Information and Access in School Choice Systems: Evidence from New York City *

Viola Corradini †

This version: October 2, 2024

Abstract

Disadvantaged urban students attend lower-quality schools, on average, than their more advantaged urban peers. This paper asks how information about school quality affects this gap. Specifically, I examine the effects of New York City's introduction of a letter-grade system rating the quality of its high schools. The ratings shifted Black and Hispanic students' choices more than those of white and Asian students, narrowing racial gaps both in enrollment at high-quality schools and in academic achievement. Using a structural model of school choice and surveys of families, I find that race differences in the response to quality information stem both by differing beliefs and, more importantly, by different preferences for school attributes. The model estimates suggest that Black and Hispanic students have less accurate perceptions of school quality, making them more receptive to the grade-based scoring system. Additionally, white and Asian students are less influenced by information on school quality because they have strong preferences for other school attributes. Simulations suggest that better quality information narrows racial gaps in choice and achievement. A system that releases coarse quality ratings for high-quality or oversubscribed schools increases test scores among lower achieving students more than perfect information by reducing the competition for high-quality schools from higher achieving students.

*I am thankful to my advisors Josh Angrist, Parag Pathak, Simon Jäger and Nikhil Agarwal for their advice and constant support, to the Office of Enrollment Research and Policy and the Office of School Performance and Accountability of the New York City Department of Education for graciously sharing data, and to Daron Acemoglu, David Autor, Clemence Idoux, Frank Schilbach, Benjamin Vatter and participants to the MIT Labor lunch and the MIT Applied Micro lunch for helpful comments. Thanks to Eryn Heying and Jim Shen 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 George and Obie Shultz Fund and the Guido Cazzavillan Fellowship.

†Columbia Business School. Email: viola.corradini@columbia.edu

1 Introduction

School choice systems are an increasingly popular alternative to neighborhood-based assignment of students to schools (Neilson, 2019). Proponents argue that such systems can reduce achievement gaps by offering everyone equal access to high-quality education and pressuring schools to improve (Friedman, 1955; Hoxby, 2000, 2003). Critics counter that market-based education reforms fall short of their goals because the conditions for competition and fair choice are not met in practice (Ravitch, 2010). Families often make choices based on the demographics of the student body, rewarding schools that draw from wealthier and more educated communities rather than pressuring them to improve quality (Ladd, 2002; Rothstein, 2006; Barseghyan et al., 2019; Cullen et al., 2006). Additionally, some contend that school choice exacerbates socio-economic inequality because affluent families are more attentive and capable of exploiting their educational opportunities (Ladd, 2002). In contrast, disadvantaged families often apply to and attend lower-quality schools even when higher-quality choices are available (Laverde, 2020; Hastings et al., 2009; Corradini and Idoux, 2023; Hoxby and Avery, 2012; Carlana et al., 2022).

Is misinformation responsible for these choice disparities and would providing better information unleash the potential of school choice to reduce inequality and boost achievement? If so, what are the most effective ways to present this information to families? This paper studies these questions in the context of high school choice in New York City. In this setting, I document that Black and Hispanic families apply to lower-quality schools, as measured by causal estimates of school value-added, even after controlling for residential location and differences in attainable options. This gap may be explained by differences in preferences for school attributes or in information about schools. Assuming families are perfectly informed may, in fact, lead to the erroneous conclusion that some do not reward quality, when in fact they may be more misinformed about it than others. Determining which of these explanations is responsible for the choice gap is crucial to understanding whether market-inspired interventions in education are bound to fail and who benefits from information interventions in equilibrium. Because families' perceptions of schools are hard to observe, however, inferring preferences from realized choices is challenging.

In 2007, New York City introduced a system that rated high schools by grades A to F, based on factors such as student progress, standardized test scores, attendance and graduation rates, while controlling for demographic differences. The grading system was then removed in 2014. This setting presents several advantages to study my research questions. The introduction, changes, and removal of grades provide a natural experiment that can be used to address the key research challenge of separating families' prior knowledge about the quality of different schools from their preferences for school quality. Moreover, detailed data on school capacity and the rules of the centralized admission mechanism allow me to credibly simulate assignment of students to schools under counterfactual information scenarios.

Exploiting within-school changes in letter grades, I find that student choices respond to information about school quality. High grades boost demand for seats, while low grades reduce demand. This shows that families indeed value school effectiveness, apart from other school attributes, such as peer quality, but hold uncertain beliefs about it. Families were more surprised by high letter grades when these were received by a school with low achievement levels and by low letter grades when these were received by a high performing school, suggesting that they initially perceived average achievement levels as an indicator of quality.

Black and Hispanic applicants respond more strongly to the school grades than do Asian and white applicants. Minority students are 7 percentage points more likely to apply to a school that always received an A after the introduction of letter grades compared to white and Asian students, off a base of 48 percentage points. Similarly, minority students are 9 percentage points (off a base of 34 percentage points) less likely than white and Asian students to apply to a school receiving consistently low grades after their introduction. While letter grades do not substantially affect white and Asian student choices on average, they still do within the subset of schools enrolling predominantly white and high-performing students. Black and Hispanic students, instead, respond to letter grades regardless of the demographic composition of a school. These findings suggest that white students hold strong preferences over school demographics, which attenuate their responses to information about school quality outside the subset of majority-white schools.

While the grading system disproportionately increased Black and Hispanic students' applications to high grade schools, these students did not always gain in admissions. In some cases, high grade schools screened out students on the basis of test scores, disproportionately favoring white students. In other cases, the increased demand from Black and Hispanic students led to greater competition for the schools they were selecting. As a result, information reduced the racial gap in applications to high grade schools more than in admission. After the reform, minority students were only 2.6 p.p. more likely thank white applicants to receive an offer to a high grade school, but 7 p.p. more likely to apply. Nevertheless, the larger shifts in demand among minority applicants reduced the cross-race gap in offered value-added by about 0.03 test score standard deviations (σ), or 4.5 percentiles.

To better understand what drives or constrains the beneficial effects of information interventions, I specify and estimate a model of demand for schools using data on rank-ordered preference lists. Departing from the standard typically adopted in the school choice literature, the model allows imperfectly informed students to hold prior beliefs about school quality and to update them when receiving quality signals using Bayes rule. Adapting the argument used in Vatter (2022), I show how variation in school quality ratings within schools and their availability over time separately identifies student preferences and beliefs over quality. Informed by differences in responses to letter grades across students and schools, I let preferences and beliefs vary across students with different demographic characteristics and let beliefs about school quality depend on school average achievement levels.

Estimates suggest that racial differences in beliefs and, even more so, preferences explain the larger response to information among minority students. Even though white and Asian families hold beliefs that are slightly more accurate than those of Black and Hispanic families, everyone is substantially misinformed about the quality of schools. Survey data that I collected among a more recent cohort of high school applicants validates the structural belief estimates. Beliefs elicited in the survey are inaccurate across all respondent races, but they are marginally more positively correlated with value-added and achievement levels among white respondents. Racial differences in preferences for school attributes are more important than beliefs to explain differing responses to the ratings. All students similarly trade-off preferences for attending higher quality schools with distaste for commuting. White and Asian students, however, strongly prefer the few public schools that are majority white and Asian, which makes them less interested and responsive to information about the quality of other schools.

I use the model to test whether information design can increase student achievement and close opportunity gaps. I begin by simulating the effects of providing perfect information about the value-added of each single school. This policy would cause students to rank schools with 0.07σ higher value-added on average. The larger response among Black and Hispanic students would close cross-race choice gaps conditional on baseline test scores. Due to some slack in the capacity of high-quality schools, students would also be matched to schools that have 0.01σ higher value-added on average, with marginally larger gains among minority and high-achieving students. This number correspond to 24% of the maximum possible achievement gains that would be realized if school seats were filled in order of quality. The ability to accurately measure school value-added is crucial, as simply providing information about school average achievement levels yields less than half of these test score gains.

Achievement gains for Black and Hispanic students under full information are comparable with those obtained through more controversial school admission reforms often targeted at reducing racial inequalities in New York City, such as removing admission priority based on test scores or residential address (Cohen, 2021). My simulations also show that providing information and leveling the playing field in admission rules are not substitute policies but their redistributive effects are cumulative. Information amplifies the gains for Black and Hispanic students and the displacement effects of removing unequal admission rules on high-achieving white and Asian students, rather than causing this latter group of students to reallocate to better schools.

Even if Black and Hispanic students respond more to information, their disproportionate benefit in equilibrium is significant because many high-quality schools use test-based admission standards that tend to screen out disadvantaged students. My simulations highlight that the stronger preferences of white and Asian students for schools enrolling more white and advantaged students are responsible for these gains. Their strong preference for school demographics dampens their response to information about quality and reduces the displacement of Black and Hispanic students from high-quality schools in equilibrium. The distribution of achievement gains would look very different if all students exclusively valued school quality and commuting time. In this scenario, information would substantially improve the quality of school offers for white and Asian students but hurt minority students.

These counterfactuals highlight that when seat capacity is fixed and everyone is informed, some disadvantaged students may still be displaced from high quality schools. Redistribution could be more easily achieved by providing information only to a targeted group of students, although this approach may be unfeasible due to fairness concerns. In the last part of the paper, I show that information can still be made public and designed to favor one group over another. Coarser information, such as partitioning value-added into school grades, can lead to better educational outcomes for low-achieving students compared to offering more detailed information. Intuitively, providing precise information about a school that remains non-desirable does not shift choices. Therefore, offering detailed information only about schools valued relatively more by lower achieving students, would limit the shifts in choices of higher achievers and therefore the competition for high quality seats. In counterfactual simulations, I show that a system that provides more precise quality ratings for schools at the bottom of the value-added distribution and coarser for those at the top, or a system that provides more precise ratings for undersubscribed schools than for oversubscribed schools, would benefit lower achieving students more than a system which provides perfect information about the value-added of every school.

This paper contributes to the literature studying household preferences for schools (Abdulkadiroğlu et al., 2020; Beuermann and Jackson, 2018; Hastings et al., 2009; Allende, 2020; Abdulkadiroğlu et al., 2017) and the effectiveness of school choice in raising achievement (Hoxby, 2000, 2003; Ladd, 2002; Rothstein, 2006; Cullen et al., 2006; Barseghyan et al., 2019; Campos, 2023a; Walters, 2018; Abdulkadiroğlu et al., 2018). In particular, it is related to studies considering the role of imperfect information in school choice (Allende et al., 2019; Bergman et al., 2020; Hastings et al., 2015; Kapor et al., 2020; Ainsworth et al., 2023; Campos, 2023b; Corradini and Idoux, 2023). Leveraging a natural experiment that made information about school value-added more easily available, I show that students care about school quality separately from peer quality. To do so, I estimate a model that does not rely on the direct elicitation of beliefs, which may be unfeasible when the set of schools is large.[1]

[1]Papers eliciting beliefs usually focus on a small subset of applicants and schools (Kapor et al., 2020; Campos,

I show that the two methods nevertheless provide qualitatively similar results comparing the model-based estimates against beliefs elicited using survey data.

Second, the paper relates to studies evaluating the effects of information interventions in education (Hastings and Weinstein, 2008; Mizala and Urquiola, 2013; Cohodes et al., 2022; Corcoran et al., 2022; Allende et al., 2019; Andrabi et al., 2017; Rockoff and Turner, 2010). Most of these studies rely on experimental evidence in which a random subset of students receive information and do not directly observe equilibrium effects. Additionally, in many of these studies families receive information about school outcome levels rather than value-added, so it is unclear what they learn about school quality and what they are responding to. The setting I study has several distinct advantages. First, letter grades were framed as measures of school effectiveness, which lets me learn about household preferences and beliefs for quality. Second, the information was provided by the school district rather than researchers, which may be more informative about the potential effects of a large-scale policy intervention. Third, I can directly observe congestion and displacement effects in the equilibrium when everyone is informed, which is important when school seats are scarce.

The paper also connects to two other distinct literatures: empirical studies of the distributional and efficiency effects of school assignment reforms, including affirmative action (Barahona et al., 2023a; Idoux, 2021; Black et al., 2023; Tincani et al., 2021; Bleemer, 2021; Kapor, 2020; Ellison and Pathak, 2021) and changes in admission rules (Dur et al., 2018; Park and Hahm, 2023), and the literature on the design of information disclosure policies (Vatter, 2022; Kamenica, 2019). This paper contributes to both by studying how information design matters for the type of students benefiting from information disclosure and by considering how to optimally coarsen quality ratings to implement the distributional objectives of a school district.

2 Background

2.1 The High School Match and School Performance Information

Every year New York City public schools enroll roughly 80,000 ninth graders at more than 400 high schools. Rising ninth graders apply for school seats by submitting an application to the centralized assignment system, ranking up to 12 academic programs.[2] Seats are allocated using the student-proposing deferred acceptance (DA) algorithm (Abdulkadiroğlu et al., 2005, 2009). Student priorities at a program depend on different factors, which vary depending on the program admission method type. There are three types of programs.

2023b)

[2]Schools may run more than one academic program, but most schools (70%) offer only one. For the purposes of this paper, programs and schools should be treated as synonyms.

Unscreened programs give priority to students based on residential zones and in some cases to those who attend an information session. Screened programs use these factors and also rank applicants based on prior grades, standardized test scores, attendance and/or program-specific requirements, such as essays, or auditions. Educational option programs use screened criteria for some of their seats and unscreened criteria for the rest. Random numbers are used to break ties among applicants with equal priority.

Parents lament that, due to the number and variety of programs and admission methods, gathering information about where to apply is difficult, costly, and time-consuming (Corradini and Idoux, 2023; Son, 2020). To aid families in their decision-making, the NYC Department of Education (DOE) assembles every year a high school directory and maintains a website with detailed measures about school performance. Before 2019 and throughout the period I study, the directory was provided in paper copy to every 8th grader in the city. This printed booklet was the main tool used by families to choose schools, as confirmed by conversations with staff at the DOE and by interviews conducted among middle school counselors by Sattin-Bajaj et al. (2018). [3] It provided an overview of the high school admission process, and an information page for each high school, which always included the school address, total enrollment, offered programs and their admission methods, courses and extracurricular activities, and a brief statement of its mission. The school pages also provided measures of school performance and student achievement, such as graduation rates.

Over the course of the years, the NYC Department of Education (DOE) changed its way of measuring and reporting school quality metrics on the school directory and online. Table A1 summarizes these changes during the study period. The most noticeable addition to the high school directory was the inclusion, from 2010 to 2015, of letter grades that graded schools from A to F. The letter grades provided a summative assessment of school performance and were part of yearly school progress report cards that were first introduced by the DOE in the fall of 2007 on their website.[4] [5] [6] They aimed to measure the school's contribution to

[3] Today, the DOE online application portal (MySchools.nyc) hosts a virtual version of what used to be the printed high school directory booklet before 2018. Survey evidence confirms it is still the source of information most widely used by high school applicants across all demographic groups, while reliance on other information sources varies across race (Corradini and Idoux, 2023). Black and Hispanic applicants use fewer sources of information than white and Asian households, are 19 p.p. less likely to rely on their family and friend networks for information about schools, and are 9 p.p. less likely to attend individual high school information sessions.

[4] Figure A1 shows an example of a school progress report card.

[5] The letter grades were introduced as part of a broader set of education reforms adopted by the Bloomberg administration after taking mayoral control of the city schools in 2002. The emphasis on school accountability was in line with Bloomberg's approach to reforming schools that was designed around market-based principles of improving school competition and incentives for school staff.

[6] Letter grades had consequences for school closures, financing and school principals. Schools receiving low grades could face leadership changes or closure, and students enrolled in F schools were eligible to transfer out through a special application process. Schools receiving an A grade received additional funding for the following school year of roughly $33 per student, and were eligible together with B schools for payments of

student academic progress, rather than simply measure school average achievement levels, like average test scores or graduation rates, which were still included alongside the grades.[7]

Grades were assigned based on a school's percentile rank within an underlying score. The scoring rule determining the grades varied slightly by education level. For high schools, it was based on three main measures: school environment (14% of the total score), student performance on Regents exams and graduation rates (30%), and student test score progress (50%). In 2010, a fourth component — college readiness — was added, accounting for 10% of the total score, which reduced the weight of the other components.[8] The scoring rule wasn't exactly based on causal estimates of school quality, but it aimed to control for differences in the student body and showed a positive correlation with causal test score value-added. A school's score for each element, in fact, was determined not only by the school relative performance city-wide, but also relative to a group of 40 "peer schools" with similar student demographics. Performance relative to peer schools was given double the weight of citywide relative performance in an attempt to separate school quality from student selection (Rockoff and Turner, 2010).

The first cohort of applicants who could use letter grades to make high school choices was the one applying for 9th grade in 2008. However, for the first two years, grades were only available online on the DOE website, requiring users to search school by school. Beginning with the 2010 cohort, applicants could easily view letter grades directly in the school directory. In 2014, the newly elected mayor, Bill de Blasio, removed letter grades from the high school directory and from the online school quality reports, and the following year his administration introduced a new approach to school quality measurement, vowed to be more holistic and less focused on test scores.[9] These new quality metrics, however, never made it to the printed school directory but could be consulted online, on a school by school basis.

$1500 to $3000 per student per year for any student accepted as a transfer from a failing school. Principals in the top 20% of scores were eligible to receive bonuses of $7000 to $25,000 (Rockoff and Turner, 2010). Rockoff and Turner (2010) studied the effects of introducing the grades on the incentives of elementary and middle schools to raise students test score in the first year the policy was introduced, finding that only receiving an F induced schools to raise test scores.

[7] The progress report was described as similar to a measure of school value-added on the school directories: *"The Progress Report measures each school's contribution to student academic progress, no matter where each child begins his or her journey to proficiency."*

[8] Schools could also receive additional points for improving student achievement from year to year among particularly vulnerable student subgroups (English Language Learner, special education students, and Black, Hispanic or LatinX students with performance in the lowest third of all students citywide). Appendix Table A2 describes the education outcomes used to compute the score in each component and reports the component weight in each year, before and after the assignment of the extra points.

[9] These changes apply to the cohort applying to enroll in the fall of the following year (in 2015 and 2016 respectively).

2.2 Data

I combine three main sources of data. The first is publicly available data from the school directories and online school quality reports issues by the NYC DOE between 2006 and 2016. The second is administrative data provided by the DOE covering all students enrolled in New York City public high schools between the 2006-2007 and the 2016-2017 school years. These data include student demographics and residence, school enrollment, student educational outcomes, including test scores on New York State standardized tests in middle school and high school, SAT and high school graduation, along with preferences submitted to the centralized high school assignment mechanism. An additional file from the National Student Clearinghouse (NSC) reports college enrollment and is internally linked to the DOE administrative data. I obtain public transport commuting time between schools and students' addresses at 7:30AM using publicly available APIs. The third source is novel survey data collected among 3500 parents of 9th grade applicants in 2023 and analyzed more extensively in a companion paper (Corradini and Idoux, 2023).[10]

I use student achievement data to construct two key attributes of high schools: school quality and peer quality. Peer quality is the average ability of students in a school, as measured by their average 7th grade standardized state Math test scores. School quality measures the causal contribution of schools to student achievement, as captured by school value added models (VAM) of high school standardized test scores. My main value-added measure is given by OLS estimates of α_j in the following regression:

$$Y_i = \alpha_0 + \sum_{j=1}^{J} \alpha_j D_{ij} + X_i' \Gamma_{t(i)} + \epsilon_i \tag{1}$$

where D_{ij} is a dummy indicating 9th grade enrollment in school j and X_i is a vector of baseline controls including race and ethnicity, subsidized-lunch, English Language Learner (ell) status, and lagged test scores (7th grade Math and English standardized state test scores). I allow the effects of X_i to vary by cohort, as denoted by $t(i)$. To measure student achievement in high school, Y_i, I primarily use New York state standardized tests in Math, called Regents, and use SAT Math scores as an alternative outcome.[11]

This model assumes that school quality is fixed over time and across student demographics and relies on a standard conditional independence assumption (CIA) that states that

[10] Appendix D provides more detail on the data and the survey.

