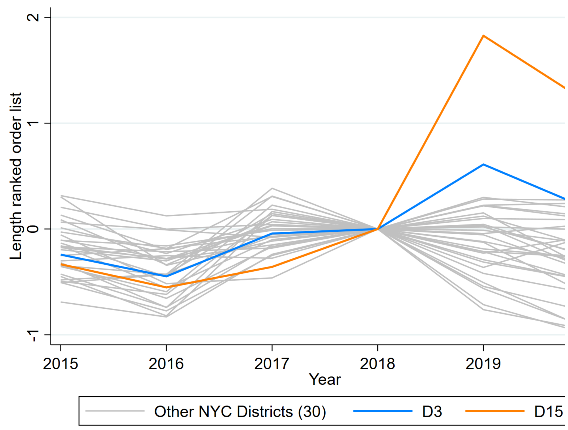
Note: These figures plot the evolution of school standardized isolation indexes for district 3, district 15 and other NYC districts between 2015 and 2020. Panel A displays the standardized index for Asian applicants. Panel B displays the standardized index for Black applicants. Standardized school isolation indexes are computed as in figure 2.5.1.

Figure A2: Comparison of Changes in Ranked Order List Length in all NYC Districts

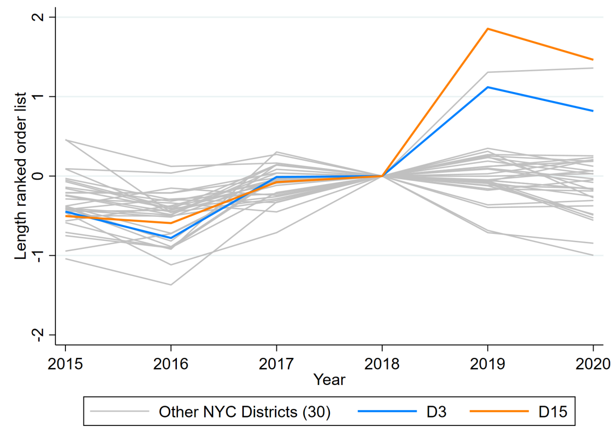


Note: These figures plot changes in the number of programs listed with respect to 2018 for all NYC districts.It includes all applicants with non-missing district of residence.

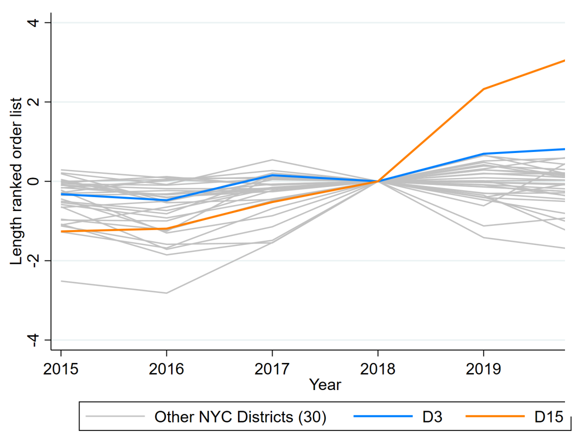
Figure A3: Comparison of Changes in Ranked Order List Length in all NYC Districts



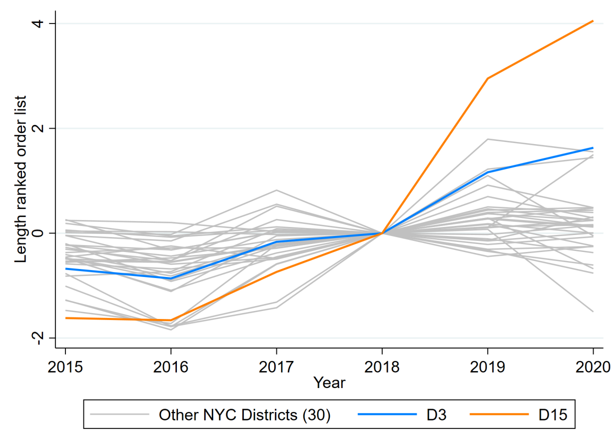
(a) FRPL low baseline test score students



(b) FRPL high baseline test score students



(c) non-FRPL low baseline test score students



(d) non-FRPL high baseline test score students

Note: These figures plot changes in the number of programs listed with respect to 2018 for all NYC districts. Panel A considers the list length for FRPL with low baseline test score applicants, panel B for FRPL with high baseline test score applicants, Panel C for non-FRPL with low baseline test score applicants, and panel D for non-FRPL with low baseline test score. Each plot only includes districts in which at least 10% and no more than 90% of the student population belongs to the group considered. The data is also winsorized by excluding districts that display variations in standardized school isolation index larger than 10 % prior to 2018.

Appendix 2.B Additional Tables

Table A1: Description of NYC Diversity In Admission policies

	Fall of Middle School enrollment year					
	2017		2018		2019 and 2020	
	Programs	Policy	Programs	Policy	Programs	Policy
District 3					All	FRPL lowest performing - 10%
						FRPL lower performing - 15%
District 15	15K447	FRPL - 30%	15K447	FRPL - 30%	All	Low income, ELL, temp housing - 52%
	15K839	FRPL - 40%	15K839	FRPL - 40%		No screening of applicants
			15K497	FRPL - 40%		
Other districts						
District 1	01M450	FRPL - 62%	01M450	FRPL - 62%	01M450	FRPL - 62%
District 2	02M114	FRPL - 10%	02M114	FRPL - 10%	02M114	FRPL - 17%
			02M260	FRPL - 17%	02M260	FRPL - 17%
					02M255	FRPL - 17%
					02M422	FRPL - 60%
District 7			07X343	Feeder schools - 40%	07X343	Feeder schools - 40%

This table describes the diversity in admission policies implemented in each school year in NYC between 2017 and 2020. The school year refers to the year in which applicants enrolled in middle school. The diversity in admission policies consisted of reserving a share of seats to which eligible applicants had priority over other applicants. For each year and each program, the policy column indicates which applicants were reserve-eligible and the size of the reserve in seats percentage.

Table A2: Characteristics of district 3 and district 15 students by race

District	Race	N	% FRPL	% ELL	English score	Math score	Reserve eligible
District 3	Asian	196	27%	4%	3.9	4.0	4 %
	Black	360	86%	2%	2.6	2.5	57 %
	Hispanic	621	74%	10%	2.8	2.8	47 %
	White	966	11%	2%	3.8	3.8	3 %
District 15	Asian	1097	84%	29%	3.2	3.4	69 %
	Black	328	74%	0%	2.9	2.6	66 %
	Hispanic	1874	78%	22%	2.8	2.8	74 %
	White	1766	16%	4%	3.7	3.6	15 %

This table describes the characteristics of 2019-2020 applicants residing in district 3 and district 15 by race. English and Math test score corresponds to the proficiency rating at the 4^{th} grade State test. Reserve eligibility is based on an indicator variable included in the 2019 and 2020 assignment files. The sample drops 206 applicants that do not list any school, as the reserve eligibility variable is missing for 2020 applicants with no school listed.

Chapter 3

Heuristic school choice: a simple model of applicants' strategic behavior[1]

3.1 Introduction

In centralized school choice systems, policy makers can influence students' assignment to schools by changing assignment rules and schools' admission criteria. Structural modelling can help forecast the effects of alternative policies and inform policy makers' decisions. The method relies on the estimation of an underlying model of families decision-making and the simulation of equilibrium outcomes under new policies.

The estimation of families preferences uses the rich information embedded in the ranked order lists that families submit to the mechanism. When the assignment mechanism is strategy-proof, historical choice rankings can be considered as corresponding to applicants' underlying preferences for schools. Assuming truthfulness of rankings, Pathak and Shi (2021)'s forecast experiment on Boston new school choice

[1]Thanks to the NYC Department of Education for graciously sharing data. The views in this paper are those of the author and do not necessarily reflect the official policy or position of the NYC Department of Education.

plan presents encouraging evidence for the reliability of these forecasts. Their structural school demand models can accurately predict counterfactual outcomes, when conditioning on the actual applicants' characteristics.

Nonetheless, the assumption that rank order lists reflect underlying preferences for schools is not always realistic. While this assumption is naturally violated in mechanisms that are not incentive compatible,[2] it also appears not to hold in mechanisms that are usually considered strategy-proof, like deferred acceptance (DA). For example, Rees-Jones and Skowronek (2018) present experimental evidence that medical students misrepresent their preferences in the National Resident Matching Program. Moreover, Artemov et al. (2017) and Larroucau and Rios (2020) find evidence consistent with students omitting programs to which they're unlikely to be admitted. Finally, Fack et al. (2019) analyze the high-school system in Paris, and reject strategy-proofness in favor of stability.[3]

The relaxation of the truthfulness of rankings assumption complicates the estimation of preferences. A recent approach models ranked order list as solving an expected utility maximization problem given students' beliefs over admission probabilities. Agarwal and Somaini (2018) propose a general methodology to estimate preferences based on the optimality of the chosen raked order list, whenever the mechanism can be represented with a cutoff structure. Nonetheless, a major challenge to implement this methodology is to avoid running into the curse of dimensionality, as the number of potential ranked order lists grows exponentially with the number of schools and the method requires checking the optimality of applicant's choice among all potential lists. To circumvent this dimensionality problem, Larroucau and Rios (2020) and Calsamiglia et al. (2020) propose strategies to limit the choice space considered for estimation.

[2] Agarwal and Somaini (2018), Kapor et al. (2020) and Calsamiglia et al. (2020) show that students engage in strategic behaviors in assignment systems that use the Boston mechanism.

[3] Relatedly, Luflade (2017) estimates the value of giving information to students on their probabilities of admission in deferred acceptance mechanisms, when the length of ranked order list is constrained and students cannot always express their true preferences.

