Blueprint Labs

Discussion Paper #2022.14

Integrating New York City Schools: The Role of Admission Criteria and Family Preferences

Clémence Idoux

March 2023

MIT Department of Economics
77 Massachusetts Avenue, Bldg. E53-390
Cambridge, MA 02139

National Bureau of Economic Research
1050 Massachusetts Avenue, 3rd Floor
Cambridge, MA 02138

Integrating New York City Schools: The Role of Admission Criteria and Family Preferences [*]

Clemence Idoux[†]

This version: March 11, 2023

Abstract

I use recent screened-school admission reforms and a structural model to gauge the contribution of admission criteria to segregation in New York City middle schools. A difference-in-differences analysis shows that two local admission reforms decreased school segregation, while prompting changes in application patterns and an increase in white and high-income student exit from the public-school sector. Using a school demand model which allows for strategic application behavior to predict the consequences of hypothetical city-wide reforms, I estimate that about half of NYC middle school segregation is due to admission criteria, with the rest due to family preferences and residential sorting.

[*] I am thankful to my advisors Joshua Angrist, Parag Pathak and Nikhil Agarwal for their advice and constant support, to the New York City Department of Education's Enrollment Research and Policy office for graciously sharing data, and to Eric Budish, Christopher Campos, YingHua He, Corinne Low, Christopher Neilson, Vendela Norman, Alex Rees-Jones, Gianluca Rinaldi and Juuso Toikka for helpful comments. Thanks to Chetan Patel for exceptional research assistance and to Eryn Heying and Anna Vallee for dependable administrative support. This paper reports on research conducted under data-use agreements between MIT, the project's principal investigator, and the New York City Department of Education. This paper reflects the views of the author alone. This research was made possible by grants from the Spencer Foundation (#202000205), the William T. Grant Foundation (#190295), and the MIT Integrated Learning Initiative (MITili). The views expressed are those of the author and do not necessarily reflect the views of the foundations.

[†] The Wharton School, University of Pennsylvania. Email: cidoux@wharton.upenn.edu

1 Introduction

In the landmark case of *Brown v. Board of Education*, the U.S. Supreme Court declared that racial segregation in public schools is "inherently unequal" and mandated that schools integrate "with all deliberate speed" (*Brown v. Board of Educ. II*, 1955). Nonetheless, seventy years later, many American schools are still divided by race, ethnicity, and class (Lutz, 2011; Reardon et al., 2012). Stark racial and socioeconomic gaps are especially apparent in large urban school districts such as New York City (NYC), where 77% of Black and Hispanic students attend schools enrolling fewer than 10% of white students while only 11% of white students and 43% of Asian students attend such schools (Cohen, 2021). The persistence of school segregation may contribute to widening racial and economic achievement gaps (Reardon, 2013), as attending a more diverse school can have both academic and psychological benefits (Billings and Hoekstra, 2019; Guryan, 2004; Johnson, 2019).

Contemporary school segregation arises in a different institutional context than in the 1960s and 1970s. Today, many large urban districts in the U.S. assign students through centralized assignment plans which allow families to choose schools, decoupling students' home addresses from the schools they can attend. Within this framework, policy makers across the country have recently adopted admission reforms aimed at reducing school segregation.[1] The political effort is particularly visible in NYC where Mayor de Blasio campaigned on the promise to make the schools "reflect the city better" (Harris, 2018). Under his mandate, the city implemented several "diversity in admissions" initiatives aimed at reducing segregation.

In the context of centralized assignment systems, segregation is an equilibrium outcome of applicant preferences (demand-side) and school admission criteria (supply-side). On the student side, applicants may prefer enrolling in schools with similar classmates (Bjerre-Nielsen and Gandil, 2020) or in schools close to home, in which case school segregation could arise from residential sorting (Laverde, 2020). At the same time, schools may also implement admission criteria based on academic achievement or residence, criteria which tend to make schools more homogeneous. Determining the respective roles of these two factors is key as it informs the extent to which policy makers can influence school segregation through admission reforms.

This paper evaluates the contribution of school admission criteria to the observed pattern of segregation in NYC middle schools. The first part of my analysis studies the effects of admission reforms in two local NYC school districts. The analysis of these unique natural experiments highlights the importance of student behavioral responses to admission reforms both at the application and enrollment stages. Based on the reduced form analysis, I develop a model of demand for schools which reflects the observed behaviors. I then simulate how counterfactual admission schemes would affect the city-wide level of school segregation,

[1] For instance, Fairfax County (VA) and San Francisco replaced the admission tests at their exam schools with lotteries in 2020 (Warikoo, 2021). In Boston, exam schools still screen students based on academic achievement but now give priorities to students living in disadvantaged zip codes (Barry, 2021). Philadelphia switched to lottery admission for its magnet schools and introduced some zip code priorities in 2021 (Mezzacappa, 2021). Chicago expanded access to its elite high school entrance test in 2021. Under the new policy, all public school students take the test, which is administered directly at their schools (Karp, 2021).

accounting for applicant responses. These counterfactual scenarios allow me to isolate the contribution of admission criteria to school segregation from demand-side factors.

In 2019, two of the 32 NYC school districts, Northwest Brooklyn (NWB) and Upper West Side (UWS), launched integration reforms which reduced the role of academic screens in middle school admission. I use a difference-in-differences strategy to estimate the impact of these reforms on school segregation and to analyze changes in applicant behaviors. To measure changes in school segregation, I compute a segregation index which measures the percentage deviation of same-group exposure at school as compared to a reference geographical level. I find that NWB's integration reform substantially decreased economic and racial segregation at the district's schools. The segregation index of low-income applicants decreased by 30%, while the segregation index of Black and Hispanic (minority) applicants decreased by 13%. UWS's integration reform, in contrast, only decreased economic segregation index by 8% and did not affect racial segregation.

Despite the difference in final impact between the NWB and UWS integration reforms, both reforms elicited large behavioral responses from applicants. At the application stage, families changed the rank-order list (ROL) of schools submitted to the central assignment system. Applicants who faced the largest decreases in admission odds due to the reforms decided to lengthen their ROLs, while applicants with increased admission odds at competitive programs ranked those programs more highly. These changes amplified the reforms' effects on integration, resulting in school offers that were roughly three times less segregated.

During the enrollment phases, I find that White and higher-income applicants were more likely to turn down their match offer and exit the public school system. These changes in school offer take-up rates may be explained by changes in the achievement of potential peers as White and higher-income applicants were assigned to schools which had, on average, lower-achieving potential peers after the implementation of the reforms. Overall, this "white flight" halved the effects of the integration reforms on racial and economic segregation in both districts.

The extent of these behavioral responses underscores the importance of understanding applicant decision-making processes both during the enrollment and application stages to predict the effects of changes in admission regime. To fully characterize the application response, I develop and estimate a model of ROL formation which captures the changes induced by the integration reforms. The observed responses are hard to reconcile with the assumption that applicants list schools in order of true preference, without considering their admission chances.[2] Thus, my model departs from the truthful benchmark in two ways. Firstly, applicants can submit short lists, i.e. stop adding schools to their list once they are confident of securing admission to one of the listed programs. Secondly, applicants can omit schools for which they have a low probability of admission.

[2] Artemov et al. (2017) and Larroucau and Rios (2020) also find evidence consistent with students omitting programs to which they are unlikely to be admitted in two other school systems that use deferred acceptance to assign students. Relatedly, Rees-Jones and Skowronek (2018) present experimental evidence that medical students misrepresent their preferences in the National Resident Matching Program. Luflade (2017) estimates the value of giving information to students on their probabilities of admission in DA matches, when the length of the ROL is constrained and students cannot always express their true preferences.

To allow for strategic application behavior, I model the applicant choice of ROL as an optimal portfolio choice problem as in Chade and Smith (2006). In this framework, the ROL can viewed as a set of lotteries over schools, where the weights assigned to each school depend on the applicant's admission probabilities and the school's position in the applicant's list. Applicants choose their ROL to maximize their expected utility, which is determined by their utility over schools, the lottery over schools induced by the ROL, and the cost of submitting the ROL. In my model, I simplify the applicant choice problem by assuming applicants use a heuristic method to construct their ROL. Instead of solving a one-shot utility maximization problem that requires considering all possible ROLs, applicants choose which schools to rank sequentially with a limited understanding of the dynamic consequences of each choice.

Besides being able to fit the observed behavioral changes, my model simplifies the estimation of preferences under strategic reports. Agarwal and Somaini (2018) propose a general methodology to estimate preferences based on the optimality of the chosen ROL, whenever the mechanism can be represented with a cutoff structure. A major challenge when implementing this methodology is the curse of dimensionality, as the number of potential ROLs grows exponentially with the number of schools and the method requires checking the optimality of each applicant's choice among all possible lists. To circumvent this dimensionality problem, Larroucau and Rios (2020) proposes a strategy to limit the choice space considered for estimation, while Calsamiglia et al. (2020) shows how the optimality of a list can be checked by backward induction. My model instead assumes that applicants do not consider all lists and implies a series of bounds on indirect utilities for schools. These bounds yield a reduction in dimensionality which allows me to estimate preferences, even in the case where admissions are not independent events.

Using the preference estimates from my model, I assess the contribution of admission criteria to school segregation by simulating the equilibrium effect of different admission regimes. The first counterfactual drops all academic selection criteria, which are used by 30% of NYC middle schools and are often deemed responsible for school segregation.[3] The second counterfactual drops all admission criteria based on residence as well as academic criteria. In this scenario, students are allowed to apply to any NYC school, do not get priority based on residence, and are not ranked on academic achievement. This would represent a sweeping change, as 97% of middle schools use some or all of these criteria. Since this second counterfactual removes all school admission criteria, it isolates how much of segregation is due to applicant preferences and residential sorting.

The simulation results suggest that academic screens play only a small role in city-wide school segregation, while geographic screens are much more important. In the first simulation which drops only academic screens, I find a modest decrease in racial and economic segregation: the segregation index of low-income students drops from 5% to 4% while the segregation index of minorities drops from 14% to 13%. In contrast, when dropping also geographic screens, the segregation index of low-income students drops from 5% to 2%, while the segregation index of minority students decreases from 14% to 9%. Hence, about half of NYC middle school segregation is due to school admission criteria while the remaining

[3]This counterfactual scenario was implemented temporarily in 2021 due to the Covid-19 pandemic.

half is due to demand-side factors. As at least half of the demand-induced school segregation is driven by residential sorting, admission reforms are unlikely to reduce school segregation in NYC by more than 50% from its current level.

The rest of the paper is organized as follows. After a brief review of the related literature in the remainder of this section, Section 2 describes segregation in NYC schools and residential areas and the NYC school assignment system. Section 3 discusses the effects of the integration reforms on school diversity. This section also details applicant responses to the plans. Section 4 discusses the estimation of a strategic model of school demand. Section 5 presents the effects of counterfactual admission regimes on NYC school segregation. Section 6 concludes.

Related literature

This paper contributes to the contemporary policy debate on school segregation. The effect of school segregation on disadvantaged student academic achievement has been documented by research which has found that the Black-white test score gap is higher in more segregated cities (Card and Rothstein, 2007; Vigdor and Ludwig, 2007) and has increased in school districts that suspended race-based admission (Billings et al., 2014; Cook, 2018). Studies on the effects of integration policies also suggest that integration may confer both academic and psychological benefits to students (Angrist and Lang, 2004; Guryan, 2004; Zebrowitz et al., 2008; Johnson, 2011; Bergman, 2018; Johnson, 2019).[4] My paper shows that admission criteria and academic screens play a role in school segregation, but that their elimination would not result in a fully integrated school system in NYC. These results are consistent with the previous literature on segregation under school choice, which has emphasized the role of informational frictions (Son, 2020), applicant preferences, and residential sorting (Laverde, 2020; Bjerre-Nielsen and Gandil, 2020).

This paper also documents enrollment responses to contemporary integration policies. I find that both districts' integration reforms resulted in white and high-income student enrollment losses, changes which are quantitatively similar to the ones observed during the adoption of court-ordered desegregation plans in the 1970s (Reber, 2005; Lutz, 2011). Building on the "white flight" literature, I am able to scrutinize the underlying cause of exit and provide evidence that changes in public-school enrollment are mediated through changes in peer achievement at the assigned school. Once peer achievement is controlled for, other peer characteristics do not affect the enrollment decision. This finding relates to the analysis of NYC application patterns in Abdulkadiroğlu et al. (2020), which shows that an important determinant of applicant preferences for schools is peer quality.

Finally, this paper contributes to the literature on estimating demand for schools in deferred acceptance (DA) under strategic behavior (Fack et al., 2019; Ajayi and Sidibe, 2020; Larroucau and Rios, 2020; Son, 2020). First, I provide evidence from a natural experiment that applicants respond to changes in the admission criteria, and that their responses are not compatible with the assumption of truthfulness of expressed preferences. Second, I offer

[4] Moreover, exposure to more diverse peers at school might also impact preferences and racial attitudes in adulthood (Boisjoly et al., 2006; Carrell et al., 2019; Billings et al., 2020).

an alternative method to estimate preferences in DA where students take into account their admission probabilities. My method modifies Agarwal and Somaini (2018) to avoid the curse of dimensionality by assuming that applicants use a heuristic method to choose their list, rather than solving a one-shot utility maximization problem that requires considering all possible rank-order lists (Larroucau and Rios, 2020). The model allows applicants to submit short lists and ignore low-probability choices, while being computationally feasible to estimate even in a large market and when admissions are not independent events.[5]

2 Background

2.1 Institutional context and data

The NYC middle school match

The NYC public school system has around 450 middle schools which enroll approximately 70,000 new 6$^{\text{th}}$ graders each fall. In the preceding winter, current 5$^{\text{th}}$ 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 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 using the deferred acceptance (DA) algorithm.[6] About 92% of students who completed their application are matched in this main round. Applicants that are unassigned at the end of DA are manually placed in programs with unfilled seats based on geographic proximity and expressed interests.

Applicants report their preferences to the mechanism through a rank-order list, which has been limited to 12 choices since 2017. To support families in the application process, the DOE provides both a physical admission guide and access to a personalized website. Each personalized website only includes schools to which the applicant is eligible.[7] Both the website and the guide include an information page about each school, which comprises of a brief statement of the school's mission; a list of offered programs, courses, and extracurricular activities; the performance of enrolled students on standardized tests; admission priorities and selection criteria for each of its programs; the number of applicants per seat and the priority of the last admitted applicant in the prior year. The DOE also issues annual school reports that list enrolled student demographics, teacher characteristics, and statistics about student performance and school environment.

[5]Ajayi and Sidibe (2020) and Son (2020) also consider a model of list formation in which applicants only consider a subset of programs. However, in their models, strategic considerations based on admission probabilities do not affect which schools are ranked whenever an applicant does not fill her list.

[6]The detailed steps of the algorithm are described in Appendix F.

[7]The application guide is district-specific and only includes schools that accept students residing in the district.

Each academic program ranks applicants according to program-specific eligibility and admission criteria.[8] While most programs prioritize students on the basis of their residential zones, the school at the time of application and attendance to information sessions are also sometimes used in rankings. 33% of programs also rank individual students based on prior grades, standardized test scores, talent test scores, and behavioral measures. These programs that rank students on academic criteria are often referred to as screened programs.

NYC middle schools are intended to serve students living in their neighborhoods. The school district is divided into 32 local school districts whose boundaries are shown in Figure 1 and most middle schools only consider applicants residing in the district or in smaller specific residential areas. 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 in or attending schools in specific districts. In part because of these rules, 85% of students attend a middle school in their district.

Data

The data 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 2020-2021. The application and match files contain information on applicants' choices, applicants' priorities and rankings at the programs they applied to, main round offers, manual placements, and final offers. 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. By 2020-2021, the replication rate increases to 99%.

The application data can be matched through a unique identifier to data on school enrollment, student demographics, standardized test scores, and residential location. The DOE collects school enrollment data each year in June. Besides the grade and school enrolled, the data also contains information about the ethnicity, poverty status (which proxies for free or reduced-price lunch (FRPL) eligibility), and English language learner (ELL) status of each NYC student.[9] The test score files include the results to NY State ELA and math standardized tests administered in grades 3 through 8. A performance level of 1, 2, 3, or 4 is associated with each scaled score. Students that score above 3, which corresponds roughly to the 60$^{\text{th}}$ percentile, are considered high performers. Finally, the DOE provided students' residential census tracts and zip codes. School distances are computed as the linear distance between the centroid of a student's census tract and each school. Appendix details the construction of the samples used in the reduced form analysis and in the demand estimation.

[8] Ties between applicants with the same rank are broken using a unique tiebreaker.

[9] The enrolled school 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.

2.2 Segregation in NYC

Measure of segregation

Segregation is measures based on the concept of same-group exposure (SGE), which corresponds to the probability that a member of one group meets a member of the same group within a given geographical unit (an area or a school). The segregation index (SI) standardizes same-group exposure by subtracting the group's marginal probability in the population of interest. For a given demographic group identified by a dummy variable X, the segregation index for that group is calculated as:

$$SI_X = SGE[X] - E[X] \tag{1}$$

This segregation index represents the percentage deviation from perfect integration, given the population's demographics, and can be used to compare segregation across time and space while taking into account differences in the demographic makeup of the population. The index takes the value of 0 when students attend schools whose student bodies mirror the population of reference, its maximum value corresponds to 1 minus the benchmark same-group exposure.

There are many other measures of the unevenness of a distribution, but this index has several useful properties. First, it can be applied to both residential and school segregation and to measures with two or more groupings (e.g. multiple races). Second, the index can be used to compute district-level school segregation by subtracting the demographic makeup of the district from the average school SGE among students residing in the district.[10] Since most NYC middle schools only accept in-district students, the district-level measures constitute relevant benchmarks to evaluate school segregation. Finally, the normalization by the demographic composition of applicants and not of enrolled students accounts for the effect of public school exits on segregation.[11]

NYC residential and school segregation

School choice may ameliorate the segregation inherent in neighborhood schools by allowing students to attend more distant schools. The policy has more leverage when segregation occurs at a smaller geographic scale, as in NYC, since students do not need to travel long distances to mix with different peers.[12]. The scope for district-level choice to increase integration is also limited by the number of non-minority students enrolled in the district public schools. In 2018, one third of NYC 6^{th} graders were white or Asian, leaving some potential for school choice to affect racial mixing.

[10] A district-level school segregation index can take a negative value if district students attend less segregated schools than the district itself thanks to out-of-district attendance.

[11] This distinguishes this index from the one used in Margolis et al. (2020), which compares the demographic makeup of district schools to the makeup of students enrolled in the district.

[12] Reardon et al. (2008) finds that NYC is amongst the five most segregated metropolises when considering small neighborhoods with a 500-meter radius, but falls to around the 20^{th} when segregation is measured over a 4,000-meter radius

Table 1 documents the extent of residential and school segregation for middle school applicants in 2018. As the second column shows, residential segregation at the census tract level is substantial. The segregation index at the census tract level for Black, white and Asian students is approximately 0.30, while it is 0.18 for Hispanic students. Residential segregation falls when measured at the district level in column (3). For insrance, the segregation index for Black and white students drops from around 0.30 at the census tract level to 0.18 and 0.12 at the school district level. Segregation is also visible along economic lines, as low-income applicants are slightly over-represented in some census tracts and districts.

Schools appear to facilitate integration as segregation is lower in schools compared to the census tract level. Nevertheless, schools are still not fully integrated, even when compared to school districts. Column (5) of Table 1 shows that city-wide segregation indexes at the enrolled school are halfway between the census tract and district level values for all groups.[13] Overall, the NYC centralized seat assignment results in middle schools which are less diverse than school districts, meaning that the match process falls short of achieving the maximum level of integration allowed by geographic eligibility and priority rules.

3 Reduced form analysis of NYC integration reforms

3.1 NYC integration reforms

In 2019, the Northwest Brooklyn (NWB) and Upper West Side (UWS) school districts implemented district-wide "Diversity in Admissions" initiatives that substantially altered program admission criteria for all district schools.[14] The integration reforms impacted 11 middle schools in NWB and 16 middle schools in UWS. These reforms provide two unique natural experiments to evaluate the impacts of admission criteria on school segregation. Since out-of-district enrollment is quite limited, district-level plans allow for an evaluation of the general equilibrium effect of admission reforms and the effect on public school exits.

Both integration reforms aimed to promote integration by reducing academic screening and increasing access to selective schools. NWB eliminated academic screens at all of its middle schools, and prioritized 52% of the seats in each school for low-income, English language learners, or homeless students. UWS's reform was less far-reaching: it maintained academic screens, but prioritized 25% percent of seats at each school for low-achieving low-income students.[15] As a result, the reforms increased admission chances for low-income students with low baseline test scores and reduced admission chances for high-income applicants with high baseline test scores who received lower priority for the reserved seats.

[13]Segregation indexes at the school offered by the match in column (4) are very similar to the ones at the enrolled school in column (5).

[14]These district-based initiatives followed school-level programs launched in previous years described in Appendix Table A2.

[15]10% of UWS middle schools' seats were prioritized for free and reduced-price lunch (FRPL) eligible students who scored an average of below 2 on a composite of 4th grade math and ELA scores ranging from 1 to 4.5; and, 15% of additional seats were prioritized for FRPL-eligible students who earned an average between 2 and 3 on the same composite score.

These two districts are suitable settings for examining the impact of admission reform effects on school segregation for three reasons. Firstly, both districts are more diverse than the city, as shown in Table 2. For instance, 40% and 31% of UWS and NWB students are white, compared to 16% city-wide. Secondly, prior to 2019, both districts had higher racial and economic school segregation compared to most other NYC districts. Black and Hispanic student district-level school segregation indexed reached 0.22 in UWS and 0.14 in NWB, against 0.04, on average, in NYC.[16] Thirdly, academic screens were prevalent in both districts before the reforms. Prior to 2019, 57% of UWS programs and 80% of NWB programs screened their applicants on grades and behavioral measures, while only 33% of NYC programs did so.

Although the direct target of the reforms was economic segregation, they also affected racial segregation. Black and Hispanic students are more likely to benefit from such reforms as they are more likely to be FRPL-eligible and have lower baseline test scores.[17] Approximately 70% of UWS Black and Hispanic students and 50% of NWB Black and Hispanic students were eligible for the reserves created under the new admission schemes. By contrast, only 3% and 15% of white students qualified for the reserves in UWS and NWB, respectively. Section 3.2 examines the changes in racial and economic school segregation after the reforms.

The diversity effect of such admission reforms hinges on how students respond to them. Figure 2 illustrates how admission criteria impact enrollment. Student behavior influences the final impacts of a change in admission criteria by affecting which school offers are made and which of those offers translate into enrollment. Section 3.3 highlights how exit from the traditional public school sector increased for white and high-income students, reducing the impact of the reforms on segregation. Finally, Section 3.4 shows that applicants altered their school choices in response to the reforms.

3.2 Effects on school segregation

The estimation of the reforms' effects on school segregation is challenging as segregation is a district-level outcome. I assess the effects of the plans through a district-level difference-in-differences (DiD).[18] This research design compares differences in measured segregation between UWS and NWB and other NYC districts before and after the implementation of the integration reforms. It identifies the causal effects of the reforms under the assumption that, in the absence of any policy changes, trends in segregation would have been similar across districts. I examine whether districts followed different trends prior to the reforms to check the validity of this assumption. To assess the statistical significance of the observed changes, I use permutation methods similar to the ones introduced in Abadie et al. (2010).

The integration reforms led to a decrease in economic segregation in both UWS and NWB. Figure 3 plots school segregation indices over time for low-income and minority students

[16] district-level school segregation indexes are standardized by the district-level shares.

[17] See appendix Table A3 for a description of student characteristics by race.

[18] A synthetic control approach is not suitable as the treated units present extreme pre-treatment outcome levels (Abadie, 2021): UWS and NWB had the highest segregation index for most demographic groups prior to the implementation of the plans.

in UWS, NWB, and other NYC districts. In 2019, low-income segregation at the school of enrollment in UWS dropped by 8%. In the same year, economic segregation in NWB dropped by 30%. The larger change in school segregation in NWB is consistent with the more far-reaching nature of its integration reform, as described in Section 3.1. On the other hand, economic segregation did not change, on average, for other NYC districts.