[11] Regents exams are New York state standardized tests in core high school subjects required for graduation. As a result, most students take Regents exams, unlike the SAT, and school accountability measures in NYC are based on Regents test scores rather than on SAT. Another reason to prefer Regents to measure schools quality is that traditional OLS VAM based on SAT scores are a more biased measure of school quality, as I show in Table B1.

potential outcomes are independent of school fixed effects after controlling for the vector of student covariates X_i. In robustness checks in appendix B, I relax these assumptions in two ways. First, I allow school effectiveness to vary by student race. I use random variation in school offers embedded in the centralized school match to test how well VA estimates that do not vary by race ("Pooled VA") predict student outcomes and how they compare to the VA measures varying across student race (Angrist et al., 2016, 2021, 2022).[12] Appendix Table B1 confirms that pooled VAM have a good predictive validity for student Regents scores of both races. School value-added in fact does not vary much across student race, as confirmed by the strong within-school correlation of value-added measures for non-white and white students (Figure B1). Next, I relax the CIA by estimating risk-controlled (RC) VAM, as introduced by Angrist et al. (2021). RC VAM supplements the vector of controls with applicant characteristics integral to school matching, such as where they apply and the priority status that a school assigns them.[13] I use random variation in school offers embedded in the centralized school match to test how well conventional and RC VA estimates predict student outcomes (Angrist et al., 2016, 2021, 2022). These tests show that conventional and RC measures are equally unbiased and well predictive of student Regents test scores.

Table 1 describes the students in my sample, their choices and achievement outcomes. These students applied to enroll in 9th grade in NYC public schools and had non-missing data about their demographics, address, and middle-school test scores.[14] The district serves a racially mixed and disadvantaged urban population, with over 77% of students eligible for free or subsidized lunch. Throughout the analysis, I compare Black and Hispanic students (labeled as "Minority") to white and Asian students (labeled as "Non-Minority"). Panel B shows that school choice attributes are very similar within this binary race definition and significantly different across the two groups. On average, white and Asian students choose schools enrolling higher achieving peers, a higher share of non-minority students and with 5 p.p. higher graduation rates. They also choose higher quality schools, as measured by value-added: their choices rank 14 and 19 percentiles higher in the distribution of Regents and SAT VA. High school education achievement also varies greatly by race (Panel C). White and Asian students are respectively 17 p.p. and 23 p.p. more likely to graduate in time and enroll in college, and have Regents Math (English) scores 0.7σ (0.5σ) higher than those of minority students.[15]

[12] More details about the test statistics and the test implementation are provided in Appendix B.

[13] RC VAM estimates are unavailable for some schools in my sample and rely on data from a shorter time frame because they depend on the ability to replicate the high school match. I have the necessary data starting from the 2012 cohort of applicants, as some schools in my sample were phased out before then. For these reasons, I rely on conventional OLS VAM estimates of school quality, which are largely equivalent to the risk-controlled measures in this context.

[14] I exclude special education students because they participate in a fully separate school match with a different set of programs.

[15] The samples used to study outcomes exclude students enrolled at the nine specialized high schools because

Table 2 describes NYC high schools, focusing on the years when letter grades were issued. Each observation correspond to a school-year. The first two columns pool all observations together, while columns (3) to (6) split them by letter grade. Schools receiving higher letter grades on average enroll more advantaged students and are of higher quality, even if grades are only imperfectly correlated with value added as shown in Appendix Figure A2.[16] Schools receiving an A have a 0.25σ higher Regents VA than schools receiving a C, D or an F. If schools were classified correctly using causal estimates of quality as simulated in the last three columns of Table 2, however, the differences in Regents value-added between grade A and low-grade schools would have been twice as large.[17]

2.3 Documenting the Race Quality Gap

On average, Black and Hispanic high school applicants choose lower quality schools. In theory, this gap may be explained by differences in residential address or baseline achievement. If Black and Hispanic students lived in neighborhoods with lower quality schools traveling to better schools could be too costly. Additionally, if they did not meet the test score criteria for high-quality schools, applying to such schools would be futile. I show, however, that differences in schools attainable for students of different races cannot account for the choice gap.

For each student in my sample, I construct a feasible choice set formed by the schools located within 38 minutes by public transport from the student's home - the mean student commute - in which the student had a non-zero probability of admission.[18] Panel (a) of figure 1 illustrates racial differences in student top choices and in the best schools within their attainable options. Regardless of baseline achievement, Black and Hispanic students choose schools with 8 percentiles lower quality than white and Asian students. Differences across race and baseline achievement in the quality of the best three schools in students' choice sets are negligible, therefore minority and lower achieving students are leaving more value-added on the table.

To better quantify the importance of differing student characteristics and school avail-

they admit students via a separate process.

[16]The correlation is higher for Regents VA than for SAT VA, since the progress report score was primarily based on Regents performance and did not take SAT scores into account.

[17]I re-assign letter grades to schools based on the school Regents Math value added ranking, keeping the distribution (the count) of letter grades within years constant.

[18]A student has a non-zero probability of admission if (1) she has below marginal priority at the school, or (2) she has marginal priority and the schools uses lotteries to break ties or (3) she has marginal priority at a school using academic screens but where the test score cutoff was not binding for the student in that year, meaning her test scores was higher than the minimum score among admitted students with marginal priority.

ability in explaining the gap, I estimate the following regression:

$$Q_i = \alpha + \beta M_i + X_i'\gamma + \epsilon_i \tag{2}$$

. Q_i is the mean value added of applicant i's top three school choices (or in the school of enrollment, for comparison), M_i indicates Black and Hispanic applicants, and X_i is a vector of controls, such as baseline test scores and residential neighborhood.

Table 3 reports estimates of the coefficient β. The first column presents raw differences: Black and Hispanic students, on average, choose schools that have 14 (18) percentiles lower Regents (SAT) VA. These differences translate into enrollment gaps and contribute to achievement disparities: if minority students attended the same schools as their white peers, they would have 0.1σ higher test scores.[19] In column (2) I control for the mean quality and the quality of the best three schools in students' feasible choice sets as sufficient statistics for differences in geographic proximity to schools and differences in feasible options due to academic screening. Differences in the quality of attainable schooling options only explain between 25% and 30% of the gap. Directly controlling for students' zip codes reduces racial differences in school choices only by a third and further controlling for baseline achievement (column (8)) still leaves more than a third of the choice gap unexplained.[20] The gap unexplained by disparities in available schooling options is however likely to be larger, because residential zip code and baseline test score may be associated with two other candidate explanations for choice disparities: differences in information and in preferences. White and Asian students might value school quality, or other school attributes correlated with it, more than Black and Hispanic students. Higher quality schools in my sample, for instance, enroll more white and higher achieving students.[21] An alternative hypothesis is differences in information about school quality.

Figure 1 suggests that lack of information about quality might indeed play a role in this setting. Panel (b) plots the change in the quality of applicants' top three choices relative to the mean of the 2007 cohort, separately by race.[22] Choices improve over time, with larger changes among Black and Hispanic students: the raw racial choice gap shrinks from 19 percentiles of Regents VA in 2007, to 13 percentiles in 2013. The increase in the quality of

[19]The measure of school value-added used here does not vary across race and it may be potentially missing whether students are choosing schools that are a better match for their demographic group. The discussion in section 2.2 suggests that these concerns should be limited. Table B2 confirms that choice differences are remarkably similar, and if anything larger, when using a measure of value-added that varies by race.

[20]Zip codes in NYC correspond to relatively small geographies. There are 204 different zip codes values in my sample.

[21]The rank-rank correlation coefficient between school quality and share of non-minority students is 0.38 and the one between school and peer quality is 0.51.

[22]The percentile position is measured using the school relative ranking within the high schools participating in the high school match in that year to keep the measure comparable across years.

school choices is more marked during 2010-2014, the years when letter grades were printed on the high school directory. The trend reverses in 2015, when letter grades where removed.[23] [24] These patterns suggest that letter grades might have played an impact in directing choice towards higher quality schools, especially among Black and Hispanic applicants.

To gauge how well informed are families in this setting, I survey families who had just applied to NYC high schools asking them to situate real schools within the quality distribution of their residential borough.[25] Answers could vary from 1, corresponding to the worst 25% of schools, to 4, for the best 25%.[26] Table 4 shows the relationship between elicited beliefs and school quality and how this varies across respondents' race. The correlation between beliefs and school value-added is positive but low, and all families appear substantially misinformed. Even tough white and Asian students's beliefs are more positively correlated with value-added, the difference with Black and Hispanic respondents is not statistically significant. Beliefs are however more strongly correlated with school achievement levels, and significantly more among white and Asian respondents. When controlling for both achievement levels and VA, beliefs are positively correlated only with the first, which is consistent with the findings of Abdulkadiroğlu et al. (2020) who document similar patterns in measures of revealed preferences for schools. These correlations suggest that families rely on easily observed school attributes, such as average student achievement, to form opinions about a school's quality as it might be difficult for them to separate value-added from the composition of a school's student body (Rothstein, 2006; Ainsworth et al., 2023). White and Asian students appear to interpret achievement levels as a stronger signal of school effectiveness, which results in a slightly stronger correlation between their beliefs and actual quality.[27] Figure 2 plots the distribution of responses by respondent race for schools with achievement levels above and below the median. Most parents select the middle response, in line with a Bayesian model of belief formation in which families shade their evaluations towards the

[23] In 2017 the DOE introduced an online search engine - the *School Finder* - on the high school admissions website that simplified the information search process. The 2017 and 2018 cohorts also took part in a large RCT that provided information about high school graduation rates conducted by Cohodes et al. (2022). To keep the information environment comparable across years, I truncate my study period to 2016.

[24] Changes in SAT VA of top choices, shown in Figure A3, follow a similar but less pronounced pattern.

[25] The exact text of the question read: "How well does *school name - (school code)* prepare students for their Regents exams compared to other schools in your borough?". The distribution of school VA within each borough is essentially a replica of the distribution of VA within the city and most students rank schools in their borough among their first three choices.

[26] I randomize schools across respondents, sampling among relatively well known schools situated close to the respondent's address. More information on the survey and the selection of schools for this specific question is provided in appendix D.

[27] By design, the schools that respondents have to assess are not statistically different across respondent race after controlling for district of residence. The different responses by race are thus only the result of differences in perceptions about identical schools. Appendix table D2 confirms that schools populating respondents' questions are observably identical by showing balance of school value-added and achievement levels across respondent race, also conditional on the other school attribute.

city mean when observing imprecise signals about school quality. Through the lenses of this framework, white and Asian respondents seem to receive signals of school quality that are either more strongly correlated with peer achievement, more precise, or both.

3 Information and Choice

3.1 Applicants Respond to the Introduction, Changes and Removal of School Letter Grades

Grades might be correlated with demand for reasons unrelated to quality. By exploiting within-school changes in grades and demand, I establish that information about quality has a causal effect on demand for schools. I use two empirical strategies.

The first compares changes in demand before and after the introduction or the removal of quality ratings across schools that consistently receive the same letter grade. I divide schools into 4 categories: *Type A* schools are those receiving a grade of A in at least 5 out of the 7 years of school quality reports; *Type low* schools receive a low grade, C D or F, in at least 5 out of the 7 years; *Never graded* schools are those that were never graded; all remaining schools are pooled in the residual category of *Type Average* schools.[28] This classification into types is fixed over years, which allows me to compare choices for the same set of schools over time even in years when letter grades were not issued.

Figure 3 plots the raw trends in the average share of students ranking a school in their top three choices ("school share") by school category between 2006 and 2016. There is a marked substitution away from Type Low schools in favor of Type A schools after the introduction of letter grades.[29] The shares diverge substantially especially when letter grades are introduced on the school directory in 2010, while the trend reverses immediately after their removal in 2015.

An event-study model isolates grade effects over time. This can be written:

$$s_{cjt} = \sum_{L} \sum_{\tau=2006}^{2014} \beta_L^{t=\tau}(D_{jL} \times \lambda_{t=\tau}) + X'_{jt}\gamma + \mu_{ct} + \alpha_{cj} + \epsilon_{cjt} \quad (3)$$

. s_{cjt} is the share of students ranking school j as a top 3 choice in year t, who belong to demographic cell c, defined by a combination of student race, baseline test score tercile and

[28] There are 74 schools in the Type A category, 38 in the Type low, 273 are Average schools and 135 are never graded.

[29] The 38 schools that were always receiving low grades were large low performing schools evenly distributed across the four main boroughs of the city. Their total enrollment share in the city was around 19% in 2007. 10 of them were closed between 2013 and 2016, while enrollment in those that remained open dropped by 63% on average, bringing their total enrollment share in the city to 8%.

residential borough. The right hand side includes interactions between year dummies $\lambda_{t=\tau}$ with dummies D_{jL} indicating whether school j belongs to letter category $L \in \{$*Type A*, *Type Average*, *Type Low*, *Never graded*$\}$. I normalize $\beta_L^{t=2007}$ to zero for all letter categories, so that the coefficients of interest, β_L^t, captures the average within-school change in school share for schools of category L in year t relative to 2007. For the effect of removing letter grades from school directories in 2015, I estimate similar regressions using years between 2011 and 2016, normalizing the coefficients $\beta_L^{t=2014}$ to zero. I control for a time-varying vector of school attributes X_{jt} that includes measures of student achievement (average student performance on English and Math Regents exams, graduation and college rates) and the share of white and Asian students enrolled at the school. To account for differences in the visibility of these statistics, I also control for their interaction with an indicator for years when they were printed on the school directory.[30] α_{cj} represents school and cell fixed effects and μ_{ct} year and cell fixed effects to account respectively for differences in unobserved preferences for school characteristics that are fixed over time within demographics and time-varying changes in demand across demographics, for instance due to opening and closure of schools in different neighborhoods. Standard errors are clustered at the school-year level.

Figure 4 plots the coefficients β_L^t while table A3 reports pooled pre-post coefficients. The introduction of positive quality signals boosts demand and the introduction of negative signals depresses it, while their removal has opposite effects. Demand for schools consistently receiving an A increased by 26% while that for Type Low schools dropped by 66% when considering pooled estimates. Trends in demand of the different school categories were parallel prior to the introduction of letter grades (and are similarly parallel after their removal), supporting the view that changes in choices are caused by the introduction and removal of grades.[31]

The second empirical strategy studies the effects on demand of year-to-year changes in grades. It also exploits within-school changes in signals and demand, but focuses on the variation coming from schools receiving different grades over the years. I regress school shares among student top choices on letter grade dummies, maintaining the same controls as in equation (3):

$$s_{cjt} = \sum_g \beta_g D_{jtg} + X'_{jt}\gamma + \mu_{ct} + \alpha_{cj} + \epsilon_{cjt} \tag{4}$$

.

D_{jtg} indicates that school j received a grade of g in year t, while the rest of the notation

[30] Table A1 summarizes the information visible on the school directories and online during the study sample.

[31] The parallel trends and the timing of the changes support the hypothesis that they are due to the introduction of school grades, rather than to the bundle of reforms introduced by the Bloomberg administration after it took mayoral control of the city schools in 2002. Moreover, there is no reason to believe that these reforms, including changes in the school district management and structure and teacher pay reforms, affected schools differently by the letter grade their received.

remains the same. The identifying assumption that allows to interpret β_g as the causal effect of receiving a grade of g is that conditional on school and time fixed effects and on observable time-varying characteristics, change in letter grades are independent of changes in unobserved preferences for schools.

Table 5 shows the estimates of β_g based on the sample of applicants who enrolled in 9th grade between 2010 and 2014 and who had access to letter grades printed in the school directories.[32] The omitted category in columns (1) and (3) is school-years not receiving a letter grade, while columns (2) and (4) restrict the subset of school-years to those with a letter grade, which comprises older and larger schools.[33]

Year-to-year changes in grades shift demand for schools substantially: an A increases the share of students choosing a school as top choice by 0.15 p.p. on average, an increase of about 25% with respect to the average school share. An F reduces the probability a school is ranked as a top choice by 0.21 p.p., or 34% of the average school share. Receiving a grade of C is approximately equivalent to receiving no grade. Column (3) uses school log shares as left hand side variable, which yields consistent estimates. These estimates are consistent with the magnitude of the pooled pre-post changes following the introduction of high letter grades, but smaller than those following the introduction of persistently low grades. This suggests that bad reputation following a low grade may be more sticky over time than positive publicity from a high grade. The table also reports the effect of graduation and college rates on demand, which is positive only for graduation rates and only when these are printed on the school directories, suggesting the existence of significant costs of searching for information about schools when this is not made easily available by institutional sources.

Letters were presented to families as measures of causal school value-added. Sophisticated families, however, might have realized that grades reflected in part student selection or other school features different from value-added. Their responses to changes in letter grades might therefore not only be indicative of preferences for school quality and families' beliefs about it, but rather of preferences and beliefs for a mix of school attributes. I argue this is not very plausible. If families were sophisticated and knew what goes in the quality score, they would be aware that it is made of different components and should respond to changes in the score sub-components depending on their taste for these different school attributes. In

[32] Each school page in the directory usually displays two letter grades from the previous two school years. In this table, I only estimate the effect of the most recent letter grade, while the effects of the two-letter combination is addressed in Appendix Tables A5 and A6. Estimates in Table A5 suggest the two consecutive grades have an additive effect and that overall families put more weight on the most recent signal. The more flexible model of Table A6, which considers the effects of any combination pair of letters yields similar estimates as the additive model.

[33] Not all schools received letter grades in all years. This happened if the school had recently opened and/or the student sample size with achievement data was deemed too small to compute reliable quality score estimates.

Appendix Table A7 I show that changes in demand are not correlated with changes in the quality score components after controlling for changes in letter grades. In fact, choices do not even respond to changes in the main underlying quality score after accounting for changes in grades, suggesting that applicants only paid attention to the latter.

In sum, the evidence presented in this subsection shows that school choices responded to information about school quality. This means that families must value school value-added but hold uncertain and possibly inaccurate beliefs about it. Misinformation about school quality, then, likely explains why Abdulkadiroğlu et al. (2020) in the same context find that families choices are only correlated with peer quality and not with school value added.

3.2 Black and Hispanic and Less Informed Students Respond More

Some studies, including in education, have found larger responses to information provision among richer households, suggesting that barriers to information take-up may be larger among the least affluent or least well connected (Corcoran et al., 2022; Bhargava and Manoli, 2015; Bergman, 2020). A standard model of Bayesian updating would instead predict larger responses among disadvantaged families, who are likely to be less connected and less well informed. Survey respondents in fact complaint that having the means to hire admission consultants or the time to attend information sessions results in unequal access to information across income and race. This discussion motivates studying heterogeneity in responses to letter grades across student demographic and socio-economic background.

Panel (b) of figure 3, plots raw trends in demand for different school types, by race. White and Asian students were already choosing schools that would have received higher grades prior to 2008, ranking Type A schools (Type low schools) among their top choices 50% more often (60% less often) than minority students. These patterns suggest that Black and Hispanic students may have been less well informed before the introduction of grades, and therefore they might respond more to information about school quality. To measure whether responses to the introduction and removal of grades are statistically different across race, I extend equation (3) and estimate a triple difference model in which I interact a dummy M_c indicating Black or Hispanic student covariate cells with year and school category indicators:

$$s_{jct} = \sum_{L} \sum_{\tau=2006}^{2014} \left(\delta_L^{t=\tau} (D_{jL} \times \lambda_{t=\tau} \times M_c) + \beta_L^{t=\tau} (D_{jL} \times \lambda_{t=\tau}) \right) + X_{jt}' \gamma + \mu_{ct} + \alpha_{cj} + \epsilon_{jct} \quad (5)$$

.

I normalize the baseline difference in share by race $\delta_L^{t=2007}$ to zero for all school categories L. Figure 4 plots the estimates of δ_L^t, confirming there is substantial racial heterogeneity in responses to information.[34] Black and Hispanic students' responses to the introduction and removal of letter grades are at least twice as large as those of white and Asian students on average. Figure 6 and Table A4 explore heterogeneity in responses to year-to-year changes in grades estimating equation (4) on race-specific school shares, which yields consistent findings. Receiving an A increases demand by 30% (of the average school share) among Black and Hispanic students, but only by 14% among white students. Symmetrically, receiving an F decreases a school share among Black and Hispanic choices by 48% while it has non statistically significant effect on the choices of white and Asian students.

Cross-race differences in letter grade premia on log shares (columns (7)-(12)) are much smaller compared to the effects on share levels. This means that white and Asian student choices are responsive to changes in the grades of the schools that they were choosing at higher rates but not responsive on average. In other words, these estimates indicate that positive grades alone may not be sufficient to persuade white students to select schools they had not previously considered. This suggests that white students may have stronger preferences for the attributes of a select group of schools, making them less responsive to changes in perceived school quality on average.