In this paper, I offer an alternative approach to estimate applicants' preferences, when applicants optimize their choice of ranked order list considering probability of admissions. I assume that applicants use a heuristic method to choose their list, instead of solving a one-shot utility maximization problem that requires considering all possible ranked order lists.[4] In my model, applicants choose their list sequentially with a limited understanding on the dynamic consequences of each choice.

Specifically, applicants consider first which school to place in position 1 of their lists, then they turn to filling in position 2 etc. Applicants do understand that the probability of being assigned to non-listed schools goes down as they go through their list. Nonetheless, they do not understand that they might later decide not to list some of the schools they had considered for earlier positions, as the probability of assignment becomes negligible. My model is consistent with two empirical regularities: (1) applicants do not list schools for which their probability of admission is small (2) applicants stop their list if they are certain of being assigned to one of their first choices.

This model of application behavior implies a series of inequalities on applicants indirect utilities for schools. These inequalities allow to estimate applicants' preferences without running into the curse of dimensionality. The estimation procedure follows the two-step estimation procedure from Agarwal and Somaini (2018). In a first step, I estimate the probability of admission at each school for each applicant depending on their baseline score. In a second step, I estimate the parameters governing the distribution of preferences through a Gibbs sampler adapted from McCulloch and Rossi (1994), given the first step estimated admission probabilities. The Gibbs sampler leverages the set of restrictions on indirect utilities derived from applicants' choice model.

To illustrate the method, I leverage an admission reform in one of NYC school

[4]Ajayi and Sidibe (2020) and Son (2020) also consider a model of applicants' list formation in which applicants only consider a subset of programs. However, in their models, strategic consideration based on admission probabilities do not affect which schools are ranked.

districts to estimate preferences for middle school characteristics. Chapter 2 of this dissertation found that the reform entailed substantial changes in admission probabilities to which students responded by changing their ranked order lists.

Preliminary results show that relaxing the assumption of truthfulness of ranked order lists substantially affects the estimates of applicants' preferences. Compared to a model that assumes truthful ranked order lists, my model's estimates reveal stronger preferences for high achieving peers. This finding is consistent with students omitting more competitive programs for which they believe their admission odds are small before the reform. Indeed, under the truthfullness assumption, the omission of competitive schools, which also enroll high achieving students on average, results in low estimated taste for high achieving peers. The difference between models arises even though most students appear to adopt safe listing strategy, omitting only programs for which their admission odds are extremely low.

The rest of the paper is organized as follows. Section 2.2 describes the NYC admission reform exploited for estimation of the model. Section 3.3 presents the model of applicants' list formation. Section 3.4 discusses the variations that identify the model and describes the method to estimated the model. Section 3.5 presents preliminary results from the estimation of preferences for school characteristics in the NYC middle school match. Section 3.6 concludes.

3.2 NYC admission reform

3.2.1 The NYC middle school match

Each year, NYC uses DA to assign 5^{th} graders to middle schools. The match is run at the academic program level, with each school being able to operate more than one program. Although the city is split into 32 local school districts, applicants may apply to schools in any district. Nonetheless, many programs restrict eligibility to only applicants residing in the district or enrolled in a district elementary school.

Moreover, most borough-wide and city-wide programs also give priority to district applicants. As a result, 85% of students are assigned to a middle school in their district of residence.

Each academic program has specific eligibility, priorities and admission criteria. Based on the factors used to rank applicants, programs may be classified in two broad categories: screened and unscreened. While both types of programs may give priority to groups of applicants, only screened programs can rank individual applicants based on test scores and past behavior. On the other hand, lottery programs can only break ties between applicants with the same priority using a unique tiebreaker. Priorities are related to applicants' residential zones, school at the time of application and, rarely, attendance to an information session. Screening criteria are based on prior grades, standardized test scores, talent tests scores and attendance at the previous school.

Applicants can rank up to 12 programs but only 5% do so. To help applicants in their choice, the NYC Department of Education provides applicants with a paper admission guide and a personalised application website. This application website only shows programs for which the applicant is eligible for. Each program page includes the admission priorities and the admission criteria that the program uses to rank applicants, if any. In addition, the page indicates the number of applicants per seat and the priority of last admitted applicant. Together with information from their school counselor, the application website can help applicants know their admission chances.

3.2.2 Northwest Brooklyn admission reform

In 2019, two NYC local school districts started implementing diversity in admission plans. These plans aimed at increasing economic and racial diversity at districts' middle schools by changing admission criteria. Of the two districts, the west Brooklyn district adopted the most far-reaching plan. West Brooklyn's plan eliminated screening criteria at all middle schools. In addition, the district set aside 52 percent of the seats in each school for disadvantaged students.[5] Prior to the adoption of the plan, 11 out of the 13 programs in the district were screened, thus the plan substantially affected applicants' admission chances at the schools they were more likely to attend.

The plan had heterogeneous effects on applicants' admission probabilities. The effect of the plan on admission odds depended on applicants' poverty status and baseline test scores. Indeed, the reform decreases admission odds for applicants with high baseline test scores that were not reserve-eligible. On the other hand, reserve-eligible applicants with low baseline test scores saw an increase in their probability of admission at the most selective school after the plan.

In Chapter 2 of my thesis, I show that applicants responded to these changes in admission odds by adapting the ranked order list they submit to the mechanism. In particular, applicants who faced the largest decrease in admission odds were more likely to include more programs in their ranked order list. Table 3.2.1 illustrates this finding by showing the mean rank of the choice obtained and the length of the ranked order list for applicants conditional on their reserve-eligibility status and their past achievement. On average, applicants who obtained a lower-ranked choice after the reform, reflecting their decrease in admission odds, increased the number of programs they ranked the most. Overall, this suggests that ranked order lists were specific to the admission regime and that applicants take into account their admission odds.

3.3 A model of list formation

I consider a school choice mechanism in which students are indexed by $i \in 1, .., n$ and schools are indexed by $s \in \{1, ..., S\}$. School 0 denotes being unmatched and manually placed into a school by the central mechanism. Each school has q_s seats, with $q_0 = \infty$ and $\sum_{s=1}^{S} q_s \geq n$. I now describe how students are assigned to these

[5] Students are reserve-eligible if they are either low-income (eligible for free or reduced-price lunch (FRPL)), English language learners, or homeless.

Table 3.2.1: Effects of the admission reform

		2018 (1)	2019 (2)	2020 (3)
Panel A: Rank of match offer				
Reserve-eligible	Low baseline	1.9	1.4	1.4
	High baseline	2.0	1.9	1.8
Non-Reserve-eligible	Low baseline	2.5	3.2	3.4
	High baseline	2.3	4.1	4.5
Panel B: Number of schools listed				
Reserve-eligible	low baseline	4.4	6.2	5.8
	High baseline	4.8	6.7	6.6
Non-Reserve-eligible	low baseline	5.6	7.9	8.9
	High baseline	6.4	9.3	10.4
N		1522	1328	1211

Note: This table reports the effect of the admission reform for four groups of applicants. The sample includes all non-SWD applicants residing and attending an elementary school in district 15 during years 2018-2020. Applicants are sorted in four distinct groups depending on their FRPL status and whether they had a 4^{th} grade math test score below or above the median grade of their cohort. Each column corresponds to an application year. Panel A reports the mean rank in applicants' lists of the school they are offered in the match. Panel B reports the mean number of schools listed by each group of applicants.

seats, the choice problem faced by students, students' preferences over schools and their beliefs.

3.3.1 Assignment mechanism: deferred acceptance (DA)

Centralized school choice mechanisms reconcile student and schools preferences to generate a single school assignment for each student. In this paper, I focus on the student-proposing deferred acceptance (DA) algorithm. Student-proposing DA has been adopted in many school districts around the world[6] because of its attractive the-

[6] Among others: Amsterdam, Boston, New York City, Chicago and Paris are assigning students using deferred acceptance (see Tables 1 of Pathak and Sönmez (2013) and Agarwal and Somaini (2018))

oretical properties. In particular, the mechanism produces a student-optimal stable match and is strategy-proof when students do not face application cost (Abdulkadiroğlu and Sönmez (2003)).

The mechanism is based on three inputs: students' rankings of school, schools' rankings of students and schools' capacities. First, students submit a ranked order list of schools to the mechanism. The length of this list may be capped in some applications of the mechanism. Second, schools rank all their applicants. These rankings are strict and may be school specific or common to all schools. Finally, school capacities are entered in the mechanism.

Once all three inputs are inputted into the mechanism, DA works as follows:

- Step 1: Each student proposes to her first choice. Schools' seats are assigned tentatively to proposers one at a time, following their rank. Students are rejected if no seats are available at the time of consideration.

- Step $k > 1$: Each student who was rejected in the previous step proposes to her next best school. Each school considers the students tentatively assigned in previous steps together with new proposers and tentatively assigns its seats to these students one at a time following the school's ranking. Students are rejected if no seats are available at the time they are considered.

The algorithm terminates either when all students are assigned or when all unassigned students have exhausted their lists of schools.

An important feature of DA is that applicants' ranks at each school are independent of students' ranked order lists. As such, each student probability of admission at each school only depends on other applicants' ranked order lists and schools' ranking of applicants. Thus, applicants may take as given their admission probabilities at each school when forming their ranked order lists.

3.3.2 Optimal portfolio choice problem

The choice of ranked order list by applicants can be viewed as an optimal portfolio choice problem. This framing was first introduced by Chade and Smith (2006) for college application and have been used more recently in other applications of school choice (Fack et al. (2019) and Larroucau and Rios (2020)). The idea behind this framework is that ranked order lists of schools can be mapped to lotteries over schools, whose weights depends both on applicants' admission probabilities and the ordering of schools in the list. Hence, applicants choose their ranked order list to maximize their expected utility, which depends on their preferences over schools, the lottery over schools induced by the ranked order list and the cost of submitting the ranked order list.