The integration reforms also led to a decrease in racial segregation in NWB, albeit to a lesser extent. The Black and Hispanic segregation index decreased by 13% in NWB between 2018 and 2019. The larger decline in economic segregation compared to racial segregation is consistent with the fact that the integration reforms targeted low-income students. The decrease in racial segregation in NWB is mostly a result of Hispanic and white students attending more diverse schools. [19]

I assess the statistical significance of changes in segregation using permutation methods. To obtain a distribution of "placebo effects" on segregation, I estimate the following dynamic DiD specification where each of 32 NYC districts is assigned iteratively to be the treated unit $\tilde{d}$:

$$Y_{td} = \lambda_t + \delta_d + \sum_{j=-3}^{-1} \beta_j I(d = \tilde{d}) \times I(t = j) + \sum_{j=1}^{2} \beta_j I(d = \tilde{d}) \times I(t = j) + \varepsilon_{td} \tag{2}$$

where λ_t and δ_d are year and district fixed effects. Y_{td} is the segregation index in district d in year t. In each regression, segregation indices are normalized to zero for 2018, the year prior to the implementation of the integration reforms. β_j captures the estimated "placebo" treatment effects for the district assigned to treatment in each year prior to and after 2018.

The comparison of the changes in the segregation index observed in UWB and NWB to the "placebo effects" observed for the other NYC districts that did not implement district-wide admission plans allows us to assess the statistical significance of the plans' effects. Figure 4 plots the coefficients from the 32 dynamic DiD regressions in which each NYC district is iteratively assumed to have implemented a policy change in 2019. The estimated effects for UWS and NWB are deemed significant when their magnitudes are extreme relative to the "placebo effects". Conversely, UWS and NWB do not show pre-trends in the outcome if estimated effects for years prior to 2018 are comparable to the placebo estimates.

The permutation test reported in Figure 4 suggests that the declines in economic and racial segregation observed in NWB were driven by its integration reform. Compared to other NYC districts, NWB had amongst the largest estimated drop in segregation in 2019 for all demographic groups.[20] Moreover, NWB had amongst the smallest estimated coefficients for years prior to the reform's implementation. Hence, the drop in NWB segregation indices is not likely due to chance or pre-existing trends.

[19] The NWB school segregation index fell by 72% for white students and by 19% for Hispanic students after the reform, while it remained stable for Asian students and declined by a small percentage for Black students, as shown in appendix Figures A4. Note that the UWS segregation index also decreased by 49% for white students, mainly due to the significant increase in the number of white students leaving the traditional public school system.

[20] Appendix figure A4 reports permutation tests for white and Hispanic students.

The evidence is less clear for UWS, where the segregation indices for both low-income and minority students were declining prior to the reform in comparison to other NYC districts. Moreover, the 2019 estimated decline in segregation for minority students is only modest compared to the declines estimated for other districts. While it is plausible that the decrease in economic segregation in UWS resulted from its integration reform, the decrease in racial segregation is small and within the bounds of typical year-to-year fluctuations.

Overall, these results are consistent with Margolis et al. (2020) but much smaller in magnitude. The difference is largely due to the accounting of exit from the traditional public school system.[21] Indeed, economic and racial segregation in match offers, which corresponds to the level of school segregation that occurs if all applicants enroll in the school they are offered, exhibit much larger declines in Figures 3 and A4. Moreover, the corresponding permutation tests reported in Figures 4 and A5 provide with strong evidence that the declines in segregation at the match offer stage are due to the reforms.

3.3 Effects on exit from the traditional public school system

The results in Figure 3 indicates that changes in enrollment behavior weakened the impact of the integration reforms. The comparison of the plans' impacts in final enrollment and in school offers reveals that shifts in offer take-up approximately halved the effects of the plans on economic and racial segregation in both districts. This section analyzes the enrollment responses to the reforms.

To estimate the effect of the integration reforms on take-up behavior, I implement a student-level DiD regression that controls for district and year fixed effects. Specifically, I estimate the following regression:

$$Y_{itd} = \lambda_t + \delta_d + \beta_1 I(d = UWS) \times I(t = 2019) + \beta_2 I(d = NWB) \times I(t = 2019) + \varepsilon_{itd} \quad (3)$$

where λ_t and δ_d are year and district fixed effects. Y_{itd} is a dummy that takes the value of one 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 change in the probability that students attend a school outside of the traditional public school sector after the implementation of the integration reforms.[22] For inference, I report robust standard errors and I confirm the statistical significance of DiD results using the permutation method introduced in the previous section and described in MacKinnon and Webb (2020).[23] The permutation test specification and results are reported

[21] The analysis in Margolis et al. (2020) is limited by the use of school-level enrollment data. Since they assess school segregation with respect to the population of students that enrolled in district schools, they cannot account for the impact of changes in out-of-district enrollment on the effects of the plans.

[22] Charter schools are outside the traditional public school sector as they do not take part in the match.

[23] Assessing the statistical significance of the estimates from this DiD specification is complicated in the context of the policy change. Since treatment is assigned at the district level, standard errors need to be clustered if treatment effects are not homogeneous (Abadie et al., 2017). Unfortunately, as NYC only comprises 32 districts, cluster-robust standard errors are not consistent (MacKinnon and Webb, 2017). Moreover, only one cluster is treated for each experiment in the DiD, which precludes the use of wild bootstrapping for clustering (MacKinnon and Webb, 2018).

in the Appendix G. The permutation tests also allow me to assess the presence of pre-trends in the outcome.

The DiD estimates presented in Panel A of Table 3 suggest that both integration reforms resulted in a large increase in the share of white and high-income students leaving the public school system. Both shares went up by 6-7 percentage points in UWS, a 60% increase from the pre-reform levels, and by 8 percentage points in NWB, a 45% increase. The integration reforms had limited effects on public school exit for Black, Hispanic, Asian, and low-income students, as the corresponding point estimates were close to zero and insignificant in both districts.[24] The drop in white enrollment in traditional public schools is quantitatively similar to the reductions seen after nationwide court-ordered desegregation plans in the 1970s and 1980s. White enrollment in UWS decreased by 7% and in NWB by 10% as a result of the reforms, which is comparable in magnitude to the 9% decline in the first two years after court-ordered desegregation plans documented by Reber (2005).

The information on student offers sheds light on the reasons behind these enrollment changes. Panel B of Table 3 shows the effect of the integration reforms on a proxy for school desirability: the mean math baseline test score of offered students.[25]. Taken together, Panels A and B of Table 3 offer suggestive evidence that the changes in student enrollment due to the plans are associated with changes in the mean peer math score obtained through the match. Changes in offered peer achievement are inversely related to the estimated changes in exit in Panel A. For instance, white applicants in NWB, who are the most likely to exit the public school system, receive offers with 0.25 standard deviations lower peer mean math scores than in previous years.

To quantify the extent to which changes in public school exit are mediated through changes in offered peer characteristics, table 4 reports IV estimates of the effect of offered peer achievement on exit from traditional public schools using different sets of reform instruments. Columns (1) and (2) show IV estimates computed using UWS and NWB integration reforms separately. In both cases, the integration reforms are interacted with 8 covariates, in a model that controls for covariate main effects as well as their interactions with district and year fixed effects.[26] Column (3) shows IV estimates that combine both sets of instruments, using the variation generated by the two district integration reforms. If all of the effect of the integration reforms on public school exit is mediated through a decrease in offered peer achievement, IV estimates using each reform instrument and across covariate-defined subgroups should be equal.

The similarity of the different IV estimates across columns provides support for peer achievement being a primary mediator of the reforms' effects on student exit. The resulting

[24]The only exceptions are Asian students in UWS and Black students in NWB, but these students represent only 7% and 6% of applicants in UWS and NWB, respectively.

[25]I use 5^{th} grade math scores as fewer students have a missing math test scores than English test scores. The peer mean math baseline test score is computed leaving out each corresponding student to avoid simultaneity bias. This choice of proxy is motivated by Abdulkadiroğlu et al. (2020), who find that peer achievement is the main determinant of NYC high school popularity.

[26]The covariates used in the estimation are dummies for English language learner status, race, and the interactions of low baseline tests score status with low-income status.

estimates are at most 0.02 standard deviation (SD) apart and not statistically distinguishable.[27] Being offered a school where peer baseline performance is 0.1 SD better on average decreases the probability of exiting the public school system by approximately 3 percentage points.

The overidentification statistics reported at the bottom of Table 4 provide a further test for treatment effect homogeneity across reform instrument and across covariate-defined subgroups. In columns (1) and (2), which use reform-covariate interactions for each integration reform separately, the overidentification test statistics implicitly test equality of IV estimates computed separately for covariate subgroups. With p-values of 0.26 and 0.13, respectively, the tests provide little evidence of differences in impact across covariate subgroups.[28] Column (3)'s overidentification test statistic, in which both district plans are used as instruments, tests additionally for the equality of IV estimates computed using district plans as instruments one at a time. The p-value for the overidentification test is 0.17, offering little evidence against treatment homogeneity.

Since other peer characteristics might be correlated with math baseline achievement and may therefore explain the changes in public school exit probability, I next consider other mediators. Specifically, Table 4 considers, in addition to math achievement, the proportion of minority students and the proportion of low-income students among applicants that receive the same offer. Columns (4) through (6) report IV estimates for models that include each pair of peer characteristics together. Models with two endogenous variables capture two causal effects at the same time. These models, identified by differences in the two integration reforms' effects on school composition, allow for the possibility that different sorts of causal effects are either reinforcing or offsetting.

The models with two endogenous variables suggest that peer achievement is the most relevant mediator of student exit decisions. For the models in columns (4) and (5), which include both peer math achievement and one of the two competing peer characteristics, the estimates for the proportion of minority students and the proportion low-income students are not statistically significant. The coefficient on peer math achievement retains its statistical significance and is similar to the one estimated using the single endogeneous variable model: -0.36 compared to -0.28.[29]

3.4 Effects on expressed preferences

Applicants responded to the reform-induced changes in offered school characteristics by changing their enrollment behavior. It is therefore plausible that they also adapted their

[27] Using as mediator the previous year's baseline achievement at the offered school, which assumes that students do not anticipate how integration reforms will change school composition, leads to slightly more variable estimates

[28] With 8 covariate interactions and an integration reform main effect in the instrument list, the resulting overidentification test has 8 degrees of freedom.

[29] Moreover, the overidentification p-value for the model which includes proportion of minority students and proportion of low-income students as endogeneous variables takes a value of 0.01, rejecting the equality of IV estimates.

application behavior, anticipating the effects of the reform. In this section, I characterize how applicants changed their rank-order lists (ROLs) in response to the plans.

I first assess the contribution of changes in ROLs to the impact of the reforms on segregation. To this aim, I simulate match offers under UWS and NWB plans using application data from 2018. Note that because this year preceded the reforms, applicant ROLs were unaffected. The differences in segregation between the simulated 2018 match and the real 2018 match capture the mechanical effect of the changes in admission criteria. The impact of behavioral responses is then the difference between the observed declines in segregation and the mechanical effects. These calculations are accurate under the assumption that expressed preferences would have remained stable between 2018 and 2019 if there had been no reforms.

Figure 5, which reports the results of these simulations, shows that changes in application behavior explain more than two-thirds of the decrease in economic and racial segregation at the match offer stage. The decline in low-income and minority student segregation indices would have been more than three times smaller if applicants had not adapted their ROLs to the integration reforms. This change in impact due to applicant behavioral responses is unlikely to be due to chance, since the actual effects of both integration reforms are not within the 95% confidence intervals surrounding the simulated effects with no behavioral response. These empirical confidence intervals account for the variability in segregation indices that arises from differences in applicant population and lottery numbers.[30]

Changes in applicants' ROLs determined the impact of the reform on segregation. Therefore, I now turn to characterizing these changes. I first consider whether the length of the ROLs was affected. As the integration reforms increased admission uncertainty for applicants, applicants should list more choices if they wish to avoid being manually placed in schools they might not like. Previously, students could estimate their admission odds based on their elementary school test scores. By replacing screened admissions with lotteries, NWB made it harder for applicants to predict their assignments. Similarly, the UWS reform raised admission uncertainty by increasing the chances of students that were less likely to qualify and by decreasing the chances of students that were more likely to qualify. Using the same difference-in-differences specification as in equation (3), Panel A of Table 5 shows the effects of the integration reforms on the length of the list submitted.

In both districts, the reforms resulted in an increase in the number of choices listed. The increase was more substantial for applicants in NWB, where admission uncertainty increased the most. On average, UWS applicants listed 1 additional choice in 2019 and 1.2 additional choices in 2020, while NWB applicants listed 2.6 additional choices in 2019 and 2.8 additional choices in 2020. All of these changes are significant at the 1% level. As shown in appendix Table A4, the increased length of the ROLs almost perfectly offset the increase in uncertainty as the share of unassigned students remained stable after the reform in NWB, while only increasing a little in UWS.

The heterogeneity in responses for applicants with different baseline achievement and low-

[30] I obtain the empirical distribution by bootstrapping 100 times the 2018 match under the two integration reforms, redrawing each time a sample of applicants and a sequence of tie-breakers. Applicants are sampled with replacement from each district independently. Students with disability (SWD) and non-SWD applicants are also sampled separately as they participate in two distinct matches.

income status in columns (2) through (5) is consistent with applicants listing more choices to reduce the probability of manual placement.[31] High-income applicants with high baseline test scores, whose admission odds were negatively impacted by the reforms, added the most choices to their lists. Low-income applicants with low baseline test scores, who benefited the most from the reforms, added fewer programs to their lists. Finally, the small and mostly insignificant changes in the probability of manual placement for all groups, shown in appendix Table A4, suggest that each group of students adjusted the length of their ROLs in proportion to the change in admission uncertainty induced by the integration reforms.

The increase in length of the ROLs was accompanied by changes in the order and composition of the lists. Panel B of Table 5 shows that the mean baseline achievement of the most preferred school changed heterogeneously for each applicant group.[32] The mean baseline achievement of each school is computed in years prior to the reform, so that it captures pre-reform school selectivity. Low-income applicants with low baseline scores ranked first more competitive schools than in previous years in both UWS and NWB. On average, the schools they preferred most enrolled students with respectively between 0.05 and 0.12 higher baseline math scores. On the other hand, high-income NWB applicants with high baseline scores ranked schools that enrolled students with between 0.10 and 0.12 smaller baseline math scores first after the reform.

These changes in expressed preferences are hard to reconcile with the assumption that applicants list schools in order of true preference, without taking into account their admission chances.In contrast, they seem consistent with applicants submitting short lists, i.e. not adding schools to their list once they are sure of getting into one of the listed programs. The increase in applicant ROL length in response to the integration reforms suggests that applicants add more schools to their list when the probability of non-assignment increases. Applicants also appear to omit schools for which they have a low probability of admission, as applicants whose admission odds increased thanks to the reforms rank more selective programs after their implementation.

4 Model of school choice under strategic behavior

The analysis of UWS and NWB integration reforms show that applicants adapt their enrollment and application behaviors to the admission environment. These changes in behavior substantially affect the impact of a change in school admission criteria. Therefore, it is important to predict behavioral changes when forecasting the effects of an admission reform. The prediction of enrollment changes can be done by examining the changes in peer characteristics of the assigned school. The prediction of the updated rank order list (ROL) requires a model of list formation, as it involves relative comparisons across programs.

This section introduces a model of ROL formation that is compatible with the deviations

[31] Students with 5$^{\text{th}}$ grade math proficiency score above 3 are defined as high-baseline. This corresponds to the threshold for UWS' reserves.

[32] Considering only the school ranked first avoids the selection bias that would arise if the plans also affected the decision of ranking more choices.

from truthfulness identified in the previous section. Specifically, this model of strategic behavior is consistent with applicants submitting short lists and omitting choices for which the probability of admission is low.

4.1 Model setup

Let students be indexed by $i \in \{1, ..., n\}$ and programs be indexed by $s \in \{1, ..., S\}$. Each student submits a rank-order list (ROL) of programs, denoted by $R_i \in \mathcal{R}$ where $\mathcal{R}$ comprises the sets resulting from all the k-permutations of $\{1, ..., S\}$ for $k = 1, ..., S$. Each R_i is a strictly ordered set where the ordering of elements in R_i corresponds to student i's expressed-preference order.[33] For instance, if student i submits rank-order list $R_i = (3, 15)$, it means that only programs 3 and 15 are indicated as acceptable by i (i.e. preferred to being unassigned) and that i indicates she would prefer to be assigned to program 3 over program 15. For any rank-order list R, I denote the ordered set including the first k elements of R as $R_k \subseteq R$ where $k \in \{0, ..., |R|\}$. In other words, R_k corresponds to the order-preserving truncation of R in position k; thus, if $R = (3, 15)$, then $R_1 = 3$. Note that $R_0 = \emptyset$ for all $R \in \mathcal{R}$.

Each program also ranks applicants and students are assigned to programs using the deferred acceptance (DA) algorithm defined in the Appendix F. At the end of the algorithm, unmatched applicants are manually placed into a program which I denote by 0.[34] Thus, program 0 corresponds to the student outside option within the traditional public sector. I next describe the choice problem faced by applicants, applicant preferences over programs, and applicant beliefs.

School portfolio choice problem

The choice of ROL by applicants is an optimal portfolio choice problem. This framework was first introduced by Chade and Smith (2006) for college applications and has been used more recently in other applications of school choice (Agarwal and Somaini, 2018; Larroucau and Rios, 2020). The key observation is that ROLs can be mapped to lotteries over programs whose weights depend both on the ordering of programs in the list and on applicant beliefs about admission probabilities.[35] Hence, applicants choose their ROL to maximize their expected utility, which depends on their preferences over programs, the lottery over programs induced by the ROL and their beliefs about admission probabilities, and the cost of submitting the ROL.

[33] I define the union of two ordered sets $A \cup B$ to be the ordered set listing elements of A in their order followed by elements of B in their order.

[34] Manually placed students are usually assigned to the under-subscribed school closest to their home.

[35] A property of DA is that students' admission probabilities at programs are independent of their rank-order lists. Assuming that applicants understand this property, applicant subjective beliefs about admission probability depend only on their beliefs about other applicants' ROLs and programs' rankings of applicants. This allows us to consider the effects of changes in ROL and subjective admission probabilities separately.

Formally, each applicant i chooses her ROL $R \in \mathcal{R}$ to maximize her expected utility:

$$\max_R \sum_{s=1}^{S} v_{is} \times p_{is}(R, \mathbf{q}_i) - C_i(R). \tag{4}$$

The expected utility from each list depends on the indirect utilities $\{v_{is}\}_{s=1}^{S}$ that applicant i derives from assignment to each program s. In the expectation, the weight given to each program is denoted by p_{is}, which captures the subjective probability of assignment to program s. Finally, each list is associated with some cost $C_i(R)$ that restricts applicant choice.

The weight assigned to each program $p_{is}(R, \mathbf{q}_i)$ is jointly determined by applicant choice of ROL R and by applicant vector of subjective beliefs about admission probabilities $\mathbf{q}_i := (q_{i1}, ..., q_{iS})$.[36] When DA is used to assign students to programs, the function $p_{is}(\cdot, \cdot)$ has three properties. First, p_{is} is increasing in q_{is}, which depends only on applicant beliefs about program rankings of applicants and other applicants' ROLs. Second, p_{is} is weakly increasing in the program s position in R as the algorithm goes down sequentially through applicant lists. Third, under the same logic, p_{is} is decreasing in q_{ij} for all programs j ranked before s in R. To keep track of these dependencies, I define $p_{isk}(R_{k-1}) := p_{is}(R_{k-1} \cup \{s\}, \mathbf{q}_i)$, applicant i's subjective probability of being offered program s when ranked in position k, given programs listed in higher positions R_{k-1}.

Finally, student i incurs a cost $C_i(R)$ when forming her ROL. $C_i(R)$ can be interpreted as capturing any psychological or monetary cost that a student might face when forming her list. For instance, students might expend time and effort to learn about programs that are not in their neighborhoods. Alternatively, listing highly-selective programs may induce a psychological cost if students anticipate being disappointed if they are not granted admission.

This application cost covers different implementations of DA. When $C_i(R) = 0$ for all R, this model coincides with the traditional setting in which DA is strategy-proof (Gale and Shapley, 1962). When applicants do not face any application cost, it is weakly strategy-proof to list programs in true order of preference. When the cost depends only on list length $|R|$, Haeringer and Klijn (2009) show that it is a weakly dominated strategy for students to submit a ROL that is not a true partial preference ordering. It follows that the ordering of programs included in a student's list should reflect true preferences among those ranked. Nonetheless, applicants may omit programs for which the odds of assignment are too low to be worth listing them. The probability of assignment to a program is low if the admission odds at that program are low or if the applicant is sure of getting in a program she prefers.

[36] By definition, admission probability at the outside option is 1 ($q_{i0} = 1$). Note that this notation is not fully general as it does not cover all possible dependencies between admission events. Nonetheless, it covers the case where admissions are either independent or based on a single score, as specified in Section 4.2.

Preferences: utilities and cost

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

$$v_{is} = U(X_{is}, \xi_s, \varepsilon_i) - d_{is}, \tag{5}$$
$$v_{i0} = 0, \tag{6}$$

where X_{is} denotes observable characteristics of student i and program s, d_{is} is the linear distance between student i's home and the school hosting program s, and ξ_s and ε_i are unobserved characteristics of schools and individuals respectively. Following a common approach in the school choice literature (Agarwal and Somaini, 2020), I assume that $\varepsilon_i \perp (D_i | X_i, \xi)$ where $X_i := (X_{i1}, ..., X_{iS})$, $D_i := (d_{i1}, ..., d_{iS})$, and $\xi = (\xi_1, ..., \xi_S)$. This assumption is violated if student residential choices are determined by students' idiosyncratic tastes for programs, conditional on student and program observable characteristics as well as program unobservables.

This representation of utility includes both a location normalization and a scale normalization. The utility of the outside option is set to zero, which normalizes the location of utilities. The coefficient on distance d_{is} is set to -1 which normalizes the scale of utilities, while assuming that distance is undesirable all else equal.

For the empirical application, I allow the utility to vary over time and I divide the vector X_{is} into a vector of applicant characteristics (Z_i) and a vector of time-varying program characteristics (Z_{st}). I further parametrize the utility as:

$$v_{ist} = \delta_{c(z_i)s} + \sum_{h=1}^{H} \beta_{c(z_i)h}(c(z_i) \times z_{sht}) - d_{is} + \varepsilon_{ist}, \tag{7}$$

where $c(z_i)$ assigns students to covariate cells based on the variables in the vector of applicant characteristics Z_i, $\delta_{c(z_i)s}$ is the mean utility of program s for students in cell $c(z_i)$,[37] $\{z_{sht}\}_{h=1}^{H}$ is a set of year-specific school characteristics, $\beta_{c(z_i)h}$ captures the effect of schools' variation in peer composition across years for students in each covariate cell, and $\varepsilon_{ist} \sim N(0, \sigma_\varepsilon)$.

In the estimation, the mutually exclusive covariate cells, $c(z_i)$, correspond to the interactions of high baseline math status and minority status (4 cells). The time-varying school variables, $\{z_{sht}\}_{h=1}^{H}$, consist of the share of high baseline math students and the share of minority students among students enrolled at the school. These shares are computed for each school in the year preceding application. As such, I assume that applicants observe the composition of the student body at each school at the time of application and make their application decisions assuming school composition will be stable over time.[38] To ensure the plausibility of this assumption, 2019 is omitted in the estimation since it is the first year in which the reforms were implemented.

Applicant i's cost of forming a ROL is assumed to be linear in list length, i.e. $C_i(R) = c_i|R|$. The cost of adding an additional program to the ROL is applicant-specific. For the

[37]The program-specific dummies interacted with student cells subsume the program unobservables ξ_s.

[38]As discussed in Section 2.1, applicants have access to detailed information on enrolled students' demographic characteristics and performance at standardized tests at each school.

empirical application, the per-unit cost is parametrized as:

$$c_i = c + \zeta_i, \tag{8}$$

where $c \geq 0$ is the average per-unit cost and $\zeta_i \sim TN(0, \sigma_\zeta, -c, \infty)$ is applicant i's unobserved deviation from the average per-unit cost.[39] I assume that idiosyncratic utility shocks and per-unit cost shocks are independent, i.e. $\zeta_i \perp \varepsilon_i$.

Beliefs: rational expectations

When submitting their ROLs to the centralized system, applicants also consider the probability of being admitted to each program. Admission to each program may not be independent if program rankings of applicants are based on a common tiebreaker or the same set of scores. In the previous section, Table 5 presented evidence consistent with applicants understanding how changes in admission criteria due to the integration reforms affected their admission probabilities. I therefore assume that students hold rational expectations over their admission chances. Applicants are able to infer their admission probabilities at each program based on their own score τ_i and the distributions of tiebreakers and scores. Moreover, they understand the dependencies across admission probabilities.

A recent literature shows that applicants tend to hold mistaken beliefs about their admission probabilities (Corcoran et al., 2018; Kapor et al., 2020; Arteaga et al., 2021). Specifically, applicants are more optimistic about their admission chances than what rational expectations would entail. My framework could accommodate biased beliefs, but this would require to take a stance on the exact nature of the bias. Unfortunately, in the absence of a survey on applicant beliefs, it is not possible to separately identify beliefs and preferences. Optimistic beliefs would bias the estimated per-unit cost of application and preferences over schools. Nonetheless, if the extent to which beliefs are mistaken is not affected by changes in admission regime, simulations based on the model's estimates would still generate accurate long-run predictions.