Even within students of the same race, some may be better informed and respond less to information disclosure. A model of Bayesian updating predicts larger responses among applicants most surprised by grades. Using student choices before the introduction of grades, I classify students who should have been more surprised by the new information based on covariates related to their demographics, middle school and neighborhood of residence. Alignment of choices with the information included in letter grades is measured by the variable $Information\_index_i$, which takes values in {-1, 0 , 1}. An applicant starts from a value of 0 and gets a point if she ranked in her top three choices a Type A school, and is subtracted a point if she ranked a Type low school.

I use the following lasso regression and a random 50% subsample of applicants in 2007 to estimate which student covariates predict concordance of choices with school grades prior to their introduction:

$$Information\_index_i = \alpha + \lambda_{MS(i)} + \lambda_{z(i)} + \lambda_{d(i)} + X_i'\beta + \epsilon_i \tag{6}$$

The right hand side covariates capture potential reasons why households might have been more or less informed about school quality before the introduction of letter grades, such as exposure to different social networks. They include middle school fixed effects $\lambda_{MS(i)}$, residential school district fixed effects $\lambda_{d(i)}$, and zip code fixed effects $\lambda_{z(i)}$. The vector X_i also includes a gender indicator, an indicator for subsidized lunch eligibility and one for

[34] Appendix figure A4 plots event study estimates of coefficients β_L^t in equation (3) separately by race.

English language learners, the share of students in the same middle school from the previous cohort ranking Type A schools and Type Low schools among their first choices.

I use the estimates to predict $Information\_index_i$ for all remaining students in the applicant sample, and split the sample in two by the median predicted value in 2007. Students whose predicted information index is below the median should be more surprised by the introduction of letter grades and I will denote them as *Treated*. Appendix table A8 shows that treated students are more likely to be Black or Hispanic, but that there is substantial heterogeneity in exposure to the information conveyed by grades even within race. Treated students are more likely to be English language learners, eligible for subsidized lunch, and have lower baseline test scores. They are more likely to live in the Bronx, less likely to live in Manhattan, and attended middle schools were students in 2006 were more likely to apply to schools receiving lower grades.

Appendix Figure A6 plots estimates of δ_L^t for a version of equation 5 that considers heterogeneity along treatment status rather than across race.[35] The results confirm that demand responses to the introduction and removal of letter grades, for both white and minority applicants, were larger among treated students. These patterns suggest that misinformation may be more prevalent among certain students, and that misinformation may partly explain initial differences in student choices. If the differences in the alignment of choices and grades before 2008 were solely indicative of differing school preferences of perfectly informed applicants, then the introduction and removal of grades should not have influenced the choices of families whose characteristics predict a preference for low-grade schools. [36]

Differences in responses across race or other demographic characteristics, however, need not be explained only by difference in information. Informing families that a school has high value-added may not be sufficient to encourage them to apply, especially if they place little value on school effectiveness or if the school has other undesirable characteristics. That is, lower preferences for school quality relative to other school attributes may also explain why certain families are less responsive to information about quality. In the next sub-section, I present evidence suggesting that this may help explain why white and Asian families were less responsive to grades.

[35] Appendix Figure A5 instead reports estimates of β_L^t for versions of equation (3) that separately use students more or less exposed to new information within race.

[36] Regression to the mean cannot explain these results both because I am using a different sample for prediction and estimation and because we see a reversal of demand back to the initial application patterns right after the removal of grades.

3.3 What School Attributes Influence Reactions to Quality Signal?

Preferences for Other School Attributes Mediate Intensity of Responses To gain insight on how information about quality interacts with preferences for other school attributes, I consider how the magnitude of demand responses varied across schools with different characteristics. Rather than separately estimating the effect of each letter grade, I estimate the effect of changing grades by constructing a quality signal index S_{jt} that varies from 1 (corresponding to an F) to 5 (corresponding to an A). To test if responses to information vary across schools, I interact this index with school attributes X_j, indicating that school j enrolls a high share of white students, has high peer quality, or has high average Regents test scores, depending on the specification.[37] I further interact these regressors with a minority dummy M_c, to measure cross-race differences in responses to grades and across schools. The resulting estimating equation is:

$$s_{cjt} = \beta_0 S_{jt} + \delta_0(S_{jt} \times X_j) + \beta_1(S_{jt} \times M_c) + \delta_1(S_{jt} \times X_j \times M_c) + \mu_{ct} + \alpha_{cj} + \epsilon_{cjt} \qquad (7)$$

. I maintain school fixed effects, so that demand responses are identified off of within-school changes in letter grades.

Table 7 shows that white and Asian students' choices responds to changes in letter grades, S_{jt}, only within schools enrolling high shares of white and high achieving peers. Choices of Black and Hispanic students are both more responsive on average and equally responsive to changes in the grades of schools with different attributes. These findings align with the view that white and Asian student choices are concentrated on a small subset of schools and that they respond to information about school quality only within these schools. This also helps explain why, in table A4, heterogeneity in log share responses to grades across race was smaller than in the specification using share levels. Preferences for the demographic composition of a school might be responsible for these patterns, limiting white students' response to information within a specific set of schools that have desirable demographics.

Beliefs Mediate Choice Responses to Positive and Negative Signals Responses to information about quality should be influenced not only by preferences for other school characteristics but also by families' prior beliefs about school quality. Survey data described in Section 2.3 indicates that families' beliefs about school quality are positively correlated with a school's average Regents test scores. If applicants update their beliefs according to Bayes' rule, we should observe larger increases in demand for schools with lower achievement

[37] These are indicators for being in the top third of the city distribution of these three dimensions (share white, peer quality, achievement levels).

levels after they receive a high grade compared to receiving no grade, as applicants initially perceive them to be of lower quality. Conversely, we would expect to see larger decreases in demand for schools with higher achievement levels after they receive a low grade.

Table 6 presents estimates of equation (4) separately for schools with above and below median achievement levels at baseline, showing they are consistent with these predictions.[38] Positive responses to higher letter grades (A or B) are larger among lower performing schools, while the negative responses to letters D and F are larger for higher performing schools. Moreover, while the estimate of β_C is positive and that of β_D is non significantly different from zero for lower performing schools, they are negative for higher performing schools, suggesting that receiving a letter grade of C or D is perceived as a negative surprise only if the school had high achievement levels. Regressions using log shares on the left hand side (columns (3) and (4)) yield qualitatively similar results.

Appendix figure A7 explores the same type of heterogeneity in responses to quality signals using the introduction and the removal of letter grades. It plots the average percent change in demand for Type A schools, distinguishing between schools with high or low achievement levels. Once again, the increase (decrease) in demand following the introduction (removal) of positive letter grades is larger for lower (higher) performing schools. There is no differential change, however, in responses to the introduction or removal of a persistent negative signal. Taken together, these estimates suggest that taking into account how households form and update beliefs seems important to predict the effects of counterfactual information disclosure policies. This discussion motivates the need to estimate the joint distribution of school preferences and beliefs, along with its heterogeneity across students, an exercise I undertake in section 5.

4 Consequences for Racial Inequality

This section studies the consequences of the larger response to grades of minority students on racial inequality in education outcomes. I estimate the following event study regression, comparing student outcomes across race and cohorts:

$$Y_i = \sum_{\tau} \delta^{t=\tau}(M_i \times \lambda_{t=\tau}) + \mu_t + X_i'\gamma + \epsilon_i \tag{8}$$

.

M_i indicates Black and Hispanic students and $\lambda_t = \lambda_{t(i)}$ are cohort indicators. The vector of controls X_i includes ethnicity, gender, ell status, subsidized lunch status and fixed effects

[38] I define a school baseline achievement level as the average yearly performance on Regents math exams of its students, taking a cross-year average for years before the introduction of letter grades on directories (2006-2009).

for combinations of student borough and baseline test score terciles to account for potential changes in the composition of students over time.

I first compare changes in school choices to changes in school offers in Figure 7, and show pooled diff-in-diff coefficients in Table 8. I plot estimates of δ^t for four different outcomes: ranking a Type A or Type low school among one's top three choices, or receiving an offer to these schools. After the introduction of grades, both the choices and offers for minority students improved compared to those of non-minority applicants. Minority students were up to 10 percentage points more likely than white students to rank Type A schools and up to 15 percentage points less likely to rank low-grade schools among their top choices.[39] These significant relative changes are not due to a lack of potential for improvement in white students' choices; only 64% of white students were ranking Type A schools before the introduction of letter grades. After the removal of grades, the racial gaps in choices partially revert to their pre-letter grade levels.

The effect on racial gaps in offers, however, is smaller than that on choice gaps. Minority students were only 3 p.p. more likely to receive an offer to a Type A school and only 4 p.p. less likely to receive an offer ot a Type Low school. That is, while the racial gap in Type A choices closed by 43%, the corresponding enrollment gap was reduced only by 23%.[40] These results highlight how in markets with binding capacity constraints information interventions may shift everyone's choices but need not translate into large average achievement boosts. Understanding how assignment rules and capacity constraints clear the market in the presence of increased demand for high quality schools becomes important to gain insight on who are the winners and losers of information disclosure policies.

In this context, while school admission rules based on middle school test scores help explain why Black and Hispanic students had relatively lower chances of receiving offers from Type A schools, binding capacity constraints play an even more important role. Appendix Table A9 compares the effect of introducing grades on the probability of receiving an offer to Type A or Low schools to the corresponding effect in simulated assignments. These simulations assume that schools prioritize students solely based on their random lottery number, rather than using middle school test scores or residential address priorities. The simulated relative increase of minority students' changes of receiving an offer to Type A schools is larger by 1 p.p. than in reality, yet it remains much smaller than the relative change in choices. Higher congestion in the grade-A schools chosen by Black and Hispanic students then accounts for the majority of the discrepancy in the effects on choices and offers.

[39] Changes in first choices, presented in Panel A of Table 8, reveal that after the introduction of letter grades Black and Hispanic students would have been 5 p.p. more likely than white students to enroll in a type A schools if schools had unlimited capacity.

[40] Effects on enrollment schools follow closely the effects on school offers because compliance is high. 73% of those offered a school in the first round of the match are enrolled at that school in June of their 9th grade.

What are the consequences of these changes in school assignment? The remaining columns of Table 8 show the effects on introducing and removing grades on different attributes of students' choices, school offers or enrollment schools. Everyone chooses and is matched to schools with higher value-added and peer quality after the introduction of letter grades.[41] Minority applicants, however, rank and are matched to schools with 4.5 percentiles higher quality (or 0.03σ) compared to white students, which corresponds to a 25% reduction in the baseline cross-race gap.

Appendix Table A10 confirms that, as a result, there is a reduction in racial disparities in high school achievement for cohorts applying after the introduction of grades. Panel A pooled difference-in-differences estimates of equation (8), comparing changes in achievement across race. Panels B and C provide similar analyses within racial groups, focusing on levels of the information exposure dummy introduced in section 3.2. After the introduction of grades, minority students' Regents Math test scores improve by 0.06σ more than those of non-minority students. Additionally, racial gaps in on-time graduation and college enrollment rates are reduced by 5 p.p. and 7 p.p., respectively. There is no differential effect on SAT Math test scores, which is consistent with the higher correlation of the quality score with Regents VA compared to SAT VA.[42] Similar patterns are observed in the difference-in-differences effects within racial groups when examining heterogeneity in exposure to information: the Regents scores, graduation rates, and college enrollment rates of less informed students improve relative to those of more informed students.

The reduction in racial inequality following the introduction of grades is likely not solely attributable to the reallocation of minority students to better schools, as the decrease in achievement gaps exceeds the reduction in differences in the value-added of school offers. This suggests a combination of reallocation effects and increased competitive pressure on schools to improve quality, which may have impacted more the schools where minority students received offers. I leave the study of these potential changes in supply-side incentives to future work.

[41] Figure A8 displays trends in the average characteristics of students' top three choices, illustrating that the timing of these changes coincides with the introduction of grades in the school directory. The bottom right panel also indicates that, when letter grades were in effect, students were more likely to rank schools outside their borough, suggesting they may be willing to travel farther to attend schools they perceive as higher quality.

[42] Effects on racial differences in offered SAT VA, which are not reported here, are in fact null.

5 A Model of School Choice with Imperfect Information About School Quality

This section presents a model of school choice with imperfectly informed students, which I estimate leveraging the variation in quality signals introduced by the letter grade system. I use the model to shed light on which features of supply and demand for schools, including families' preferences and beliefs, school location, admission rules, and capacity constraints, are important to explain differences in application patterns and the equilibrium effects of providing more information.

5.1 Set Up

Applicant i's indirect utility from attending school j is additive separable in distance and linear in school characteristic. It can be written as:

$$u_{ij} = \underbrace{X'_{jt}\beta_{c(Z_i)} + \gamma_{c(Z_i)} E_{f_{c(Z_i)j}}[q_j|s_{jt}] + \xi_{c(Z_i)jt}}_{\delta_{cjt}} - \lambda_{c(Z_i)} d_{ij} + \epsilon_{ij} \tag{9}$$

, where δ_{cjt} denotes the average utility from attending school j for students of demographic cell $c(Z_i)$, applying to high school in year $t = t(i)$. Student demographic cells are defined based on the vector of covariates Z_i. In the empirical estimation, cells correspond to combinations of student baseline test score terciles and race (Black, Hispanic and white or Asian).

Students' utility depends on their preferences for school quality, which is not perfectly observed by students, and for other known school attributes. These include characteristics of school j in year t observable to both the student and the econometrician, denoted by X_{jt}, the distance d_{ij} between student i's home and school j, and ξ_{cjt}, which captures preferences for school characteristics unobserved by the econometrician. X_{jt} includes, in particular, peer quality and the share of white and Asian students enrolled at the school, both measured in the year before applicants submit their applications. ϵ_{ij} captures idiosyncratic tastes for schools, which I assume are distributed according to a type-1 extreme value distribution with a location parameter equal to 0 and scale parameter $\sigma^{\epsilon}_{c(Z_i)}$. The coefficient $\lambda_{c(Z_i)}$ is normalized to 1 for all students to specify a distance-metric utility function, an approach often adopted in the school choice literature (Agarwal and Somaini, 2020). [43] The utility of the outside option is normalized to zero $u_{i0} = 0$ for all students.

Departing from standard models of school choice that assume perfect information, I

[43]Because the scale parameter of the structural error distribution can vary across student cells c, this normalization does not impose restrictions on how students with different demographics trade off various school characteristics.

assume that students' utility depends on their expectation of the quality of school j, which I denote with $E_{f_{c(Z_i)j}}[q_j|s_{jt}]$. Students form expectations about school quality using their prior beliefs, which are distributed according to $f_{c(Z_i)j}(q)$, and a signal s_{jt} about the quality of the school, when provided by the school district in the form of a letter grade. Because grades may change from year to year, different cohorts of students may receive different signals about the quality of the same school. Grades partition the quality space into intervals, $[\underline{c}_L, \bar{c}_L]$ for $L \in \{A, B, C, D, F\}$. I assume students know the quantile cutoffs used to assign grades and update their beliefs according to Bayes rule after observing a school letter grade.[44]

In the main specification, I assume that priors are distributed as a truncated normal with mean μ_{cj} and standard deviation σ_{cj} over the space $[\underline{q}, \bar{q}]$ of value-added in the city.[45] [46] To reduce the number of parameters in the estimation, in the main specification I model prior moments as linear functions of school value-added and school achievement levels. In robustness checks, I let them vary non-parametrically across discrete school types. Finally, I assume that within-school changes in letter grades over time are orthogonal to within-school changes in preferences for school characteristics unobserved by the econometrician. [47]

This model can explain patterns observed in the reduced form analysis. If students value school quality but there is uncertainty in their beliefs, their choices would positively respond to high letter grades and negatively to low ones. If they update their beliefs according to Bayes' rule, their choices would shift more when quality signals are surprising, such as when high performing schools receive low grades. Finally, differences in responses to grades across student race could be explained by differences in their preferences for quality relative to other school attributes, or differences in belief precision and bias.

Microfounding Heterogeneity in Prior Beliefs Across Students Why would different students hold different priors? One source of heterogeneity across families of different backgrounds is differences in access to information through sources like social networks,

[44]Previous literature using similar methods often maintains the same assumption (Vatter, 2022; Barahona et al., 2023b; Dranove and Sfekas, 2008; Chernew et al., 2008). In this setting the assumption is credible because the cutoffs in terms of percentiles and quantiles of the quality score distribution were clearly communicated, as can be seen from the progress report in figure A1.

[45]Since the empirical distribution of value-added across schools in the city is approximately normal, as shown in appendix figure C2, this functional form restriction implies that students' beliefs align with the overall distribution of school quality in the city, with scale and location adjustments that are school-specific.

[46]Thanks to this functional form, the expected quality of school j given the letter signal L is conveniently the mean of a twice truncated normal given by: $E_{f_{cj}}[q_j|s_{jt} = L] = \int_{\underline{c}_L}^{\bar{c}_L} \frac{f_{cj}(q)}{F_{cj}(\bar{c}_L) - F_{cj}(\underline{c}_L)} dq$.

[47]That is, assuming that preferences for school unobserved characteristics can be decomposed into a component that is fixed over time and one that varies over years, $\xi_{jt} = \tilde{\xi}_j + e_{jt}$, I can re-writing δ_{cjt} as: $\delta_{cjt} = X'_{jt}\beta_c + \underbrace{\gamma_c E_{f_{cj}}[q_j|s_{jt}]}_{\eta_{cjs}} + \tilde{\xi}_{cj} + e_{cjt}$ I am assuming, $E[e_{cjt}|X_{jt}, \eta_{cj}, \tilde{\xi}_{cj}] = 0$, where η_{cjs} are fixed effects for combinations of schools and quality signal (letter grades or their absence), and η_{cj} stacks the fixed effects relative to one school.

school admission consultants, guidance counselors or different use of online websites. For simplicity, I will refer to these additional sources as social networks. Formally, assume that before receiving signals from their social networks, all students hold similarly uninformed priors $f_{ij}(q) = f^U(q)\ \forall i, j$, identical to the distribution of quality in the city. Each student i receives a signal from her social network about the quality of school j, n_{ij}, before observing any rating of quality provided by the policy maker. The parameters governing the distribution of students' beliefs will depend on the mean and precision of the signals they receive from their social network. Considering the simple case in which the distribution of school quality in NYC is well approximated by a normal $f^U(q) \sim N(\mu_q, \sigma_q)$, and students receive social network signals that are also normally distributed, $n_{ij} = \tilde{\mu}_{ij} + e_{ij},\ e_{ij} \sim N(0, \sigma_{e_{ij}})$, the resulting belief about the quality of school j, f_{ij}, will be also normally distributed. Its mean μ_{ij} and variance σ^2_{ij} depend on the social network signal as follows:

$$\mu_{ij} = \mu_q + \frac{\sigma_q^2}{\sigma_q^2 + \sigma_{e_{ij}}^2}(\tilde{\mu}_{ij} - \mu_q)\ , \qquad \sigma_{ij}^2 = \frac{1}{\frac{1}{\sigma_q^2} + \frac{1}{\sigma_{e_{ij}}^2}}$$

. Students receiving more precise social network signals (smaller $\sigma_{e_{ij}}$) will have less uncertain beliefs about the quality of school j. The social network signal mean and precision will also govern how far students believe the quality of j is from the average school in the city.

5.2 Estimation and Identification

I adopt a two-step estimation procedure similar to the one in Goolsbee and Petrin (2004). In the first step I use students' rank-order lists to estimate δ_{cjt} with maximum likelihood. In the second step, I use a minimum distance estimator to decompose the mean utility $\hat{\delta}_{cjt}$ in its main components. In the second step, I leverage within-school variation in letter grades and in student preferences, to account for systematic preferences for specific schools that are fixed over time. The resulting estimator is:

$$\min_{\theta_c} \sum_j \sum_t \sum_{\tau > t} (\Delta\hat{\delta}_{cjt,\tau} - \Delta X_{jt,\tau}\beta_c - \gamma_c \Delta E_{f_{cj}}[q_j | s_{jt}, s_{j\tau}])^2$$

The parameter vector $\theta_c = \{\beta_c, \gamma_c, \boldsymbol{\mu_c}, \boldsymbol{\sigma_c}\}$ varies across student demographic cells c and includes the parameters governing preferences for school characteristics and the vector of prior moments. I recover estimates of $\tilde{\xi}_{cj}$, the time-invariant component of preferences for school unobserved attributes, from average residuals: $\hat{\tilde{\xi}}_{cj} = (\sum_t \hat{\delta}_{cjt} - X_{jt}\hat{\beta}_c - \hat{\gamma}_c \widehat{E[q_{jt,\tau}]})/T$.