Formally, each applicant i chooses her ranked order list ℓ_i to maximize her expected utility

$$\max_{\ell} U_{i\ell} = \sum_{s=1}^{S} u_{is} \times a_{is}(\ell, \hat{p}_{is}) - C_i(\ell) \tag{3.1}$$

The expected utility from each list depends on the utilities $\{u_{is}\}_{s=1}^{S}$ that applicant i derives from assignment to each school s. In the expectation, the weight given to each school is denoted by a_{is}, which captures the reduced form probability applicant i is offered school s. Finally, each list is associated to some cost $C_i(\ell)$ that restricts applicant's choice.

The weight assigned to each school $a_{is}(\ell, \hat{p}_{is})$ is jointly determined by applicant's ranked order list ℓ and applicant's belief about her admission probability $\hat{p}_{is}$. a_{is} depends on applicant's ranked order list since DA attempts to assign each applicant to her highest ranked school. Thus, a_{is} is mechanically increasing in school s' ranking in ℓ_i. On the other hand, applicant i's subjective probability of qualifying at s only depends in DA on his beliefs about others applicants' ranked order lists $\hat{\ell}_{-i}$ and about schools ranking of applicants.

Finally, student i incurs a cost $C_i(\ell)$ when forming their ranked order list. This cost could depend on the number of choice in ℓ and also on the specific schools ranked. In the absence of a monetary cost of application, this could be viewed as capturing any psychological cost that a student might face. For instance, students might need to learn about schools that are not in their neighborhood. Or, they could prefer avoid disappointment and thus face a cost of listing choices to which they have very little chance of being admitted.

Such a cost flexibly covers different applications of DA. When $C(\ell) = 0$ for all ℓ, this model coincides with the traditional setting without costs in which DA is strategy proof (Abdulkadiroğlu and Sönmez (2003)). When applicants do not face any application cost, they are better off listing schools in their true order of preference. Moreover, it also encompasses the constrained DA where applicants cannot rank more than K schools (Haeringer and Klijn (2009)) by setting $C(\ell) = \infty$ when $|\ell| > K$. Finally, when $C(\ell) = c|\ell|$, applicants have to pay an unit cost for each additional school listed and the cost of a list is a linear function of its length.

What does it mean for DA not to be strategy proof with application cost? When the cost depends only on list length ($C(|\ell|)$, Fack et al. (2019) shows that it is a weakly dominated strategy for students to submit a ranked order list that is not a true partial preference order, although DA is not strategy proof. It follows that the ordering of the schools included in a student's list reflects her true preferences. Nonetheless, applicants may omit unlikely choices, i.e. schools for which the subjective probability of admission ($\hat{p}_{is}$) is low. Additionally, they might stop ranking their list if they are sure to get in one of the schools they have already ranked.

3.3.3 Preferences: utilities and cost

I assume that applicant i's indirect utility from assignment into school s is given by

$$u_{is} = U(X_{is}, \xi_s, \epsilon_i) - d_{is} \tag{3.2}$$

$$u_{i0} = 0 \tag{3.3}$$

where X_{is} are observable characteristics of student i and school s, d_{is} is the distance between student i's home and school s, and ζ_s and ϵ_i are unobserved characteristics. Following a common approach in the school choice literature, I assume that $\epsilon_i \perp (d_{i1}, ..., d_{is})$. This assumption is violated if students' residential choice is determined by unobserved components of preferences for schools.

This representation of utility includes both a location normalization and scale normalization. The utility of the outside option is normalized to 0, which normalizes the location of utilities. The coefficient on distance d_{is} is set to -1 which normalizes the scale of utilities.

For the empirical application, I further parametrize the utility as

$$u_{is} = \sum_{k=1}^{K} \beta_k x_{isk} - d_{is} + \varepsilon_{is} \tag{3.4}$$

$\{x_{isk}\}_{k=1}^{K}$ are a set of student-school interacted characteristics, β_k are interaction specific parameter to be estimated, and $\varepsilon_i = (\varepsilon_{i1}, ...\varepsilon_{iS}) \sim N(0, \Sigma_\varepsilon)$ are independent of x, d. In the estimation, I construct x_{isk} by interacting indicators for free and reduced price lunch eligibility, ethnicity, high baseline math test score with the shares of minority students and of students with high math baseline test score in each school.[7]

The cost for applicant i for each ranked order list is assumed to be linear in list length, i.e. $C_i(\ell) = c_i|\ell|$. The cost of adding an additional program to the ranked order list is applicant-specific. For the empirical application, the unit cost is parametrized as

[7] I assume that applicants are able to forecast the composition of the student body at each school. This assumption is small for years preceding the admission reform as During applicants have access to detailed information about the ethnicity representation at each school and the academic results of students currently enrolled at each school. Nonetheless, the assumption is more substantial for the year immediately after the reform as the changes in admission scheme affects the demographic makeup of schools.

$$c_i = c + \zeta_i \tag{3.5}$$

where c is the average unit-cost and $\zeta_i \sim N(0, \sigma_\zeta)$ is applicant i's unobserved deviation from the average cost.

3.3.4 Beliefs: rational expectations

When submitting their ranked order lists to the centralized system, applicants do not only care about the utility they derive from each assignment but also about the probability of being admitted in each program. Admission to each school may not be independent if schools' rankings of applicants are based on a common tiebreaker or the same set of scores. I assume that students hold rational expectations over their admission chances. Applicants are able to infer their admission probabilities at each school based on their own score and the distribution of tiebreakers and scores. Moreover, they understand the dependency between admission probabilities.

3.3.5 Limited rationality assumption

To solve the general student optimal portfolio, one needs to select the best permutation out of the set of schools. In general, this is an NP-hard problem (Chade and Smith (2006)). Indeed, the choice space increases exponentially with the number of schools. For instance, if applicants can list 12 among 60 potential schools as in NYC, they may choose among more than 10^{20} lists. It follows that inferring preferences for schools from students' ranked order lists also runs into the curse of dimensionality. Recent attempts circumvent the problem by limiting the number of lists considered for estimation (Calsamiglia et al. (2020), Larroucau and Rios (2020))

Nonetheless, applicants cannot plausibly choose their ranked order lists by solving such a complex problem. Indeed, it is impossible for applicants to consider all potential lists. Likewise, it is unlikely that applicants actually compare the expected utilities they would get from ranking different sets of the schools they are considering. Applicants most probably use some heuristics criteria to choose the best possible lists that might differ from the optimal portfolio.

I model applicants' heuristics approach by assuming that applicants choose their list sequentially without fully internalizing how their choice affects the continuation value of their list. Specifically, applicants consider first which school to place in position 1 of their lists, then they turn to filling in position 2 etc. Applicants do understand that the probability of being assigned to non-listed schools goes down as they go through their list. Nonetheless, they do not understand that this means that they might decide not to list some of the schools they had considered for early positions, as the probability of assignment becomes negligible. The following assumption spells formally this intuition:

Assumption 3 (Limited Rationality of applicants)

1. *Applicants choose their list sequentially, i.e. for each position k in the list, an applicant chooses $s \in S \cup 0$ to maximize its value function*

$$V_k(s) = u_i(s)p_{i,k}(s) - c_i(s) + \tilde{V}_{k+1}(s) \tag{3.6}$$

$p_{i,k}(s)$ denotes the probability applicant i is offered school s when ranked in position k.

2. *But, applicants do not take into account how s affects the set of choices whose expected utility is larger than its cost at $k+1$, i.e.*

$$\tilde{V}_{k+1}(s) = \max_{s' \in S_k} u_i(s')p_{i,k+1(s)}(s') + \tilde{V}_{k+2}(s, s') \tag{3.7}$$

where $S_k = \{s' \in S \cup 0 : u_i(s')p_{i,k}(s') - c_i(s') > 0\}$.

The first part of the assumption states that students solve the portfolio choice problem sequentially. Under the correct specification of the continuation value (if $\tilde{V}_{k+1}(s) = V_{k+1}(s)$, solving the sequential problem of equation 3.6 is equivalent to solving the one-shot optimization problem of equation 3.1. As such the first part of the assumption is not imposing any constraint on the optimization.

The second part of the assumption specifies the cognitive limit of applicants that affects the optimization. At each step, applicants only consider schools for which $u(s)p_k(s) \geq c(s)$. Applicants take into account the cost of listing an additional school so that they do not list schools whose probability or utility are small. Nonetheless, they do not anticipate that a school that satisfies these conditions at step k might not satisfy at step $k+1$. In other words, applicants believe they can always rank an alternative school further down in their list.[8]

When should such assumption matters? The assumption precludes students to rank earlier a school because it will not be acceptable further down in the list. In practice, this means that the expected utility ($u(.)p_k(.) - c(.)$) of the unranked school when it was available was larger than the expected utility of the chosen school. But, at the same time the utility (u) of the chosen school is larger. The following example outlines such a context.

Example 1 (Effect of limited rationality on choice)
Assume a student has only two acceptable schools for the first position, i.e. two schools for which $u_s p_{s1} > c$. *Let* $u_1 = 10c$ *and* $u_2 = 9c$, $p_1 = 0.89$ *and* $p_2 = 1$. *Assuming admissions at each schools are independent, the student will not list two schools as* $10c \cdot 0 < 9c \cdot 0.11 < c$. *This is true both under full and limited rationality. But, under full rationality, the student lists only school* 2, *while under limited rationality, he only lists school* 1.

[8] This assumption is more realistic when the number of schools is not capped or when applicants typically do not exhaust their list.