Limited rationality assumption

Given the large number of programs available, it is unrealistic for applicants to find their optimal ROL by evaluating all possible program combinations.[40]. Indeed, if applicants can list up to 12 among 60 potential programs, as in NYC, they have more than 10^{20} lists to choose from.[41] To simplify the decision-making process, I assume applicants use heuristics

[39] $TN(\cdot)$ denotes the truncated normal distribution.

[40] This also entails that inferring preference for programs by comparing ROL utilities runs into the curse of dimensionality. To circumvent this problem, (Larroucau and Rios, 2020) shows how to limit the number of lists considered for estimation when admission probabilities are independent.

[41] Although the NYC middle school match includes almost 700 programs, each applicant is only eligible for approximately 60 programs.

instead of comparing every possible list.[42] Applicants choose their list sequentially without fully internalizing how their choices affect the continuation value of their list through changes in assignment probabilities. Assumption 1 formalizes this heuristic.

Assumption 1 (Limited rationality of applicants)

For each position k in the list, applicant i chooses $s \in \{0\} \cup (\{1, ..., S\} \setminus R_{k-1})$ to maximize:

$$v_{is} p_{isk} \left(R_{k-1} \right) - I(s \neq 0) c_i + \tilde{V}_k \left(\{s\} \right), \tag{9}$$

where the continuation value for position k is defined for each ordered set of choices C as:

$$\tilde{V}_k \left(C \right) := \max_{s' \in \{0\} \cup (S_k \setminus C)} v_{is'} p_{is',k+|C|} \left(R_{k-1} \cup C \right) + \tilde{V}_k \left(C \cup \{s'\} \right), \tag{10}$$

where $S_k = \{s' \in \{0\} \cup (\{1, ..., S\} \setminus R_{k-1}) : v_{is'} p_{is'k} \left(R_{k-1} \right) - I(s' \neq 0) c_i \geq 0\}$, i.e. S_k is the "consideration set" consisting of remaining programs which would clear the per-unit cost c_i if they were added in position k.

This assumption states that an applicant solves the portfolio choice problem sequentially but does not fully anticipate the consequences of her choice at each step. Her mistake is reflected in the specification of the continuation value $\tilde{V}_k$ in Equation (10). Specifically, the definition of $\tilde{V}_k$ fixes a "consideration set" S_k which contains the programs that would clear the per-unit cost if they were added in position k.[43] Fixing S_k in $\tilde{V}_k$ is equivalent to assuming that, when choosing the k-th ranked program, the applicant thinks that she will eventually add all programs in S_k to her ranked order list. This assumption is realistic when the number of programs is not capped or when applicants typically do not exhaust their list, which is the case for over 92% of applicants to NYC middle schools.

This choice heuristic is consistent with applicants listing a mix of "reach", "match" and "safety" programs as in Ali and Shorrer (2021) and choosing their most preferred option within each set of options. Given the applicant's mistaken belief that she will list all programs in S_k, and since the assignment mechanism uses deferred acceptance, it is optimal for her to list the highest utility program in S_k in position k. The fact that the applicant only considers programs in S_k implies a departure from truthfulness of expressed preferences. This is because S_k depends on the unit cost of listing and the perceived probability of

[42] This approach is similar to the one used by Kennan and Walker (2011) who face a similar dimensional problem when modelling dynamic migrations decisions. They restrict the choice space and simplify the dynamic problem by assuming that agents partially forget past information.

[43] On the other hand, the fully rational specification for the continuation value does not feature a consideration set S_k. The maximization problem solved by a fully rational applicant can be obtained by replacing $\tilde{V}_k(\{s\})$ in Equation (9) by $V(\{s\})$ where

$$V(C) = \max_{s' \in \{0\} \cup (\{1, ..., S\} \setminus R_{k-1} \cup C)} v_{is'} p_{is'k+|C|} \left(R_{k-1} \cup C \right) - I(s' \neq 0) c_i + V(C \cup \{s'\}).$$

Using this continuation value specification solving the sequential problem is equivalent to solving the optimization problem in Equation (4).

assignment at each schools. Thus, applicants take into account the cost of listing additional programs and do not list programs with small assignment probabilities or utilities.

This assumption precludes students from ranking a program earlier in their list in anticipation of it not being acceptable further down the list. Such a case arises when the expected utility of the unlisted program $v_{is}p_{isk}(R_{k-1})$ is larger than the expected utility of the chosen program whenever the unlisted program is acceptable, but the utility $v_{is'}$ of the chosen program is larger than that of the unlisted program. The following example fleshes out a case in which this happens.

Example 1 (Effect of limited rationality on choice)
Assume a student has only two acceptable programs for the first position, i.e. two programs for which $v_s p_{s1} > c$. Let $v_1 = 10c$ and $v_2 = 9c$, $q_1 = 0.89$, and $q_2 = 1$. Assuming program admissions are independent, the student does not list two programs as $v_1 \times q_1(1 - q_2) < v_2 \times q_2(1 - q_1) < c$. This is true under both full and limited rationality. But, under full rationality, the student lists only program 2 as $v_1q_1 < v_2q_2$; while under limited rationality, she lists only program 1 as $v_1 > v_2$ and $c < v_1q_1 < v_2q_2$.

Bounds on utilities and cost

Under the limited rationality assumption, each applicant ROL entails bounds on $\{v_{is}\}_{s=1}^{S}$, the utilities for each program, and on c_i, the per-unit cost of adding a program to the ROL. These bounds depend on the admission probabilities. Proposition 1 spells them out:

Proposition 1
Under Assumption 1, and for any cost $c \geq 0$, the rank-order list R chosen by applicant i maximizes her overall expected utility V_{iR} if and only if the following conditions hold:

1. *Any program s listed k^{th} in list R has weakly higher expected utility ($p_{isk}v_{is}$) than the cost c.*

2. *Any unlisted program s has less expected utility than the cost if it were added at the end of the list.*

3. *Each program s in R has indirect utility (v_{is}) lower than all programs listed above it and higher than all those listed below it.*

4. *Each program s listed k^{th} in R has higher utility than all programs not listed in R which would have delivered weakly higher subjective utility than the cost c if listed in the k^{th} position.*

Proof. Proof in Appendix A. □

This proposition allows me to estimate preferences, as it defines the vectors of program indirect utilities and per-unit cost that are consistent with the optimality of the observed list. This insight can be used to construct the likelihood of observing a given ROL as a function of the distributions of utilities and per-unit cost. Indeed, given admission probabilities, any vector of program indirect utilities and cost corresponds to a uniquely optimal ROL.

Under this proposition, an applicant ROL is not only a true partial order but also carries information about unlisted programs. If the admission probability of an unlisted program is high enough, it entails that listed programs are preferred to the unlisted program. Nevertheless, the utility of an unlisted program for which the probability of admission is low cannot be bound, as the program might just have been skipped because of its low admission odds.

This proposition is consistent with applicants omitting programs for which their admission probabilities are low and submitting short lists. If the unconditional probability of admission at a program q_{is} is sufficiently low, applicant i omits that program from any position on her list since her expected utility of listing the program is always lower than the corresponding cost. The expected gain from listing any program also goes down with the list position. Indeed, $p_{isk}(\cdot)$ is decreasing in k so $p_{isk}(\cdot)$ may be small even when $p_{is1}(\cdot)$ is non-negligible. This entails that applicants stop their list once they are sure of being admitted to one of the programs they have already listed. Finally, when $c = 0$, DA is strategy proof and these inequalities correspond to the ones implied by truthful reports.

Identification of the model

I now intuitively discuss the role that shifts in distance and probabilities play in learning about the distributions of both indirect utilities and per-unit cost. Agarwal and Somaini (2018) derive the conditions on the utility regions for non-parametric identification of the distribution of indirect utilities using 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.[44] Indeed, variations in ROL for different distance vectors D_i, given student characteristics Z_i and vector of admission probabilities P_i[45], cannot separately identify indirect utilities and per-unit cost. The second source of identification comes from variation in admission probabilities. Fixing D_i and Z_i, changes in P_i identify the distribution of per-unit cost, as long as the variation in admission probabilities does not affect the distributions of indirect utilities and per-unit cost. Together with the variation in distance, shifts in admission probabilities faced by similar students allow us to separately identify utilities and per-unit cost.[46]

Shifts in admission probabilities can be found both within year and across years. Within-year variations in admission probabilities arise from program priorities and seat reserves.[47] Admission probabilities vary discontinuously at school zone and school district boundaries

[44] In addition, the identification of preferences for time-varying school characteristics requires at least two years of data.

[45] $P_i := (p_{i1}, ..., p_{iS})$

[46] Note also that applicants ranking programs for which $p_{isk} = 0$ have per-unit cost $c_i = 0$, by Proposition 1. Thus, the distribution of utilities is fully identified by these applicants as long as $v_i \perp c_i | X_i, D_i, \xi$.

[47] In the empirical application, the model is also identified by variations in admission probability along continuous variations in applicant test scores since the vector of applicant characteristics only includes a

due to program priorities, while FRPL eligibility, elementary school of enrollment, and ELL status jointly determine reserve eligibility. Moreover, year-to-year changes in admission criteria generate variation in admission probabilities for similar students. NYC's admission system and admission reforms provide an ideal setting as they entailed substantial variation in admission odds for different types of applicants.

4.2 Model estimation

The estimation of the preference parameters, $\theta = (\{\delta_{c(z_i)s}\}_{s=1}^{S}, \{\beta_{c(z_i)h}\}_{h=1}^{H}, \sigma_\varepsilon, c, \sigma_\zeta)$, follows a two-step method. In the first step, I estimate admission probabilities for each applicant at each program. The second step uses the bounds on program indirect utilities derived in Proposition 1 to estimate θ, taking as given the first step's 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 equivalent to a maximum likelihood estimator.

First step: estimation of admission probabilities

I estimate admission probabilities q_{is} by bootstrapping student assignments. The bootstrap procedure samples applicants with ROLs and test scores, draws a new lottery tiebreaker, and runs DA to obtain an assignment. For each bootstrapped assignment, I compute admission cutoffs as the largest lottery number among admitted applicants for programs that rank applicants based on lottery number, or as the lowest score among admitted applicants for programs that rank applicants based on prior academic performance.[48] The admission probabilities are estimated based on these bootstrapped cutoffs, which capture the uncertainty in admission due to variation in the lottery draw and year-to-year variation in the applicant population. Agarwal and Somaini (2018) show that this estimator is consistent if applicants hold rational expectations about their admission probabilities. Appendix H.1 describes the estimator in detail.

Given the estimated probability of admission at each program, I then compute $p_{isk}(R_{k-1})$, the conditional offer probability for each program s if placed in position k of the applicant's chosen list. This conditional probability takes into account the dependencies between admission events, and varies depending on whether the program admits applicants based on their score or their lottery number.

I assume that admission score cutoffs are approximately independent, thus admissions at each score program are independent events given applicant scores.[49] On the other hand,

binary indicator for high-achieving applicants.

[48]The score for each applicant τ_i corresponds to the average 4^{th} grade math and ELA state test score. In practice, programs also use additional discretionary criteria to rank applicants. Nonetheless, this composite score is a good predictor of admission given student information at the time of application.

[49]This is a plausible assumption if the market becomes large, as variations in cutoffs are due to changes in the applicant population. In the limit, admission cutoffs are constant in DA (Azevedo and Leshno, 2016). In the empirical application, variations in cutoffs are also driven by changes in the discretionary weights programs give to different criteria when deciding student rankings.

admissions at lottery programs are not independent events. Since all programs use the same lottery and lottery numbers are not known at the time of application, an applicant's rejection at a lottery program reveals information on the applicant's tiebreaker. Indeed, the probability of being offered each lottery program if placed in a given position depends on the lottery cutoffs of the lottery programs ranked higher in the list. The formulas used to capture these dependencies can be found in Appendix H.1.

Second step: estimation of preference parameters

The second step takes as given the admission probabilities computed in the first step. Specifically, Assumption 1 implies that the vector of parameters maximize the likelihood that the observed ROL solves the student's sequential optimization problem:

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmax}} \sum_{i=1}^{n} \log \mathbb{P}(R_i = \arg \max_{R \in \mathcal{R}} V(R) | X_i, D_i, \tau_i; \theta) \tag{11}$$

This likelihood does not have a closed-form solution. Thus, I implement a Gibbs sampler adapted from McCulloch and Rossi (1994), which yields estimates that are asymptotically equivalent to the maximum likelihood estimator. The Gibbs sampler obtains draws of δ, β, c, σ_ε, and σ_ζ from the posterior distribution by constructing a Markov chain from any initial set of parameters θ^0. The posterior given the data and the prior corresponds to the invariant distribution obtained through the Markov chain.

The chain is constructed by sampling from the conditional posteriors of the parameters, utility vectors, and per-unit cost given the previous draws. The sampler iterates between sampling the parameters of the model conditional on the simulated utilities and per-unit cost and sampling the utilities and per-unit cost conditional on the parameters. Specifically, the estimator computes the following sequence of conditional posteriors at each new iteration $t+1$:

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

The first four steps of the sampler follow McCulloch and Rossi (1994). The two last data augmentation steps are different as the set of constraints differ. I modify the constraints on utilities to be consistent with the applicant optimization problem. I use the bounds on indirect utilities and per-unit cost derived in Proposition 1 to pick utilities and per-unit costs that are consistent with applicants choosing their ROL to solve the sequential optimization problem described in Equation (10). Appendix H.2 reports details for the Gibbs sampler.

This approach differs from Agarwal and Somaini (2018) as it does not compare the utilities of potential lists to pick program utility vectors that are consistent with observed choices. This substantially reduces the number of constraints. To draw each program utility, it is enough to check $S + 1$ constraints instead of $|\mathcal{R}|$ constraints, the cardinality of the set of all potential lists.[50]

4.3 Model estimation results

Table 6 reports the estimates of preferences for time-varying characteristics. The model in column (1) allows for strategic reports by including a per-unit cost, while the model in column (2) assumes truthful reports by setting the per-unit cost to zero. To reduce the computational burden, these parameters are estimated on the sample of 2018 and 2020 NWB applicants only.[51] Table 7 reports estimates for the remaining model parameters for both the NWB sample and the sample including all 2018 applicants.

The two models of applicant behavior generate substantially different preference estimates. The starker contrasts appear when considering preferences for time-varying school characteristics. Estimated preferences for changes in Hispanic and Black shares are null and insignificant in the model allowing strategic reports, while they are negative and significant in the model that assumes truthful reports. In addition, an increase in the share of high-achieving students at a program is much more valued by applicants in the model that allows for strategic reports. In the model assuming truthful reports, applicants with low baseline math test scores are actually found to like a school less after it increases its share of high-achieving students.

These two observations are consistent with omission of programs affecting the validity of the estimates from the model which assumes truthful reports. A model without cost might estimate a stronger distaste for programs that enroll more minority students, as it ignores that applicants tend to omit less-preferred programs when they are certain of being assigned to one of the programs already listed. Moreover, such a model interprets a decrease in program selectivity as a preference for lower-achieving peers if applicants are less likely to list programs with proportionally more high-achieving students for which their admission odds are lower.

Table 7 also reveals heterogeneity in the mean preferences for schools estimated by the two models. Specifically, if students only include programs when it is strictly beneficial to do so, assuming truth-telling overstates the value of being unassigned. The truthful model also overestimates the degree of preference heterogeneity across students since it attributes differences in admission probabilities to differences in taste. As a result, both estimated mean program utilities and their correlation across student cells are larger in the model allowing for strategic reports. More than 50% of the mean program utility estimates for

[50] For estimation purposes, the set of programs for each student only includes programs for which the student is eligible given her demographic characteristics.

[51] NWB applicants faced the most substantial changes in school composition and application probabilities over this period because of the scope of NWB integration reform.

low-achieving minority students are not statistically different from these of high-achieving non-minority students.

The inclusion of an application per-unit cost affects the preference estimates despite the relatively small estimated mean per-unit cost. On average, the cost of including one more program in a list corresponds to traveling 0.005 miles. Moreover, 76% of applicants have a per-unit cost lower than 0.01, which suggests that most applicants adopt very safe listing strategies by omitting only very unlikely assignments.

5 Simulations of counterfactual admission regimes

Based on the model estimates, I can simulate student assignments under alternative admission regimes. These simulations allow us to asses how different admission criteria contribute to school segregation. The change in school segregation induced by dropping a given admission criterion measures that criterion's contribution to segregation. In addition, residual segregation in a counterfactual assignment where schools do not screen applicants captures the contribution of family preferences and residential segregation to school segregation.

I simulate the impact of two counterfactual admission regimes: the first counterfactual removes all academic admission criteria while the second removes both academic admission criteria and admission criteria based on residence. These simulations are motivated by ongoing policy discussions in NYC. Academic screening at middle schools was temporarily suspended in 2021 due to the Covid-19 pandemic, and geographic criteria were partially phased out of NYC high schools.[52] The second counterfactual isolates the contribution to segregation of applicant preferences and residential sorting since it eliminates almost all school admission criteria.

Before showing the results from these counterfactual exercises, I discuss the technical difficulties with the simulation of an equilibrium outcome and how I assess the fit of the model.

5.1 Simulation setup and model fit

To determine the equilibrium level of segregation under a counterfactual admission policy, I iteratively simulate the match until a fixed point for school segregation is reached. This iterative procedure is necessary as changes in admission policy lead to changes in admission probabilities which in turn induce different application patterns and school characteristics. At each iteration of the match, applicants update their ROL by optimizing given the previous step's school attributes and probabilities of admission.[53]

[52] On December 14$^{\text{th}}$ 2021, the phase-out was suspended: high schools were authorized to maintain borough and zone priorities, while district priorities had been eliminated in the previous admission cycle.

[53] Only non-SWD applicants with baseline demographics update their ROL (90% of non-SWD applicants). The ROLs of all other applicants (SWD applicants and the remaining non-SWD applicants) are fixed across iterations.

The baseline utility for each school is given by the estimated mean program utilities and student characteristics, while cost and utility shocks are randomly drawn so that applicants' original lists of programs are optimal.[54] The utility that applicants derive from each school is updated at each iteration , depending on the changes in the composition of each school, while the utility of the outside option is fixed. At each iteration, the match is bootstrapped 100 times by re-drawing a new set of applicants and lottery numbers. This bootstrap procedure allows us to compute probabilities of admission and average school composition that account for the uncertainty in the student assignment.

The results presented in this section ignore the effect that counterfactual admission regimes might have on exit from the public school system. In Appendix B, I extend the simulations to account for potential exit. Applicants' binary decision to quit the public school system is predicted using the probit model detailed in Appendix B. This model is based on offered school composition and follows the multiple endogenous variables IV specification of Table 4. The share of applicants exiting the traditional public school sector increases from 11.5% to 13.2% on average in the simulated match with no academic screens and to 13.7% on average in the simulated match with neither academic nor geographic screens. The increase in public school exit generally diminishes the impact of changes in admission criteria, but all the main conclusions are robust to accounting for it.

To assess the fit of the model, I first check whether the model can predict the segregation effect of the NWB diversity reform. Section 3 shows that a simulation that fixes ROLs falls short of predicting the full decrease in both economic and racial segregation. Figure 6 shows that the simulation using pre-reform data but updating applicant lists is able to match the observed effect of the reform more closely. For minority and low-income students, the observed decline in segregation index lies within the 95% confidence interval of the simulated decline that updates applicant behavior based on changes in admission probabilities and school composition.[55]

The replication of the NWB diversity reform's effects is an out-of-sample validation of the simulation's accuracy, as it almost solely relies on pre-reform data. All parameters are estimated on the 2018 applicant sample, except the taste for changes in school composition which leverages the reform. Nonetheless, changes in school composition only modestly affect student utilities compared to time-invariant school characteristics. Hence, ignoring the effect of school characteristic changes does not substantially affect the accuracy of the replication, as shown in appendix Figure A6.

5.2 Effect of dropping academic screens

The first counterfactual scenario drops all academic screens while maintaining school priorities and eligibility criteria. This counterfactual is motivated by the current public debate

[54]This method of drawing errors leverages the information on student utilities embodied in the observed ROLs. For programs an applicant was not eligible for in the original 2018 match, the utility error is drawn from a Normal distribution using the estimated variance.

[55]Appendix Figure A1 incorporates enrollment decisions into the simulation. The simulation replicates the changes in school segregation induced by the integration reform for all groups except Asian applicants.

over segregation in NYC which often focuses on the practice of screened admission. Selective enrollment schools are accused of perpetuating racial and economic segregation by allowing white and upper-income families to avoid mostly-minority and low-income public schools (Hu and Harris, 2018).[56] Recently, the city has suspended the use of academic screens for the 2020-2021 application cycle, as some of the measures used — including grades, test scores, attendance, student interviews, and auditions for performing arts programs — were upended by the Covid-19 pandemic.

Despite the public emphasis, dropping all academic screens would only slightly decrease city-wide school segregation, as shown in Figure 7. In the absence of academic screens, economic and racial school segregation would be 1 percentage point lower, 4% instead of 5% and 13% instead of 14%, respectively.[57] These moderate drops in segregation are consistent with the ones observed during the 2020-2021 application cycle, shown in appendix Figure A7, although these changes might also have been affected by the pandemic.

Figure 7 also displays segregation levels for a simulation with no academic screens in which applicants rank programs randomly with unrestricted list size. This simulation provides a benchmark for the contribution of academic admission criteria to school segregation, given additional non-academic school preferences embedded in program priorities and eligibility criteria. Random application behavior effectively removes the demand-side contribution to school segregation, isolating the effect of non-academic school preferences. In this benchmark scenario school segregation closely matches residential segregation at the school district level, which is a reflection of 83% of programs being reserved for district applicants.[58] Relative to this benchmark, dropping academic screens would achieve between a quarter and a third of the possible drop in segregation.

As academic screens are not equally prevalent in all NYC districts, these city-wide figures could aggregate substantial heterogeneity. Figure 8 compares the effect of dropping admission screens on district-level school segregation in NYC boroughs.[59] In Brooklyn and Manhattan, dropping admission screens would almost halve the district-level economic and racial school segregation. For instance, a low-income student in Manhattan would attend a school 4% more segregated than her district in the absence of academic admission criteria, compared to 6% currently. Similarly, district-level low-income school segregation would drop from 3% to 1% in Brooklyn. On the other hand, dropping academic screens would have almost no effect in the Bronx or Queens.

5.3 Effect of dropping all admission screens

Results from the previous section highlighted that city-wide school segregation is largely determined by the current system of geographic eligibility and priority rules. Hence, the second counterfactual simulates a match that drops all geographic priorities and eligibility criteria, in addition to academic screens. As such, this second scenario is akin to a city-wide

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

[57] Appendix Figure A2 shows similar effects on school segregation after enrollment.

[58] As a result, 16.6% of applicants are offered a school in their district of residence in the simulation.

[59] The exhibit omits Staten Island, which enrolls less than 5% of NYC students.

lottery-based admission, in which only gender-specific and language-specific programs are able to select students.[60] Figure 10 compares school segregation under this counterfactual admission regime to the observed levels of segregation. As a benchmark, the figure also displays school segregation if applicants rank programs randomly under the same admission regime.[61]

Geographic eligibility criteria and priorities appear to play a more substantial role in city-wide school segregation. Dropping geographic barriers in addition to academic screens would halve school segregation, as measured by the segregation index, for all demographic groups. For instance, low-income student segregation would drop from 5% to 2%, while minority student segregation would decrease from 14% to 9%. The largest drop in school segregation would be for Black students, whose segregation index would fall from 26% to 13%.[62] Appendix Figure A8 shows that all NYC boroughs would be impacted by the change in admission regime, with the Bronx nonetheless experiencing the smallest decreases in school segregation.

5.4 What drives school segregation?

School segregation does not drop to the corresponding benchmark in either counterfactual exercise. The residual differences correspond to the contribution of applicant preferences to observed school segregation. Applicant preferences may increase school segregation either because of differences in taste for specific schools and peers or because of the combination of preference for nearby schools and residential segregation. As a last step, I attempt to disentangle these two channels.

Preference for nearby schools is reflected by the fact that even in the absence of admission barriers, 50% and 84% of applicants still attend a school within their district and borough, respectively. To quantify the contribution of residential segregation combined with preference for nearby schools to school segregation, I simulate a match in which applicants do not take distance into account school when forming their rank-order lists.[63] Results from this simulation compared to previous estimates are displayed in Figure 7.