In the estimation, I focus on applicants enrolling in 9th grade between 2011-2015, relying on variation in letter grades within years (2011-2014) and on their removal in 2015.[48] I use

[48] I could also rely on the introduction of letter grades as an additional source of variation, but choice shifts

2016 applicants to assess model fit out-of-sample and for counterfactuals.

The identification of the mean utility δ_{cjt} in the first step relies on a standard revealed preference argument, which is valid under the assumption that students rank schools in order of true preference and that the ranked schools are preferred to the outside option. Truthful reporting is often assumed when DA is used to allocate students to schools, because it is strategy-proof when applicants are allowed to rank every school. Even if the number of school choices is capped at 12 in NYC, most students do not fill their list, and truthful ranking is a dominant strategy also in this situation. I relax the trutful reporting assumption in a robustness check in table A15.

The identification of preferences for school quality and prior beliefs in the second step requires instead a more careful discussion. Letter-grade demand premia for different schools can be estimated as part of fixed effects for combinations of grades and schools, which are identified from willingness to trade distance for higher letter-grades, all else equal. However, simulating the equilibrium effects of counterfactual information disclosure policies demands additional structure in order to understand how families update beliefs under scoring designs that use different quality cutoffs. The challenge is to tell whether students are willing to travel further to enroll in schools receiving higher letter grades because they believe the quality difference is small but very valuable (i.e., γ is large) or because they value quality little but they are updating their quality belief a lot, for instance due to large uncertainty or large biases (i.e., γ is small). I adapt the argument used in Vatter (2022) to my setting and show in appendix C.1 that within-school changes in letter grades over the years would lead these two configurations to generate systematically different choices. As in Vatter (2022), I maintain the assumption that the payoff from quality enters the utility linearly and, as already discussed, that students understand the school letter-grade cutoff structure.

The intuition behind this argument is that the assumption that families know the letter grade cutoffs implies bounds on belief updating. In turn, this implies bounds on preferences for quality, given students' willingness to commute for increments of letter grades. For instance, the quality score cutoffs used to assign letter grades imply that the quality of B-schools is bounded between $[-0.4\sigma_q, 0.5\sigma_q]$ while the quality of D-schools is in $[-1.4\sigma_q, -1\sigma_q]$, where σ_q denotes standard deviations of the distribution of quality across schools. If the willingness to commute for a B-school, relative to a D-school is 5 minutes, simple algebra shows

following the grades removal in 2015 are more informative of student beliefs in 2016. Differences in choice responses to the introduction and the removal of grades suggest that some learning occurred, potentially due to some stickiness in reputation. Using 2006-2014 to estimate the model yields similar estimates, except for prior means. Minority students' beliefs before the introduction of grades were negatively correlated with the school achievement levels. This is no longer the case today, as validated by the survey evidence. I also exclude 2010 because the directory in 2010 does not show graduation rates and I want the information environment to be the same in all years, except for changes in grades. Nevertheless, adding 2010 to the sample changes estimates very little.

that the change in expected quality is bounded between $\Delta E[q] \in [0.6\sigma_q, 1.9\sigma_q]$. This implies that γ can be bounded between $[2.6, 8.3]$ minutes. Variation in letter grades (and their absence) within and across schools generates additional bounds. With sufficient variation in the quality signals that a school receives, the priors could be non-parametrically identified. In practice, because the variation is limited, the functional form of the priors allows me to achieve point identification. Identification of preferences for time-varying school characteristics, namely peer quality and the school racial composition, comes from within-school changes in these characteristics over the years.[49]

5.3 Estimates and Model Fit

Table 9 summarizes the model estimates. It shows weighted averages of cell-specific estimates across cells sharing the same covariate (race or baseline test scores), using weights proportional to cell size. The cell-specific estimates and their asymptotic standard errors are reported in Appendix Table A11.[50] Panel A reports summary statistics of the first step estimates of the mean school utility δ_{cjt}. Mean school utilities are positively correlated with peer and school quality, more so for white and higher achieving students. Their within-cell standard deviation ranges from 21 to 28 minutes of public transport commute.

Panel B presents summary statistics for the second step estimates of the preference parameters $\gamma_c, \beta_c, \tilde{\xi}_{cj}$ and panel C for prior beliefs. In this benchmark version, I model the prior mean and precision about the quality of a school as a linear function of its average Regents test scores R_{jt} and its value-added Q_j: $\mu_{jt} = \mu_0 + \mu_1 \cdot R_{jt} + \mu_2 \cdot Q_j$, $\sigma_{jt}^{-1} = \sigma_0 + \sigma_1 \cdot R_{jt} + \sigma_2 \cdot Q_j$. This specification allows for the case in which students have accurate beliefs about school quality (namely $\mu_0 = 0, \mu_1 = 0, \mu_2 = 1$). In robustness checks, I instead let prior moments vary non-parametrically across discrete school types. All models yield largely consistent estimates.

Estimates provide two explanations for the differences in responses to information across students. The first relates to differences in beliefs, characterized by smaller biases and lower uncertainty in the priors of white students. While all students recognize that schools with higher value-added and achievement levels are of higher quality, the prior means for white

[49] Peer quality is highly correlated with a school average achievement level, R_{jt}. If peer quality and achievement levels change together from year to year, both school preferences and beliefs about its value-added change. Thus one might be worried about the separate identification of the two. However, preferences are identified from within-school changes in X_{jt} over time, while beliefs thanks to changes in grades. Even if X_{jt} and R_{jt} were perfectly correlated, changes in preferences due to changes in the peer quality of two schools with the same change in X_{jt} would be identical, but changes in beliefs of their quality would be different if the two schools receive different letter grades.

[50] Asymptotic standard errors of the minimum distance second-step estimates take into account the sampling error of the first stage estimates. They apply the delta-method to the first-step estimates of the variance-covariance matrix of δ_{cjt} and rely on numerical approximations when necessary.

and higher-achieving students are more strongly correlated with mean achievement levels. Since value-added is positively correlated with a school's average test scores in this context, white and high-achieving students hold beliefs that are somewhat less biased. These findings support the idea that white and higher-achieving students receive more precise signals about school quality from their social networks, which are based on student achievement levels at the school. Overall, however, all students tend to be quite uncertain and misinformed about school quality. The prior means across schools and students are closer to the average quality in the city than they would be if students were perfectly informed. These estimates are consistent with the survey evidence presented in Section 2.3, which showed that all respondents were significantly misinformed and reluctant to categorize schools as extremely good or extremely bad.

The second explanation, which is quantitatively more important, is differences across students in their preferences for school quality relative to other school characteristics. The willingness to travel for an additional cross-school standard deviation in school quality is similar across student demographics and ranges between 3 and 7 minutes. It is somewhat higher among students with higher baseline achievement, particularly among Black and Hispanic applicants. Different racial groups have also similar preferences for changes in peer quality and school demographics over time. White and Asian students, however, hold much stronger preferences for the school-specific attributes of a small selected sample of schools. Their preferences for the school-specific attributes that are fixed over time, $\tilde{\xi}_{cj}$, are in fact right-skewed and concentrated on few schools with high peer quality and enrolling many white and Asian students. This means that, relative to other school-specific attributes, school quality is less important for white and Asian students. As a consequence, changes in their expectations of school quality affect their choices primarily within the few schools that are majority white and Asian and have higher peer quality.

Table A12 assesses the model fit out of sample using the choices and offers of the 2016 cohort. Model-based changes in the probability of ranking schools receiving an A or a low grade after the removal of grades are benchmarked against choices made by the 2014 cohorts. Overall, the model predicts well heterogeneity in choices, choice changes, and offers across applicants' demographics.[51]

Robustness Checks I consider robustness of my estimates to alternative functional forms of student priors. Appendix Table A13 presents second-step estimates from a model in which priors vary non-parametrically across four discrete school types, defined by whether a school

[51]School offers in panel C are simulated using the model-based rank ordered lists, real school capacities and the student priorities assigned based on admission rules announced for the 2016 admission cycle. The equilibrium simulations require some restrictions and assumptions that make the fit of offered school characteristics worse than that of choices. I discuss the details in appendix C.2.

has above or below median value-added and above or below median average achievement levels. Appendix Table A14 instead considers the cases in which priors are distributed as a log-normal or as the empirical distribution of quality with a location and a scale shifts that vary across school discrete types. Regardless of the specification, white and higher achieving student mean beliefs are more strongly correlated with achievement levels, and belief precision is higher on average for white students. Nevertheless, as in the benchmark model specification, prior means are close to the mean quality in the city and differences in beliefs across race are not very large. The other preference parameters are virtually unchanged and always show a small degree of heterogeneity in willingness to travel for quality across race on average, and a somewhat larger degree of heterogeneity across baseline achievement.

Finally, I relax the assumption that students report preferences truthfully in table A15. To do that, I assume that students do not consider schools where they have a zero probability of admission but report preferences truthfully among the remaining schools.[52] While Black and Hispanic applicants' preferences for peer quality and for the share of white and Asian students in a school are slightly larger in the strategic reporting scenario than under truthful reporting, estimates of preferences for quality are unchanged. The main heterogeneity patterns in preferences and beliefs across students and schools are also largely unchanged.[53]

6 Counterfactuals

In this section, I use my model to evaluate the effects of counterfactual information disclosure policies on student welfare. My definition of welfare is the average of students' test scores Y_i, which may be weighted by welfare weights ω_i.[54] According to the model of student achievement in equation (1), welfare depends on the allocation of students to schools, μ,

[52] To compute probabilities of admission I bootstrap each school match 100 times redrawing each time a sample of applicants and a sequence of tie-breakers. Applicants are sampled with replacement independently. For each assignment and school, I obtain admission cutoffs from the priority and tiebreaker of the marginal student admitted to each school. The relevant tiebreaker is 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. 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.

[53] More sophisticated approaches to school demand estimation under strategic reporting rely on stability as in Fack et al. (2019) or view applicants' rank ordered lists as the outcome of an optimal portfolio problem as in Agarwal and Somaini (2018); Larroucau and Rios (2020); Idoux (2021); Calsamiglia et al. (2020). These methods could also be applied to estimate my model in the future.

[54] In the education literature, evaluating interventions and changes in market designs using a notion of welfare that depends directly on student outcomes is often the standard (Kapor, 2020; Barahona et al., 2023a), although it is also possible to prefer revealed-preference measures of student utility (Abdulkadiroğlu et al., 2017; Kapor et al., 2020).

through school value added:

$$W(\mu) = \sum_i \omega_i Y_i(\mu) = \sum_i \omega_i(\alpha_{\mu(i)} + X_i'\Gamma + \epsilon_i)$$

where $\alpha_{\mu(i)}$ denotes the value added of the school that student i gets under allocation μ.

I assess welfare gains or losses relative to a simulated status-quo scenario in which the policymaker provides no information about school effectiveness. I have shown in the previous section that this scenario accurately replicates realized choices and offers from 2016. Denoting the status-quo allocation with μ^0, the total change in welfare associated with an information policy that induces the allocation μ is given by the average change in the value-added of the schools to which students are allocated:

$$\Delta W(\mu, \mu^0) = \sum_i \omega_i(\alpha_{\mu(i)} - \alpha_{\mu^0(i)})$$

.

A slack in the capacity constraint of high-quality schools is necessary to obtain improvements in average student welfare, because gains are realized only when students reallocate to vacant school seats that are of higher quality than their current allocation.

As a benchmark of what are feasible welfare gains given school capacities in 2016, I quantify the maximum possible gains as the difference between the average student achievement under the allocation that matches students to the best available school and average achievement in the status-quo allocation. I call this difference "first-best" achievement gains. They would be realized in the student-proposing DA allocation if students only valued school quality and ranked schools in order of value-added. Reallocating students to vacant high quality school seats can increase average test scores at most by 0.039σ. In what follows, I often express welfare gains under different allocations as a percentage of this number.

6.1 Effects of Providing Perfect Information

Full Information Benchmark This section studies the effects of providing perfect information about the value added of each school on choices and offers. This serves as a natural benchmark to quantify the impact of misinformation on missed achievement gains. Panel A of figure 8 compares the average quality of top three choices under full information and in the simulated status quo. On average, chosen quality increases from the 68th to the 74th value-added school percentile ranking. These simulated changes are larger than those observed in the Bloomberg era, when letter grades were only an imperfect proxy for value-added. In the absence of capacity constraints, these changes in choices would result in average achievement gains of 0.07σ. Thanks to the larger response of Black and Hispanic student choices, information entirely closes the racial choice gap conditional on test scores and reduces 62% of the racial choice gap unconditional on test scores. Students, however, do not "max out" on value-added even under perfect information because of preferences for other school attributes.

Binding capacity constraints further reduce the extent to which information provision can improve allocative efficiency and boost test scores.[55] Panel B of figure 8 compares changes in chosen quality to changes in offered quality for different groups of students, defined on the basis of the student race (minority or non-minority) and baseline achievement (above or below median), showing that the average offered VA improves only by 0.01σ. Nevertheless, perfect information yields 23% of the first best average achievement gains, which are presented in the last group of bars on the right. Information also disproportionately improves Black and Hispanic students offers, not just their choices.

Why Do Black and Hispanic Students Benefit More From Information? Model estimates indicate that information can disproportionately affect Black and Hispanic students both because this group is relatively less well informed than non-minorities and because their preferences for school quality are stronger relative to those for other school traits. To quantify the importance of each channel, I compare the simulated effects of providing perfect information about school VA on school choices and offers in three hypotetical scenarios. In the first, called "Uninformed Priors" (UP), I remove differences in prior information about school quality. I assume everyone holds the same uninformed prior for all schools, equal to the empirical distribution of VA in NYC. In the second, "No preferences for Peers" (NP), I assume applicants have no preferences over the composition of students enrolled in a school, and other school-specific characteristics different from quality. The third (UP+NP) combines the first two.

Panel (a) of figure 9 shows the effect of providing information on choices in these scenarios. The first bar in each subgroup reports the information effects estimated under the actual preferences and beliefs as a benchmark. Information effects in the uninformed-prior simulation are similar to the actual full-information benchmark across student demographics. In contrast, white and Asian students would respond to information about school quality much more in the no-peer-preferences scenario than in the benchmark. These results indicate that student priors are overall quite misinformed across demographics and that differences in preferences are more important than differences in beliefs to explain the larger response to information of minority students.

Panel (b) reveals that the larger response of white and Asian students in the simulations that remove preferences for school traits different from quality and distance changes

[55] 72% of the best 20% of schools are oversubscribed in the status quo, while only 21% of the worst 20% are.

who benefits from information. In these hypothetical scenarios, their stronger reaction to information would displace low achieving Black and Hispanic students out of high-quality seats, even if their choices remain the same. This exercise reveals that Black and Hispanic students disproportionately benefit from information thanks to white and Asian students' strong preferences for school traits different from quality.

Finally, I show that differences in distance to high quality schools across race play no role in explaining the choice gap and the effects of information on choices and offers. Appendix figure A9 compares the benchmark effects of providing information (FI) on choices and offers to the effects of removing distaste for commuting (ND) and the effects of providing information in this hypothetical scenario (FI if ND). Removing any role for differences in distance to schools has essentially no effect on student choices and does not change the effects of providing information on chosen and offered school quality. This also shows that the discrepancy between the effects on Black and Hispanic students' choices and their offers is not explained by a lack of good schools nearby where they live that creates congestion only in a small number of high quality schools.

Information About Value-Added vs. Achievement Levels Information interventions in education often inform families of the average achievement of students enrolled at a school rather than of causal school value-added (Hastings and Weinstein, 2008; Cohodes et al., 2022; Corcoran et al., 2022; Allende et al., 2019; Andrabi et al., 2017). On the one hand, if achievement levels and value-added are positively correlated, naive policies providing information about the former can still induce students to reallocate to better schools. On the other, this information might be mostly redundant and fail to shift households' choices towards higher quality options because families' perceptions of quality are already based on school achievement levels.

Figure A10 compares the welfare gains obtained by information about school value-added relative to information about school achievement levels. In these simulations, information interventions either perfectly disclose differences across schools (denoted by "FI" or Full-Information) or take the form of coarse ratings corresponding to quintiles of value-added or achievement levels (denoted by "5L" or 5 Letters).[56] Information interventions based on achievement levels obtain half of the gains of those based on causal value-added for Black and Hispanic students' test scores, and produce no gains for white and Asian students. White and Asian quality beliefs are more strongly correlated with achievement levels and their

[56]The first bar for each student group shows gains under the full-information benchmark. The second bar corresponds to gains realized under the disclosure intervention that assigns schools letter grades based on VA quintiles. The third and four counterfactuals simulated the effects of providing information about achievement levels in these two forms while presenting it as if it were about school VA. That is, if the difference in achievement levels of two schools is one school-level standard deviation, households are told that the difference in quality between the two schools is one school-level standard deviation.

preferences concentrated on the subset of schools with high performing students, therefore information about school performance levels does not change their choices much. By virtue of the positive correlation of VA and achievement levels, however, information about the latter can still partly re-direct Black and Hispanic students to choose better schools.

Comparing Information Provision to Changes in Admission Rules The previous results suggest that information disproportionately improves Black and Hispanic student achievement. An alternative set of policies currently considered to help Black and Hispanic students match to higher quality schools are changes in school admission rules ranking students based on their residential address or their middle school test scores.[57] These admission rules may also substantially limit the beneficial effects of information interventions, which motivates not only comparing the effects of information provision to those of admission reforms but also studying their combined effects.

Figure 10 compares the welfare effects of providing full-information to those of removing all geographic priorities and academic screens in admissions, denoted by "NS" in the figure.[58] The effects of combining changes in admission rules with perfect information about value added are denoted by "FI+NS". This exercise offers three main insights. First, changing admission rules redistributes school quality from high to low achieving students within student race, while providing information primarily redistributes across race and benefits Black and Hispanic students across all achievement levels. Low achieving white and Asian students are instead hurt by information on average under the current admission rules, while high achieving white and Asian students are hurt by information if schools remove screens. Second, providing information and leveling the playing field in admission rules are not substitute policies. Their effects are cumulative and, if anything, they seem to act as complements in raising Black and Hispanic student test scores. Intuitively, in that scenario not only minority students would know where to find higher quality schools, but they would also have fair chances to get in. Combining the two, instead, does not help white students reallocate to better schools when they lose their priority advantage, but it further displaces them out of high-value added schools. Third, providing information yields achievement gains among Black and Hispanic students comparable to those obtained by reforming admission rules, which is a more politically controversial policy.[59] In the full-information benchmark, the

[57] These rules are often thought to disproportionately favor white and Asian students and are often at the center of debates about the equity of the school match (Cohen, 2021; Idoux, 2021; Park and Hahm, 2023).

[58] Removing geographic priorites and academic screens ensures that any two students with the same rank ordered list have the same admission chances to any school before uncertainty in their lottery number is resolved.

[59] The assumption I maintain here is that application behavior does not change in response to changes in admission rules. Past research finds that this feedback effect might amplify the effects of these policies, suggesting that the changes I simulate in the no-screen counterfactual provide a lower bound for the

test scores of Black and Hispanic students improve on average by 80% of what they would increase if school admission rules treated all students equally. When focusing on minority students with achievement levels below the median, providing information enables them to achieve one-third of the potential de-screening gains. Additionally, average achievement is higher under full information, as white students also benefit from the information on average.

Targeted Outreach Redistribution of high-quality school seats in favor of lower achieving students might be in line with the policy maker's objective, for instance if there are critical levels of achievement that need to be reached (e.g. failing to graduate is more costly than failing to graduate with the highest honors). Due to capacity constraints, one strategy to achieve this objective could involve selectively providing information to a targeted group of students to mitigate congestion effects that can arise if information is disseminated to everyone. To mimic the effects of a realistic policy, I simulate the effects of an outreach intervention that provides information to all students at middle schools with average test score levels below the median, rather than differentiating students within middle schools. This targets 30% of students, coming from a disadvantaged population: targeted students' 7th grade Math test scores are lower by 0.9σ on average, 84% are eligible for free and reduced price lunch (compared to 64% among the non-targeted) and 90% are Black or Hispanic (compared to 51% among the non-targeted). Table A17 shows that targeted students choose schools with 0.074σ higher value added and their offers would improve by 0.033σ, 85% of the average first-best gains, substantially more than if everyone were informed. The cost of preventing targeted outreach on the outcomes of the most disadvantaged students can be quantified as the welfare change with respect to the full-information benchmark, corresponding to achievement losses for targeted students of 0.02σ on average.