3.3.6 Bounds on utility

Under the limited rationality assumption, the ranked order list of each applicant entails relative bounds on $\{u_{is}\}_{s=1}^{N}$, the utilities for each school, and on c_i, the unit-cost of adding a school to the ranked order list. These bounds depend on the admission probabilities. Proposition 3 spell them out:

Proposition 3

Let r_{is} denotes the rank of school s in applicant i ranked order list. If school s is not listed by applicant i, $r_{is} = \infty$. $r_{i0} = \underline{k}$ denotes the rank of the outside option.

Under assumption 3, the limited rationality assumption, the applicant chooses the rank order list that maximizes her utility if and only if the following conditions on applicant's utilities and cost hold:

$$1.\ \textit{For schools s.t. } r_{is} \neq \infty : \tag{3.8}$$

$$u_{is} \geq \frac{c_i}{p_{isk}}. \tag{3.9}$$

$$u_{is} \geq u_{ij} \quad \forall j \textit{ s.t. } r_{ij} \neq \infty \textit{ and } r_{is} < r_{ij} \textit{ and } \forall j \textit{ s.t. } r_{ij} = \infty \textit{ and } u_{ij} > \frac{c_i}{p_{ijk}} \tag{3.10}$$

$$u_{is} \leq u_{ij} \quad \forall j \textit{ s.t. } r_{is} > r_{ij} \tag{3.11}$$

$$2.\ \textit{For schools s.t. } r_{is} = \infty : \tag{3.12}$$

$$u_{is} \leq \frac{c_i}{p_{is\underline{k}}} \tag{3.13}$$

$$u_{is} \leq \max(u_{ij}, \frac{c_i}{p_{isk}}) \quad \forall j \textit{ s.t. } r_{ij} = k \tag{3.14}$$

$$3.\ \textit{For the cost } c_i : \tag{3.15}$$

$$c_i \leq p_{ijk} u_{ijk} \quad \forall j \textit{ s.t. } r_{ij} = k \tag{3.16}$$

$$c_i \geq 0 \tag{3.17}$$

$$c_i \geq u_{ij} p_{ij\underline{k}} \quad \forall j \textit{ s.t. } r_{ij} = \infty \tag{3.18}$$

$$c_i \geq u_{ij} p_{ij\underline{k}} \quad \forall j \textit{ s.t. } r_{ij} = \infty \textit{ and } u_{ij} > u_{ij'} \textit{ where } r_{ij'} = k \tag{3.19}$$

Proof. Proof in Appendix. □

This proposition states that applicant's ranked order list is not only a true partial order but that it also carries information about unlisted schools. If the admission probability of an unlisted school is large enough, it entails that the listed schools have larger utilities. On the other hand, it is not possible to rank the utility of an unlisted school for which admission odds are very low as an applicant could have decided to skip the school for this reason.

In general, this proposition is consistent with applicants omitting unlikely choices at each step when filling their list. A choice might be unlikely at some rank because the unconditional probability of admission at the school p_{is} is low. But, a choice might also become unlikely as the applicant fills her list. Indeed, p_{isk} is decreasing in k so it is possible that p_{isk} is small even when p_{is1} is non-negligible. This second case is consistent with applicants stopping their list when they are sure of being admitted to one of the schools they listed. Finally, these inequalities correspond to the ones implied by truth full reports, when $c = 0$ and DA is a strategy-proof.

This proposition will be key for estimation of preferences. Indeed, it entails that it is possible to construct a vector of school indirect utilities and an unit-cost that is consistent with optimally of the list observed, without considering all possible lists.

3.4 Identification and estimation of preferences and cost

3.4.1 Identification

Most school choice models are identified by a "special regressor" that is additively separable from the utilities. In my setting, distance satisfies this property and can be used as a special regressor for identification. Nonetheless, as the model also includes a cost parameter a second source of identification is required. Indeed, variations in ranked order list for different distance vector D_i, given students' characteristics X_i and admission probabilities P_i, cannot separately identify indirect utilities and cost.

The second source of identification comes from variation in admission probabilities. Fixing D_i and X_i, changes in P_i identifies the distribution of unit-cost, as long as the variation in admission probabilities do not affect the distribution of indirect utilities. Together with the variation in distances, the shifts in admission probabilities faced by similar students allow to separately identify indirect utilities and cost.

Such shifts in admission probabilities may be hard to find in a cross-section as admission probabilities are mostly determined by students' characteristics (e.g. scores). Nonetheless, year-to-year changes in admission criteria generates variation in admission probabilities for similar students. The NYC admission reform provides a perfect setting as it entailed substantial variation in admission odds for different type of applicants.

3.4.2 Estimation

The estimation of the preference parameters, $\theta = (\{\beta_k\}_{k=1}^{K}, \Sigma_\varepsilon, c, \sigma_\zeta)$, follows the two-step method outlined in Agarwal and Somaini (2018). In the first step, I estimate admission probabilities for each applicant at each school. The second step estimates θ taking as given the first step estimates of admission probabilities. If the admission probabilities estimators are consistent and asymptotically normal, the two-step estimator of θ is also consistent and asymptotically normal as the second step is is equivalent to a maximum likelihood estimator.

First step: estimation of admission probabilities

Assuming that applicants hold rational expectations about their admission probabilities, Agarwal and Somaini (2018) shows that a consistent estimator of applicants beliefs $\hat{p_i}s$ can be obtained by bootstrapping students' assignments. This bootstrap procedure captures the uncertainty in the probability of admission both due to the lottery and to the year-to-year variation in the applicants population . Specifically, for a match in which some schools use a lottery number t_s to rank students while other schools rank students based on a score τ_i, the estimation of admission probabilities at each school unfolds as follows.

- for each bootstrap simulation $b = 1, ..., B$:
  - sample with replacement n applicants with their corresponding scores and ranked order list.
  - draw a new lottery number for each applicant
  - run DA to obtain an assignment
  - obtain the lottery number of the last admitted applicant t_s^b for lottery schools and the score τ_s^b of the last admitted applicant for score schools.
- estimate the probability of admission of student i at each lottery school s as

$$\hat{p}_{is} = \frac{1}{B} \sum_{b=1}^{B} t_s^b \tag{3.20}$$

- estimate the probability of admission of student i at each score school s as

$$\hat{p}_{is} = \frac{1}{B} \sum_{b=1}^{B} \mathbb{I}(\tau_i \geq \tau_s^b) \tag{3.21}$$

I assume that admissions' cutoff are approximately independent [9]. Hence, given an applicant's score, her admission at each school are independent events. On the other hand, admissions at lottery schools are not independent events if multiple schools use the same lottery number and lottery numbers are not known to applicant at

[9] This is a plausible assumption if the market becomes large, as variations in cutoff are due to changes in the applicants' population. In the limit, admission cutoffs are constant in DA (Haeringer and Klijn (2009)). In the empirical application, variations in cutoffs are also driven by changes in the discretionary weights schools give to different criteria when deciding students' ranking.

the time of application. Indeed, applicant's rejection at a school carries information on applicant's tiebreaker. As such, the probability of being offered each school is a function of the school position in the ranked order list and of the maximum tiebreaker at the schools ranked before. The list and position specific probability of each lottery school can be computed by taking off the largest cutoff of schools ranked before. Formally, this entails that for each school s that uses a lottery and is ranked k

$$p_{isk} = \max \left[0 \ , \ p_{is} - \max_{r_{ij} < k} p_{ij} \right] \tag{3.22}$$

Second step: estimation of preference parameters

The second step is a maximum likelihood estimation that takes as given the admission probabilities computed in the first step. Specifically, assumption 3.6 implies that the vector of parameters maximize the likelihood that the observed ranked order list solve students' sequential optimization problem:

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmax}} \sum_{i=1}^{n} \log \mathbb{P}(\ell_i = \arg \max_{\ell \in \mathcal{L}} V(\ell) | X_i, D_i, t_i, \tau_i; \theta) \tag{3.23}$$

This likelihood does not have a closed-form solution as the model imposes no restriction on the covariance of indirect utilities from schools. Moreover, the large number of potential ranked order lists makes it practically impossible to use simulated likelihood methods, as the number of draws needed to compute the probability of observing each choice would be extremely large.

To solve these problems, I adapt the Gibbs sampler developed in McCulloch and Rossi (1994) to estimate a discrete choice multinomial probit model. The main difference with their approach is that I modify the constraints on utilities in the data-augmentation step to be consistent with applicants' optimization problem. Specifically, I use the bounds on indirect utilities and cost derived in proposition 3 to pick utilities and unit-cost that are consistent with applicants choosing their ranked order list to solve the sequential optimization problem described in assumption 3.7.

The Gibbs sampler obtains draws of β, c, Σ_ε and σ_ζ from the posterior distribution by constructing a Markhov Chain from any initial set of parameters. The chain is constructed by sampling from the conditional posteriors of the parameters and the utility vectors and unit-cost given the previous draws. The sampler iterates through the following sequence of conditional posteriors:

$$
\begin{aligned}
&\beta^{s+1}|U_i^s, \Sigma_\varepsilon^s \\
&\Sigma_\varepsilon^{s+1}|U_i^s, \beta^{s+1} \\
&c^{s+1}|c_i^s, \sigma_\zeta^s \\
&\sigma_\zeta^{s+1}|c_i^s, c^{s+1} \\
&U_i^{s+1}|\Sigma_\varepsilon^{s+1}, \beta^{s+1}, c_i^s \\
&c_i^{s+1}|\sigma_\zeta^{s+1}, c^{s+1}, U_i^{s+1}
\end{aligned}
$$

The first four steps of the sampler follows McCulloch and Rossi (1994). The two last data augmentation steps are different as the set of constrains differ. Details for the Gibbs sampler can be found in Appendix.

This approach differs from Agarwal and Somaini (2018) as it does not compare potential lists' utilities to pick schools' utility vectors that are consistent with observed choice. This reduces substantially the number of constraints and avoids running into the curse of dimensionality, Indeed, to draw each utility, it is enough to check at most $S+1$ constraints instead of $|\mathcal{L}|$, the cardinality of the set of all potential lists in this case.