The level of segregation for students indifferent to school distance falls midway between that of the two other counterfactual simulations and their corresponding benchmarks. This suggests that the combination of preference for nearby schools and residential segregation explains approximately half of the residual school segregation. Consequently, applicant preferences for similar peers and for specific schools account for only a quarter of the observed school segregation.

[60] In the simulation, applicants are eligible to a dual language program only if they applied to the program in the actual match.

[61] The benchmark levels of school segregation are slightly larger than zero, reflecting the remaining eligibility criteria as well as sample restrictions in the simulation.

[62] Appendix Figure A2 shows similar effects on school segregation after enrollment.

[63] In practice, this corresponds to setting the coefficient on distance to zero when computing updated utilities for schools in order to predict optimal rank-order lists under the new admission regime.

6 Conclusion

School district leadership, politicians, and the public are increasingly concerned that admission schemes are responsible for school segregation. To understand the role of school admission criteria in determining segregation, I studied two middle school admission reforms in NYC northwest Brooklyn and upper west side districts. My analysis of these district-level reforms showed that reducing academic screening can lead to a decrease in segregation. Nonetheless, the impact of such reforms hinges on student responses both at the enrollment and application stages. In particular, I found that the NYC admission reforms entailed an increase in white and high-income student exit from the traditional public school system, which partially offset the effects on segregation. Interestingly, this "white flight" appears to be driven by an increase in exposure to lower-achieving peers, rather than to racial minorities. Changes in application behavior in response to the reforms, on the other hand, reinforced rather than diminished the their effects.

The magnitude of the behavioral responses I documented highlights the need to understand applicant decision-making as a first step toward evaluating alternative admission regimes. To fully characterize the application response, I developed a model of school choice which allows for strategic behavior but avoids the curse of dimensionality and can thus be estimated in settings with many schools, like NYC. This model rationalizes two important aspects of applicants' behavior. First, applicants submit shorter lists than necessary. Second, they omit schools to which they are unlikely to be admitted.

Based on the model's estimates, I simulate the effect of alternative city-wide admission regimes in NYC. The simulations suggest that academic screens play only a modest role in city-wide school segregation. The role played by geographic screens, however, is substantial. The simulation exercise also shows that demand-side factors account for around half of school segregation. Thus, policy makers might find it hard to reduce school segregation in NYC by more than 50% from its current level, as at least half of the demand-induced school segregation is driven by residential sorting.

This paper highlights the limit of what admission reforms can accomplish as their impact is limited by demand-side factors. An open direction of research would consist in understanding how family preferences for schools are formed and whether these preferences can be affected through policies. Can family preferences be affected by changing the information provided at the time of application? Since students apply to school several time over their school life, do previous assignments affect expressed preferences when applying to later grades?

References

Abadie, A. (2021). Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*, 59(2):391–425.

Abadie, A., Athey, S., Imbens, G., and Wooldridge, J. (2017). When Should You Adjust Standard Errors for Clustering? Technical Report w24003, National Bureau of Economic Research, Cambridge, MA.

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105(490):493–505.

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

Abdulkadiroğlu, A. and Sönmez, T. (2003). School Choice: A Mechanism Design Approach. *American Economic Review*, 93(3):729–747.

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

Agarwal, N. and Somaini, P. (2020). Revealed Preference Analysis of School Choice Models. *Annual Review of Economics*, 12:471–501.

Ajayi, K. and Sidibe, M. (2020). School Choice Under Imperfect Information. *SSRN Electronic Journal*.

Ali, S. N. and Shorrer, R. I. (2021). The College Portfolio Problem.

Angrist, J. D. and Lang, K. (2004). Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *The American Economic Review*, 94(5):1613–1634.

Arteaga, F., Kapor, A., Neilson, C., and Zimmerman, S. (2021). Smart Matching Platforms and Heterogeneous Beliefs in Centralized School Choice. Technical Report w28946, National Bureau of Economic Research, Cambridge, MA.

Artemov, G., Che, Y.-K., and He, Y. (2017). Strategic 'Mistakes': Implications for Market Design Research. page 59.

Azevedo, E. M. and Leshno, J. D. (2016). A Supply and Demand Framework for Two-Sided Matching Markets. *journal of political economy*, page 34.

Barry, E. (2021). Boston Overhauls Admissions to Exclusive Exam Schools. *The New York Times*.

Bergman, P. (2018). The Risks and Benefits of School Integration for Participating Students: Evidence from a Randomized Desegregation Program. *Working paper*.

Billings, S., Chyn, E., and Haggag, K. (2020). The Long-Run Effects of School Racial Diversity on Political Identity. Technical Report w27302, National Bureau of Economic Research, Cambridge, MA.

Billings, S. and Hoekstra, M. (2019). Schools, Neighborhoods, and the Long-Run Effect of Crime-Prone Peers. Technical Report w25730, National Bureau of Economic Research, Cambridge, MA.

Billings, S. B., Deming, D. J., and Rockoff, J. (2014). School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg*. *The Quarterly Journal of Economics*, 129(1):435–476.

Bjerre-Nielsen, A. and Gandil, M. H. (2020). Attendance Boundary Policies and the Limits to Combating School Segregation.

Boisjoly, J., Duncan, G. J., Kremer, M., Levy, D. M., and Eccles, J. (2006). Empathy or Antipathy? The Impact of Diversity. *American Economic Review*, 96(5):1890–1905.

Calsamiglia, C., Fu, C., and Güell, M. (2020). Structural Estimation of a Model of School Choices: The Boston Mechanism versus Its Alternatives. *Journal of Political Economy*, 128(2):642–680.

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

Carrell, S. E., Hoekstra, M., and West, J. E. (2019). The Impact of College Diversity on Behavior toward Minorities. *American Economic Journal: Economic Policy*, 11(4):159–182.

Chade, H. and Smith, L. (2006). Simultaneous Search. *Econometrica*, 74(5):1293–1307.

Cohen, D. (2021). NYC School Segregation Report Card: Still Last, Action Needed Now.

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

Corcoran, S. P., Jennings, J. L., Cohodes, S. R., and Sattin-Bajaj, C. (2018). Leveling the playing field for high school choice: Results from a field experiment of informational interventions. Technical report, National Bureau of Economic Research.

Fack, G., Grenet, J., and He, Y. (2019). Beyond Truth-Telling: Preference Estimation with Centralized School Choice. page 79.

Gale, D. and Shapley, L. S. (1962). College Admissions and the Stability of Marriage. *American Mathematical Monthly*, 69(1):9–15.

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

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

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

Hu, W. and Harris, E. A. (2018). A Shadow System Feeds Segregation in New York City Schools. *The New York Times*.

Johnson, R. (2011). Long-run Impacts of School Desegregation & School Quality on Adult Attainments. Technical Report w16664, National Bureau of Economic Research, Cambridge, MA.

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

Kapor, A. J., Neilson, C. A., and Zimmerman, S. D. (2020). Heterogeneous Beliefs and School Choice Mechanisms. *American Economic Review*, 110(5):1274–1315.

Karp, S. (2021). Big changes in how students are picked for CPS' elite high schools start today. *WBEZ Chicago*.

Kennan, J. and Walker, J. R. (2011). The Effect of Expected Income on Individual Migration Decisions. *Econometrica*, 79(1):211–251.

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

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

Luflade, M. (2017). The value of information in centralized school choice systems. page 38.

Lutz, B. (2011). The End of Court-Ordered Desegregation. *American Economic Journal: Economic Policy*, 3(2):130–168.

MacKinnon, J. G. and Webb, M. D. (2017). Wild Bootstrap Inference for Wildly Different Cluster Sizes. *Journal of Applied Econometrics*, 32(2):233–254.

MacKinnon, J. G. and Webb, M. D. (2018). The wild bootstrap for few (treated) clusters. *The Econometrics Journal*, 21(2):114–135.

MacKinnon, J. G. and Webb, M. D. (2020). Randomization inference for difference-in-differences with few treated clusters. *Journal of Econometrics*, 218(2):435–450.

Margolis, J., Dench, D., and Hashim, S. (2020). The Impact of Middle School Integration Efforts on Segregation in Two New York City Districts. page 44.

McCulloch, R. and Rossi, P. E. (1994). An exact likelihood analysis of the multinomial probit model. *Journal of Econometrics*, 64(1-2):207–240.

Mezzacappa, D. (2021). Philly overhauls selective admissions policy in bid to be antiracist. *Chalkbeat*.

Pathak, P. A. and Sönmez, T. (2013). School Admissions Reform in Chicago and England: Comparing Mechanisms by their Vulnerability to Manipulation. *American Economic Review*, 103(1):80–106.

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

Reardon, S. F., Grewal, E. T., Kalogrides, D., and Greenberg, E. (2012). 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.

Reardon, S. F., Matthews, S. A., O'Sullivan, D., Lee, B. A., Firebaugh, G., Farrell, C. R., and Bischoff, K. (2008). The geographic scale of Metropolitan racial segregation. *Demography*, 45(3):489–514.

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

Rees-Jones, A. and Skowronek, S. (2018). An experimental investigation of preference misrepresentation in the residency match. *Proceedings of the National Academy of Sciences*, 115(45):11471–11476.

Son, S. J. (2020). Distributional Impacts of Centralized School Choice. page 62.

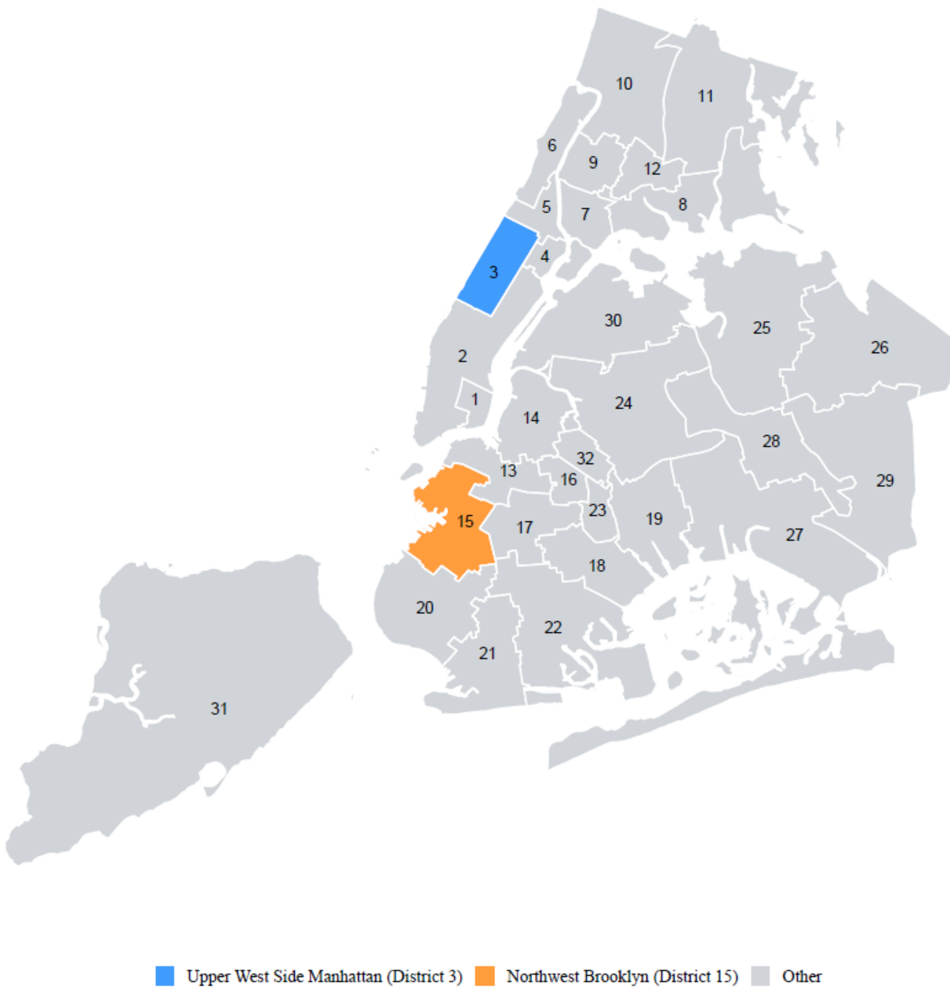
Vigdor, J. and Ludwig, J. (2007). Segregation and the Black-White Test Score Gap. Technical Report w12988, National Bureau of Economic Research, Cambridge, MA.

Warikoo, N. (2021). Elite public schools won't become more diverse just by ditching exams. *The Washington Post*.

Zebrowitz, L. A., White, B., and Wieneke, K. (2008). Mere Exposure and Racial Prejudice: Exposure to Other-Race Faces Increases Liking for Strangers of that Race. *Social Cognition*, 26(3):259–275.

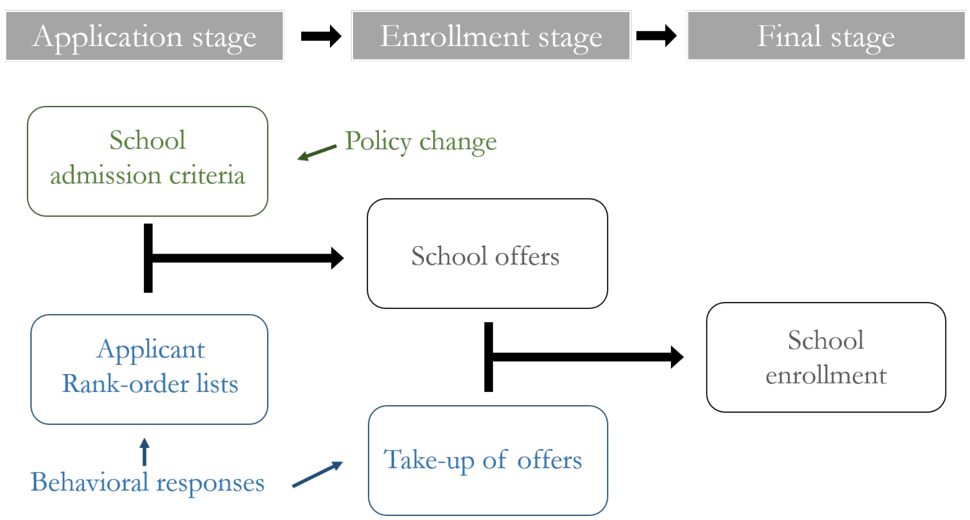
7 Figures

Figure 1: NYC's Local School Districts



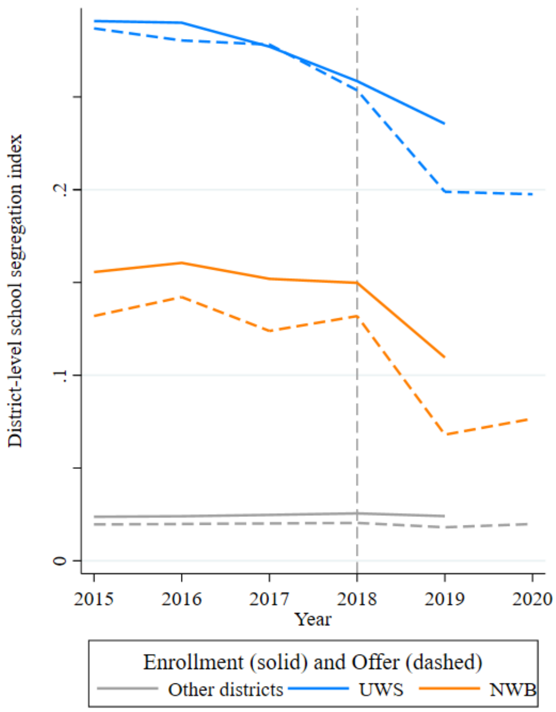
Note: This map shows NYC's 32 local school districts. The two districts that implemented middle school integration reforms in 2019 are highlighted in color.

Figure 2: From Admission Criteria to Final Enrollment

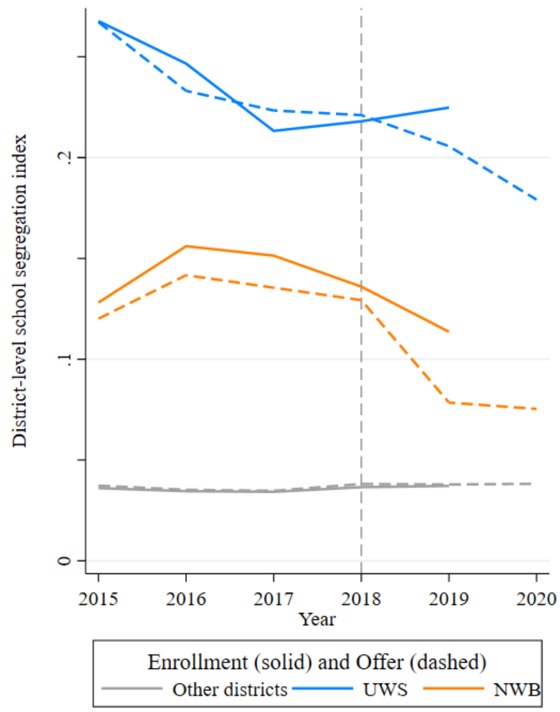


Note: This flow chart describes the mechanisms through which a change in admission criteria affects students' final enrollment and by extension school segregation.

Figure 3: Evolution of School Segregation Indices for Low-Income and Minority Students



(a) Low-income students

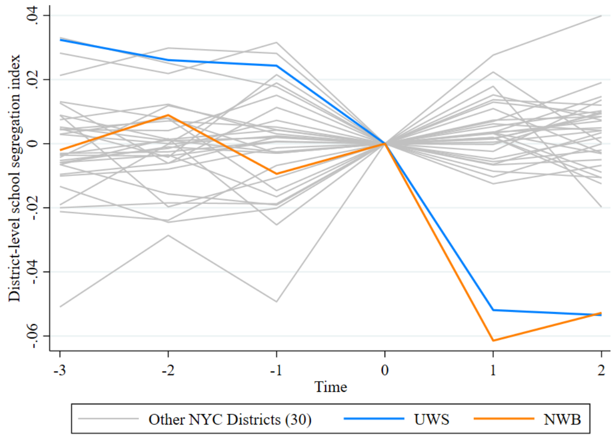


(b) Minority students

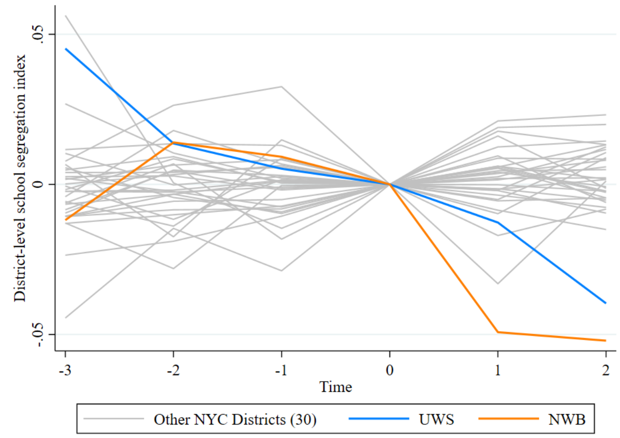
Note: These figures plot the evolution of district school segregation indices for UWS, NWB, and other NYC districts between 2015 and 2020. Figure (a) displays the segregation index for applicants classified as disadvantaged by the DOE, a proxy for low-income status. Figure (b) displays the index for Black and Hispanic applicants. To obtain the plotted values, school segregation indices are standardized by the share of students belonging to the group among applicants residing in the district. The school segregation index for other NYC districts corresponds to the weighted average of district-level indices, 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 school segregation index at the offer stage, that is if all students were to enroll in the school they were offered in the match. Solid lines correspond to the value of the index after enrollment. Students that leave the NYC public school system are included in the standardization.

Figure 4: Comparison of Changes in School Segregation Indices in all NYC Districts

Panel A: Offered School

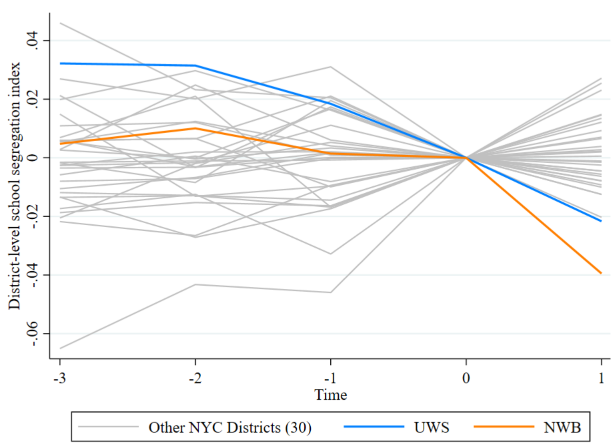


(a) Low-income students

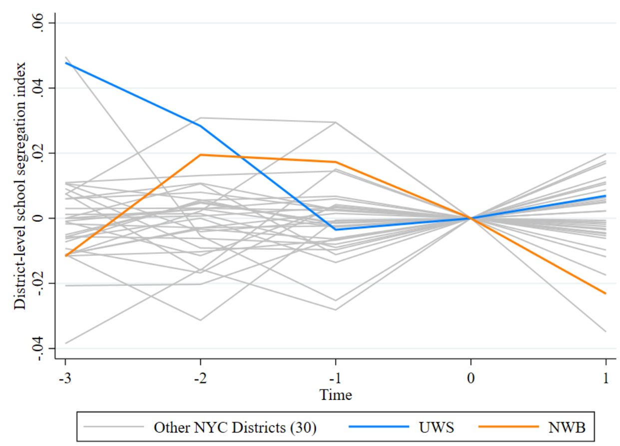


(b) Minority students

Panel B: School of Enrollment



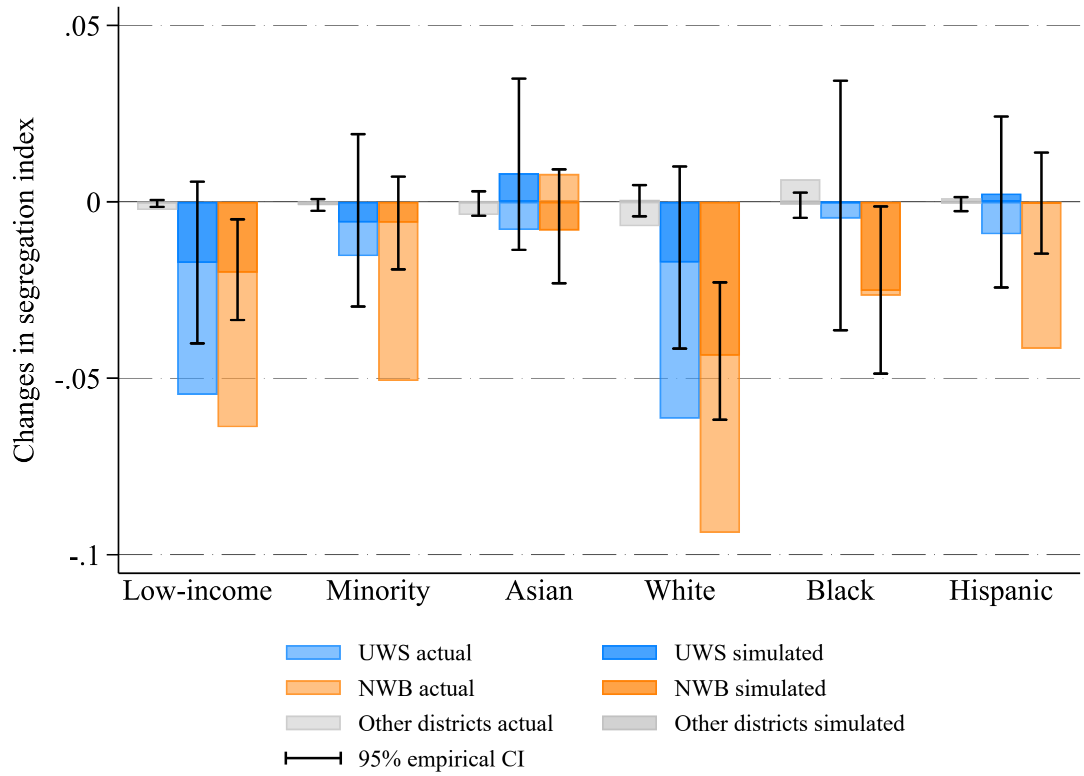
(c) Low-income students



(d) Minority students

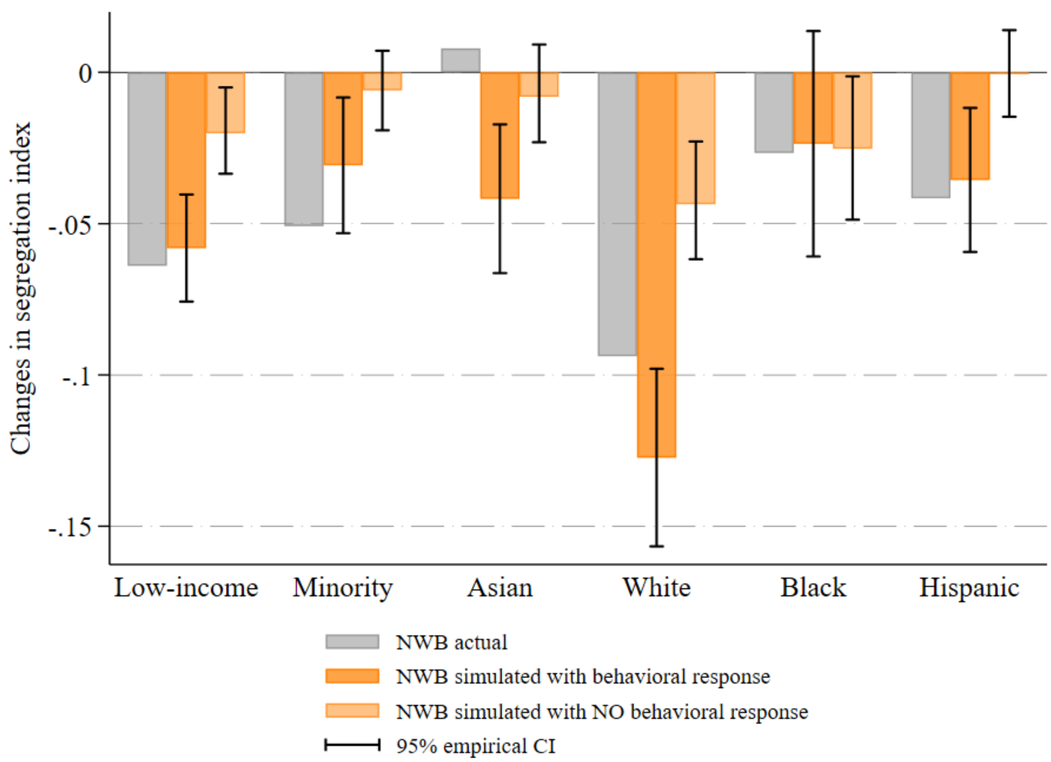
Note: These figures plot changes in the district-level school segregation indexes with respect to 2018 for all NYC districts. In Panel A, SIs are computed using the school offered. In Panel B, SIs are computed using the school of enrollment. Each plot only includes districts in which at least 5% of the student population belongs to the group considered.