Panel B of table A17, instead, presents the effects of supplying quality information exclusively to Black and Hispanic students whose test scores fall within the top tercile of the city's distribution. Such outreach initiatives are currently under consideration as policy alternatives to affirmative action, with the aim of enhancing the representation of racial minorities in high-quality schools. Due to the large preferences for quality of this subgroup of students, and because few students receive information, achievement gains for targeted students are three times as large as when everyone is fully informed. However, the share of Black and Hispanic students in the best 20% of schools increases only by 2.3 p.p. Providing full information increases this share by 2.6 p.p. and removing geographic priorities and screening by 9 p.p., suggesting that targeted outreach is less effective than other policies to increase representation of non-white students in top quality schools.

achievement gains of low achieving students (Idoux, 2021).

6.2 Optimal Information Design

Providing information solely to a selected group of students is one obvious way to redistribute high quality school seats when capacity constraints are binding, but it may be politically unfeasible. In this subsection, I explore how policymakers could instead design school ratings to improve achievement among minority or low-achieving students while keeping the information publicly accessible. The effects of coarse quality ratings may also offer a more realistic reference point for the effects of information disclosure, as this is often the prevailing format in real-world settings, such as healthcare, education, nutrition, and finance.

Coarsening information may be preferable to full disclosure for maximizing student achievement for two main reasons. First, when the social planner objectives are different from the maximization of participants' utility, as is the case if it only cared about education achievement, coarsening information may persuade agents to take actions that would not be optimal from the agent's point of view but may be from the point of view of the sender of the signal (Kamenica and Gentzkow, 2011). In this setting, coarsening information could induce students to choose schools of higher quality compared to the full-information benchmark, and translate into higher average achievement. Second, coarser information may allow to redistribute quality to less advantaged students more than precise information in equilibrium.

To build some intuition, consider the following two stylized examples.

Example 1 - Coarser information is Pareto improving from the planner's perspective. There are 4 schools a, b, c, d, each with 1 seat, and 2 students i_1, i_2. Student i_1 always has higher priority than student i_2 at all schools, for instance because she has higher test scores. Students care about school quality q and other school attributes p. School characteristics and student utilities are as follows:

	a	b	c	d
q	3.5	1.5	1	0.5
p	3	2	4	3.5

$$u_{1j} = E[q_j] + p_j$$
$$u_{2j} = E[q_j] + \frac{1}{2}p_j$$

Students hold uninformed priors that correspond to the distribution of quality in the city. Denote with F the state in which students know the quality of each school, with N the state in which they receive no information, and with C the state in which they receive information about which of the four schools are the best two and which are the worst two. Student preferences in these three states are:

$$c \succ_1^N d \succ_1^N a \succ_1^N b \qquad c \succ_2^N d \succ_2^N a \succ_2^N b$$
$$a \succ_1^F c \succ_1^F d \succ_1^F b \qquad a \succ_2^F c \succ_2^F b \succ_2^F d$$
$$a \succ_1^C b \succ_1^C c \succ_1^C d \qquad a \succ_2^C b \succ_2^C c \succ_2^C d$$

And the resulting allocations from student-proposing DA are:

$$\mu^N = \begin{pmatrix} i_1 & i_2 \\ c & d \end{pmatrix} \qquad \mu^F = \begin{pmatrix} i_1 & i_2 \\ a & c \end{pmatrix} \qquad \mu^C = \begin{pmatrix} i_1 & i_2 \\ a & b \end{pmatrix}$$

Under coarse information, i_2 receives strictly higher value-added than under full information, and i_1 is no worse-off. Intuitively, because students care about quality, pooling the best two schools convinces them to rank b higher, which has higher quality but not high enough to be preferred to c under full information.

Example 2 - Coarser information is redistributive. Now instead, let student i_1 have higher preferences for other school attributes p, so that her utility becomes $u_1 = E[q] + 2p$. Preferences and allocations now are:

$$\begin{aligned} &c \succ_1^N d \succ_1^N a \succ_1^N b, \qquad && c \succ_2^N d \succ_2^N a \succ_2^N b \\ &a \succ_1^F c \succ_1^F d \succ_1^F b, && a \succ_2^F c \succ_2^F d \succ_2^F b \\ &c \succ_1^C a \succ_1^C d \succ_1^C b, && a \succ_2^C c \succ_2^C b \succ_2^C d \end{aligned}$$

$$\mu^N = \begin{pmatrix} i_1 & i_2 \\ c & d \end{pmatrix} \qquad \mu^F = \begin{pmatrix} i_1 & i_2 \\ a & c \end{pmatrix} \qquad \mu^C = \begin{pmatrix} i_1 & i_2 \\ c & a \end{pmatrix}$$

The average test scores under both the full and the coarse information scenario are the same, and are higher than under no information. However, coarsening information affects who is matched to the highest value-added school, a. Because student i_1 cares a lot about other school attributes, coarse information does not provide a strong enough signal to induce her to choose the highest quality school. Coarsening information therefore removes the competition for school a, which student i_2 prefers, but that gives priority to i_1.

Similar patterns could be observed in reality due to heterogeneity in preferences for school attributes across students. Intuitively, information about the quality of schools that are considered non-desirable for other reasons is not much valuable. Therefore, precise information about schools with lower peer quality or enrolling higher shares of minority students will not induce large responses among white and higher achieving students. Conversely, precise information about schools that white and higher achieving students like increases their sorting to high-quality seats, displacing disadvantaged students. Moreover, the precision of priors of white and high-achieving students is higher, which may induce them to respond less than other students when quality signals are coarse but not when they are precise. The policy maker, therefore, may face a trade-off between providing information and convincing students to rank less preferred but higher quality schools (as in example 1) or discouraging some particular students from applying to higher quality schools (as in example 2).

To assess the potential benefits of coarsening information, I simulate the effects of two types of realistic school rating policies. I first consider where to place the quality cutoffs of 5 letters grades, and then let information precision vary depending on the slackness of the capacity constraint. Table A18 reports the best and worst 5 letter policy for improving welfare among different groups of students and the associated average welfare change, expressed as a share of first best mean achievement gains. The best 5 letter policy in terms of average welfare can achieve 20% of the first-best gains, and a naive policy that places cutoffs evenly distributed along the value-added distribution around 17%. The different spacing of letter cutoffs, however, determines who gains the most from information. The policy that helps lower achieving students the most, even more than full-information, is one that provides very precise signals at the bottom of the quality distribution. On the contrary, the worst policy for this subgroup is one providing precise signals at the top. The opposite is true for high-achieving students, who would benefit from more precise signals at the top than at the bottom.

Examining the simple case of designing two quality ratings - high and low - helps clarify the mechanism at play. Figure A11 plots quality of school choices and offers as a function of the high rating cutoff. As the cutoff increases, all students choose higher quality schools. Chosen quality, however, increases faster among higher achieving students as the signal at the top gets more precise, resulting in tougher competition for high quality seats. Intuitively, because high-achieving students hold strong preferences for the attributes of schools that on average have higher quality, increasing the cutoff provides them more precise information about the set of schools they like. They find this information more valuable than precise information about low quality schools and react to it more strongly. As a result, low-achieving students receive higher quality offers in equilibrium when the cutoff is placed at the 30th percentile. As the cutoff increases they are displaced out of higher quality schools at higher rates.

These results indicate that policies providing students with a list of top performing schools need not help equally everyone (Cohodes et al., 2022). As in example 2, such policies might increase competition for high-quality schools as high-achieving students react more when seeing precise signals of quality at the top. Disadvantaged students might be screened out of top performing schools based on their achievement, and might be at risk of ending up in worst schools if they cannot distinguish the bad schools from the average ones.

Next, I simulate the effects of providing students with a less precise quality signal (above or below median) about schools oversubscribed in the status-quo and an infinitely precise signal for schools that are undersubscribed. This policy mimicks an advertising campaign that provides exact information only about undersubscribed schools. The ultimate goal for the policy-maker is in fact to convince students to rank higher in their list good schools that are, however, not yet at capacity. I call this counterfactual "pooling", because the quality of oversubscribed schools is pooled, while that of undersubscribed schools is not.

Figure 11 compares welfare gains from the pooled policy with those of the full-information benchmark and of the best and worst five letter grades policies for low achieving students. Pooling yields the same average gains as full information but allows for better redistribution of quality to low-achieving students. As in the previous counterfactual, low-achieving students strictly benefit in equilibrium from the policy with less information compared to the full information scenario. As before, high-achieving students do not respond much to precise information about under-subscribed schools because they like them relatively less than low-achieving students, while coarse quality signals for oversubscribed schools are not informative enough to induce them to strongly sort to the best oversubscribed schools. This limits the displacement of low-achieving students out of high quality oversubscribed schools. The two counterfactual exercises presented in this sub-section thus suggest that the policy maker can partly leverage information design to redistribute value added across student groups, less so for pushing the Pareto frontier of test score achievement.

7 Conclusions

School choice can achieve equity, allocative, and efficiency gains provided that families reward school effectiveness. This assumption, however, has proven to fail in many settings. This paper shows that a lack of accessible information about school quality is partly to blame and explains a portion of the disparities in access to high-quality education across races. To do this, the paper leverages a natural policy experiment that varied the information about school quality available to students in NYC. It finds that Black and Hispanic students are more responsive to school ratings, allowing information to reduce achievement inequality.

Based on a structural model of demand for schools, choice responses to information reveal differences in both beliefs and preferences for quality across students of different races and with different baseline achievement. Everyone is misinformed about which schools are of higher quality, and minority students more than non-minority students. Even if misinformed, all applicants care about value-added, separately from the composition of the students at a school. White and Asian families, however, value other school attributes relatively more than minority families. Their strong preferences for schools enrolling white and high-achieving students limit their responsiveness to quality information. As a consequence, information interventions and their design can partly redistribute school quality even under a fixed supply of school seats.

The findings of this paper become even more relevant in light of recent developments in school accountability policy. Following the passage of the No-Child Left Behind Act, school accountability received significant attention. Accountability reforms aimed not only to incentivize schools to improve student achievement but also to empower families with informed decision-making tools. In many school districts, these objectives were achieved through the provision of easily understandable school performance ratings, such as letter grades. In recent years, however, school performance measurement and accountability appear less prominent on the education policy agenda. Many school districts have shifted away from summative assessments based on student achievement. In the specific context of my study, letter grades have been replaced with multi-dimensional school quality measures that may be less visible and harder to parse.

While school letter ratings in NYC were far from perfect and, to some extent, reflected student selection rather than true causal value-added, they had a significant impact on the choices of less advantaged families. This underscores the importance of providing accessible information about school quality to all. Designing ratings that more accurately represent causal estimates of school effectiveness and can be tweaked to help the most disadvantaged could be a more effective policy approach to reducing achievement inequalities. This approach may prove superior to both the earlier simplistic ratings and current policies that are veering away from ratings altogether.

References

Abdulkadiroğlu, A., Agarwal, N., and Pathak, P. A. (2017). The welfare effects of coordinated assignment: Evidence from the new york city high school match. *American Economic Review*, 107(12):3635–89.

Abdulkadiroğlu, A., Angrist, J., Narita, Y., and Pathak, P. A. (2022). Breaking Ties: Regression Discontinuity Design Meets Market Design. *Econometrica*, 90(1):117–151.

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

Abdulkadiroğlu, A., Pathak, P. A., and Roth, A. E. (2009). Strategy-proofness versus efficiency in matching with indifferences: Redesigning the nyc high school match. *The American Economic Review*, 99(5):1954–1978.

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., Pathak, P. A., and Walters, C. R. (2018). Free to choose: Can school choice reduce student achievement? *American Economic Journal: Applied Economics*, 10(1):175–206.

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(1):471–501.

Ainsworth, R., Dehejia, R., Pop-Eleches, C., and Urquiola, M. (2023). Why do households leave school value added on the table? the roles of information and preferences. *American Economic Review*, 113(4):1049–82.

Allende, C. (2020). Competition under social interactions and the design of information policies. Technical report.

Allende, C., Gallego, F., and Neilson, C. (2019). Approximating the equilibrium effects of information interventions. Technical report.

Amin, R. (2019). The iconic nyc high school directory is going largely virtual. will that improve the daunting process or disadvantage some families?

Andrabi, T., Das, J., and Khwaja, A. I. (2017). Report cards: The impact of providing school and child test scores on educational markets. *American Economic Review*, 107(6):1535–63.

Angrist, J., Hull, P., Pathak, P., and Walters, C. (2016). Interpreting tests of school vam validity. *American Economic Review*, 106(5):388–92.

Angrist, J., Hull, P., Pathak, P. A., and Walters, C. (2021). Credible school value-added with undersubscribed school lotteries. *The Review of Economics and Statistics*, pages 1–46.

Angrist, J., Hull, P., and Walters, C. R. (2022). Methods for measuring school effectiveness. Working Paper 30803.

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

Barahona, N., Dobbin, C., and Otero, S. (2023a). Affirmative action in centralized college admission systems. Technical report.

Barahona, N., Otero, C., and Otero, S. (2023b). Equilibrium effects of food labeling policies. *Econometrica*, 91(3):839–868.

Barseghyan, L., Clark, D., and Coate, S. (2019). Peer preferences, school competition, and the effects of public school choice. *American Economic Journal: Economic Policy*, 11(4):124–58.

Bergman, P. (2020). Nudging Technology Use: Descriptive and Experimental Evidence from School Information Systems. *Education Finance and Policy*, 15(4):623–647.

Bergman, P., Chan, E. W., and Kapor, A. (2020). Housing search frictions: Evidence from detailed search data and a field experiment. Working Paper 27209.

Beuermann, D. W. and Jackson, C. (2018). Do parents know best? the short and long-run effects of attending the schools that parents prefer. IDB Working Paper Series IDB-WP-914, Washington, DC.

Bhargava, S. and Manoli, D. (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an irs field experiment. *American Economic Review*, 105(11):3489–3529.

Black, S. E., Denning, J. T., and Rothstein, J. (2023). Winners and losers? the effect of gaining and losing access to selective colleges on education and labor market outcomes. *American Economic Journal: Applied Economics*, 15(1):26–67.

Bleemer, Z. (2021). Affirmative Action, Mismatch, and Economic Mobility after California's Proposition 209*. *The Quarterly Journal of Economics*, 137(1):115–160.

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

Campos, C. (2023a). The impact of public school choice: Evidence from los angeles' zones of choice. Technical report.

Campos, C. (2023b). Social interactions and preferences for schools:experimental evidence from los angeles. Technical report.

Carlana, M., La Ferrara, E., and Pinotti, P. (2022). Goals and gaps: Educational careers of immigrant children. *Econometrica*, 90(1):1–29.

Chernew, M., Gowrisankaran, G., and Scanlon, D. P. (2008). Learning and the value of information: Evidence from health plan report cards. *Journal of Econometrics*, 144(1):156–174.

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

Cohodes, S., Corcoran, S., Jennings, J., and Sattin-Bajaj, C. (2022). When do informational interventions work? experimental evidence from new york city high school choice. Working Paper 29690.

Corcoran, S. P., Jennings, J. L., Cohodes, S. R., and Sattin-Bajaj, C. (2022). Leveling the playing field for high school choice: results from a field experiment of informational interventions. Working paper.

Corradini, V. and Idoux, C. (2023). Untangling the paradox of choice: Why minority families choose lower-quality schools (and what can be done about it). Technical report.

Cullen, J. B., Jacob, B. A., and Levitt, S. (2006). The effect of school choice on participants: Evidence from randomized lotteries. *Econometrica*, 74(5):1191–1230.

Dranove, D. and Sfekas, A. (2008). Start spreading the news: A structural estimate of the effects of new york hospital report cards. *Journal of Health Economics*, 27(5):1201–1207.

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

Ellison, G. and Pathak, P. A. (2021). The efficiency of race-neutral alternatives to race-based affirmative action: Evidence from chicago's exam schools. *American Economic Review*, 111(3):943–75.

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

Friedman, M. (1955). The role of government in education. In *Economics and the Public Interest.*

Goolsbee, A. and Petrin, A. (2004). The consumer gains from direct broadcast satellites and the competition with cable tv. *Econometrica*, 72(2):351–381.

Hailey, C. A. (2022). Racial Preferences for Schools: Evidence from an Experiment with White, Black, Latinx, and Asian Parents and Students. *Sociology of Education*, 95(2):110–132. _eprint: https://doi.org/10.1177/00380407211065179.

Hastings, J., Neilson, C. A., and Zimmerman, S. D. (2015). The effects of earnings disclosure on college enrollment decisions. Working Paper 21300.

Hastings, J. S., Kane, T. J., and Staiger, D. O. (2009). Heterogeneous preferences and the efficacy of public school choice. Working paper.

Hastings, J. S. and Weinstein, J. M. (2008). Information, School Choice, and Academic Achievement: Evidence from Two Experiments*. *The Quarterly Journal of Economics*, 123(4):1373–1414.

Hoxby, C. M. (2000). Does competition among public schools benefit students and taxpayers? *American Economic Review*, 90(5):1209–1238.

Hoxby, C. M. (2003). School choice and school productivity (or could school choice be a tide that lifts all boats?). In *The Economics of School Choice*, number 8873 in Working Paper Series, pages 287–342. University of Chicago Press.

Hoxby, C. M. and Avery, C. (2012). The missing : The hidden supply of high-achieving, low income students. Working Paper 18586.

Idoux, C. (2021). Integrating new york city schools: The role of admission criteria and family preferences.

Kamenica, E. (2019). Bayesian persuasion and information design. *Annual Review of Economics*, 11(1):249–272.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6):2590–2615.

Kapor, A. (2020). Distributional effects of race-blind affirmative action. Technical report.

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

Ladd, H. (2002). Market-based reforms in urban education. Technical report, Economic Policy Institute.

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.

Mizala, A. and Urquiola, M. (2013). School markets: The impact of information approximating schools' effectiveness. *Journal of Development Economics*, 103:313–335.

Neilson, C. (2019). The rise of centralized choice and assignment mechanisms in education markets around the world. Technical report.

Park, M. and Hahm, D. (2023). Location choice, commuting, and school choice. Technical report.

Ravitch, D. (2010). *The Death and Life of the Great American School System: How Testing and Choice are Undermining Education*. ReadHowYouWant.

Rockoff, J. and Turner, L. J. (2010). Short-run impacts of accountability on school quality. *American Economic Journal: Economic Policy*, 2(4):119–47.

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

Sattin-Bajaj, C., Jennings, J. L., Corcoran, S. P., Baker-Smith, E. C., and Hailey, C. (2018). Surviving at the street level: How counselors' implementation of school choice policy shapes students' high school destinations. *Sociology of Education*, 91(1):46–71.

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

Tincani, M., Kosse, F., and Miglino, E. (2021). Subjective beliefs and inclusion policies: Evidence from college admissions.

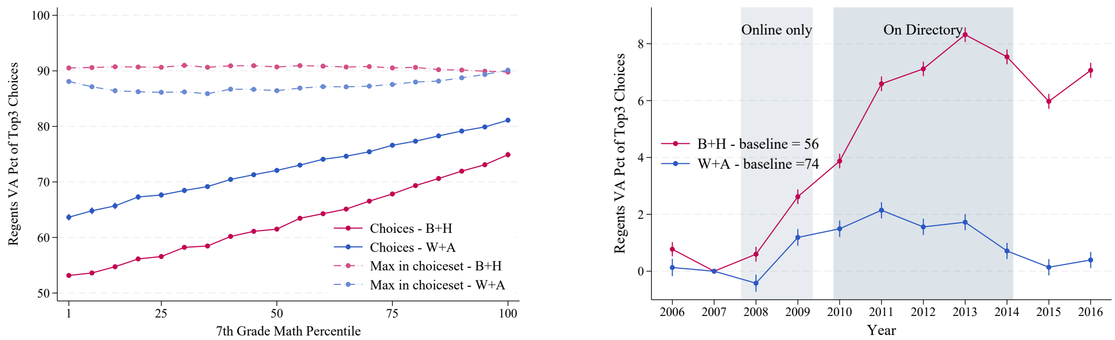
Vatter, B. (2022). Quality disclosure and regulation: Scoring design in medicare advantage. Technical report.

Walters, C. R. (2018). The demand for effective charter schools. *Journal of Political Economy*, 126(6):2179–2223.