3.5 Preliminary results

I illustrate the method by estimating preferences for school characteristics and the unit-cost for adding a program in the NYC middle school match. To identify both set of parameters, I exploit the 2019 admission reform in NYC district 15. I focus on application years 2018, 2019 and 2020 and I use data from applicants both residing and attending an elementary school in district 15 at the time of application.

Table 3.5.1 presents the characteristics of the sample used in the demand estimation. The sample only includes non-SWD 2018-2020 applicants, as SWD applicants are matched separately and were less impacted by the admission reform. Based on programs' eligibility rule, a district 15 applicant may be eligible to 55 different programs, 13 of which located in the district. For simplicity, the sample is further restricted to applicants that only listed programs among these.[10]

The demographic composition of the sample varies year-to-year reflecting the variation in the district population. Nonetheless, the share of students offered a district school that was affected by the admission reform is stable at 90 % over time. The effect of the reform on applicants' alternatives is illustrated by the fall in the share of screened programs between 2018 and 2019. While 50 % of programs available to applicants screened applicants in 2018, less than 33% continued to do so in 2019 and 2020.

I first estimate preferences for the model with a cost of application, which entails that applicants take into consideration admission probability when forming their ranked order lists. The second model I estimate sets the cost of application to zero for all students. Imposing this restriction is equivalent to estimating demand under the assumption that ranked order lists are truthful. The estimates based on the last 50 000 out of 1 million draws are reported in table 3.5.2 for both models. The coefficients' trace plots in appendix figures 3..1 show that the sampler has almost converged by the 1 millionth draw. Initial values and priors used for estimation are specified in Appendix B.

Both models' estimates are consistent with applicants preferring schools enrolling high-achieving peers and White and Asian peers. Nonetheless, high-achieving peers

[10]A few students list programs they are not eligible for according to the program's rules.

Table 3.5.1: Characteristics of sample

	2018 (1)	2019 (2)	2020 (3)
% Asian applicants	0.12	0.11	0.10
% Black applicants	0.07	0.06	0.06
% Hispanic applicants	0.41	0.38	0.35
% White applicants	0.36	0.40	0.43
% FRPL applicants	0.49	0.47	0.44
% ELL applicants	0.10	0.10	0.08
Applicants mean math proficiency	3.2	3.3	3.4
Applicants mean english proficiency	3.2	3.4	3.3
Nb of Applicants	1522	1328	1211
% Applicants offered in-district	0.90	0.91	0.89
Nb of programs	55	55	55
% of screened programs	0.51	0.33	0.31

Note: This table presents the characteristics of the sample for the demand estimation. The sample includes non-SWD applicants residing and attending an elementary school in district 15 during years 2018-2020. The sample is restricted to applicants applying to programs for which applicants in district 15 are eligible. Each column corresponds to an application year. Mean math and English proficiency are computed based on the proficiency level obtained in 4^{th} grade state tests. The number of programs counts all programs for which at least one student in the sample is eligible, according to NYC school directory. Screened programs are defined as programs that ranked applicants not only using the tiebreaker.

are valued twice more than white and Asian peers. In the model allowing strategic reports, applicants are willing to travel 0.55 more miles more to attend a school with 10 percent more high-achieving peers but only 0.25 more miles to attend a school with 10 percent more White and Asians peers.

The estimates also reveal some extent of heterogeneity in preferences for school characteristics. Applicants tend to prefer schools that enroll students more similar to them. As such, higher-achieving applicants are more likely to apply to farther schools that enroll a higher share of high-achieving students. On the other hand, minority applicants value less schools enrolling many White and Asian students.

Estimated preferences for White and Asian peers and for higher-achieving peers

Table 3.5.2: Preference Estimates for Models With and Without Application Cost

		Prog. characteristics x Stud. characteristics	
		Model with cost	Model with no cost
	Share minority		
β_1	Main effect	-2.579 (0.001)	-1.651 (0.001)
β_2	Minority applicants	0.989 (0.001)	0.680 (0.000)
β_3	FRPL applicants	1.728 (0.000)	1.111 (0.001)
β_4	High baseline applicants	-1.007 (0.001)	-1.061 (0.001)
	Share high baseline		
β_5	Main effect	5.545 (0.000)	4.383 (0.000)
β_6	Minority applicants	-0.558 (0.001)	-0.462 (0.000)
β_7	FRPL applicants	-1.272 (0.001)	-1.205 (0.001)
β_8	High baseline applicants	1.048 (0.001)	1.130 (0.001)
Σ_ε	Mean variance of utility error	2.768	2.584
σ_ζ	variance of cost error	0.076	-
c	Avg. cost	0.073 (0.000)	-
	Share w/ cost $c_i < 0.01$	0.865	-
	N	4,061	4,061

Notes: Select demand estimates from Gibbs Sampling. Column 1 shows estimates for a model including an applicant cost. Column 2 shows estimates assuming such cost does not exist. Sample is of middle school applicants to New York City public schools from 2018-2020. Estimates reflect interactions between program share minority, share high baseline and dummies for student-level FRPL, minority, and high-baseline. Standard errors are in parentheses and are calculated using the trailing 50,000 draws.

are respectively 50% and 25% larger in the model with a cost than in the model that assumes truthfulness of ranked order lists. In addition, the heterogeneity in preferences is of about the same magnitude. These two observations are consistent with the omission of unlikely programs affecting the validity of the estimates from the model assuming truthful reports. Indeed, applicants are less likely to list more selective programs that enroll on average higher-achieving students, as their admission odds at these programs tend to be lower. On the other hand, applicants also omit listing less preferred programs as they might be certain of getting into one of the programs they listed first.

The substantial impact of including an application cost arises despite the relatively small estimated mean unit cost. On average, the cost of including one more program in a student's list corresponds to travelling 0.07 mile. Moreover, 87 percent of applicants have an unit-cost lower than 0.01, which suggests that most applicants adopt very safe listing strategy by omitting only very unlikely assignment.

3.6 Conclusion

In this paper, I develop a model of applicants' formation of schools' ranked order lists which the truthfulness assumption. My model assumes applicants bear some cost of application and consider their chances of admission at each school when forming their lists. In my model, applicants follow an heuristic method to choose their list. Applicants fill in their list sequentially with a limited understanding on the dynamic consequences of each choice. Specifically, applicants choose which school to list at each step under the assumption that any omitted school may be ranked further down in their list, without internalizing that such school might not be worth listing later on. This model of applicants' listing behavior is consistent with applicants submitting "short-list" to the mechanism either because they have good admission odds at only few schools or because they have a high probability of being admitted to their most preferred schools.

Besides being maybe a better approximation to applicants' decision making, this model of application behavior facilitates the estimation of applicants' preferences for schools. Indeed, I show that this model implies a series of inequalities that characterize applicants' indirect school utilities and application cost. These inequalities allows to estimate applicants' preferences without running into the curse of dimensionality, following the two-step procedure developed in Agarwal and Somaini (2018).

I illustrate this method by estimating preferences for middle schools in NYC, leveraging an admission reform in one of NYC school district. Preliminary results show that the truthfulness assumption impacts substantially the preference estimates. Compared to a model that considers students' ranked order lists as truthful, my model's estimates entail that applicants preference for White and Asian peers is 50% larger while the preference for high-achieving peers is 25% larger. This finding is consistent with students omitting more competitive programs for which they believe their admission odds are small. Indeed, the omission of competitive schools, which also enroll more high achieving students and more White and Asian students on average, results in low estimated taste for high achieving peers and White and Asian peers under the truthfulness assumption.

Appendix

Appendix A : Proof of proposition 1

I aim at showing that applicant's ranked order list maximizes applicant's utility under limited rationality if and only if the above bounds hold:

1. Bounds for listed programs:

   The following bounds follow straightforwardly from DA and the optimization problem. Indeed, if two programs are ranked, one may swap their ranking to increase the probability of the program with the highest utility:

$$
\begin{aligned}
&u_{is} \geq \frac{c_i}{p_{isk}} \\
&u_{is} \geq u_{ij} \quad \forall j \text{ s.t. } r_{is} < r_{ij} \\
&u_{is} \leq u_{ij} \quad \forall j \text{ s.t. } r_{is} > r_{ij}
\end{aligned}
$$

   Hence, I only need to prove that the following bound holds:

$$u_{is} \geq u_{ij} \quad \forall j \text{ s.t. } r_{ij} = \infty \text{ and } u_{ij} > \frac{c_i}{p_{ijk}}$$

   I drop the subscript i throughout the proof for simplicity.

   Under the limited rationality assumption, at step k, the applicant's optimization problem corresponds to choosing the ranked order list ℓ over S_k given $\{u_j\}_{j \in S_k}$

and $\{p_{jk}\}_{j \in S_k}$ so that to maximize

$$\sum_{j \in S_k} u_j a_j(\ell, p_{jk})$$

Let $u_s = \max u_{j_{j \in S_k}}$ then s must be the first element of ℓ. Indeed, let's assume by contradiction that this is not the case. Then by moving s to the top of the list, the utility of the applicant decreases by:

$$\sum_{j \in S_k : r_j < r_s} \alpha_j(\ell) u_j$$

At the same time, her utility increase by:

$$\sum_{j \in S_k : r_j < r_s} \alpha_j(\ell) u_s$$

where $\alpha_j(\ell)$ is the probability of being above the cutoff for admission at both schools s and j, conditional on not being above the cutoff for any school ranked before. Formally, if we denote A_k to be the event of being above the cutoff for school k, we have

$$\alpha_j(\ell) = Pr\left[A_s \cap A_j \mid \left(\bigcup_{j' \in S_k : r_{j'} < r_j} A_{j'} \right)^c \right].$$

2. Bounds for unlisted programs

$$u_{is} \leq \max(u_{ij}, \frac{c_i}{p_{isk}}) \quad \forall j \text{ s.t. } r_{ij} = k$$
$$u_{is} \leq \frac{c_i}{p_{is\underline{k}}}$$

The first bound follows directly from the lower bounds of listed programs. Indeed an unlisted program must either have an expected utility lower than the cost of adding a program or a lower utility than the listed program at all the ranks of the applicant's list. The second bound follows from the fact that $u_{i0} = 0$ and by considering that the outside option can be viewed as the last program listed by any applicant.