Figure 5: Decomposition of Plan Effects on School Segregation



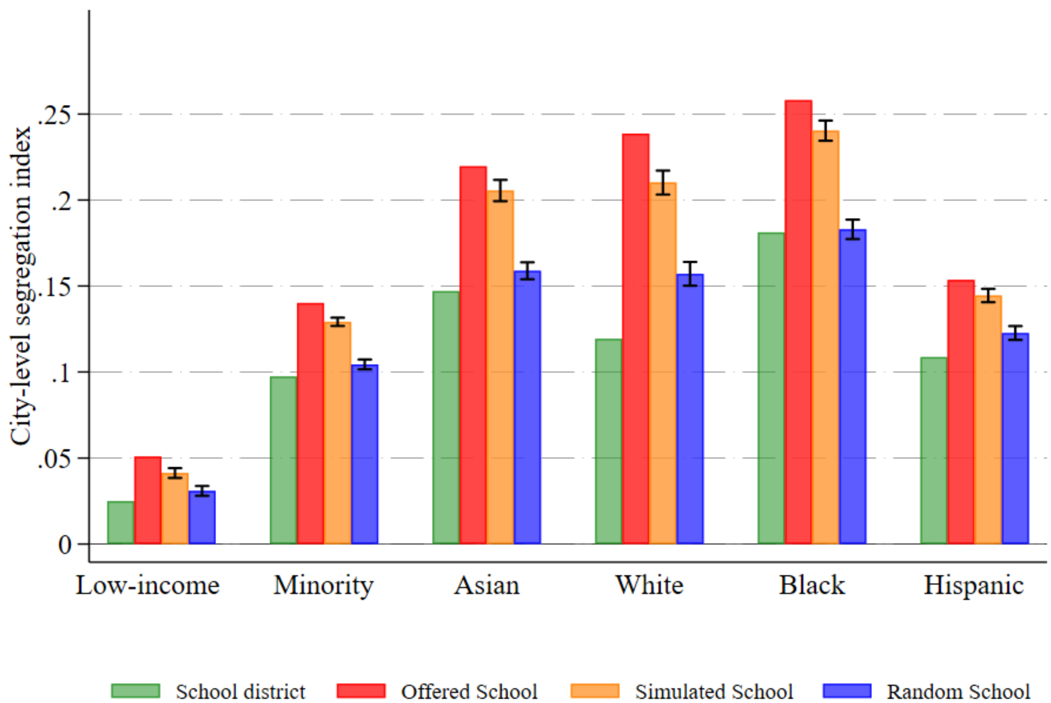
Note: This bar chart plots the changes in district school segregation indices at the match offer stage between 2018 and 2019. Each bar gives the observed change in the index between 2018 and 2019. The shaded part of each bar corresponds to the effect of the integration policies absent behavioral response,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

Figure 6: Replication of the NWB Plan's Effect on Segregation



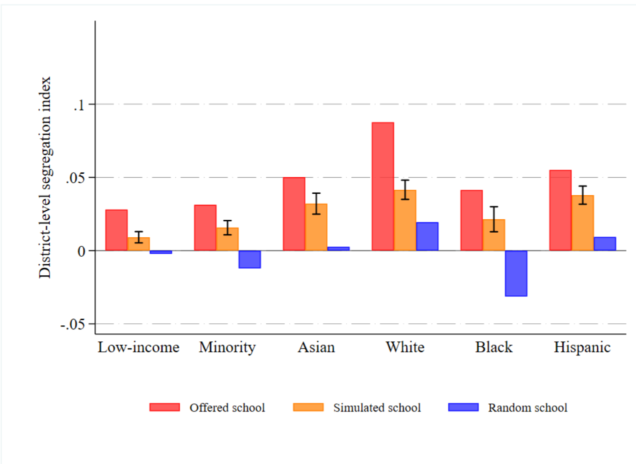
Note: This bar chart plots the changes in NWB school segregation indices at the match offer stage between 2018 and 2019, and between 2018 and two simulations of the 2018 match under the NWB integration reform. The first simulation updates 2018 applicant ROLs according to the model of list formation and preference estimates. The second simulation does not update 2018 applicant ROLs, as in Figure 5. The 95% confidence intervals for the simulated effects of the integration reforms are computed by simulating 50 times the 2018 match. For each simulation, a new sequence of tiebreakers is drawn and applicants are sampled with replacement, through sampling stratified by SWD indicator and district. The first simulation which updates applicant ROLs is iterated 5 times to obtain the equilibrium levels of segregation.

Figure 7: Effect of Dropping Academic Admission Criteria on School Segregation

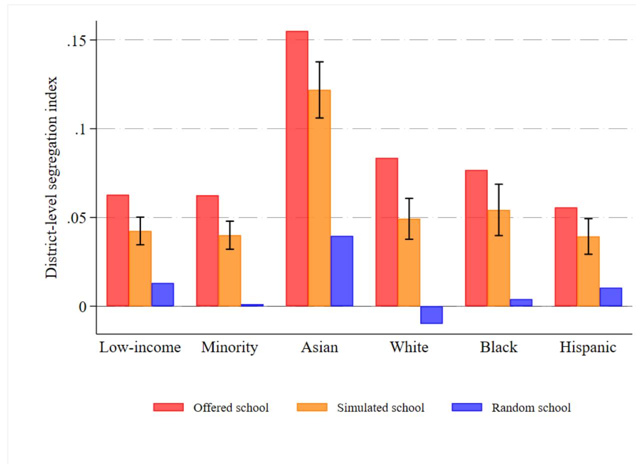


Note: This bar chart plots segregation indices in 2018 for the school district and for offered schools under different admission criteria. The offered school corresponds to the school students were offered in the 2018 match. The simulated school corresponds to the school students are offered in a simulation that drops all academic screens. The random school corresponds to the school students are offered in a simulation that drops all academic screens and where students rank programs randomly. For both simulations, students can only apply to programs to which they are eligible in the 2018 match.

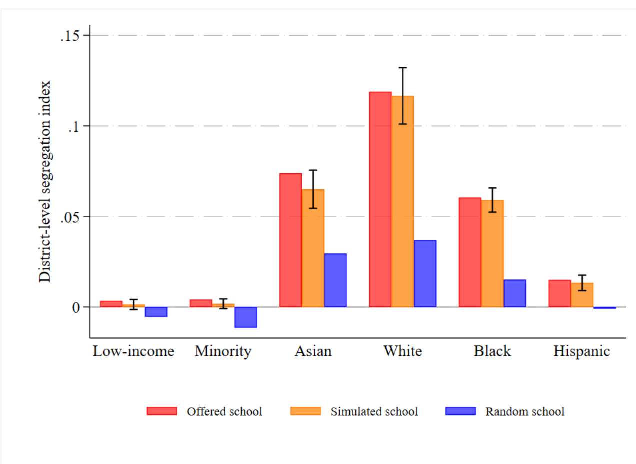
Figure 8: Effect of Dropping Academic Admission Criteria on District-Level School Segregation



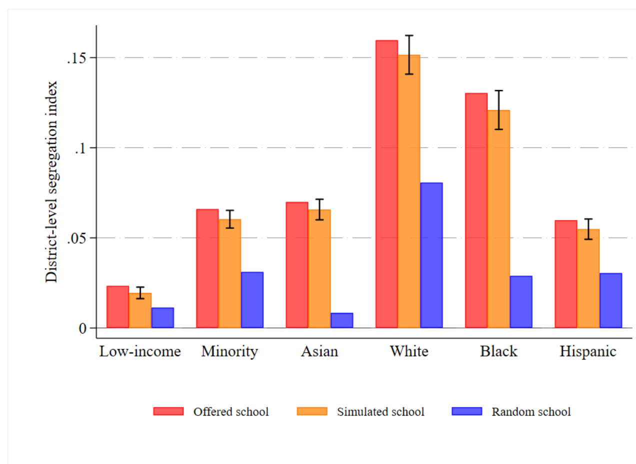
(a) Brooklyn



(b) Manhattan



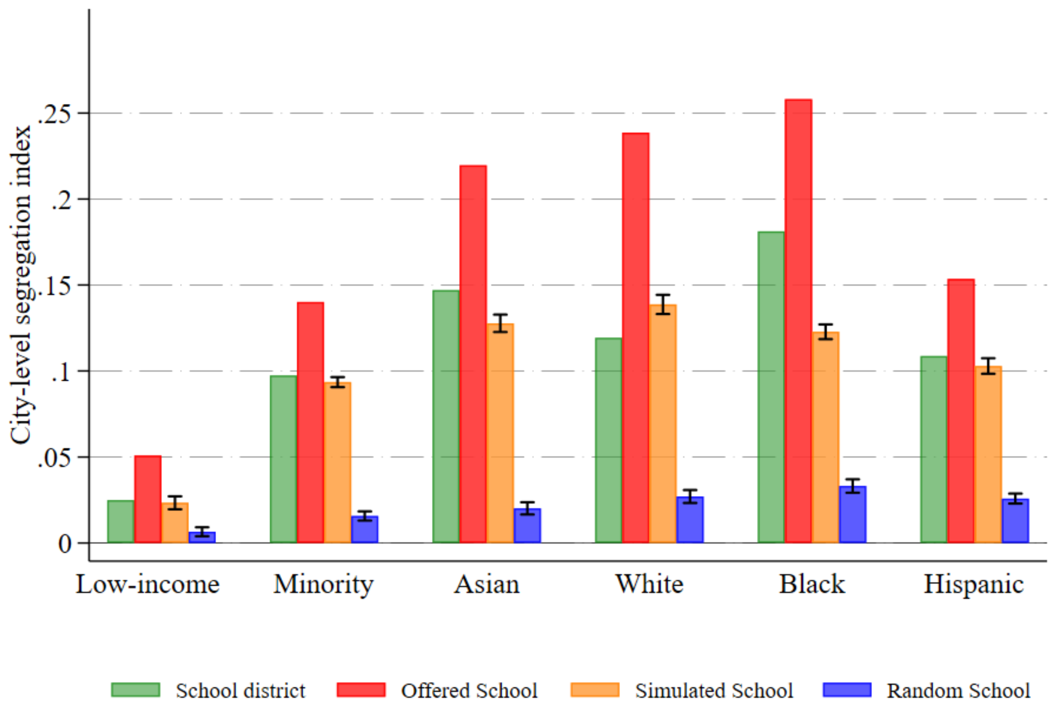
(c) Bronx



(d) Queens

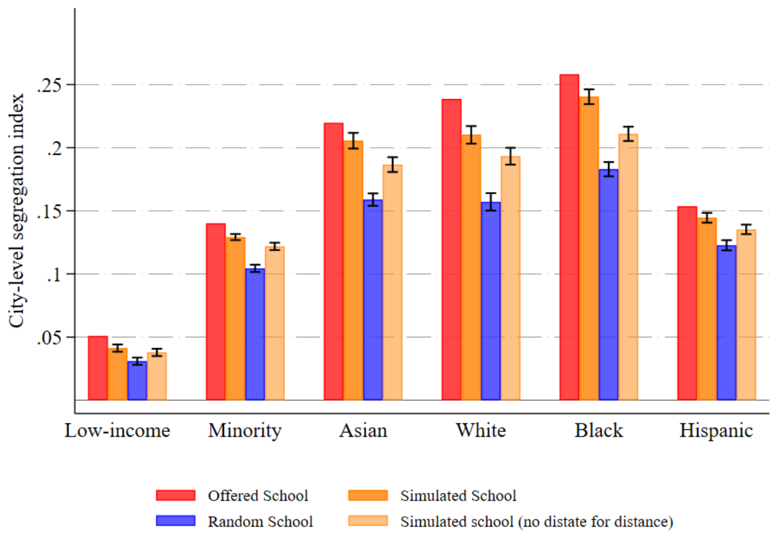
Note: This bar chart plots district-level school segregation indices in 2018 under the actual match, a match that drops all academic screens, and a match that drops all academic screens and where students rank programs randomly.

Figure 9: Effect of Dropping Academic and Geographic Admission Criteria on School Segregation

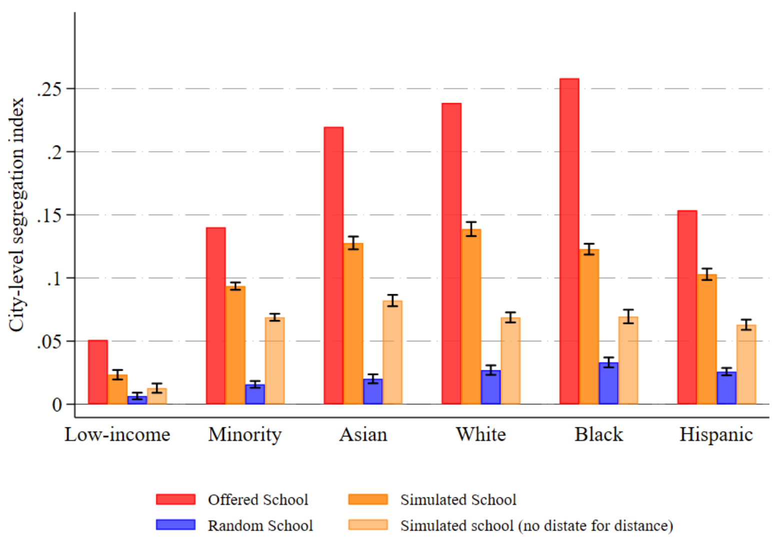


Note: This bar chart plots segregation indices in 2018 for the school district and for offered schools under different admission criteria. The offered school corresponds to the school applicants were offered in the 2018 match. The simulated school corresponds to the school applicants are offered in a simulation that drops all academic screens and all geographic eligibilities and priorities. The random school corresponds to the school applicants are offered in a simulation that drops all academic screens and all geographic eligibilities and priorities, and where applicants rank programs randomly. For both simulations, applicants can apply to any program in the 2018 match except dual-language programs and gender-specific programs, for which they need to be eligible.

Figure 10: Effect of Counterfactual Admissions Criteria on School Segregation



(a) No academic screens



(b) No academic and geographic screens

Note: This bar chart plots the school segregation index in 2018 under different admission criteria and applicant preferences. The offered school corresponds to the school applicants were offered in the 2018 match. The simulated school corresponds to the school applicants are offered in a simulation that drops all academic screens (Figure (a)) or all academic screens and geographic eligibilities and priorities (Figure (b)). The random school corresponds to the simulated school when applicants rank programs randomly. The simulated school with no distaste for distance corresponds to the simulated school when applicants do not account for distance when ranking schools.

8 Tables

Table 1: NYC Segregation in 2018

		Segregation index			
	City-wide shares (1)	Census tract (2)	School district (3)	Offered school (4)	Enrolled school (5)
Black	0.23	0.32	0.18	0.25	0.25
Hispanic	0.41	0.18	0.11	0.15	0.16
Black + Hispanic	0.64	0.17	0.10	0.14	0.14
Asian	0.18	0.30	0.15	0.22	0.23
White	0.16	0.33	0.12	0.24	0.25
Low-income	0.75	0.07	0.03	0.05	0.06
N units	1	2,099	32	479	479
N students per unit	73,600	35	2,300	154	136

Note: This table reports 2018 segregation indices for different demographic groups at different geographic levels in NYC. The sample is restricted to 6^{th} grade applicants offered or enrolled in match schools who have non-missing demographic information. Column (1) reports the share of each group among NYC middle school applicants and enrolled students. Columns (2) and (3) report residential segregation at the census tract level and school district level. Columns (4) and (5) report school segregation for middle school match offers and middle school enrollment. Column (4) considers the school the student is offered in the match when available (92% of cases) and inputs the enrolled school when missing. Column (5) only considers the school the student enrolls in; students that do not enroll in match schools are dropped.

Table 2: Characteristics of UWS and NWB

	NYC (1)	UWS (2)	NWB (3)
Panel A: Characteristics of schools and applicants			
N middle schools	479	20	11
N programs	686	21	13
% screened programs	33%	57%	80%
N applicants	71,512	1,134	2,536
% 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
% Low-income 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: School segregation index at the district-level			
Black + Hispanic	0.04	0.22	0.14
White	0.12	0.12	0.12
Low-income	0.03	0.26	0.15

Note: This table presents the characteristics of NYC school districts. Column (1) includes all NYC middle school applicants, while columns (2) and (3) only include applicants residing in UWS and NWB, respectively, at the time of application. Panel A describes the population of middle school applicants that enroll in 6^{th} grade in year 2018. Mean math and English proficiency are computed based on the proficiency level obtained in 4^{th} grade state tests. Panel B presents the district-level school segregation index for three groups of applicants (Black and Hispanic, White, and Low-income). School segregation indices are computed for each NYC district separately, standardizing by the group share among applicants residing in the district. The average school segregation index for NYC corresponds to the average of district-level school segregation indices, weighted by the share of NYC students of each group living in each district.

Table 3: DiD Estimates of Reform Effects on Public School Enrollment

	All (1)	Black (2)	Hispanic (3)	Asian (4)	White (5)	Low-income (6)	High-income (7)
	Panel A: Effects on exit from traditional public system						
UWS × 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)
NWB × 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 UWS pre-2019	0.10	0.13	0.07	0.12	0.10	0.09	0.11
Mean NWB pre-2019	0.12	0.18	0.09	0.07	0.18	0.09	0.17
	Panel B: Effects on offered peer mean math score						
UWS × 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)
NWB × 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 UWS pre-2019	0.37	-0.16	0.05	0.81	0.77	-0.11	0.72
Mean NWB pre-2019	0.31	0.19	0.11	0.36	0.550	0.15	0.51
N	332,488	73,261	136,303	60,515	56,173	239,494	92,997

Note: This table reports difference-in-differences estimates of integration reform effects for 2015-2019 middle school applicants. In all panels, the endogeneous variable is the interaction of dummies for residing in UWS or NWB and applying for admission in 2019. Panel A's dependent variable is a dummy equal to one for applicants that do not enroll in a NYC public school at any point in the school year following middle school admission. Panel B's dependent variable is the leave-out mean 5^{th} grade math test score among applicants offered the same school in 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 are reported in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. The corresponding permutation tests are reported in A9 and A10.

Table 4: 2SLS Estimates of Offered Peer Characteristics on Exit from Traditional Public Schools

	UWS (1)	NWB (2)	UWS and NWB (3)	UWS and NWB (4)	UWS and NWB (5)	UWS and NWB (6)
Peer achievement	-0.27***	-0.29***	-0.28***	-0.30***	-0.36***	
	(0.06)	(0.05)	(0.04)	(0.05)	(0.08)	
Proportion minority				-0.08		-0.41**
				(0.11)		(0.18)
Proportion low-income					-0.16	0.70***
					(0.14)	(0.15)
First stage F	19.9	74.0	44.7	29.1	12.7	14.0
Overid p-value	0.26	0.13	0.17	0.14	0.18	0.01
Overid DF	8	8	17	16	16	16
N	319,867	327,000	332,488	332,488	332,488	332,488

Note: This table reports alternative IV estimates of the effects of offered peer characteristics on the probability of exiting the traditional public school system for 2015-2019 middle school applicants. Peer achievement corresponds to the leave-out mean 5^{th} grade math test score among applicants offered the same school in the match. Proportion minority and proportion low-income correspond to the equivalent leave-out means for the respective characteristics. Estimates are computed by instrumenting potential peer characteristics by UWS and NWB integration reforms plus interactions with covariates (English learner, race, and interactions of low-baseline test scores with low-income status). 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%.

Table 5: DiD Estimates of Changes in Applicant ROLs

	All	Low-income		High-income	
		Low-baseline	High-baseline	Low-baseline	High-baseline
	(1)	(2)	(3)	(4)	(5)
		Panel A: Length of the rank-order list submitted			
UWS × 2019	1.01***	0.69***	1.18***	0.50	1.07***
	(0.09)	(0.21)	(0.24)	(0.36)	(0.12)
UWS × 2020	1.20***	0.39**	1.00***	0.93***	1.53***
	(0.09)	(0.19)	(0.28)	(0.35)	(0.13)
NWB × 2019	2.56***	2.07***	1.89***	2.74***	3.45***
	(0.07)	(0.13)	(0.14)	(0.24)	(0.12)
NWB × 2020	2.84***	1.59***	1.58***	3.84***	4.58***
	(0.07)	(0.12)	(0.14)	(0.24)	(0.11)
Mean UWS pre-2019	4.0	4.5	4.6	4.3	3.5
Mean NWB pre-2019	4.9	4.3	5.1	4.7	5.3
N	396,958	171,653	94,672	37,999	63,778
		Panel B: Mean math score in school ranked first			
UWS × 2019	0.05**	0.06	0.09	0.10**	-0.02
	(0.02)	(0.05)	(0.06)	(0.05)	(0.02)
UWS × 2020	0.11***	0.12***	0.09	0.08	0.01
	(0.02)	(0.05)	(0.07)	(0.07)	(0.02)
NWB × 2019	-0.02**	0.06***	0.00	0.02	-0.10***
	(0.011)	(0.02)	(0.03)	(0.03)	(0.02)
NWB × 2020	-0.03**	0.05***	0.03	0.00	-0.12***
	(0.01)	(0.02)	(0.03)	(0.03)	(0.02)
Mean UWS pre-2019	0.51	-0.02	0.49	0.33	0.92
Mean NWB pre-2019	0.55	0.23	0.67	0.46	0.86
N	396,656	171,523	94,618	37,973	63,724

Note: This table reports difference-in-differences estimates of integration reform effects for 2015-2020 middle school applicants. In all panels, the endogeneous variables are the interactions of dummies for residing in UWS and NWB and applying for admission in 2019 and 2020. Panel A's dependent variable is a count variable that indicates the length of applicant ROLs. Panel B's dependent variable is the 5^{th} grade mean math test score of students enrolling between 2015 and 2018 for the school ranked first by each applicant. 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 are reported in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. The corresponding permutation tests are reported in A11 and A12.