Tables and Figures

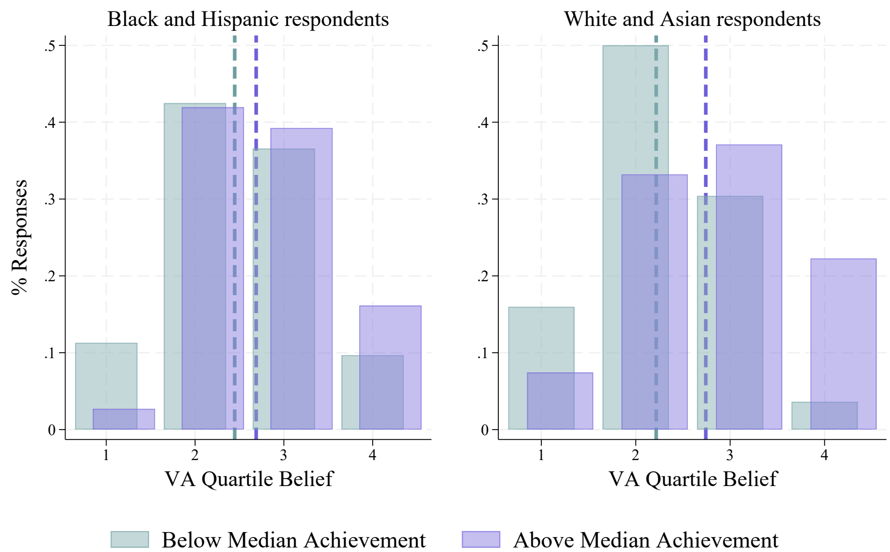
Figure 1: The Racial School Quality Choice Gap

(a) By Baseline Achievement (b) Across Cohorts



Notes: This figure describes cross-race differences in chosen school quality, as measured by Regents value-added. Panel (a) plots the relationship between the percentile rank of the student's baseline score and two variables: the average value-added of students' first three choices (solid lines) and the average value-added of the best three school options in the student's feasible set (dashed lines). Blue lines are averages for white and Asian students, pink lines are averages for Black and Hispanic students. Each line is a raw average computed within student cells defined by combinations of race and 20 baseline test score bins. Panel (b) plots the difference in average school value-added in the first three choices by race and cohort with respect to 2007. Race differences in choices are normalized to zero in 2007.

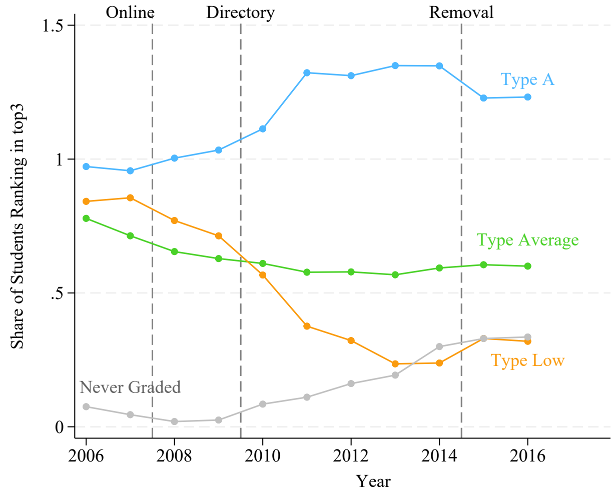
Figure 2: Distribution of Beliefs About School Quality by Race and School Achievement



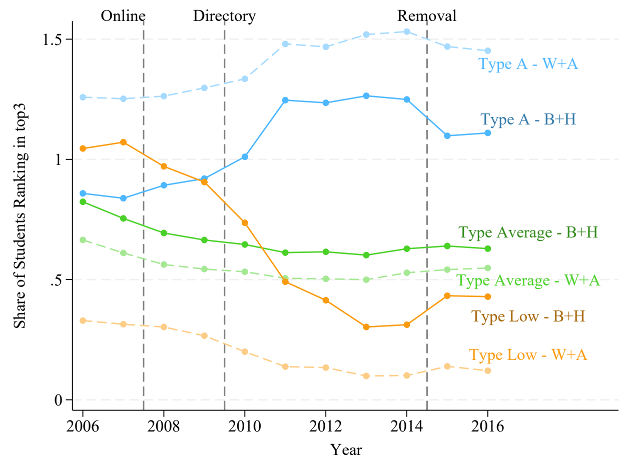
Notes: This figure shows the distribution of responses to a survey question eliciting beliefs about school quality, separately by respondents' race and the school average achievement level. The question asked: "How well does *school name - (school code)* prepare students for their Regents exams compared to other schools in your borough?". Possible responses ranged from 1 (corresponding to the bottom 25% of school quality) to 4 (best 25% of schools). Violet bars are responses to questions asking beliefs about schools with above median average Regents levels and green bars are for schools with below median Regents levels. The panel on the left shows the distribution of answers for Black and Hispanic respondents, the one on the right that of white and Asian respondents.

Figure 3: Trends in School Shares by School Letter Category

(a) All Students



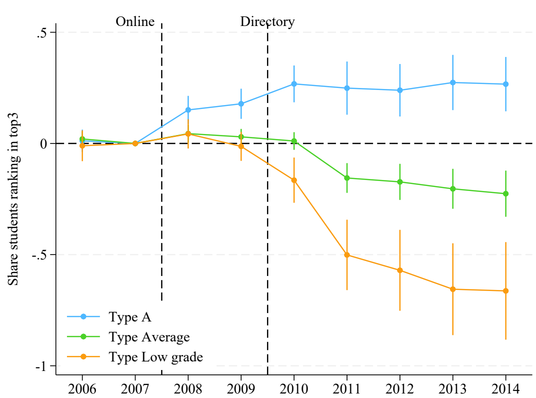
(b) Separately by Race



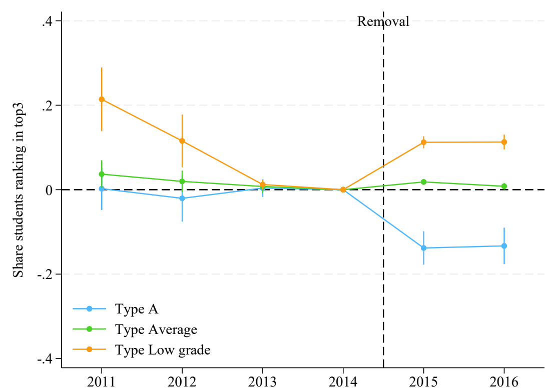
Notes: The figure shows trends in demand for schools. Demand is measured as a school share among student first three choices. The graph shows the average school share by school grade type and year. Type A schools receive an A in 5 out of the 7 years, Type Low schools receive a grade of C,D, or F in 5 out of the 7 years, Never graded schools never received a grade and Type Average schools are all remaining schools. Vertical lines indicate, in order, the introduction of letter grades online, their introduction on the directory, and their removal. Panel A pools the choices of students of all races, while panel B separately shows school shares by student race.

Figure 4: Event Study Estimates of Demand Responses to Introduction and Removal of Quality Signals

(a) Introduction of Letters



(b) Removal of Letters

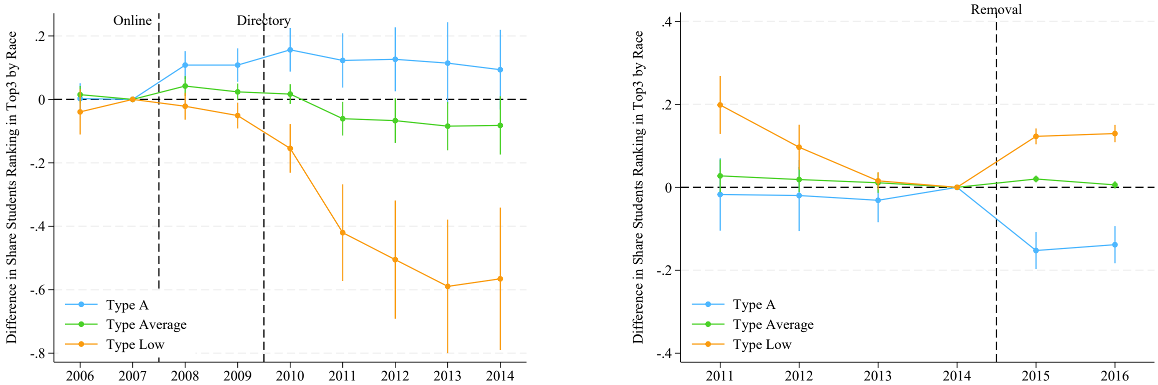


Notes: The figure plots event study estimates of the coefficient β_L^t of equation (3). Panel (a) considers changes relative to 2007, the year before the introduction of letters, using applicant cohorts of 2006-2014. Panel (b) considers changes around the removal of letters, normalizing shares differences to 0 in 2014, and using cohorts of 2011-2016. The blue line is for changes in shares of Type A schools, the orange for shares of Type Low schools and the green line is for Type Average schools.

Figure 5: Event Study Estimates - Heterogeneity by Student Race

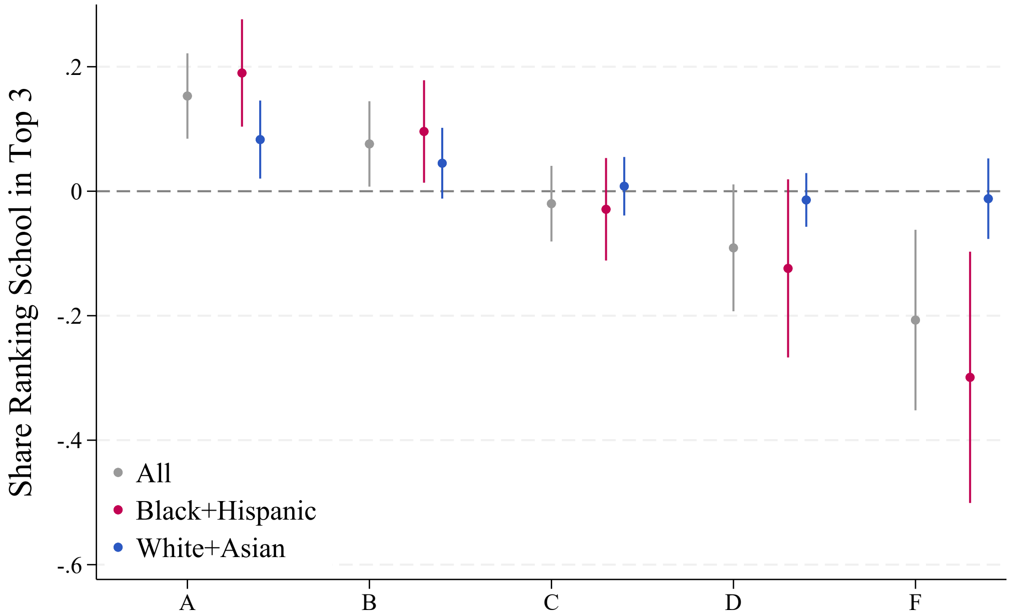
(a) Introduction of Letters

(b) Removal of Letters



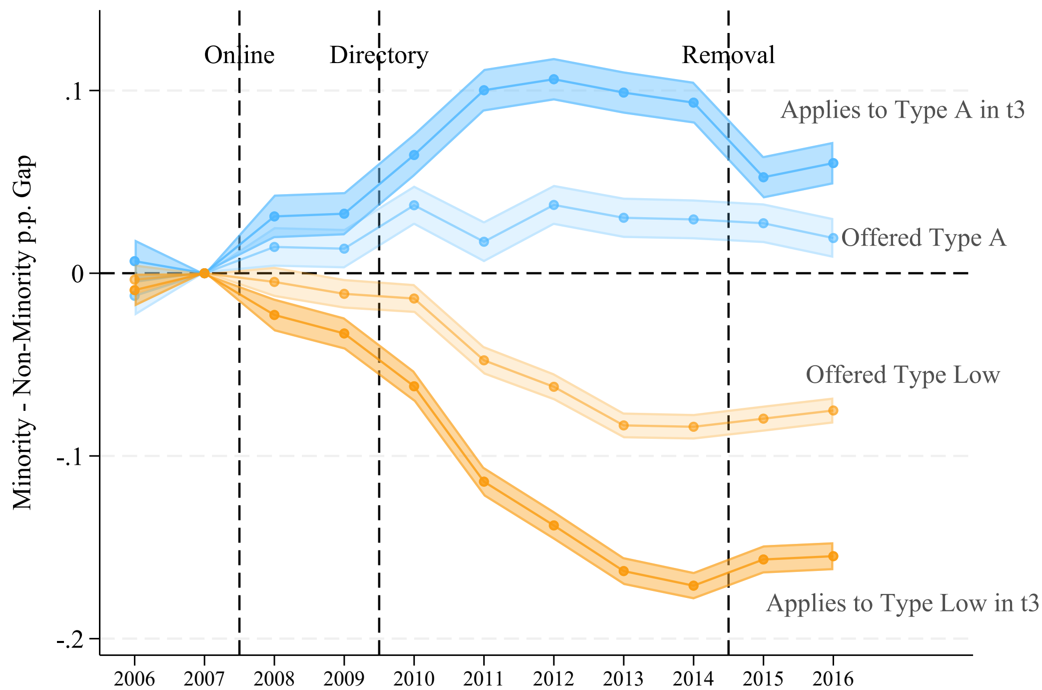
Notes: The figure plots event study estimates of the coefficient δ_L^t of equation (5), capturing cross-race differences in choice responses to the introduction and removal of letter grades. Panel (a) considers differential changes by race relative to 2007, the year before the introduction of letters, using applicant cohorts of 2006-2014. Panel (b) considers changes around the removal of letters, normalizing shares differences to 0 in 2014, and using cohorts of 2011-2016. The blue line is for changes in shares for Type A schools, the orange for shares of Type Low schools and the green line is for Type Average schools.

Figure 6: Year-to-Year Demand Responses to Letter Grades - Heterogeneity by Race



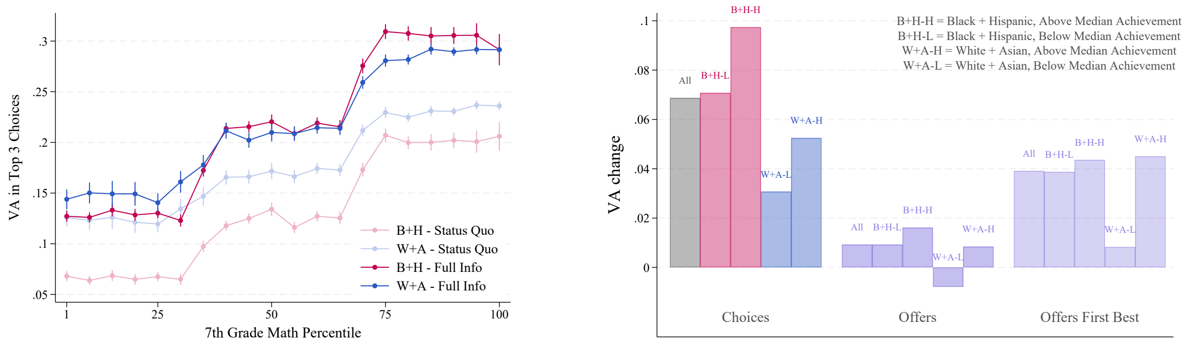
Notes: The figure plots letter grade effects on demand for schools, measured by estimates of the coefficients β_g in equation (4). The dependent variable is the share of students in demographic cell c and application cohort t ranking the school among their first three choices. Demographic cells are defined by the interaction of a student residential borough and baseline test score tercile (gray estimates) supplemented with student race (pink and blue estimates). Controls are those used in table 5 and always include school-cell fixed effects and year-cell fixed effects. The sample includes applicant cohorts from 2010 to 2014 included. Standard errors are clustered at the school-year level.

Figure 7: Event Study Estimates of Changes in Choice and Offers Probabilites by Race



Notes: The figure plots event study estimates of the coefficient δ^t of equation (8) for regressions using four different dependent variables, indicating ranking a Type A school among a student first three choices, receiving an offer to a Type A school, and similar events for Type Low schools. The coefficient δ^{2007} is normalized to zero. The sample includes students applying to enroll in 9th grade between 2006 and 2016. Controls include gender, ell status, subsidized lunch status and fixed effects for combinations of student borough and baseline test score terciles.

Figure 8: Change in VA in Top 3 Choices and Offers Under Full Information

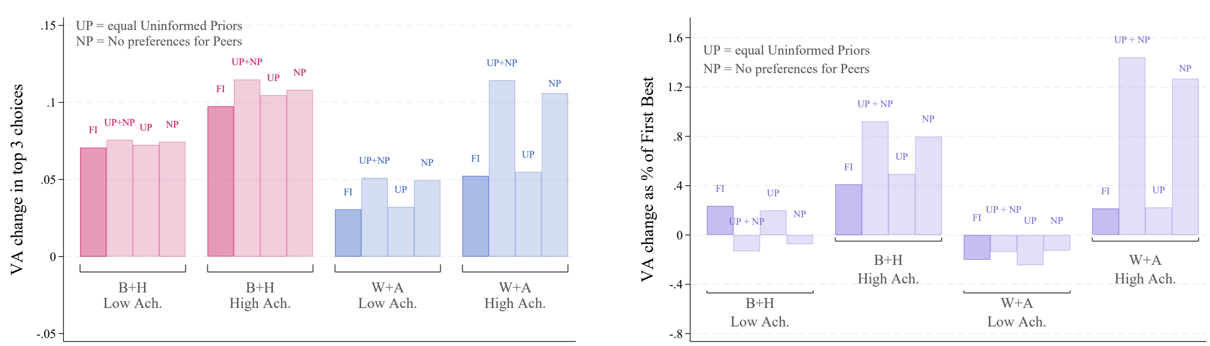


(a) ΔVA in Choices by Race and Bl. Achiev.

(b) Average ΔVA in Choices and Offers

Notes: Panel (a) plots the relationship between the percentile rank of the student's baseline score and the average value-added in the student's simulated first three school choices. The darker lines correspond to the full-information counterfactual, while the lighter ones are for choices in the (simulated) status-quo. Blue lines are averages for white and Asian students, pink lines are averages for Black and Hispanic students. Each line is made of raw averages computed within student cells defined by combinations of race and 20 baseline test scores bins. Panel (b) compares changes in VA of choices to changes in offered VA. The first group of bars plots the average change in VA of students' top 3 choices between the full-information benchmark and the status quo by student subgroups defined by combinations of race and baseline achievement (above or below median). The second group of bars plots the corresponding changes in offered VA. The last group of bars plots the average change in offered VA between the first best and the status quo.

Figure 9: Role of Beliefs and Preferences in Explaining Effects of Information

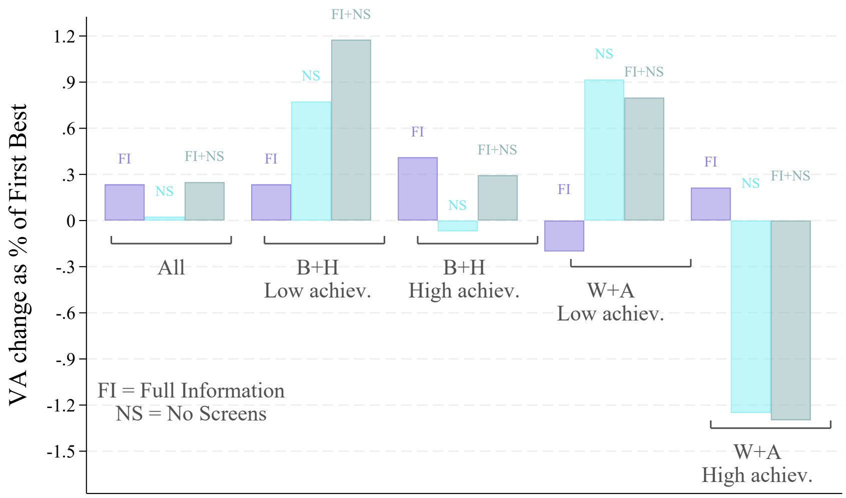


(a) Average ΔVA in Choices

(b) Welfare Gains

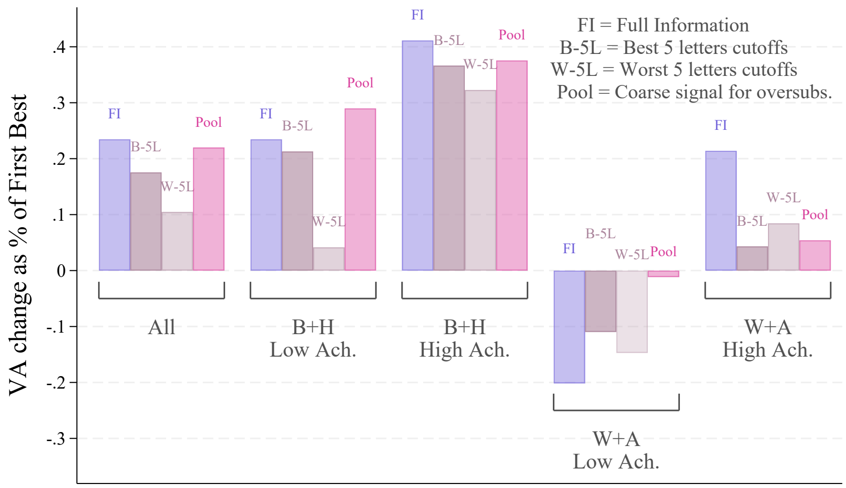
Notes: Panel (a) shows changes in VA of top 3 choices across four different simulations, taking averages within student groups defined by race and baseline achievement (above or below median). Panel (b) does the same thing for the resulting change in offered VA, expressed as a percentage of the average first-best achievement gains. Within each students subgroup, the first bar corresponds to the full-information benchmark that uses the actual model estimates. The second bar corresponds to a simulation that changes both priors and preferences: students' priors are equally uninformed and students judge schools only on the basis of their quality and their distaste for commuting. The third bar simulates choice changes only assuming students' priors are equally uninformed while the fourth bar assumes students only care about quality and distance but may hold different priors.