3. Bounds for the cost c_i:

$$c_i \leq p_{ijk} u_{ijk} \quad \forall j \text{ s.t. } r_{ij} = k$$
$$c_i \geq 0$$
$$c_i \geq u_{ij} p_{ij\underline{k}} \quad \forall j \text{ s.t. } r_{ij} = \infty$$
$$c_i \geq u_{ij} p_{ij\underline{k}} \quad \forall j \text{ s.t. } r_{ij} = \infty \text{ and } u_{ij} > u_{ij'} \text{ where } r_{ij'} = k$$

The bounds for the cost follow directly from inverting the utility bounds for all listed or unlisted programs.

Appendix B : Gibbs sampler

Given the school indirect utilities and unit-cost parameterization:

$$u_{is} = \sum_{k=1}^{K} \beta_k x_{isk} - d_{is} + \varepsilon_{is} \tag{3.24}$$

$$c_i = c + \zeta_i \tag{3.25}$$

The vector of parameters to be estimated is the following $\theta = (\{\beta_k\}_{k=1}^{K}, \Sigma_\epsilon, c, \sigma_\zeta)$. This parameters' vector is estimated through data augmentation using the following Gibbs sampler:

1. Initiate the sampler with vectors U_i^0, cost c_i^0 and priors $\beta^0 \sim N(\mu^0, V^0)$, Σ_ϵ^0, $c^0 \sim TN(\mu_c^0, \sigma_c^0, 0, \infty)$, σ_ζ^0 .

2. Sample β^1 given Σ_ϵ^0, and U^0 from $N(\mu^1, V^1)$.

   - Compute the diagonal matrix C from the Cholesky decomposition of $\Sigma_\epsilon^0 = CC'$
   - Compute $X_i^* = C'X_i$ and $R_i^* = C'(U_i + D_i)$
   - Compute $V^1 = (X^{*'}X^* + (V^0)^{-1})^{-1}$
   - Compute $\mu^1 = V^1(X^{*'}R^* + (V^0)^{-1}\mu^0)$

3. Sample Σ_ϵ^1 given β^1 and U^0 from a IW (N + 100, O + S)

   - O is an identity matrix of size $S \times S$
   - $S = \sum_{i=1}^{N} \epsilon_i \epsilon_i'$ where $\epsilon_i = U_i^0 + D_i - X_i \beta^1$

4. Sample the vectors U_i^1 by sampling iteratively from truncated normal distributions given U_i^0, c_i^0, β^1 and Σ_ϵ^1.

- Draw iteratively each $u_{is}^1 = \sum_{k=1}^K \beta_k x_{isk} - d_{is} + \epsilon_{is}$ where $\epsilon_{is} \sim TN(0, \sigma_{is}^2, l_{is}, u_{is})$ with l_{is}, u_{is} computed using the ranked order list of each applicant and the bounds specified in proposition 3, u_i^0, u_i^1, c_i^0 and $\sigma_{is}^2 = \Sigma_{\epsilon ss}^1 - \Sigma_{\epsilon s(-s)}^1 [\Sigma_{\epsilon(-s)(-s)}^1]^{-1} \Sigma_{\epsilon(-s)s}^1$.

5. Sample c^1 given σ_ζ^0, and c_i^0 from $TN(\mu_c^1, \sigma_c^1, 0, \infty)$.
   - Compute $\sigma_c^1 = (N/\sigma_\zeta^0 + 1/\sigma_c^0)^{-1}$. N is the number of students as here $X_c' X_c = N$ because $X_c = 1$.
   - Compute $\mu_c^1 = \sigma_c^1 (1' R_c/\sigma_\zeta^0 + \mu_c^0/\sigma_c^0)$ where $R_{ci} = c_i^0$
6. Sample σ_ζ^1 given c^1 and c_i^0 from a IW (N + 3, 3 + $\sum_i (c_i^0 - c^1)^2$)
7. Sample the vectors c_i^1 by sampling from truncated normal distributions given U_i^1, c^1 and σ_ζ^1.
   - Consider the bounds for each c_i defined in proposition 3 and given by U_i^1 and the rank order list.

As starting values for the indirect utility vector and the unit-cost, I set

$$
\begin{aligned}
u_{is} &= 0 \quad \forall s \text{ s.t. } r_{is} = \infty \\
u_{is} &= (13 - r_{is})/13 \quad \forall s \text{ s.t. } r_{is} \neq \infty \\
c_i &= 0
\end{aligned}
$$

I use diffuse priors to minimize their influence on our estimates. I set the prior distributions for parameters $\beta \sim N(\mu^0, V^0)$ and $c^0 \sim TN(\mu_c^0, \sigma_c^0, 0, \infty)$