Table 6: Model Estimates of Preference for Time-Varying School Characteristics

		Strategic reports model with cost (1)	Truthful reports model without cost (2)
School char. (z_{stk})	**Student's cell** ($c(z_i)$)		
Minority share	low math × minority	1.40***	-0.99**
		(0.59)	(0.48)
	high math × minority	3.89***	-1.19***
		(0.59)	(0.48)
	low math × non minority	0.31	-2.50***
		(0.76)	(0.61)
	high math × non minority	2.27***	-2.96***
		(0.37)	(0.28)
High math share	low math × minority	2.49***	-1.64*
		(1.10)	(0.84)
	high math × minority	7.99***	-0.55
		(1.37)	(1.04)
	low math × non minority	0.57	-4.44***
		(1.42)	(1.12)
	high math × non minority	9.90***	-0.57
		(0.77)	(0.54)
N students		2,343	2,343
N programs		60	60

Note: This table reports estimates of applicant preferences for time-varying school characteristics for a model that includes a per-unit cost and a model that set this cost to zero. Results are reported for the sample of middle school applicants residing in NWB and attending a NWB elementary school in 2018 and 2020. Students are assigned to one of four cells depending on their minority and math achievement status. Minority students include Hispanic and Black students. High math students include students who scored more than 3 (NYC proficiency level) on their 4^{th} grade math state exam. School minority and high math shares correspond to the shares of minority and high math students that were enrolled at the school the year prior. Reported estimates are based on the last 50,000 out of 100,000 draws. Standard errors are reported in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: Model Estimates of Preference for Schools and per-unit cost

Parameter	Student's cell ($c(z_i)$)	Strategic reports model with cost	Truthful reports model without cost
Mean program utility	low math × minority	0.82	-0.65
	high math × minority	0.97	-0.63
	low math × non minority	0.33	-1.37
	high math × non minority	-0.27	-2.00
SD program utility	low math × minority	2.94	2.70
	high math × minority	3.33	3.09
	low math × non minority	3.86	3.55
	high math × non minority	4.29	3.98
Corr. program utilities	cell 1 VS cell 4	0.73	0.58
Mean cost		0.0054	
SD cost		0.0001	
% with cost < 0.01		0.76	
Variance utility error (σ^2_ε)		8.78	10.64
Variance cost error (σ^2_ζ)		0.0001	
N students		47,226	47,226
N programs		645	645

Note: This table reports preference parameters for a model that includes a per-unit cost and a model that sets this cost to zero. Results are reported for the sample of all non-SWD NYC middle school applicants in 2018 with baseline covariates. Students are assigned to one of four cells depending on their minority and math achievement status. Minority students include Hispanic and Black students. High math students include students who scored more than 3 (NYC proficiency level) on their 4^{th} grade math state exam. School minority and high math shares correspond to the shares of minority and high math students that were enrolled at the school the year prior. Reported estimates are based on the last 50,000 out of 100,000 draws. Standard errors are reported in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%.

APPENDIX

A Proof of Proposition 1

As the order of programs in applicant ROL determines offer probabilities, I explicitly define notation which indicates the position of each program in the ROL. Specifically, I denote the rank of program s in applicant i ROL by r_{is}. If program s is not listed, $r_{is} = \infty$. I also define R_{ik} as applicant i's rank-order list up to the k^{th} position. Thus, $r_{i0} = |R|+1 = \underline{k}$ denotes the rank of the outside option. Given this notation, $p_{isk}\left(R_{k-1}\right)$ denotes the probability applicant i is offered program s when ranked in position k given choices for previous positions.

Under Assumption 1, and for any cost $c \geq 0$, the rank-order list R chosen by applicant i maximizes her overall expected utility $V_i(R)$ if and only if the following conditions hold:

1. For listed programs $(r_{is} = k \neq \infty)$:

   $v_{is} \geq \dfrac{c_i}{p_{isk}\left(R_{k-1}\right)},$

   $v_{is} \geq v_{ij} \quad \forall j$ s.t. $(r_{ij} \neq \infty$ and $r_{is} < r_{ij})$ or $(r_{ij} = \infty$ and $v_{ij} \geq \dfrac{c_i}{p_{ijk}\left(R_{k-1}\right)}),$

   $v_{is} \leq v_{ij} \quad \forall j$ s.t. $r_{is} > r_{ij}$.

2. For unlisted programs $(r_{is} = \infty)$:

   $v_{is} \leq \dfrac{c_i}{p_{is\underline{k}}\left(R_{\underline{k}-1}\right)},$

   $v_{is} \leq \max(v_{ij}, \dfrac{c_i}{p_{isk}\left(R_{k-1}\right)}) \quad \forall j$ s.t. $r_{ij} = k < \infty$.

3. For the cost c_i :

   $c_i \geq 0,$

   $c_i \leq p_{ijk}\left(R_{k-1}\right) v_{ij} \quad \forall j$ s.t. $r_{ij} = k < \infty,$

   $c_i \geq v_{ij} p_{ij\underline{k}}\left(R_{\underline{k}-1}\right) \quad \forall j$ s.t. $r_{ij} = \infty,$

   $c_i \geq v_{ij} p_{ijk}\left(R_{k-1}\right) \quad \forall j$ s.t. $r_{ij} = \infty$ and $v_{ij} > v_{ij'}$ where $r_{ij'} = k$.

When both c_i and $p_{isk}\left(R_{k-1}\right)$ are zero, I define $\frac{c_i}{p_{isk}(R_{k-1})}$ as zero.

1. Bounds for listed programs $(r_{is} = k \neq \infty)$:

   The first bound $v_{is} \geq \frac{c_i}{p_{isk}(R_{k-1})}$ follows directly from the sequential optimization problem as $V_i(R_{k-1} \cup \{s\}) = V_i(R_{k-1}) + v_{is} p_{isk}\left(R_{k-1}\right) - c_i \geq V_i(R_{k-1})$ for any R.

   The second bound follows from the sequential optimization problem as well as the properties of DA. The second bound is equivalent to:

   $v_{is} \geq v_{ij} \quad \forall j \in S_k$, where $S_k = \{j \in \{0, 1, ..., S\} \setminus R_{k-1} : v_{ij} p_{ijk}\left(R_{k-1}\right) - I(j \neq 0) c_i \geq 0\}$

To simplify the notation, I drop the subscript i throughout the proof.

I prove that if $s = \arg\max_{j \in S_k} v_j$, then program s must be the k-th element of the chosen list R. This can be proven by contradiction: let's assume that this is not the case and that instead $r_s > k$ in the optimal list R. Under the limited rationality assumption, moving s earlier in the list to the k-th position only changes expected utility by reducing the probability of being assigned to schools originally listed higher than s. Concretely, moving s to the k-th position of R reduces the utility of the applicant by:

$$\sum_{j \in S_k : r_j < r_s} \alpha_j(R) v_j$$

where $\alpha_j(R)$ 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 s and j, i.e. denoting A_h to be the event of being above the cutoff for school h, we have:

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

At the same time, her utility is increased by:

$$\sum_{j \in S_k : r_j < r_s} \alpha_j(R) v_s,$$

Since we assumed that list R is optimal and that there are no ties between programs, the decrease must be larger than the increase

$$\sum_{j \in S_k : r_j < r_s} \alpha_j(R) v_s - \sum_{j \in S_k : r_j < r_s} \alpha_j(R) v_j = \sum_{j \in S_k : r_j < r_s} \alpha_j(R)(v_s - v_j) < 0.$$

but this contradicts the fact that $v_s \geq v_j$ and therefore it has to be the case that $r_s = k$.

Intuitively, this proof follows from the fact that in DA, the position of a program in the ROL does not affect the probability of clearing its admission cutoff. As such, moving a program to a higher position in the ROL only increases the assignment probability by displacing some of the probability of being offered programs previously ranked higher in the ROL.

The last bound $v_{is} \leq v_{ij} \quad \forall j$ s.t. $r_{is} > r_{ij}$ follows from the same reasoning since program s is included in S'_k for $r_{ij} = k'$.

2. Bounds for unlisted programs ($r_{is} = \infty$):

$$v_{is} \leq \frac{c_i}{p_{is\underline{k}}\left(R_{\underline{k}-1}\right)}$$
$$v_{is} \leq \max(v_{ij}, \frac{c_i}{p_{isk}\left(R_{k-1}\right)}) \quad \forall j \text{ s.t. } r_{ij} = k < \infty$$

   The second bound follows from the second bound for listed programs. Indeed, for any program $j \neq s$ such that $r_{ij} = k < \infty$, if $v_{is} \geq \frac{c_i}{p_{isk}(R_{k-1})}$ then $v_{is} \leq v_{ij}$ by the second bound for listed programs.

   The first bound also follows from the second bound for listed programs given that the outside option (program 0) can be viewed as the last program listed by any applicant and that I assumed it has indirect utility $v_{i0} = 0$.

3. Bounds for the per-unit cost c_i:

$$c_i \geq 0,$$
$$c_i \leq p_{ijk}\left(R_{k-1}\right) v_{ij} \quad \forall j \text{ s.t. } r_{ij} = k < \infty,$$
$$c_i \geq v_{ij} p_{ij\underline{k}}\left(R_{\underline{k}-1}\right) \quad \forall j \text{ s.t. } r_{ij} = \infty,$$
$$c_i \geq v_{ij} p_{ijk}\left(R_{k-1}\right) \quad \forall j \text{ s.t. } r_{ij} = \infty \text{ and } v_{ij} > v_{ij'} \text{ where } r_{ij'} = k.$$

   Beside the first bound which is an assumption, the other bounds for the per-unit cost follow directly from inverting the utility bounds for all listed or unlisted programs.

B Robustness of simulations to applicant exits

Simulation results presented in the main text do not incorporate the potential effect of counterfactual admission regimes on applicant exits from the traditional public school sector. In this appendix, I extend the simulations to account for potential exit. Applicants' binary decision to quit the public school system is predicted using a probit model which follows the multiple endogenous variables IV specification of Table 4.

$$\begin{aligned} Pr[Y_{idt} = 1 | \lambda_{tv}, \gamma_{dv}, W_{idt}] &= Pr[Y^* > 0] \\ &= Pr\left[\sum_{v}^{V} \lambda_{tv} + \sum_{v}^{V} \gamma_{dv} + \beta W_{idt} + \epsilon_{idt} > 0\right] \\ &= \Phi\left(\sum_{v}^{V} \lambda_{tv} + \sum_{v}^{V} \gamma_{dv} + \beta W_{idt}\right) \end{aligned}$$

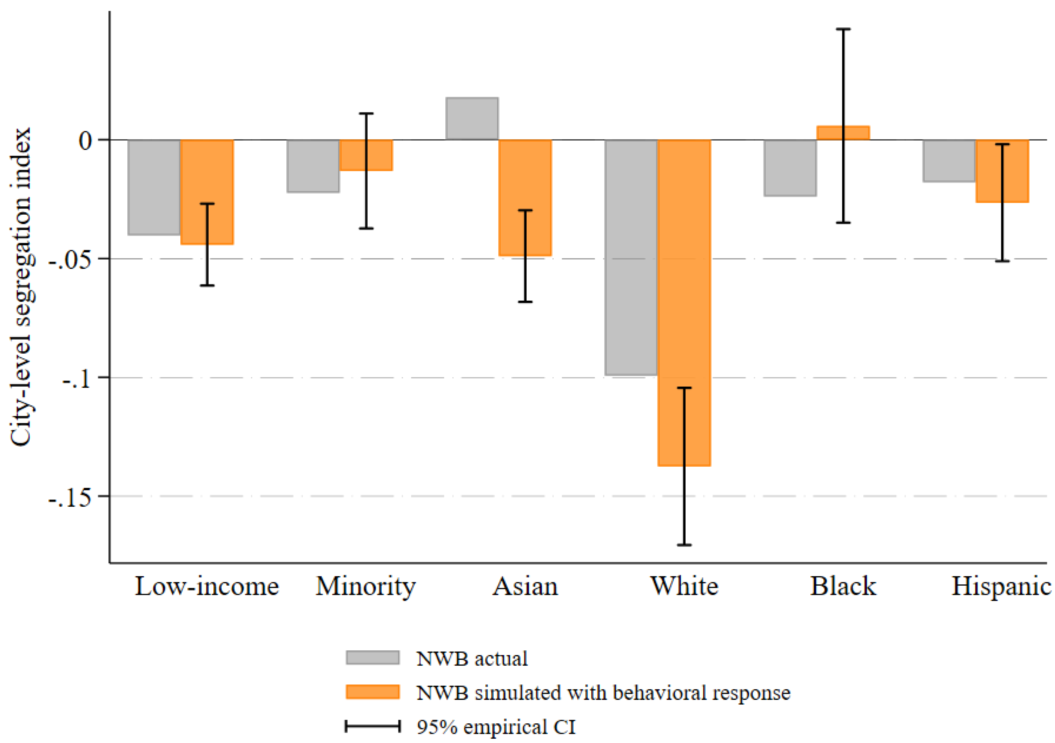
where λ_{tv} are covariate-specific time fixed effects and γ_{dv} are covariate-specific district fixed effects. W_{idt} are time-varying characteristics of the school student i is assigned to (the share of minority students and the share of high-math students). These characteristics are instrumented by the reform instruments interacted with 8 covariates. The covariates used in the estimation are dummies for English learner status, race, and the interactions of low baseline test scores with low-income status. The model is estimated using maximum likelihood on 2018 and 2019 data. Estimates for the coefficients on school time-varying characteristics are reported below.

Table A1: IV Probit Estimates of Offered Peer Characteristics on Public School Exit

	Exit (1)
Share high-math peers	-2.47***
	(0.49)
Share minority peers	-0.39
	(0.70)
Mean Exit 2018	0.115
N	133,078

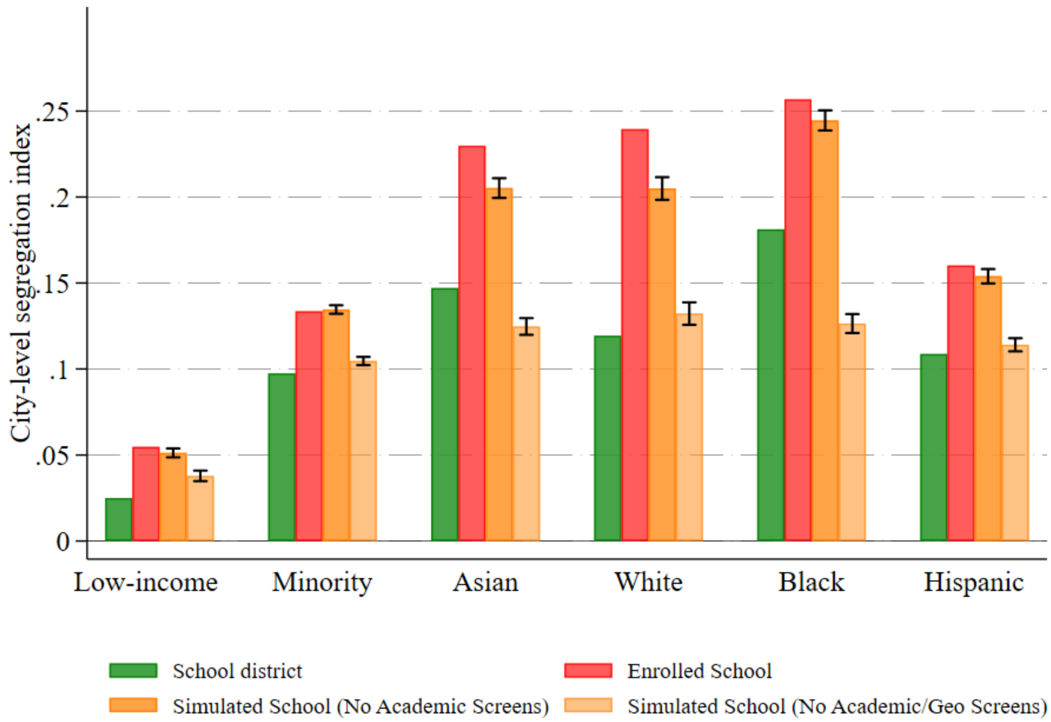
I then use these model estimates to incorporate applicant decisions on traditional public school exit into the simulations of counterfactual admission regimes. At each iteration of the match, applicants decide whether to exit public school based on the characteristics of the offer they receive. Specifically, I update the applicant latent variable Y_i^* based on the estimated $\hat{\beta}$ and the iteration-specific W_i. An applicant exits in the iteration whenever $Y_i^* > 0$. To set the baseline value of the latent variable Y_i^*, I draw ϵ_i for each applicant so that the baseline value of Y_i^* is compatible with her observed exit decision. In the simulations that incorporate applicant exits, applicants also update their ROLs based on post-exit school characteristics.

Figure A1: Replication of the NWB Plan's Effects on Enrolled School Segregation



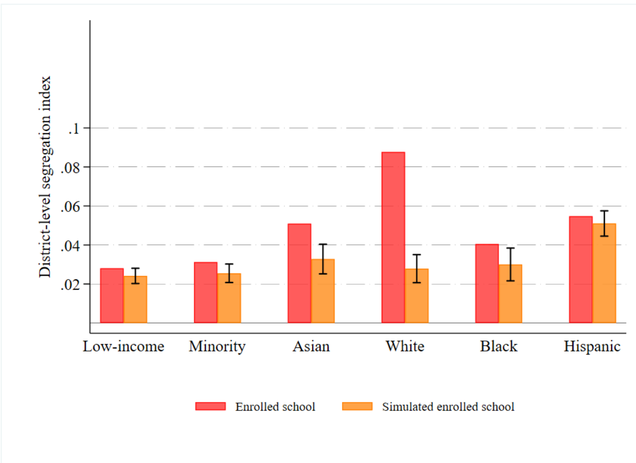
Note: This bar chart plots the changes in school segregation indices at the match offer stage between 2018 and 2019, and between 2018 and two simulations of the 2018 match under the integration reforms. The 95% confidence intervals for the effect of the integration reforms absent behavioral response are computed by simulating 100 times the 2018 match under integration reforms. For each simulation, a new sequence of tiebreakers is drawn and applicants are sampled with replacement, through sampling stratified by SWD indicator and district.

Figure A2: Effect of Counterfactual Admission Criteria on Enrolled School Segregation

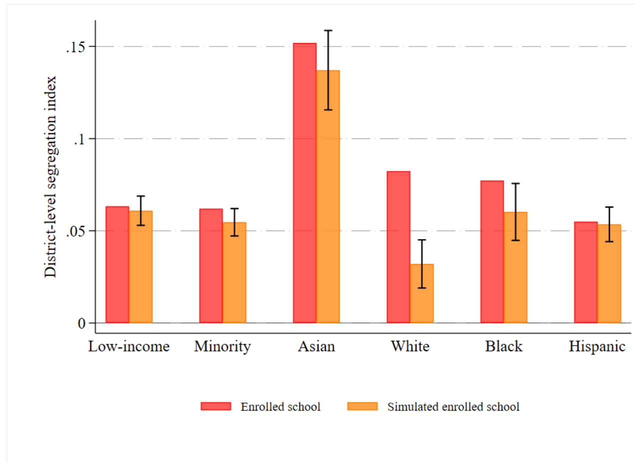


Note: This bar chart plots the segregation indices for different levels and under different admission criteria in 2018. The enrolled school corresponds to the school applicants enrolled in 2018. The simulated school corresponds to the school applicants enroll in a simulation that drops all academic screens. The random school corresponds to the school applicants enroll in a simulation that drops all academic screens and where students rank programs randomly. For both simulations, applicants can only apply to programs to which they are eligible in the 2018 match. 11.5% of applicants exit the traditional public school district in the actual match, 13.2% on average in the simulated match with no academic screens, and 13.7% on average in the simulated match with neither academic nor geographic screens.

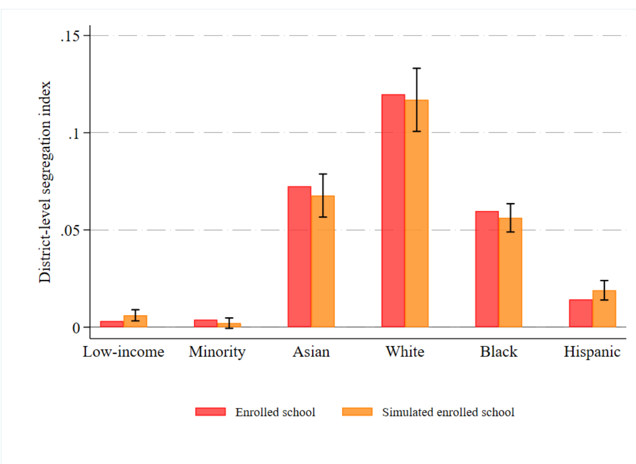
Figure A3: Effect of Dropping Academic Screens on district-level Enrolled School Segregation



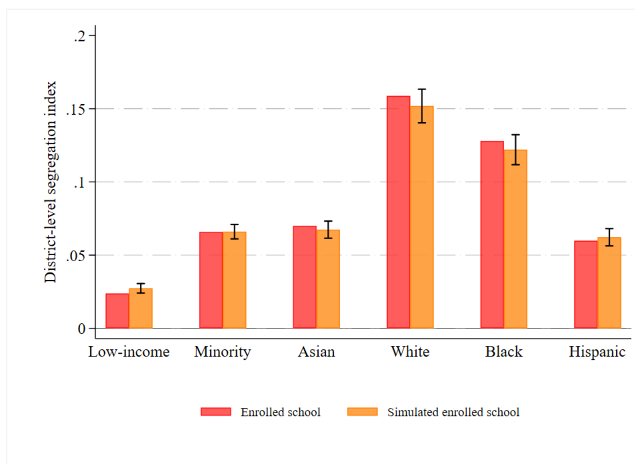
(a) Brooklyn



(b) Manhattan



(c) Bronx



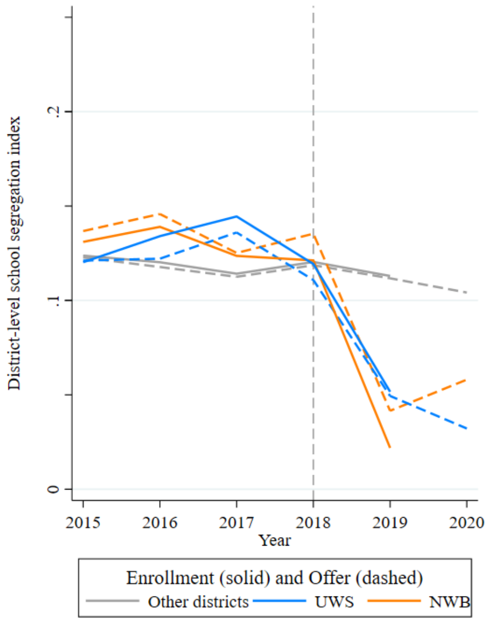
(d) Queens

Note: This bar chart plots the district-level school segregation index under the actual match and a match that drops all academic screens in 2018.