Figure 10: Welfare Changes Under Full Information and No Screening in Admissions



Notes: This figure plots the average welfare change from the status quo by student subgroups for three different counterfactual simulations of student assignment. Welfare gains are VA changes, expressed as a percentage of the average first-best achievement gains. Student subgroups are defined by combinations of race and baseline achievement. FI denotes student assignment under full-information; NS student assignment if admission rules had no academic screens or geographic priorities under the status-quo information environment; FI+NS combines full information with the removal of academic screens and geographic priorities.

Figure 11: Welfare Changes Under Coarser Information



Notes: This figure plots the average welfare change from the status quo by student subgroups for four different counterfactual simulations of student assignment. Welfare gains are expressed as a percentage of the average first-best achievement gains. Student subgroups are defined by combinations of baseline achievement and race. FI denotes student assignment under full-information, "B-5L" assignment under the best 5 letter grade rule for students with below median baseline achievement, "W-5L" assignment under the worst 5 letter grade rule for students with below median baseline achievement, while "pooled" denotes a counterfactual in which students receive a coarse signal about the quality of schools oversubscribed in the status quo and an infinitely precise signal about the quality of under-subscribed schools.

Table 1: Applicants Descriptive Statistics

	All (1)	Minority (2)	Non-minority (3)	Black (4)	Hispanic (5)	White (6)	Asian (7)
Panel A: student demographics							
N	625,868	425,579	200,289	185,658	239,921	91,272	104,405
Black	0.30	0.44	0.00	1.00	0.00	0.00	0.00
Hispanic	0.38	0.56	0.00	0.00	1.00	0.00	0.00
White	0.15	0.00	0.46	0.00	0.00	1.00	0.00
Asian	0.17	0.00	0.52	0.00	0.00	0.00	1.00
Subsidized lunch	0.78	0.84	0.64	0.81	0.87	0.48	0.78
Ell	0.09	0.09	0.08	0.02	0.15	0.04	0.12
7th grade Math	0.18	-0.07	0.72	-0.10	-0.05	0.64	0.82
7th grade English	0.13	-0.05	0.52	0.01	-0.10	0.64	0.43
Bronx	0.22	0.30	0.06	0.23	0.35	0.06	0.06
Brooklyn	0.32	0.32	0.30	0.46	0.22	0.32	0.29
Manhattan	0.11	0.12	0.09	0.08	0.14	0.11	0.07
Queens	0.29	0.24	0.42	0.20	0.26	0.28	0.54
Staten Island	0.06	0.03	0.13	0.03	0.03	0.24	0.04
Panel B: characteristics of top3 high school choices							
Commuting time (minutes)	40	41	39	45	38	38	39
Regents math VA (percentile)	65	61	75	60	62	73	77
SAT math VA (percentile)	70	65	83	64	65	82	85
Peer quality (percentile)	74	68	86	69	68	85	87
White+Asian %	0.35	0.25	0.56	0.24	0.26	0.58	0.54
Graduation rate	0.78	0.75	0.82	0.76	0.75	0.82	0.83
Panel C: student outcomes							
Regents Math σ	0.04	-0.13	0.54	-0.18	-0.10	0.47	0.62
Regents Ela σ	0.33	0.16	0.69	0.14	0.17	0.71	0.67
SAT Math σ	0.13	-0.25	0.72	-0.30	-0.21	0.54	0.86
SAT Ela σ	0.16	-0.14	0.63	-0.15	-0.13	0.69	0.60
Graduates in 4 years	0.78	0.73	0.90	0.73	0.73	0.90	0.91
Enrolls in college	0.67	0.58	0.82	0.58	0.59	0.79	0.85

Notes: This table provides descriptive statistics for the sample of 9th grade applicants applying to enroll in high school between 2006 and 2016. Panel A describes applicants' demographic composition, baseline test scores and residential boroughs. Panel B summarizes the characteristics of their first three high school choices. Panel C restricts the applicant samples to students who enrolled in the district and for whom I observe achievement outcomes. Column (1) reports averages across all students, while columns (2)-(7) consider student subgroups by race or ethnicity. Minority refers to Black and Hispanic students, while non-minority to white and Asian students.

Table 2: School - Year Descriptives by Letter Grade

	All schools Mean	Sd	A	B	C/D/F	N/A	A	B	C/D/F
			Mean by letter grade				Mean by "correct" grade		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Black %	0.39	0.26	0.32	0.39	0.44	0.41	0.31	0.38	0.49
Hispanic %	0.44	0.24	0.45	0.42	0.46	0.44	0.40	0.48	0.44
White %	0.08	0.13	0.12	0.09	0.04	0.07	0.13	0.07	0.03
Asian %	0.09	0.12	0.11	0.09	0.06	0.07	0.15	0.06	0.04
FRPL %	0.80	0.17	0.78	0.77	0.83	0.83	0.75	0.81	0.82
Ell %	0.12	0.19	0.14	0.10	0.11	0.14	0.13	0.12	0.09
Regents Math VA (percentile)	50	29	66	51	33	49	82	49	17
Regents Math VA σ	0.01	0.21	0.13	0.00	-0.11	0.00	0.23	-0.02	-0.20
SAT Math VA (percentile)	50	29	60	52	42	45	70	48	36
SAT Math VA σ	0.00	0.15	0.06	0.00	-0.05	-0.03	0.11	-0.02	-0.06
Peer quality (percentile)	50	29	60	54	41	43	71	50	38
Peer quality (avg. 7th grade Math σ)	-0.15	0.41	0.01	-0.11	-0.30	-0.24	0.13	-0.21	-0.34
Graduation rate	0.72	0.16	0.82	0.71	0.60	0.73	0.80	0.68	0.62
Average Regents Math σ	-0.18	0.45	0.07	-0.14	-0.39	-0.28	0.22	-0.26	-0.41
Average SAT Math σ	-0.42	0.42	-0.29	-0.37	-0.53	-0.49	-0.03	-0.47	-0.72
Screened	0.24	0.43	0.38	0.20	0.16	0.19	0.40	0.15	0.12
Size	679	858	641	980	928	248	949	885	691
N (school-year)	2,716		733	736	507	740	733	736	507

Notes: This table provides school descriptive statistics for the 2006-2007 to the 2012-2013 schools years. The progress reports were based on data covering these school years. An observation in this sample is a school-year. Column (1) and (2) report means and standard deviations across school-year observations. Columns (3)-(6) report means by letter grades and columns (7)-(9) by the letter grade schools would have received if grades had been actually based on causal estimates of Regents VA.

Table 3: Race Gap in Choice of School Quality

		Race gap							
	N	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: school value added percentile ranking									
Mean Regents VA in top 3 choices	620,975	-14.48***	-11.13***	-12.74***	-7.21***	-10.07***	-7.52***	-8.88***	-4.51***
		(0.04)	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.05)	(0.06)
Mean SAT VA in top 3 choices	620,975	-18.35***	-12.70***	-15.78***	-9.23***	-12.71***	-8.37***	-10.83***	-5.78***
		(0.05)	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.05)	(0.06)
Regents VA in school of enrollment	520,275	-17.97***	-12.88***	-16.58***	-8.64***	-13.32***	-9.20***	-12.41***	-5.81***
		(0.07)	(0.08)	(0.08)	(0.09)	(0.08)	(0.09)	(0.08)	(0.09)
SAT VA in school of enrollment	520,275	-20.89***	-13.70***	-18.72***	-9.75***	-15.80***	-9.97***	-14.19***	-6.71***
		(0.07)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.09)
Panel B: school value added									
Mean Regents VA in top 3 choices	620,975	-0.09***	-0.08***	-0.09***	-0.06***	-0.06***	-0.05***	-0.06***	-0.03***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Mean SAT VA in top 3 choices	620,975	-0.11***	-0.09***	-0.10***	-0.07***	-0.07***	-0.06***	-0.07***	-0.05***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Regents VA in school of enrollment	520,275	-0.12***	-0.09***	-0.11***	-0.06***	-0.08***	-0.06***	-0.08***	-0.04***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
SAT VA in school of enrollment	520,275	-0.11***	-0.08***	-0.10***	-0.06***	-0.08***	-0.06***	-0.08***	-0.04***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Mean and max in choice-set			X				X		
Borough FE				X				X	
Zipcode FE					X				X
Test score controls						X	X	X	X

Notes: This table reports race differences in the quality of the top 3 school choices and of the school of enrollment, as estimated by the coefficient β in equation (2). The regressions in the first column correspond to raw race gaps, while columns (2)-(8) progressively add controls for residential location, test scores and school quality available in the students' feasible set. Each row uses a different left-hand side outcome, that is a measure of school quality in applicant's top 3 choices or in the school where she enrolls.

Table 4: Correlation Between Elicited Beliefs, School Quality and Mean Achievement Levels

	Elicited belief					
	(1)	(2)	(3)	(4)	(5)	(6)
Value-Added (SD)	0.124***	0.096**			-0.133***	-0.037
	(0.033)	(0.043)			(0.049)	(0.068)
Value-Added (SD) · Non-minority		0.063				-0.194**
		(0.061)				(0.091)
Achievement level (SD)			0.230***	0.160***	0.351***	0.200***
			(0.035)	(0.049)	(0.053)	(0.076)
Achievement level (SD) · Non-minority				0.130**		0.278***
				(0.063)		(0.093)
Non-minority Respondent		-0.010		-0.081		-0.137
		(0.074)		(0.086)		(0.087)
N	849	849	849	849	849	849
Mean response	2.55	2.55	2.55	2.55	2.55	2.55

Notes: This table reports regression estimates of the relationship between elicited beliefs about school quality and school characteristics. Elicited school quality ranges from 1 (bottom quartile of school quality) to 4 (top quartile of school quality). The school attributes considered in the right hand side of the regressions are the average achievement in Regents exams of students enrolled in the school, and the school Regents VA, both expressed in standard deviations (SD) of the cross-school distribution. Even columns allow the relationship between left hand side variables and school attributes to vary across respondent's race, as measured by the interaction between school attributes and a dummy indicating white and Asian respondents.

Table 5: Demand Responses to School Letter Grades

	School share		School log share	
	(1)	(2)	(3)	(4)
A	0.15**	0.14***	0.22**	0.31***
	(0.03)	(0.03)	(0.07)	(0.06)
B	0.08*	0.07**	0.06	0.17***
	(0.04)	(0.02)	(0.05)	(0.03)
C	-0.02		-0.19**	
	(0.03)		(0.05)	
D	-0.09	-0.05	-0.35***	-0.12
	(0.05)	(0.03)	(0.05)	(0.06)
F	-0.21**	-0.15*	-0.36**	-0.21*
	(0.07)	(0.07)	(0.13)	(0.08)
Graduation % (SD)	-0.00	-0.07**	0.07	0.01
	(0.04)	(0.02)	(0.08)	(0.04)
College % (SD)	0.02	-0.01	0.01	-0.02
	(0.02)	(0.02)	(0.02)	(0.03)
Graduation % (SD) · Visible	0.02**	0.17***	0.05***	0.27***
	(0.00)	(0.02)	(0.01)	(0.04)
College % (SD) · Visible	0.00	0.03	0.03	0.07**
	(0.02)	(0.01)	(0.03)	(0.02)
Only graded schools		X		X
N	32,190	22,815	15,908	12,470
N schools	458	338	429	334
Average school share	0.606	0.766	0.606	0.766

Notes: This table presents regression estimates of letter grade effects on demand for schools. The dependent variable is the share (or log share in columns (3)-(4)) of students in a demographic cell c and application cohort t ranking the school in their top 3 choices. Demographic cells are defined by the interaction of a student residential borough and baseline test score tercile. The first 5 rows report estimates of the coefficients β_g in equation (4) for each letter grade. The other rows the coefficients of a school graduation or college rates in the year prior to when cohort t applies to schools and of their interaction with an indicator (*Visible*) for years in which these statistics were printed on the school directories. Other controls include school-cell fixed effects, year-cell fixed effects, a school average Regents performance and the share of white and Asian students enrolled at the school in the year before cohort t applies to school. Standard errors are clustered at the school-year level. All columns use school-years observations between 2010 and 2014 included, the years in which letters were printed on directories. Columns (2) and (4) restrict the observations to school-year with a grade, so that the omitted category is receiving a grade of C.

Table 6: Heterogeneity in Responses to Letter Grades by School Achievement Level

	School share		School log share	
School achievement level:	*Above median*	*Below median*	*Above median*	*Below median*
	(1)	(2)	(3)	(4)
A	0.10*	0.19***	0.19*	0.43***
	(0.04)	(0.03)	(0.08)	(0.09)
B	0.02	0.13***	0.07	0.24**
	(0.04)	(0.02)	(0.06)	(0.06)
C	-0.14**	0.07**	-0.14*	0.01
	(0.04)	(0.02)	(0.05)	(0.06)
D	-0.35	0.03	-0.49*	-0.09
	(0.19)	(0.02)	(0.19)	(0.06)
F	-0.62	-0.08	-0.58**	-0.16
	(0.36)	(0.06)	(0.18)	(0.12)
Graduation % (SD)	0.02	0.01	0.06	0.15
	(0.06)	(0.03)	(0.07)	(0.10)
College % (SD)	0.02	-0.03	0.02	-0.09**
	(0.03)	(0.02)	(0.03)	(0.03)
Graduation % (SD) · Visible	0.04***	-0.01***	0.06***	-0.02
	(0.00)	(0.00)	(0.01)	(0.01)
College % (SD) · Visible	0.03	0.04	0.04	0.12**
	(0.02)	(0.02)	(0.02)	(0.04)
N	14775	14445	8793	6159
N schools	199	204	197	187
Average school share	1.030	0.257	1.030	0.257

Notes: This table presents regression estimates of letter grade effects on demand for schools defined by the coefficient β_g in equation (4), distinguishing schools by the mean achievement levels of their students. Columns (1) and (3) restrict the sample to schools with above median student achievement levels and columns (2) and (4) to schools with below median achievement levels. The dependent variable is a school share among students choices (or log share in columns (3)-(4)), defined as the share of students in demographic cell c and application cohort t ranking the school among their first three choices. Demographic cells are defined by the interaction of a student residential borough and baseline test score tercile. Controls are those used in table 5 and always include school-cell fixed effects and year-cell fixed effects. The sample includes applicant cohorts from 2010 to 2014 included. Standard errors are clustered at the school-year level.

Table 7: Heterogeneity in Responses Across School Attributes

	X_j = "% White"			X_j = "Peer Quality"			X_j = "Achiev. Level"		
	B+H	W+A	All	B+H	W+A	All	B+H	W+A	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
S_{jt}	0.07***	-0.00	-0.02	0.07**	0.00	-0.02	0.06**	-0.00	-0.02
	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$S_{jt} \times X_j$	0.10	0.08**	0.09**	0.10	0.06*	0.08**	0.21*	0.08**	0.09**
	(0.07)	(0.02)	(0.02)	(0.05)	(0.02)	(0.02)	(0.10)	(0.02)	(0.03)
$S_{jt} \times M_c$			0.10**			0.10**			0.08**
			(0.02)			(0.03)			(0.02)
$S_{jt} \times X_j \times M_c$			0.02			0.01			0.11
			(0.06)			(0.03)			(0.07)
N	22,815	22,815	45,630	22,815	22,815	45,630	22,815	22,815	45,630
N schools	338	338	338	338	338	338	338	338	338
Average school share	0.788	0.753	0.771	0.788	0.753	0.771	0.788	0.753	0.771

Notes: This table presents regression estimates of equation (7) or a variant that does not include interactions between school attributes and student race. The dependent variable is a school share in students' top three choices, within student demographic cells defined by the interaction of student race, residential borough and baseline test score tercile. Right hand side variables include S_{jt}, a discrete letter grade rank varying from 1 to 5, a dummy X_j indicating whether school j is in the top third of schools in terms of white and Asian enrollment (columns (1)-(3)), peer quality (columns (4)-(6)) or mean achievement levels (columns (7)-(9)), a dummy M_c indicating Black and Hispanic students demographic cells, and their interactions. Controls are those used in table 5 and include school-cell fixed effects and year-cell fixed effects. The sample includes applicant cohorts from 2010 to 2014 included. Standard errors are clustered at the school-year level.

Table 8: Consequences of Letter Grade Introduction on Ranked and Offered School Attributes

	Grade A	Low grade	Regents VA σ	Regents VA pct	Peer quality pct	White and Asian %	Screened	P(matched) or P(enrolls)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: first choices				
$Post2010 \cdot M_i$	0.049***	-0.049***	0.035***	4.667***	1.487***	0.001	-0.000	
	(0.003)	(0.001)	(0.001)	(0.126)	(0.104)	(0.001)	(0.003)	
$Post2010$	0.045***	-0.024***	0.023***	1.200***	3.188***	0.021***	0.044***	
	(0.002)	(0.001)	(0.091)	(0.001)	(0.068)	(0.001)	(0.002)	
N	502,923	502,923	502,923	502,923	502,923	502,923	502,923	
Black+Hispanic mean	0.225	0.142	0.0495	58.08	67.68	0.249	0.318	
White+Asian mean	0.397	0.0400	0.172	76.32	86.37	0.572	0.568	
				Panel B: first 3 choices				
$Post2010 \cdot M_i$	0.074***	-0.112***	0.031***	4.488***	1.556***	0.001	0.005*	
	(0.003)	(0.002)	(0.001)	(0.090)	(0.082)	(0.001)	(0.003)	
$Post2010$	0.051***	-0.062***	0.028***	1.588***	3.555***	0.022***	0.031***	
	(0.002)	(0.001)	(0.001)	(0.067)	(0.057)	(0.001)	(0.002)	
N	502,923	502,923	502,923	502,923	502,923	502,923	502,923	
Black+Hispanic mean	0.484	0.340	0.0388	56.43	65.18	0.241	0.558	
White+Asian mean	0.645	0.107	0.155	74.43	84.16	0.550	0.729	
				Panel C: offers				
$Post2010 \cdot M_i$	0.026***	-0.052***	0.029***	4.492***	1.582***	-0.006***	-0.005*	-0.000
	(0.003)	(0.002)	(0.001)	(0.142)	(0.118)	(0.001)	(0.003)	(0.002)
$Post2010$	0.017***	-0.042***	0.025***	1.967***	4.108***	0.021***	0.036***	-0.031***
	(0.002)	(0.001)	(0.001)	(0.108)	(0.084)	(0.001)	(0.002)	(0.001)
N	459,617	459,617	459,617	459,617	459,617	459,617	459,617	502,923
Black+Hispanic mean	0.144	0.211	-0.0264	46.75	53.93	0.167	0.224	0.929
White+Asian mean	0.276	0.0857	0.123	69.78	79.49	0.503	0.443	0.919
				Panel D: enrollment				
$Post2010 \cdot M_i$	0.022***	-0.040***	0.030***	4.549***	1.488***	-0.004***	-0.002	-0.038***
	(0.003)	(0.002)	(0.001)	(0.157)	(0.130)	(0.001)	(0.003)	(0.002)
$Post2010$	0.012***	-0.041***	0.022***	1.530***	4.106***	0.017***	0.033***	0.008***
	(0.003)	(0.002)	(0.001)	(0.126)	(0.099)	(0.001)	(0.002)	(0.002)
N	422,654	422,654	422,654	422,654	422,654	422,654	422,654	502,923
Black+Hispanic mean	0.142	0.213	-0.0237	47.13	53.50	0.176	0.157	0.885
White+Asian mean	0.262	0.105	0.116	68.79	77.61	0.494	0.214	0.728

Notes: This table presents pooled differences in differences estimates of the differential changes in the attributes of school choices (panels A and B), school offers (panel C) and enrollment schools (panel D) by student race after the introduction of letter grades. The sample includes students applying to enroll in 9th grade between 2006 and 2014. Controls include gender, ell status, subsidized lunch status and fixed effects for combinations of student borough and baseline test score terciles.