$$\mu^0 = 0 \tag{3.26}$$

$$V^0 = 100 \times I \tag{3.27}$$

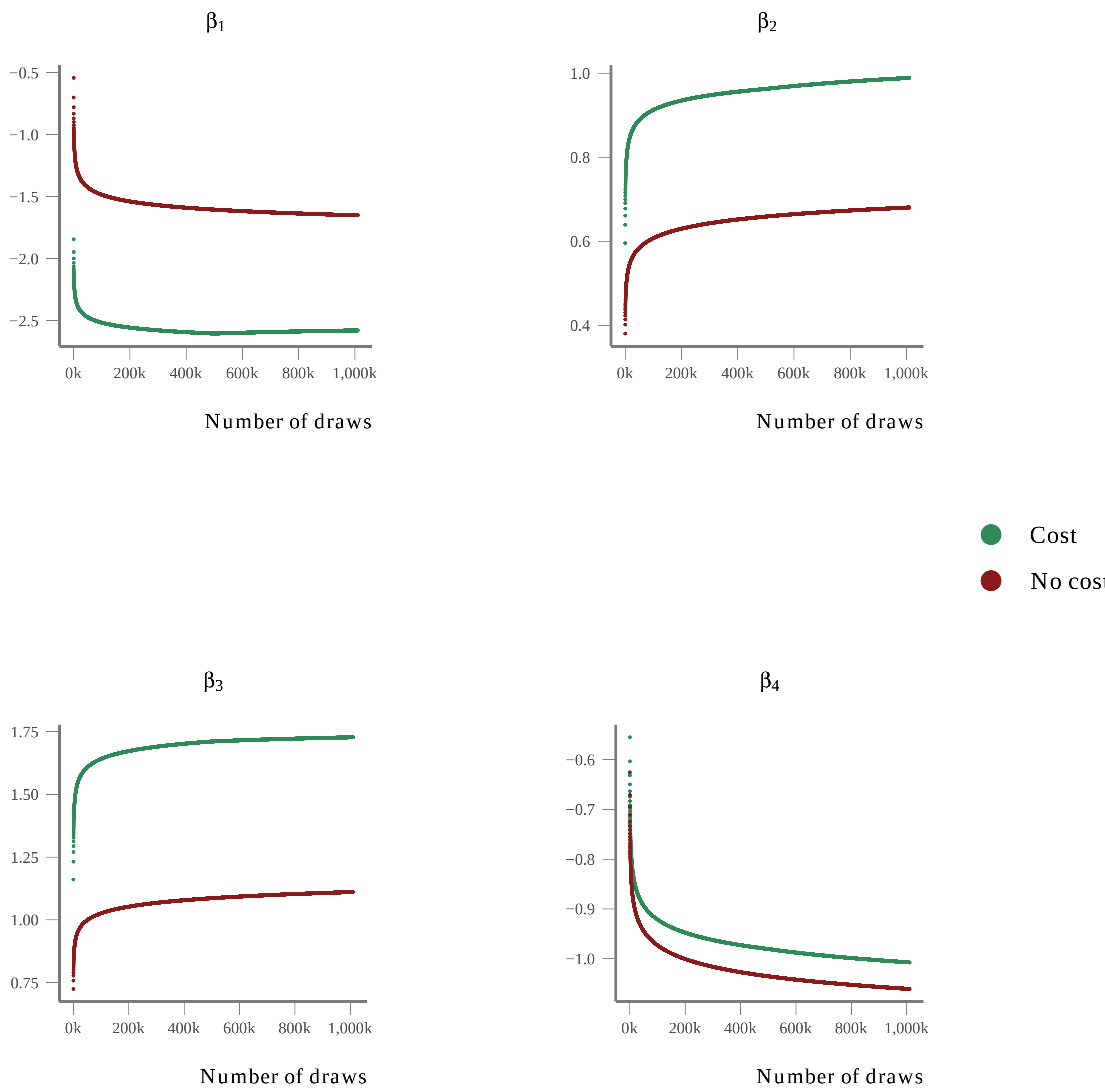
$$\mu_c^0 = 0 \tag{3.28}$$

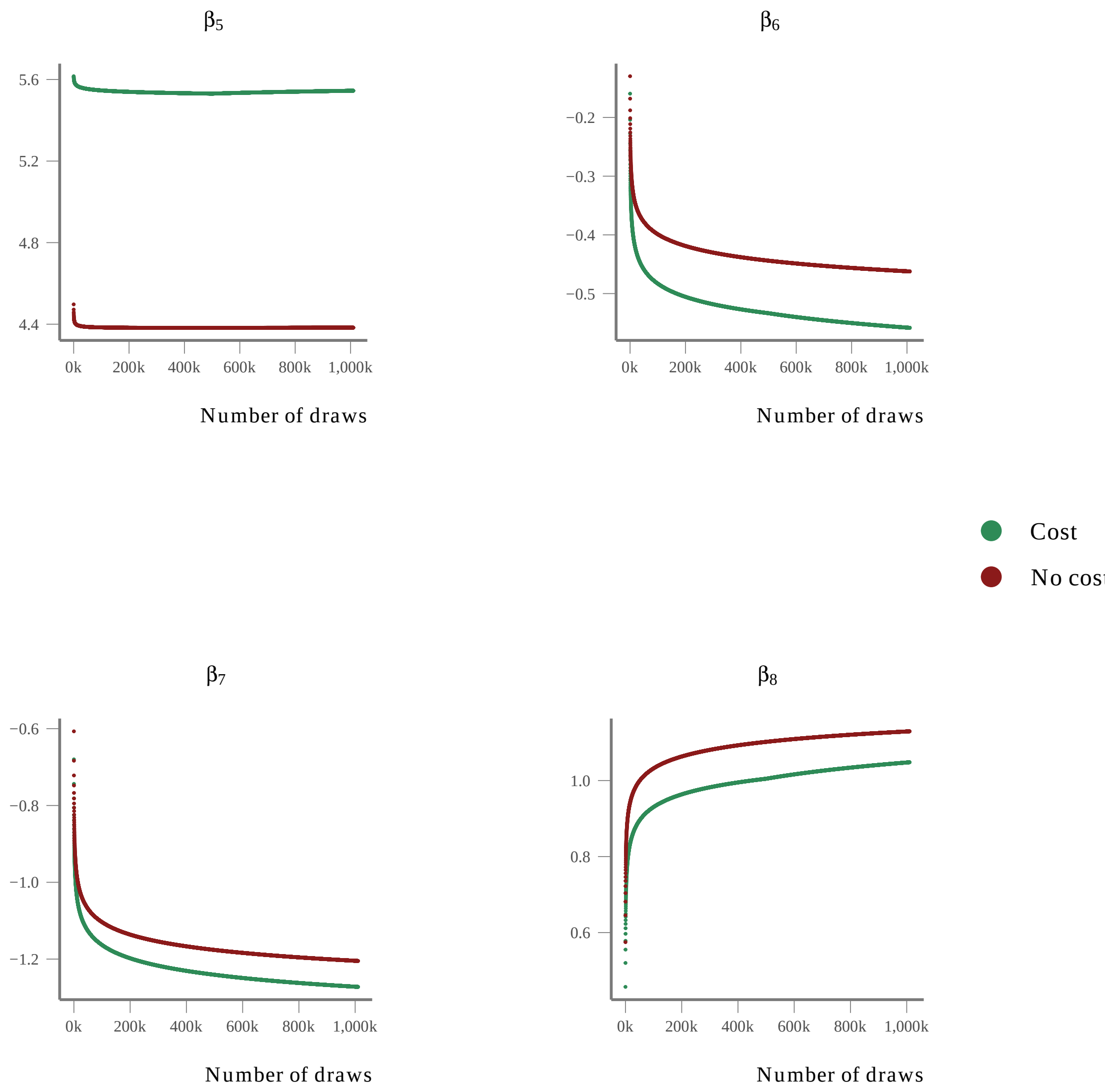
$$\sigma_c^0 = 10 \tag{3.29}$$

and the priors for the variances of indirect utilities, $\Sigma_\epsilon \sim IW(100, I)$ and the unit-cost $\sigma_\zeta \sim IW(3, 3)$.

Appendix C : Additional Figures

Figure 3..1: Evolution of Standardized School Isolation Indexes for FRPL and Minority Students





Note: Trace plots for all coefficients that parametrize students' indirect utilities for schools for the model that includes a unit-cost for adding a school to the ranked order list and the model with no cost.

Bibliography

Atila Abdulkadiroglu, Joshua Angrist, and Parag Pathak. The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. *Econometrica*, 82(1):137–196, 2014. doi: 10.3982/ECTA10266. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA10266.

Atila Abdulkadiroglu, Joshua Angrist, Yusuke Narita, and Parag A. Pathak. Breaking Ties: Regression Discontinuity Design Meets Market Design. *SSRN Electronic Journal*, 2019. ISSN 1556-5068. doi: 10.2139/ssrn.3348194. URL https://www.ssrn.com/abstract=3348194.

Atila Abdulkadiroğlu and Tayfun Sönmez. School Choice: A Mechanism Design Approach. *American Economic Review*, 93(3):729–747, May 2003. ISSN 0002-8282. doi: 10.1257/000282803322157061. URL https://pubs.aeaweb.org/doi/10.1257/000282803322157061.

Atila Abdulkadiroğlu, Joshua D. Angrist, Yusuke Narita, and Parag A. Pathak. Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation. *Econometrica*, 85(5):1373–1432, 2017. ISSN 0012-9682. doi: 10.3982/ECTA13925. URL https://www.econometricsociety.org/doi/10.3982/ECTA13925.

Atila Abdulkadiroğlu, Parag A. Pathak, Jonathan Schellenberg, and Christopher R. Walters. Do Parents Value School Effectiveness? *American Economic Review*, 110(5):1502–1539, May 2020. ISSN 0002-8282. doi: 10.1257/aer.20172040. URL https://pubs.aeaweb.org/doi/10.1257/aer.20172040.

Nikhil Agarwal and Paulo Somaini. Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism. *Econometrica*, 86(2):391–444, 2018. ISSN 0012-9682. doi: 10.3982/ECTA13615. URL https://www.econometricsociety.org/doi/10.3982/ECTA13615.

Kehinde Ajayi and Modibo Sidibe. School Choice Under Imperfect Information. *SSRN Electronic Journal*, 2020. ISSN 1556-5068. doi: 10.2139/ssrn.3524535. URL https://www.ssrn.com/abstract=3524535.

Joshua D. Angrist and Kevin Lang. Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *The American Economic Review*, 94(5): 1613–1634, 2004. URL http://www.jstor.org/stable/3592836.

Joshua D. Angrist, Peter D. Hull, Parag A. Pathak, and Christopher R. Walters. Leveraging Lotteries for School Value-Added: Testing and Estimation*. *The Quarterly Journal of Economics*, 132(2):871–919, May 2017. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjx001. URL https://academic.oup.com/qje/article-lookup/doi/10.1093/qje/qjx001.

Joshua D Angrist, Parag A Pathak, and Román Andrés Zárate. Choice and Consequence: Assessing Mismatch at Chicago Exam Schools. Working Paper 26137, National Bureau of Economic Research, August 2019. URL http://www.nber.org/papers/w26137.

Georgy Artemov, Yeon-Koo Che, and Yinghua He. Strategic 'Mistakes': Implications for Market Design Research. page 59, 2017.

Luc Behaghel, Bruno Crépon, and Marc Gurgand. Robustness of the encouragement design in a two-treatment randomized control trial. 2013.

Stephen Billings and Mark Hoekstra. Schools, Neighborhoods, and the Long-Run Effect of Crime-Prone Peers. Technical Report w25730, National Bureau of Economic Research, Cambridge, MA, April 2019. URL `http://www.nber.org/papers/w25730.pdf`.

Andreas Bjerre-Nielsen and Mikkel Host Gandil. Attendance Boundary Policies and the Limits to Combating School Segregation. February 2020.

Matthew Blackwell. Instrumental Variable Methods for Conditional Effects and Causal Interaction in Voter Mobilization Experiments. *Journal of the American Statistical Association*, 112(518):590–599, April 2017. ISSN 0162-1459, 1537-274X. doi: 10.1080/01621459.2016.1246363. URL `https://www.tandfonline.com/doi/full/10.1080/01621459.2016.1246363`.

Kevin Booker, Tim R. Sass, Brian Gill, and Ron Zimmer. The Effects of Charter High Schools on Educational Attainment. *Journal of Labor Economics*, 29(2): 377–415, April 2011. ISSN 0734-306X, 1537-5307. doi: 10.1086/658089. URL `https://www.journals.uchicago.edu/doi/10.1086/658089`.

Caterina Calsamiglia, Chao Fu, and Maia Güell. Structural Estimation of a Model of School Choices: The Boston Mechanism versus Its Alternatives. *Journal of Political Economy*, 128(2):642–680, February 2020. ISSN 0022-3808, 1537-534X. doi: 10.1086/704573. URL `https://www.journals.uchicago.edu/doi/10.1086/704573`.

David Card. Using geographic variation in college proximity to estimate the return to schooling, Aspects of labour market behaviour: essays in honour of John Vanderkamp. ed. *LN Christofides, EK Grant, and R. Swidinsky*, 1995.

David Card and Jesse Rothstein. Racial segregation and the black–white test score gap. *Journal of Public Economics*, 91(11-12):2158–2184, December 2007.

ISSN 00472727. doi: 10.1016/j.jpubeco.2007.03.006. URL https://linkinghub.elsevier.com/retrieve/pii/S0047272707000503.

Julia Chabrier, Sarah Cohodes, and Philip Oreopoulos. What Can We Learn from Charter School Lotteries? *Journal of Economic Perspectives*, 30(3):57–84, August 2016. ISSN 0895-3309. doi: 10.1257/jep.30.3.57. URL https://pubs.aeaweb.org/doi/10.1257/jep.30.3.57.

Hector Chade and Lones Smith. Simultaneous Search. *Econometrica*, 74(5):1293–1307, September 2006. ISSN 0012-9682, 1468-0262. doi: 10.1111/j.1468-0262.2006.00705.x. URL http://doi.wiley.com/10.1111/j.1468-0262.2006.00705.x.