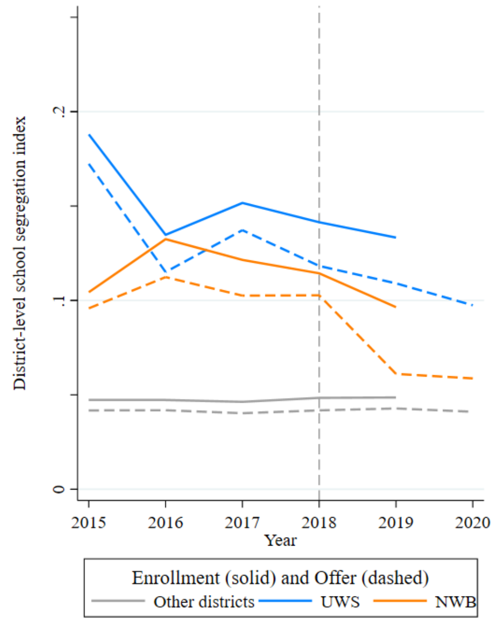
ONLINE APPENDIX
Not for Publication

C Additional Figures

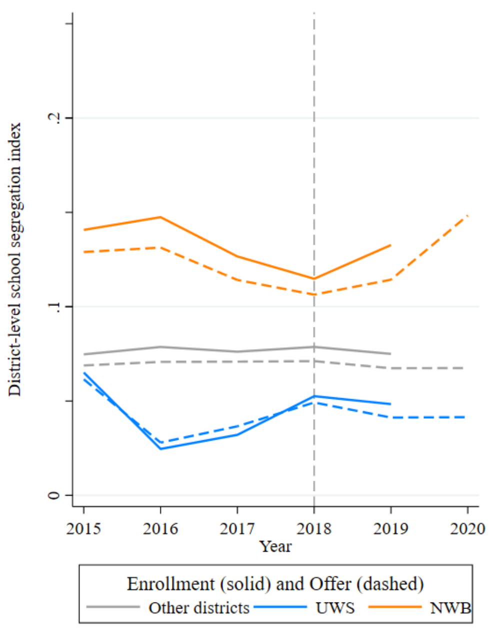
Figure A4: Evolution of School Segregation Indices for White, Hispanic, Black and Asian Students



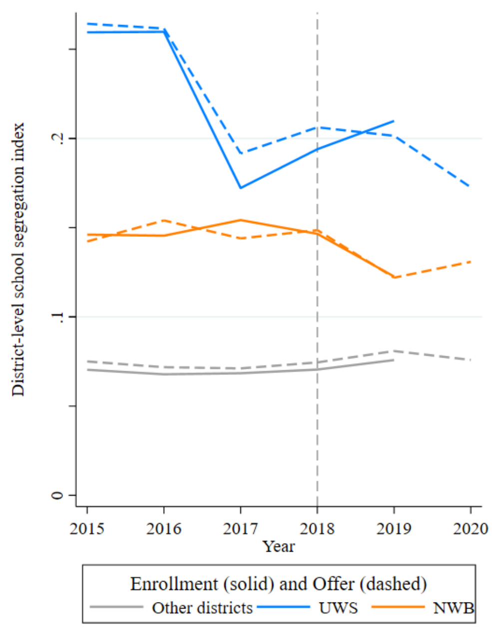
(a) White students



(b) Hispanic students



(c) Asian students

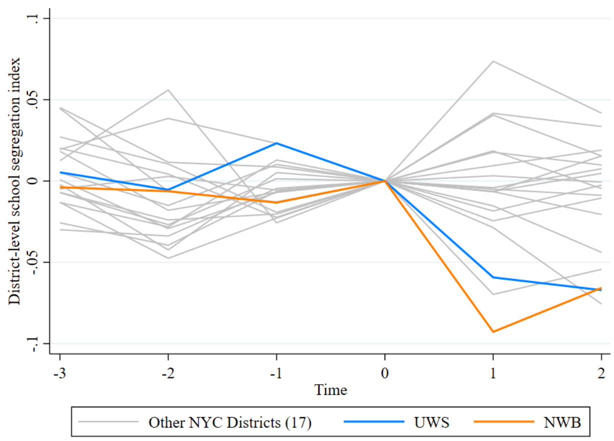


(d) Black students

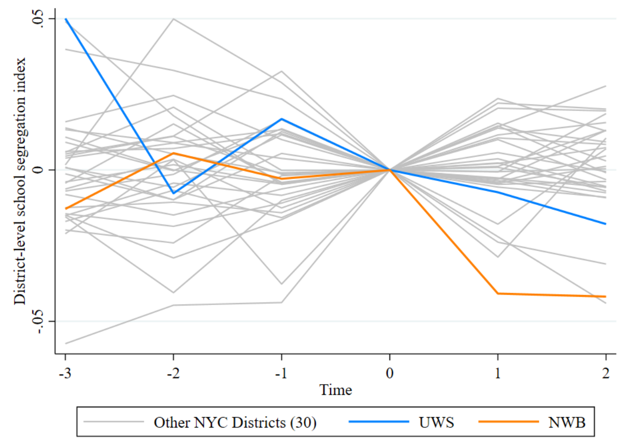
Note: These figures plot the evolution of district school segregation indices for UWS, NWB, and other NYC districts between 2015 and 2020. Figure (a) displays the segregation index for White applicants, Figure (b) for Hispanic applicants, Figure (c) for Asian applicants and Figure (e) for Black applicants. School segregation indices are computed as in Figure 3.

Figure A5: Comparison of Changes in School Segregation Indices in all NYC Districts

Panel A: Offered School

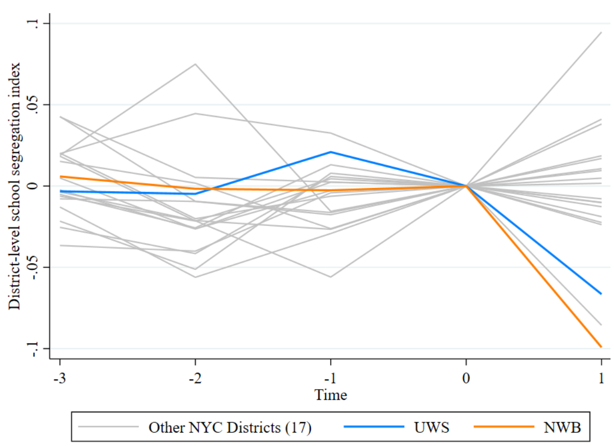


(a) White students

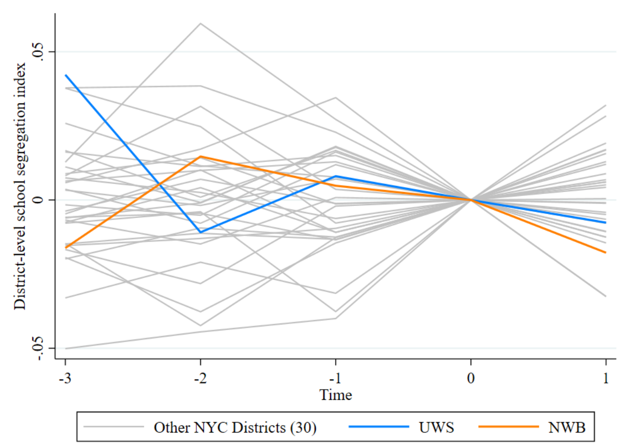


(b) Hispanic students

Panel B: School of Enrollment



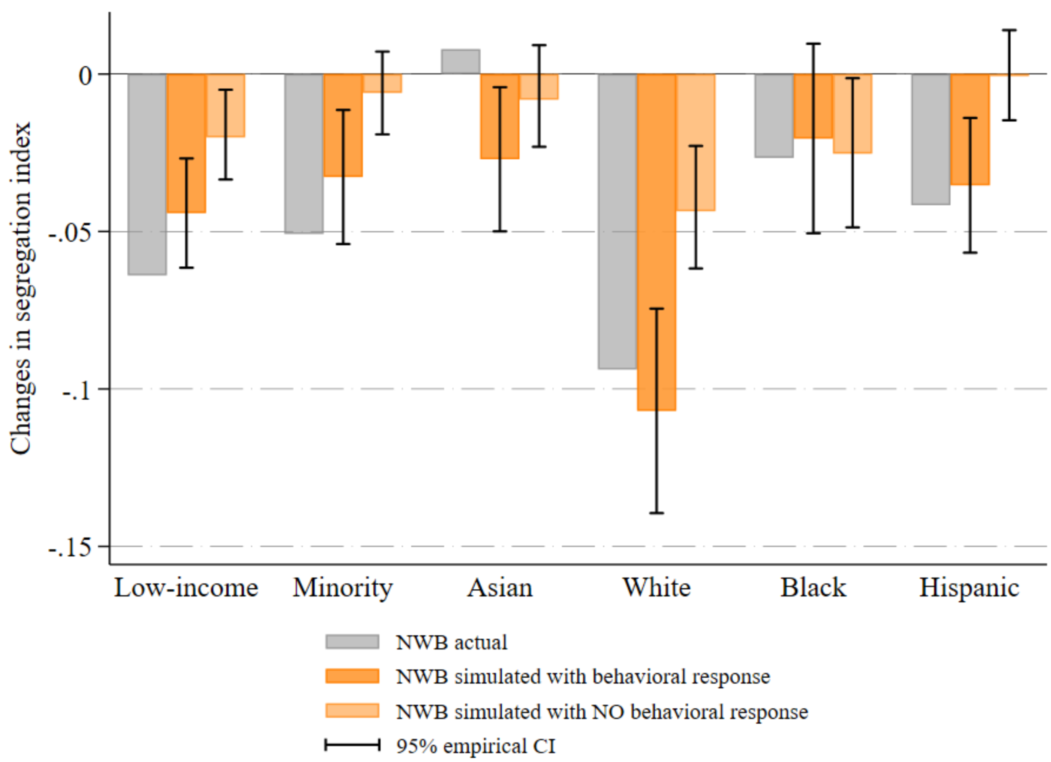
(c) White students



(d) Hispanic students

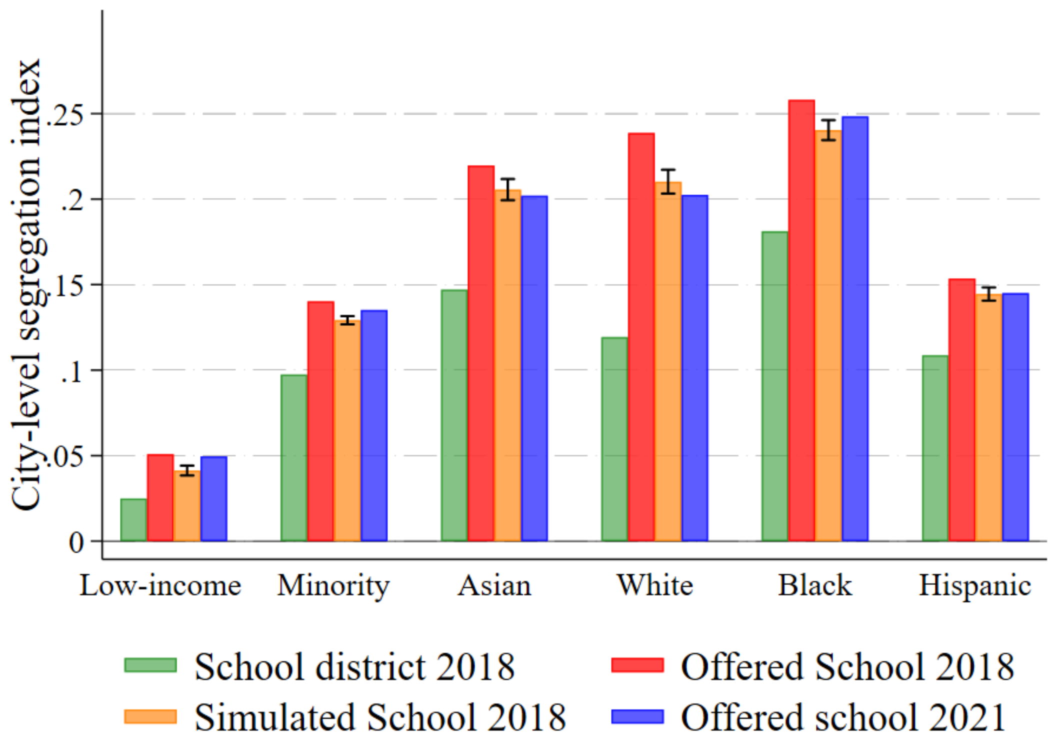
Note: These figures plot changes in the district-level school segregation indexes with respect to 2018 for all NYC districts. In Panel A, SIs are computed using the school offered. In Panel B, SIs are computed using the school of enrollment. Since the segregation index is more volatile for districts with fewer students of the group considered, each plot only includes districts in which at least 5% of the student population belongs to the group considered.

Figure A6: Replication of NWB Plan Effects on School Segregation



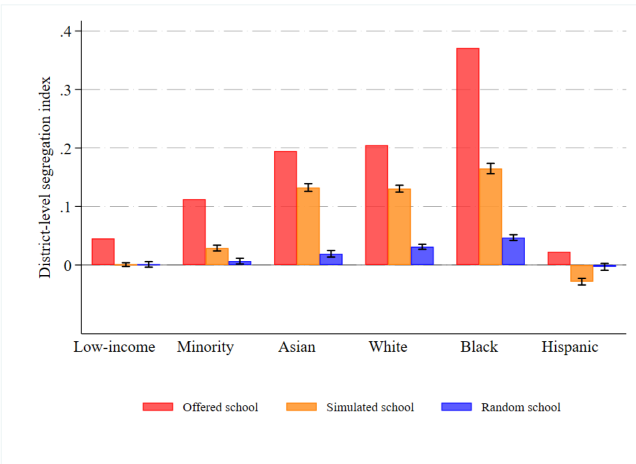
Note: This bar chart plots the changes in district school segregation indices at the match offer stage between 2018 and 2019, and between 2018 and two simulations of the 2018 match under the integration reforms. The first simulation updates 2018 applicant ROLs according to the model of list formation but without taking into account changes in school composition. The second simulation does not update 2018 applicant ROLs, as in 5. The 95% confidence intervals for the effect of the integration reforms absent behavioral response are computed by simulating 50 times the 2018 match under integration reforms. For each simulation, a new sequence of tiebreakers is drawn and applicants are sampled with replacement, through sampling stratified by SWD indicator and district.

Figure A7: School Segregation Under Alternative Admission Criteria

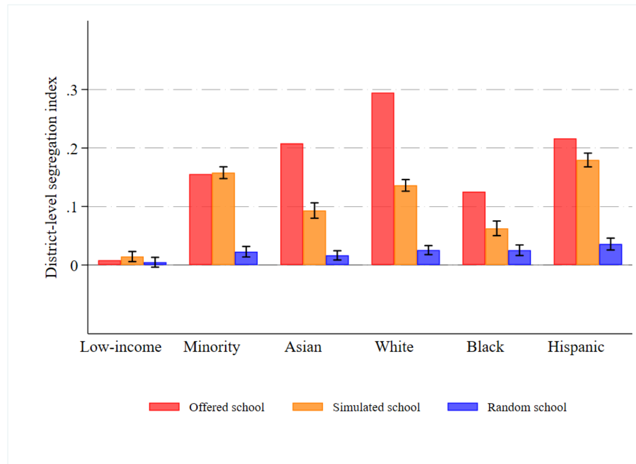


Note: This bar chart plots segregation indices in 2018 for the school district and for offered schools under different admission criteria. The 2018 offered school corresponds to the school 2018 applicants were offered in the actual match. The 2018 simulated school corresponds to the school 2018 applicants are offered in a simulation that drops all academic screens. The 2021 offered school corresponds to the school 2021 applicants were offered in the actual match. For the 2018 simulation, applicants can only apply to programs to which they are eligible in the 2018 match.

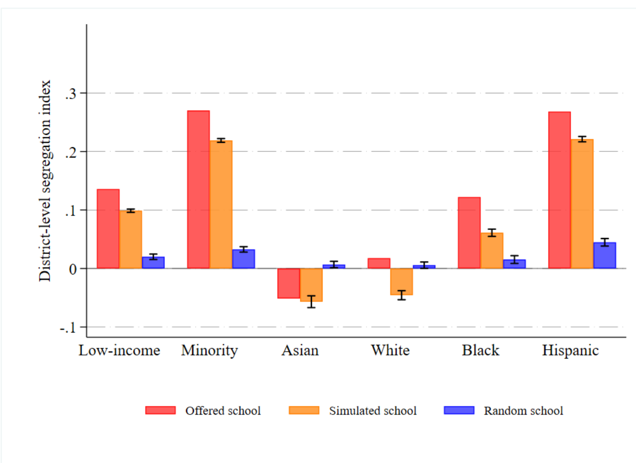
Figure A8: Effect of Dropping Academic and Geographic Screens on School Segregation



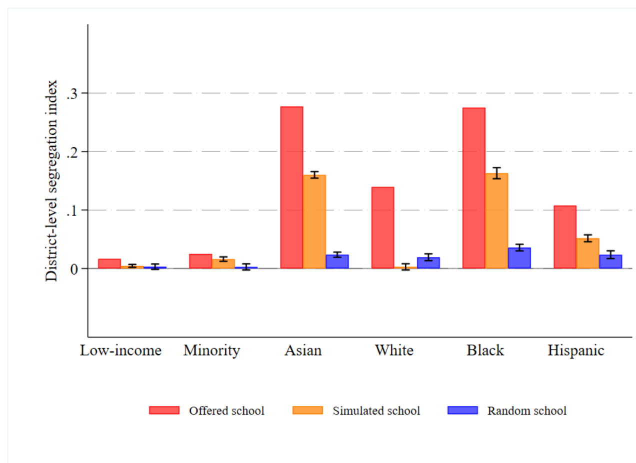
(a) Brooklyn



(b) Manhattan



(c) Bronx



(d) Queens

Note: This bar chart plots the school segregation indices in 2018 under different counterfactual admission regimes for four NYC boroughs. The offered school corresponds to the school applicants were offered in the 2018 match. The simulated school corresponds to the school applicants are offered in a simulation that drops all academic screens and all geographic eligibilities and priorities. The random school corresponds to the school applicants are offered in a simulation that drops all academic screens and all geographic eligibilities and priorities, and where applicants rank programs randomly. For both simulations, applicants can apply to any program in the 2018 match except dual-language programs and gender-specific programs, for which they need to be eligible.

D Additional Tables

Table A2: Description of NYC Diversity in Admissions Policies

	Fall of middle school enrollment year					
	2017		2018		2019 and 2020	
	Programs	Policy	Programs	Policy	Programs	Policy
UWS					All	FRPL lowest performing - 10%
						FRPL lower performing - 15%
NWB	15K447	FRPL - 30%	15K447	FRPL - 30%	All	Low-income, ELL, Temp housing - 52%
	15K839	FRPL - 40%	15K839	FRPL - 40%		
			15K497	FRPL - 40%		No screening of applicants
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%

Note: This table describes the Diversity in Admissions 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 Admissions 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 percentage seats.

Table A3: Characteristics of UWS and NWB Students by Race

District	Race	N	% Low-income	% ELL	English score	Math score	Reserve eligible
UWS	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%
NWB	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%

Note: This table describes the characteristics of 2019-2020 applicants residing in UWS and NWB by race. English and Math test scores correspond to the proficiency rating on 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.

Table A4: DiD Estimates of Changes in Probability of Unassignment

	All	Low-income		High-income	
		Low baseline	High baseline	Low baseline	High baseline
	(1)	(2)	(3)	(4)	(5)
UWS × 2019	0.01*	0.01	0.03	-0.02	0.00
	(0.01)	(0.01)	(0.03)	(0.03)	(0.01)
UWS × 2020	0.02**	0.01	0.04	0.00	0.01
	(0.01)	(0.01)	(0.03)	(0.03)	(0.01)
NWB × 2019	0.00	-0.02**	-0.02**	-0.01	0.03***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
NWB × 2020	-0.01	0.00	-0.03***	-0.01	0.01
	(0.00)	(0.01)	(0.01)	(0.02)	(0.01)
Mean UWS pre-2019	0.06	0.05	0.06	0.08	0.05
Mean NWB pre-2019	0.07	0.07	0.07	0.10	0.05
N	396,958	171,653	94,672	37,999	63,778

Note: This table reports difference-in-differences estimates of integration reform effects for 2015-2020 middle school applicants. The endogeneous variable is the interaction of dummies for residing in UWS and NWB and applying for admission in 2019 and 2020. The 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 are reported in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%.

E Data description and sample construction

The data used come from several sources. Lists of high school applicants, applicant rank-order lists, and assignments are constructed from annual records from the New York City Department of Education (NYCDOE) school assignment system. Information on student demographic characteristics and schools attended comes from the NYCDOE's Office of School Performance and Accountability (OSPA). Achievement test scores are taken from the New York State Assessment. Geographic information on students comes from Zoned DBN data. All non-public data files were provided by NYCDOE. These files include a unique student identification number that links records across files. More details on each data source are provided below.

NYCDOE Assignment Data

Data on NYC middle school applications are controlled by the Student Enrollment Office of the NYCDOE. I received all applications for the 2015-16 through 2021-2022 school years. Application records include students' rank-order lists of academic programs submitted in each round of the application process, priorities and rank at each program listed, lottery tiebreaker values, and the program to which the applicant was assigned.

OSPA Data

I received registration and enrollment files for the 2015-2020 school years from NYCDOE's Office of School Performance and Accountability (OSPA). These data include every student's grade and school District Borough Number (DBN), as of June of each school year, as well as information on student demographic variables. I use this file to code school enrollment, special education status, subsidized lunch status, and limited English proficiency.

New York State Assessment Data

The New York State Assessment is the standardized state exam for New York, taken in grades 3-8. The NYCDOE provided scores for students taking the exam from the 2005-06 to 2018-19 school years. Each observation in the dataset corresponds to a single test record. I use 4th grade test scores from 2014-2019 to assign baseline math and English Language Arts (ELA) scores. Baseline scores are normalized to have mean zero and standard deviation one within a subject-year among all 4th grade NYC public school students.

Zoned DBN Data

The Zoned DBN dataset provides geographic data for elementary, middle, and high school students in NYC based on the address provided to the DOE. In these files, there is a record for every student with an active address record during the school year. I use Zoned DBN files from school years 2010-2019 to collect data on student residential districts and census tracts.

F Description of deferred acceptance (DA)

Student-proposing DA has been adopted in many school districts around the world because of its attractive theoretical properties.[64] In particular, the mechanism produces a student-optimal stable match and is strategy-proof when students do not face application costs (Abdulkadiroğlu and Sönmez, 2003).

The mechanism is based on three inputs: student rankings of schools, school rankings of students, and school capacities. First, students submit a ROL 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 her first choice. School 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 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 applicant ranks at each school are independent of student ROLs. As such, each student's probability of admission at each school only depends on other applicant ROLs and school rankings of applicants. Thus, applicants may take as given their admission probabilities at each school when forming their ROLs.

[64] 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)).

G Robustness of DiD Results

This section reports permutation test results for each main DiD specification. These permutation tests aim at evaluating the statistical significance of DiD results (MacKinnon and Webb, 2020). They also allow us to to assess the presence of pre-trends in the outcome for the treated units. To obtain a distribution of "placebo effects" for each outcome, I estimate the following dynamic DiD specification assigning each of 32 NYC districts to be the treated unit $\tilde{d}$ iteratively:

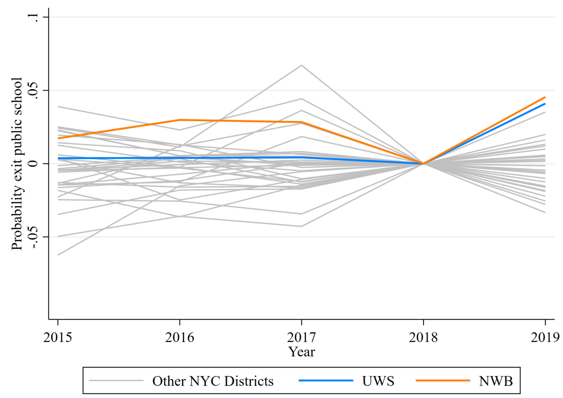
$$Y_{itd} = \lambda_t + \delta_d + \sum_{j=-3}^{-1} \beta_j I(d = \tilde{d}) \times I(t = j) + \sum_{j=1}^{2} \beta_j I(d = \tilde{d}) \times I(t = j) + \varepsilon_{itd} \tag{12}$$

where λ_t and δ_d are year and district fixed effects. Y_{itd} is the outcome of interest. β_j captures the estimated "placebo" treatment effects in each year prior to and after 2018 for the district assigned to treatment. To assess statistical significance of the plans' effect, the treatment effects estimated for UWS and NWB are compared to the placebo effects computed for the remaining 30 districts. The effects estimated for UWS and NWB are deemed significant when their magnitudes are extreme relative to the permutation distribution. Conversely, UWS and NWB do not show pre-trends in the outcome if estimated effects for years prior to 2018 are comparable to the placebo estimates.

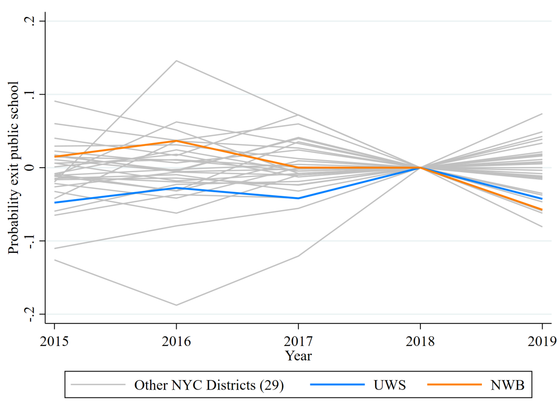
The remainder of this appendix section reports plots of the estimated coefficients β_j for each DiD specification. The plots exclude districts with less than 50 observations in any year of the sample period. Results are reported in the order they appear in the text and under the same rubric names.