Raj Chetty, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner. Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility. *American Economic Review*, 104(5):141–147, May 2014. ISSN 0002-8282. doi: 10.1257/aer.104.5.141. URL http://pubs.aeaweb.org/doi/10.1257/aer.104.5.141.

Jason Cook. Race-Blind Admissions, School Segregation, and Student Outcomes: Evidence from Race-Blind Magnet School Lotteries. page 55, 2018.

David M. Cutler, Edward L. Glaeser, and Jacob L. Vigdor. The Rise and Decline of the American Ghetto. *Journal of Political Economy*, 107(3):455–506, June 1999. ISSN 0022-3808, 1537-534X. doi: 10.1086/250069. URL https://www.journals.uchicago.edu/doi/10.1086/250069.

Will Dobbie and Roland G. Fryer. The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools. *American Economic Journal: Applied Economics*, 6(3):58–75, July 2014. ISSN 1945-7782, 1945-7790. doi: 10.1257/app.6.3.58. URL http://pubs.aeaweb.org/doi/10.1257/app.6.3.58.

Umut Dur, Parag A. Pathak, and Tayfun Sönmez. Explicit vs. statistical targeting in affirmative action: Theory and evidence from Chicago's exam schools. *Journal of Economic Theory*, 187:104996, May 2020. ISSN 00220531. doi: 10.1016/j.jet.2020.104996. URL `https://linkinghub.elsevier.com/retrieve/pii/S0022053118302801`.

Gabrielle Fack, Julien Grenet, and Yinghua He. Beyond Truth-Telling: Preference Estimation with Centralized School Choice. page 79, 2019.

Ford Fessenden. A Portrait of Segregation in New York City's Schools. *The New York Times*, May 2012. URL `https://archive.nytimes.com/www.nytimes.com/interactive/2012/05/11/nyregion/segregation-in-new-york-city-public-schools.html?_r=0&action=click&contentCollection=Education&module=RelatedCoverage`.

Joshua Goodman and Melanie Rucinski. Increasing Diversity in Boston's Exam Schools. page 12, 2018.

Jonathan Guryan. Desegregation and Black Dropout Rates. *The American Economic Review*, 94(4):33, 2004.

Guillaume Haeringer and Flip Klijn. Constrained school choice. *Journal of Economic Theory*, 144(5):1921–1947, September 2009. ISSN 00220531. doi: 10.1016/j.jet.2009.05.002. URL `https://linkinghub.elsevier.com/retrieve/pii/S002205310900057X`.

Elizabeth A. Harris. De Blasio Proposes Changes to New York's Elite High Schools. *The New York Times*, June 2018.

James J Heckman and Sergio Urzúa. Comparing IV with structural models: What simple IV can and cannot identify. *Journal of Econometrics*, page 11, 2010.

Winnie Hu and Elizabeth A. Harris. A Shadow System Feeds Segregation in New York City Schools. *The New York Times*, June 2018.

Peter Hull. IsoLATEing: Identifying Counterfactual-Specific Treatment Effects with Cross-Stratum Comparisons. *SSRN Electronic Journal*, 2015. ISSN 1556-5068. doi: 10.2139/ssrn.2705108. URL `http://www.ssrn.com/abstract=2705108`.

Guido Imbens and Karthik Kalyanaraman. Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, 79(3):933–959, 2012.

Rucker Johnson. Long-run Impacts of School Desegregation & School Quality on Adult Attainments. Technical Report w16664, National Bureau of Economic Research, Cambridge, MA, January 2011. URL `http://www.nber.org/papers/w16664.pdf`.

Rucker C. Johnson. *Children of the Dream: Why School Integration Works.* Hachette UK, 2019.

Adam J. Kapor, Christopher A. Neilson, and Seth D. Zimmerman. Heterogeneous Beliefs and School Choice Mechanisms. *American Economic Review*, 110(5):1274–1315, May 2020. ISSN 0002-8282. doi: 10.1257/aer.20170129. URL `https://pubs.aeaweb.org/doi/10.1257/aer.20170129`.

Lars Kirkebøen, Edwin Leuven, and Magne Mogstad. Field of Study, Earnings, and Self-Selection. *The Quarterly Journal of Economics*, 132(3):1551–1552, 2017.

Patrick Kline and Christopher R. Walters. Evaluating Public Programs with Close Substitutes: The Case of Head Start. *The Quarterly Journal of Economics*, 131 (4):1795–1848, November 2016. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjw027. URL `https://academic.oup.com/qje/article-lookup/doi/10.1093/qje/qjw027`.

Tomas Larroucau and Ignacio Rios. Do "Short-List" Students Report Truthfully? Strategic Behavior in the Chilean College Admissions Problem. page 65, 2020.

Mariana Laverde. Unequal Assignments to Public Schools and the Limits of School Choice. *job market paper*, page 48, 2020.

Sokbae Lee and Bernard Salanié. Identifying effects of multivalued treatments. *Econometrica*, 86(6):1939–1963, 2018.

Margaux Luflade. The value of information in centralized school choice systems. page 38, 2017.

Byron Lutz. The End of Court-Ordered Desegregation. *American Economic Journal: Economic Policy*, 3(2):130–168, May 2011. ISSN 1945-7731, 1945-774X. doi: 10.1257/pol.3.2.130. URL `https://pubs.aeaweb.org/doi/10.1257/pol.3.2.130`.

Jesse Margolis, Daniel Dench, and Shirin Hashim. The Impact of Middle School Integration Efforts on Segregation in Two New York City Districts. page 44, 2020.

Robert McCulloch and Peter E Rossi. An exact likelihood analysis of the multinomial probit model. *Journal of Econometrics*, 64(1-2):207–240, September 1994. ISSN 03044076. doi: 10.1016/0304-4076(94)90064-7. URL `https://linkinghub.elsevier.com/retrieve/pii/0304407694900647`.

Jack Mountjoy. Community Colleges and Upward Mobility. *SSRN Electronic Journal*, 2019. ISSN 1556-5068. doi: 10.2139/ssrn.3373801. URL `https://www.ssrn.com/abstract=3373801`.

Derek Neal. The Effects of Catholic Secondary Schooling on Educational Achievement. *Journal of Labor Economics*, 15(1, Part 1):98–123, 1997. doi: 10.1086/209848. URL `https://doi.org/10.1086/209848`.

Parag A. Pathak and Peng Shi. How well do structural demand models work? Counterfactual predictions in school choice. *Journal of Econometrics*, 222(1):161–195, May 2021.

Parag A Pathak and Tayfun Sönmez. School Admissions Reform in Chicago and England: Comparing Mechanisms by their Vulnerability to Manipulation. *American Economic Review*, 103(1):80–106, February 2013. ISSN 0002-8282. doi: 10.1257/aer.103.1.80. URL `https://pubs.aeaweb.org/doi/10.1257/aer.103.1.80`.

Leah Platt Boustan. School Desegregation and Urban Change: Evidence from City Boundaries. *American Economic Journal: Applied Economics*, 4(1):85–108, January 2012. ISSN 1945-7782, 1945-7790. doi: 10.1257/app.4.1.85. URL `https://pubs.aeaweb.org/doi/10.1257/app.4.1.85`.

Sean F Reardon. The Widening Income Achievement Gap. *Educational Leadership*, 70(8):7, 2013.

Sean F. Reardon, Stephen A. Matthews, David O'Sullivan, Barrett A. Lee, Glenn Firebaugh, Chad R. Farrell, and Kendra Bischoff. The geographic scale of Metropolitan racial segregation. *Demography*, 45(3):489–514, August 2008. ISSN 0070-3370, 1533-7790. doi: 10.1353/dem.0.0019. URL `https://read.dukeupress.edu/demography/article/45/3/489/169844/The-geographic-scale-of-Metropolitan-racial`.

Sean F. Reardon, Elena Tej Grewal, Demetra Kalogrides, and Erica Greenberg. Brown Fades: The End of Court-Ordered School Desegregation and the Resegregation of American Public Schools: Brown Fades. *Journal of Policy Analysis and Management*, 31(4):876–904, September 2012. ISSN 02768739. doi: 10.1002/pam.21649. URL `http://doi.wiley.com/10.1002/pam.21649`.

Sarah J Reber. Court-Ordered Desegregation: Successes and Failures in Integrating American Schools Since Brown. *The Journal of Human Resources*, page 57, 2005.

Alex Rees-Jones and Samuel Skowronek. An experimental investigation of preference misrepresentation in the residency match. *Proceedings of the National Academy of Sciences*, 115(45):11471–11476, November 2018. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1803212115. URL `http://www.pnas.org/lookup/doi/10.1073/pnas.1803212115`.

Eliza Shapiro. Segregation Has Been the Story of New York City's Schools for 50 Years. *The New York Times*, March 2019. URL `https://www.nytimes.com/2019/03/26/nyregion/school-segregation-new-york.html`.

Suk Joon Son. Distributional Impacts of Centralized School Choice. page 62, 2020.

Jacob Vigdor and Jens Ludwig. Segregation and the Black-White Test Score Gap. Technical Report w12988, National Bureau of Economic Research, Cambridge, MA, March 2007. URL `http://www.nber.org/papers/w12988.pdf`.

Christopher R Walters. The Demand for Effective Charter Schools. *Journal of Political Economy*, 126(6):2179–2223, 2018.

Leslie A. Zebrowitz, Benjamin White, and Kristin Wieneke. Mere Exposure and Racial Prejudice: Exposure to Other-Race Faces Increases Liking for Strangers of that Race. *Social Cognition*, 26(3):259–275, June 2008. ISSN 0278-016X. doi: 10.1521/soco.2008.26.3.259. URL `http://guilfordjournals.com/doi/10.1521/soco.2008.26.3.259`.