G.1 Effects on exit from the traditional public school system

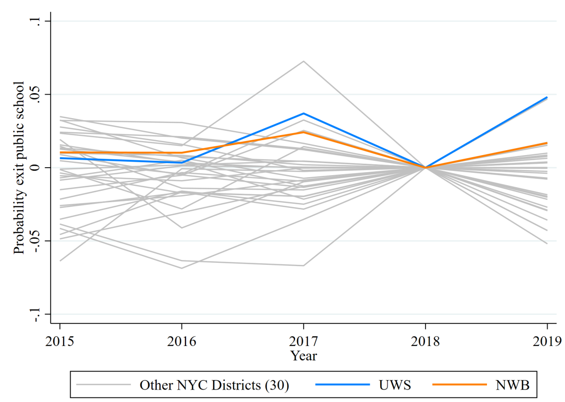
Figure A9: Treatment and Placebo Effects on Exit from Traditional Public School



(a) All students

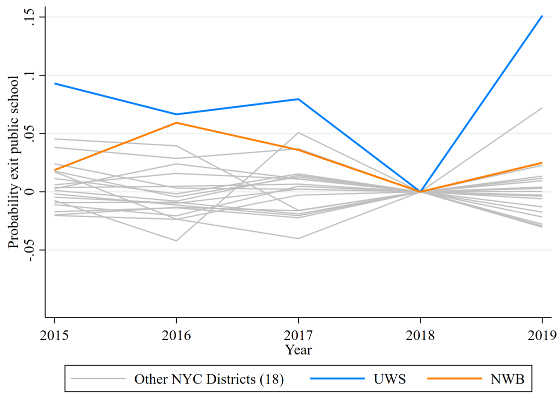


(b) Black students

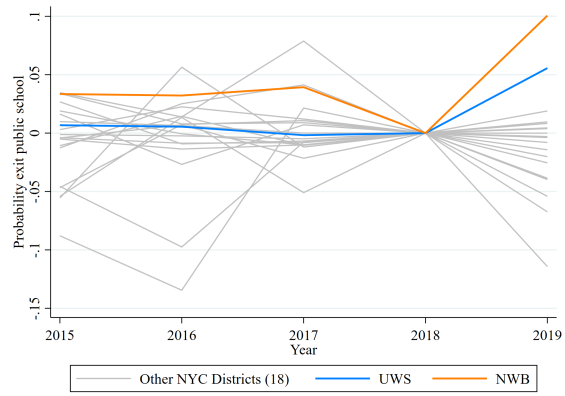


(c) Hispanic students

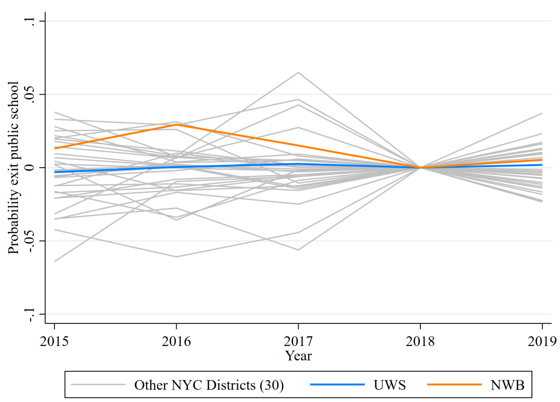
Figure A9: Treatment and Placebo Effects on Exit from Traditional Public School



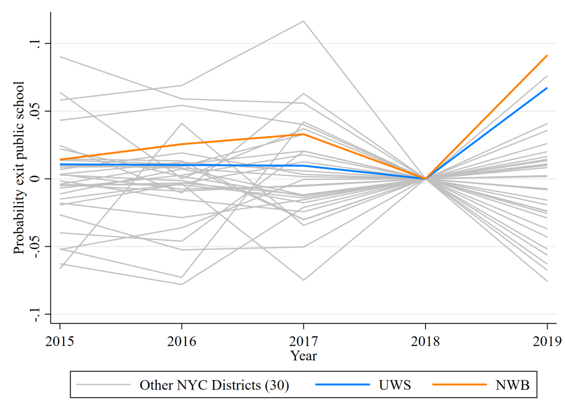
(a) Asian students



(b) White students



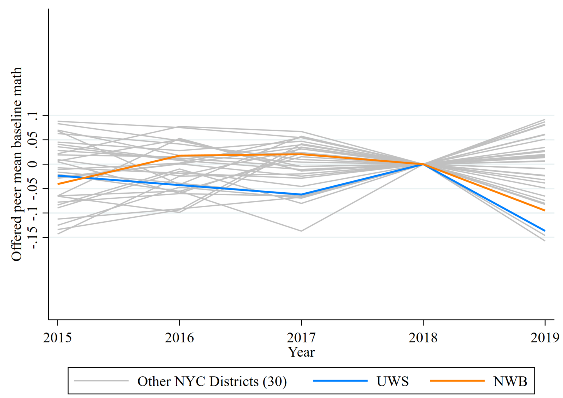
(c) Low-income students



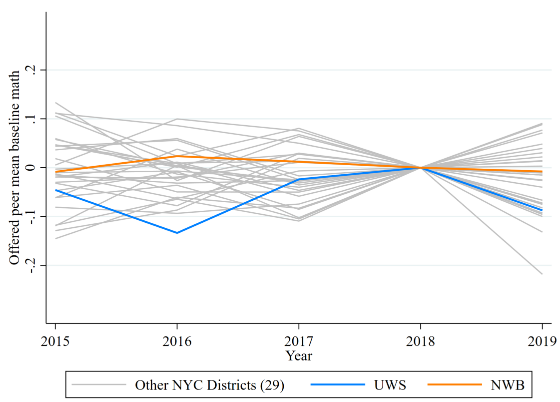
(d) High-income students

Note: These figures plot the permutation tests for Panel A of Table 3.

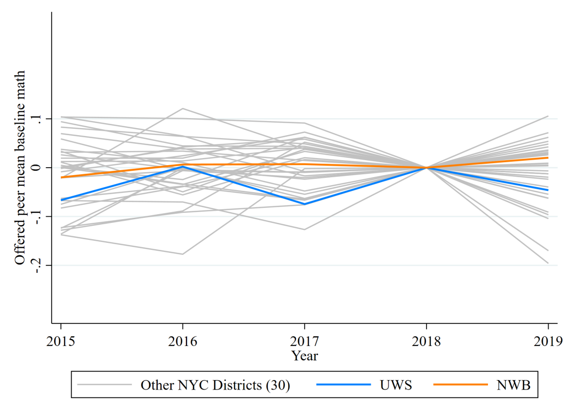
Figure A10: Treatment and Placebo Effects on Offered Peer Mean Math Score



(a) All students

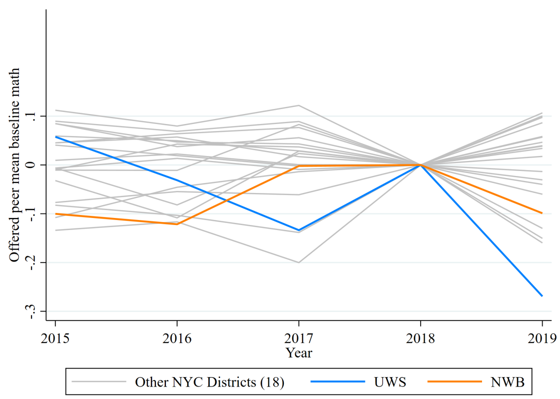


(b) Black students

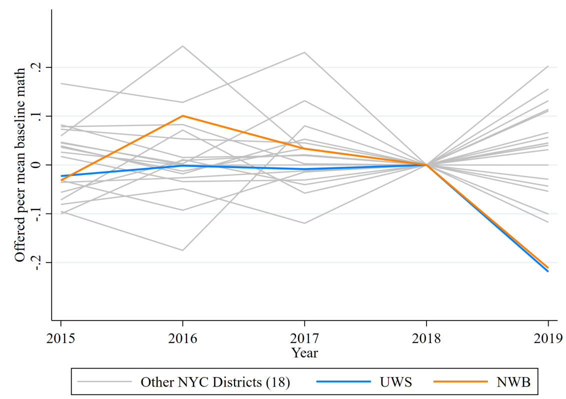


(c) Hispanic students

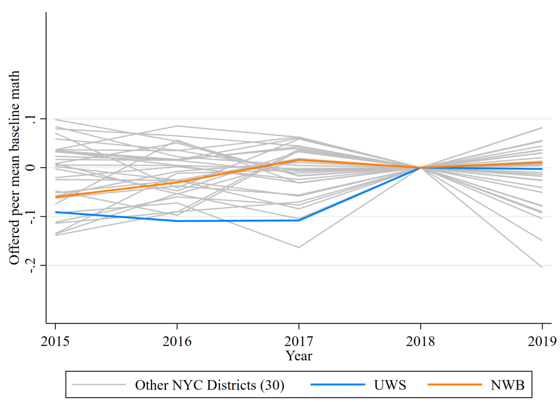
Figure A10: Treatment and Placebo Effects on Offered Peer Mean Math Score



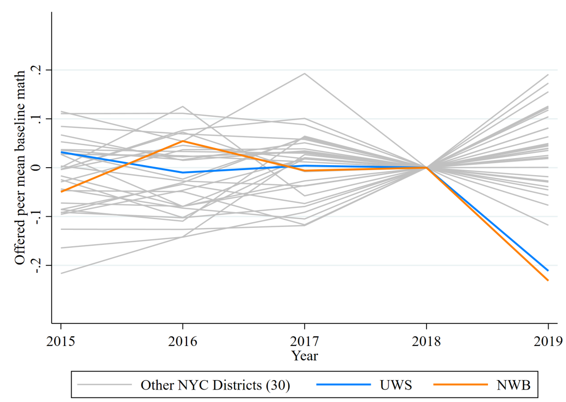
(a) Asian students



(b) White students



(c) Low-income students

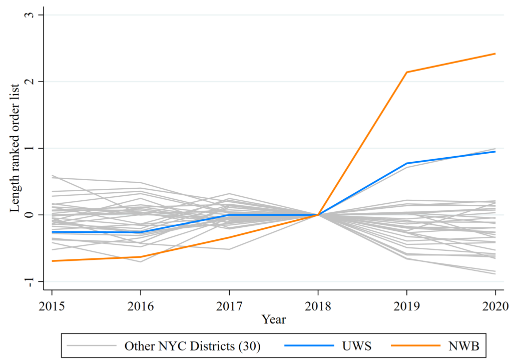


(d) High-income students

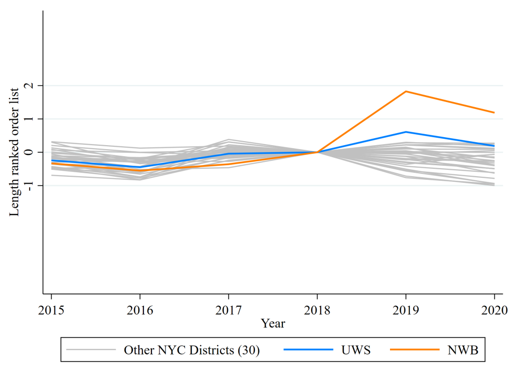
Note: These figures plot the permutation tests for Panel B of Table 3.

G.2 Effects on expressed preferences

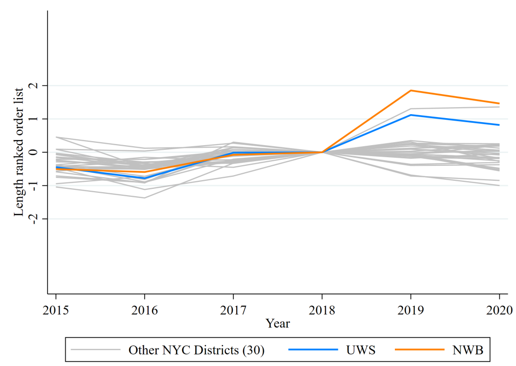
Figure A11: Treatment and Placebo Effects on Length of Rank-Order List



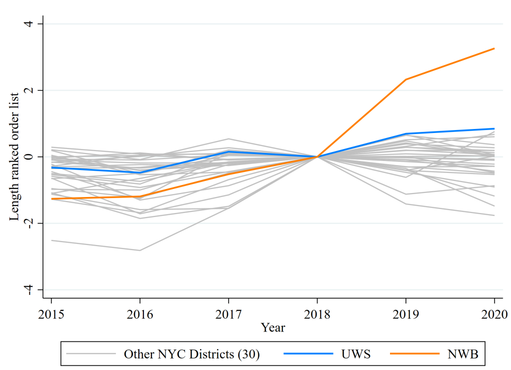
(a) All students



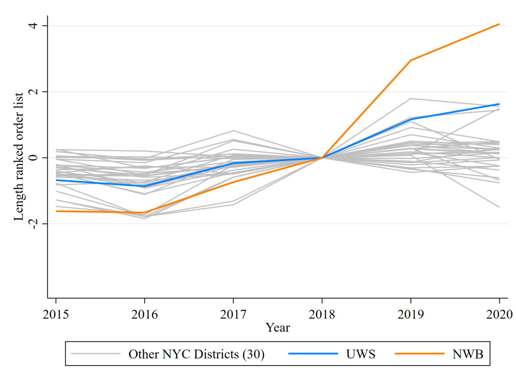
(b) Low-income low-baseline test score students



(c) Low-income high-baseline test score students



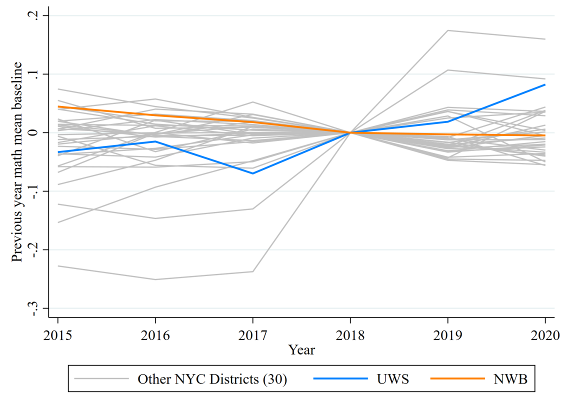
(d) High-income low-baseline test score students



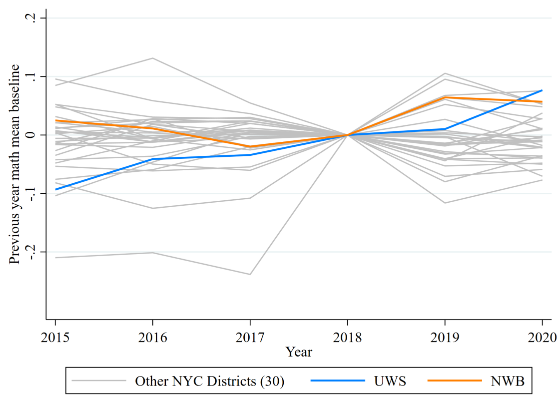
(e) High-income high-baseline test score students

Note: These figures plot the permutation tests for Panel A of Table 5.

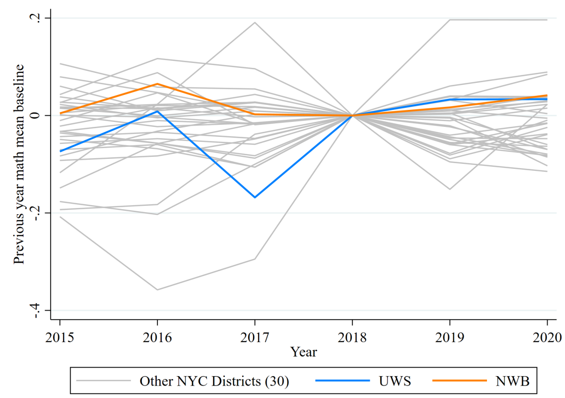
Figure A12: Treatment and Placebo Effects on Mean Math Score in School Ranked First



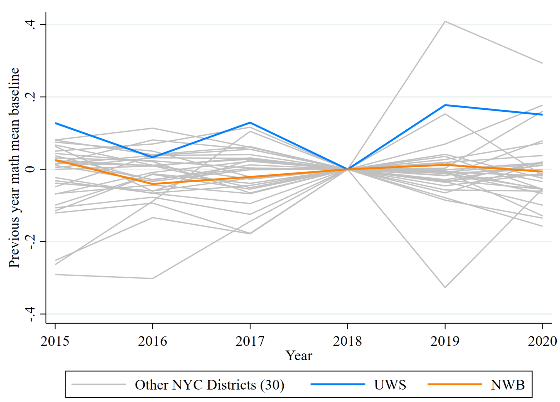
(a) All students



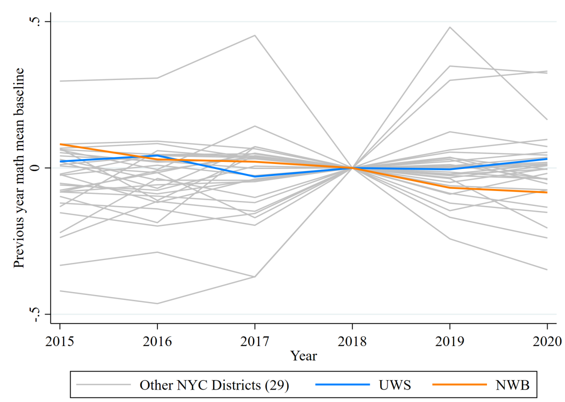
(b) Low-income low-baseline test score students



(c) Low-income high-baseline test score students



(d) High-income low-baseline test score students



(e) High-income high-baseline test score students

Note: These figures plot the permutation tests for Panel B of Table 5.

H Details of the Model Estimation

H.1 Description of the estimation of admission probabilities

Assuming that applicants hold rational expectations about their admission probabilities, a consistent estimator of applicant beliefs about admission probabilities q_{is} can be obtained by bootstrapping student assignments. This bootstrap procedure captures the uncertainty in the probability of admission both due to the lottery and year-to-year variation in the applicant population. Specifically, for a match in which a set of programs L rank students based on their lottery numbers while other programs rank students based on a scores, the estimation of admission probabilities at each program unfolds as follows.

- For each bootstrap simulation $b = 1, ..., B$, where $B = 1000$:
  - Sample with replacement n applicants with their score τ_i and ROL R_i.[65]
  - Draw a new lottery number t_i for each applicant.
  - Run DA to obtain an assignment.
  - Obtain the lottery number of the last admitted applicant t_s^b for lottery programs $s \in L$ and the score of the last admitted applicant τ_s^b for score programs $s \in L^c$.
- Estimate the probability of admission of student i at each lottery program $s \in L$ as:

$$\hat{q}_{is} = \frac{1}{B} \sum_{b=1}^{B} t_s^b$$

- Estimate the probability of admission of student i at each score program $s \in L^c$ as:

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

For any rank-order list R, it is possible to compute $p_{isk}(R_{k-1})$, the conditional offer probability for each program s if placed in position k of a student's list. This conditional probability takes into account the dependency between admission events, which differs depending on whether a program admits applicants based on their score or lottery number.

I assume that admission score cutoffs are approximately independent, thus admissions at each score program are independent events given applicant score.[66] However, admissions at lottery programs are not independent events. Since all programs use the same lottery number and lottery numbers are not known at the time of application, applicant rejection

[65] The re-sampling is stratified by school district of residence.

[66] This is a plausible assumption if the market becomes large, as variations in cutoff are due to changes in the applicant population. In the limit, admission cutoffs are constant in DA (Azevedo and Leshno, 2016). In the empirical application, variations in cutoffs are also driven by changes in the discretionary weights programs give to different criteria when deciding student rankings.

at a lottery program carries information on applicant tiebreaker. As such, the probability of being offered each lottery program if placed in a given position depends on the lottery cutoff at the lottery programs ranked higher in the list. Formally, this entails that the probability of being offered each lottery program $s \in L$ if ranked in position k is:

$$p_{isk}(R_{k-1}) = \max\left[0\ ,\ q_{is} - \max_{j \in L:\{r_{ij}<k\}} q_{ij}\right] \prod_{j \in L^c:\{r_{ij}<k\}} (1 - q_{ij})\ ,$$

and the probability of being offered each score program $s \in L^c$ if ranked in position k is:

$$p_{isk}(R_{k-1}) = q_{is} \times \left(1 - \max_{j \in L:\{r_{ij}<k\}} q_{ij}\right) \prod_{j \in L^c:\{r_{ij}<k\}} (1 - q_{ij})\ .$$

H.2 Description of the Gibbs sampler

Given the school indirect utilities and per-unit cost parameterization:

$$v_{ist} = \delta_{c(z_i)s} + \sum_{h=1}^{H} \beta_{c(z_i)h}(c(z_i) \times z_{sht}) - d_{is} + \varepsilon_{ist}, \tag{13}$$

$$c_i = c + \zeta_i \tag{14}$$

The vector of parameters to be estimated is the following $\theta = (\{\delta_{c(z)s}\}_{s=1}^{S}, \{\beta_{c(z)h}\}_{h=1}^{H}, \sigma_\varepsilon, c, \sigma_\zeta)$.

This parameter 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, $\delta_{c(z)s}^0 \sim N(\mu_{\delta_{c(z)s}^0}, V_{\delta_{c(z)s}^0})$ with $\delta_{c(z)s}^0 = 0$ and $\beta^0 = 0$.

2. Sample $\{\delta_{c(z)s}^1\}_{s=1}^{S}$ for each cell $c(z)$ given $\{\delta_{c(z)s}^0\}_{s=1}^{S}$, σ_ε^0, U^0, and β^0 from $N(\mu_{\delta_{c(z)s}^1}, V_{\delta_{c(z)s}^1})$.

   - Compute $V_{\delta_{c(z)s}^1} = (\frac{N_{c(z)}}{\sigma_\varepsilon^0} + \frac{1}{V_{\delta_{c(z)s}^0}})^{-1}$.
   - Compute $E_{is} = v_{is}^0 + d_{is} - X_i\beta^0$ (vector of residuals). Multiply this vector by the vector identifying cell $c(z)$ to obtain $E_{isc(z)}$, the residuals only for individuals of cell $c(z)$.
   - Compute $\mu_{\delta_{c(z)s}^1} = V_{\delta_{c(z)s}^1}(\frac{\sum_{i=1}^{N} E_{isc(z)}}{\sigma_\varepsilon^0} + \frac{\mu_{\delta_{c(z)s}^0}}{V_{\delta\ c(z)s}^0})$.
   - Sample $\delta_{c(z)s}^1 \sim N(\mu_{\delta_{c(z)s}^1}, V_{\delta_{c(z)s}^1})$ for all $s \in S$.

3. Sample β^1 given σ_ε^0, U^0, and δ^1 from $N(\mu^1, V^1)$.

   - Compute the vector of residuals $E_i = U_i^0 + D_i - \delta_{c(z_i)}^1$.
   - Compute $V^1 = (\frac{X'X}{\sigma_\varepsilon} + (V^0)^{-1})^{-1}$.
   - Compute $\mu^1 = V^1(\frac{X'E}{\sigma_\varepsilon} + (V^0)^{-1}\mu^0)$.

4. Sample σ_{ε}^1 given β^1, δ^1, and U^0 from a $IW(N \times S + 100, 3 + O)$.
   - N is the number of students, S is the number of programs.
   - $O = \sum_{i=1}^{N} \varepsilon_i' \varepsilon_i$ where $\varepsilon_i = U_i^0 + D_i - X_i \beta^1 - \delta^1$.
5. Sample the vectors U_i^1 by sampling iteratively from truncated normal distributions given U_i^0, c_i^0, β^1, δ^1, and σ_{ε}^1.
   - Draw iteratively each $v_{ist}^1 = \delta_{c(z_i)s}^1 + \sum_{h=1}^{H} \beta_{c(z_i)h}^1 (c(z_i) \times z_{sht}) - d_{is} + \varepsilon_{ist}$ where $\varepsilon_{ist} \sim TN(0, \sigma_{\varepsilon}^1, l_{ist}, u_{ist})$ with l_{ist}, u_{ist} computed using the rank-order list of each applicant and the bounds specified in Proposition 1, given U_i^0, U_i^1, c_i^0.
6. Sample c^1 given σ_{ζ}^0 and c_i^0 from $TN(\mu_c^1, \sigma_c^1, 0, \infty)$.
   - Compute $\sigma_c^1 = (N/\sigma_{\zeta}^0 + 1/\sigma_c^0)^{-1}$.
   - Compute $\mu_c^1 = \sigma_c^1 (1' E_c / \sigma_{\zeta}^0 + \mu_c^0 / \sigma_c^0)$ where $E_{ci} = c_i^0$.
7. Sample σ_{ζ}^1 given c^1 and c_i^0 from a $IW(N + 3, 3 + \sum_i (c_i^0 - c^1)^2)$.
8. 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 1 and given by U_i^1 and the rank-order list.

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

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

I use diffuse priors to minimize their influence on estimates. I set the prior distributions for parameters $\delta_{c(z)s}^0 \sim N(\mu_{\delta_{c(z)s}^0}, V_{\delta_{c(z)s}^0})$, $\beta \sim N(\mu^0, V^0)$ and $c^0 \sim TN(\mu_c^0, \sigma_c^0, 0, \infty)$ as:

$$
\begin{aligned}
\mu_{\delta}^0 &= 0 \\
V_{\delta}^0 &= 100 \times I \\
\mu^0 &= 0 \\
V^0 &= 100 \times I \\
\mu_c^0 &= 0 \\
\sigma_c^0 &= 100
\end{aligned}
$$

To check convergence of the sampler, I simulated three chains of length 100,000 each and burned-in the first half of each chain to ensure mixing. I monitored convergence by examining the trace plots of the various coefficients and by making sure that the potential scale reduction factor was below 1.1 for 95% of parameters before stopping the sampler.