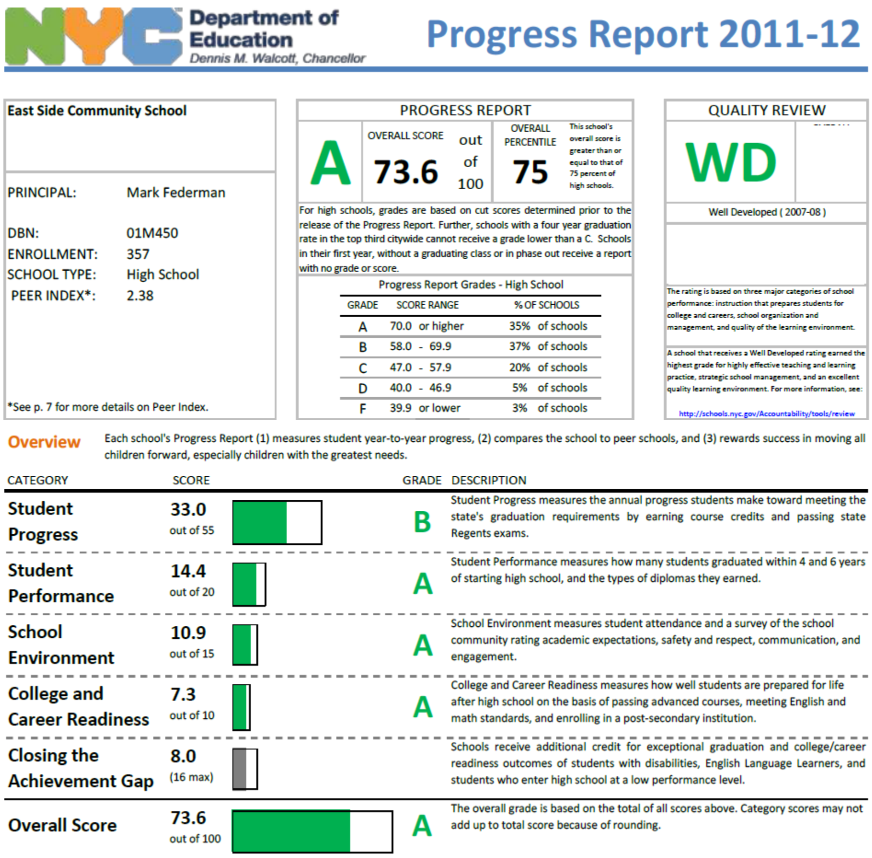
Table 9: Model Estimates - Summary Statistics

	By race			By 7th grade Math tercile		
	Black	Hispanic	White	Low	Median	High
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: first step						
δ_{cjt} SD	24.8	21.5	25.6	22.1	22.0	27.5
δ_{cjt} range	160.5	130.7	133.9	143.5	133.0	144.0
Corr(δ_{cjt}, VA)	0.33	0.37	0.60	0.24	0.47	0.61
Corr(δ_{cjt}, Peer quality)	0.48	0.49	0.74	0.34	0.63	0.77
Corr(δ_{cjt}, % white)	0.33	0.42	0.70	0.32	0.51	0.65
Panel B: second step - preferences						
γ_c	5.6	4.5	5.4	3.7	5.2	6.5
β_c^{white}	0.9	0.5	-1.0	0.5	-0.7	0.6
$\beta_c^{peerquality}$	4.9	4.3	4.3	3.7	4.7	5.1
$\tilde{\xi}_{cj}$ SD	22	17	27	15	22	29
$\tilde{\xi}_{cj}$ range	123	103	144	100	120	149
$\tilde{\xi}_{cj}$ skewness	0.00	0.26	0.45	0.24	0.26	0.26
Corr($\tilde{\xi}_{cj}$, VA)	0.15	0.21	0.47	0.09	0.32	0.45
Corr($\tilde{\xi}_{cj}$, Peer quality)	0.25	0.28	0.57	0.16	0.42	0.53
Corr($\tilde{\xi}_{jc}$, % white)	0.15	0.28	0.60	0.21	0.39	0.46
Panel C: second step - beliefs						
μ_{c0}	-0.09	-0.07	0.07	-0.11	-0.03	0.06
μ_{c1}	-0.01	0.04	0.13	0.02	0.01	0.14
μ_{c2}	0.15	0.15	0.14	0.14	0.16	0.14
σ_{c0}	2.88	2.52	3.18	2.79	3.10	2.63
σ_{c1}	-0.12	-0.10	0.25	-0.17	0.17	0.04
σ_{c2}	0.08	0.11	0.10	0.09	0.11	0.09
Absolute Bias	0.65	0.62	0.55	0.63	0.64	0.54
μ_{cj} below med. R_j, below med. Q_j	-0.20	-0.22	-0.14	-0.24	-0.15	-0.16
μ_{cj} above med. R_j, below med. Q_j	-0.19	-0.16	0.02	-0.19	-0.12	0.00
μ_{cj} below med. R_j, above med. Q_j	-0.04	-0.05	0.03	-0.08	0.01	0.00
μ_{cj} above med. R_j, above med. Q_j	0.04	0.11	0.33	0.05	0.13	0.32

Notes: This table summarizes the model estimates. Panel A reports summary statistics for the estimates of the mean school utility δ_{cjt} obtained in the first step. Panel B reports statistics for the second step estimates of the preference parameters $\gamma_c, \beta_c, \xi_{cj}$ and panel C for the prior moments $\boldsymbol{\mu_c}, \boldsymbol{\sigma_c^{-1}}$ taking a weighted average of cell-specific estimates across cells sharing the same covariate (race or baseline test score), using weights proportional to cell size.

A Appendix Tables and Figures

Figure A1: Example of An Online School Quality Report With Letter Grades

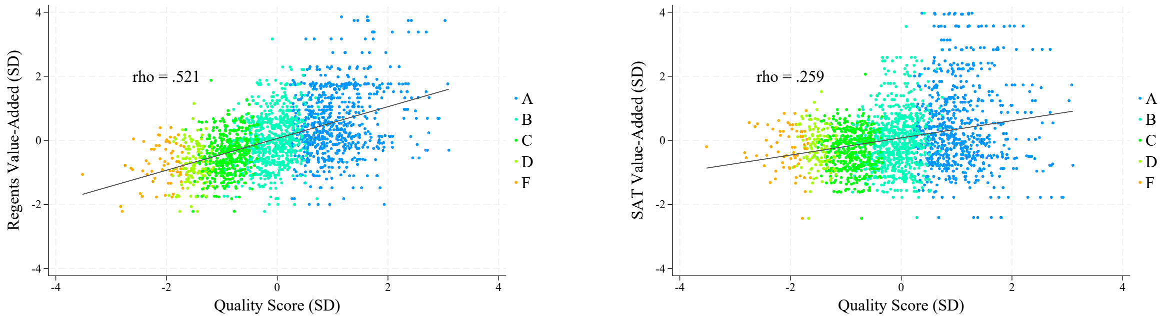


Notes: This figure shows the 2011/12 progress report for East Side Community School as an example of how a school progress report looked like. Source: www.crpe.org

Figure A2: Correlation Between School VA and the Bloomberg Quality Score

(a) Regents VA

(b) SAT VA

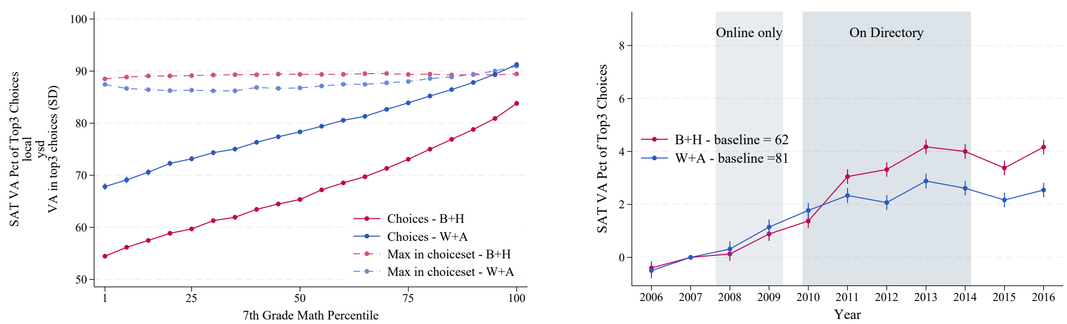


Notes: This figure shows scatter plots of the quality score used in progress reports (x-axis) against OLS measures of Regents VA (panel a)) and SAT VA (panel b)), together with the corresponding correlation coefficient. Each dot is a school-year. Different colors indicate the letter grade received by each observation.

Figure A3: The Racial School Quality Choice Gap - SAT VA

(a) By Baseline Achievement

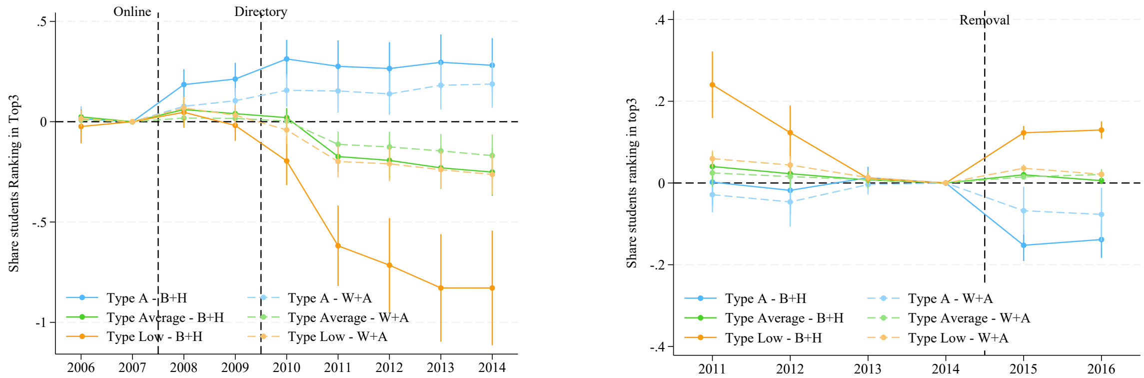
(b) Across Cohorts



Notes: This figure describes cross-race differences in chosen school quality, as measured by SAT value-added. It is analogous to Figure 1, but uses SAT value-added to measure school quality, rather than Regents value-added.

Figure A4: Event Study Estimates - Separate Regressions by Student Race

(a) Introduction of Letters (b) Removal of Letters

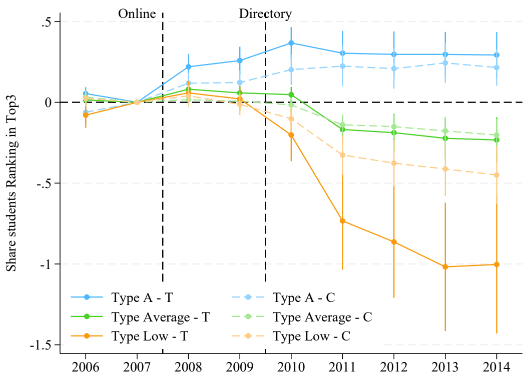


Notes: The figure plots event study estimates of the coefficient β_L^t of equation (3), from separate regressions by race. Panel (a) considers changes relative to 2007, the year before the introduction of letters, using applicant cohorts of 2006-2014. Panel (b) considers changes around the removal of letters, normalizing share differences to 0 in 2014, and using cohorts of 2011-2016. Blue lines are for changes in shares for Type A schools, orange ones for shares of Type Low schools and the green ones for Type Average schools. Dashed lighter lines are for changes in shares of white and Asian students, solid ones for changes in shares of Black and Hispanic students.

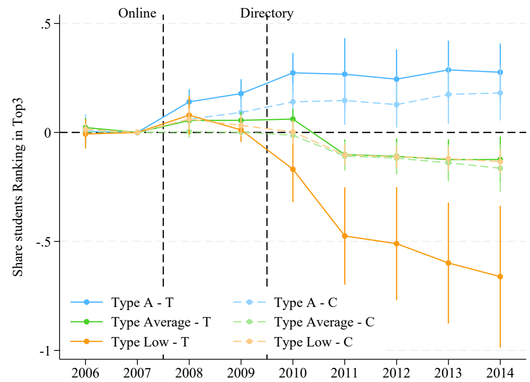
Figure A5: Demand Responses to Introduction and Removal of Quality Signals - Heterogeneity by Exposure to New Information

Introduction of Letters

(a) Minority Applicants

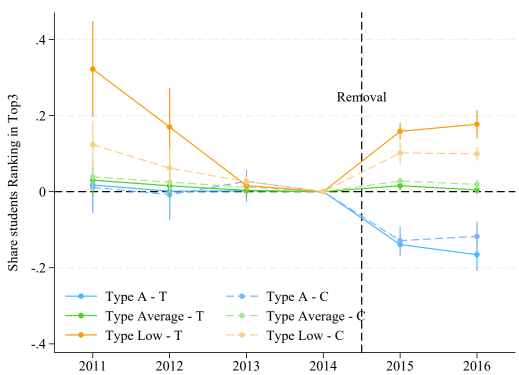


(b) Non-Minority Applicants

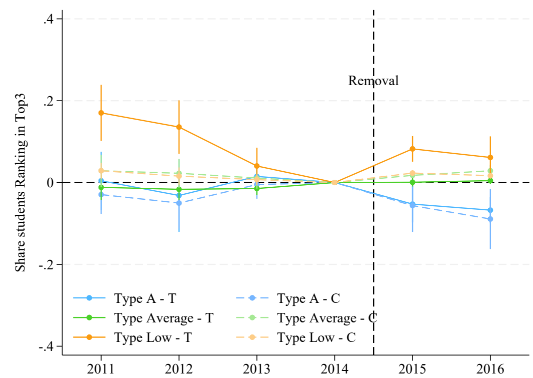


Removal of Letters

(c) Minority applicants



(d) Non-Minority applicants

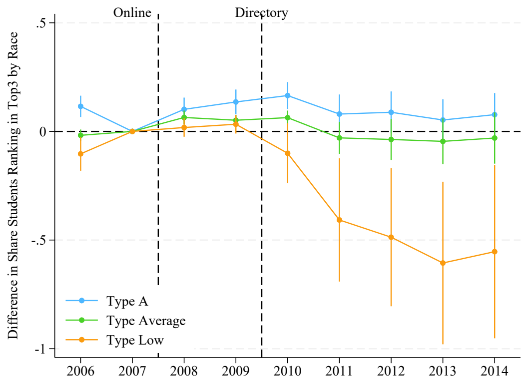


Notes: The figure plots event study estimates of the coefficient β_L^t of equation (3), from separate regressions by race and values of the dummy $Treated_i$. Panels (a) and (b) consider changes relative to 2007, the year before the introduction of letters, using applicant cohorts of 2006-2014. Panels (c) and (d) changes around the removal of letters, normalizing share differences to 0 in 2014, and using cohorts of 2011-2016. Blue lines are for changes in shares for Type A schools, orange ones for shares of Type Low schools and the green ones for Type Average schools. Dashed lighter lines are for changes in choice shares of students for whom $Treated_i = 0$, solid ones for choice shares among students with $Treated_i = 1$.

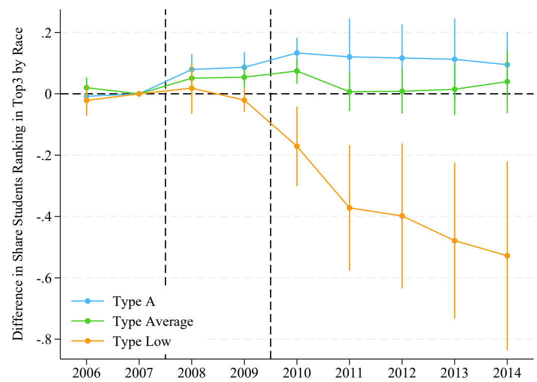
Figure A6: Event Study Estimates of Demand Responses to Introduction and Removal of Quality Signals - Heterogeneity by Exposure to New Information

Introduction of Letters

(a) Minority Applicants

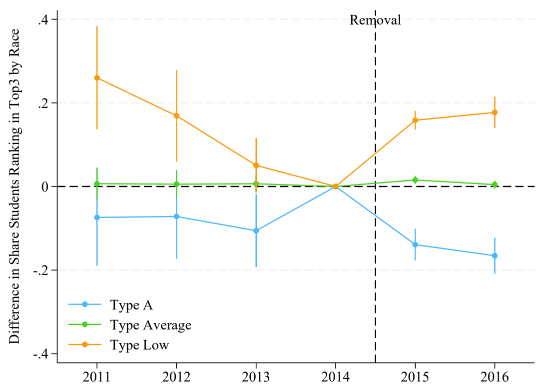


(b) Non-Minority Applicants

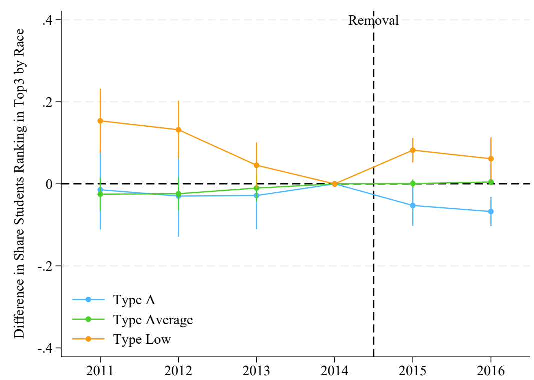


Removal of Letters

(c) Minority Applicants



(d) Non-Minority Applicants

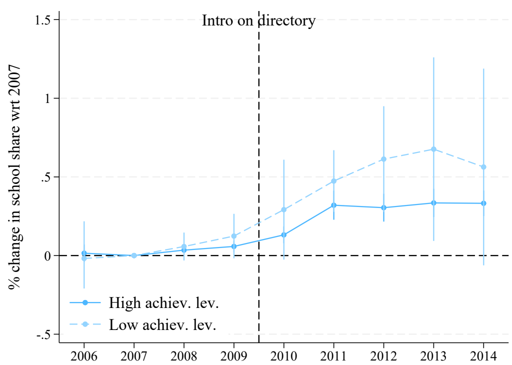


Notes: The figure plots event study estimates of the coefficient δ_L^t of a variant of equation (5) that considers differences in choice responses to the introduction and removal of letter grades along values of the dummy $Treated_i$ (rather than across race). Panels (a) and (b) consider differential changes relative to 2007, using applicant cohorts of 2006-2014, separately for minority and white students. Panels (c) and (d) consider changes around the removal of letters, normalizing share differences to 0 in 2014, and using cohorts of 2011-2016. Blue lines are for changes in shares for Type A schools, orange lines for shares of Type Low schools and the green lines are for Type Average schools.

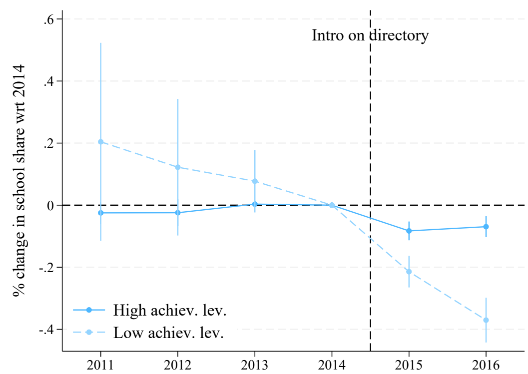
Figure A7: Heterogeneity in Responses to Introduction and Removal of Quality Signals by School Peer Quality

Introduction and Removal of Positive Quality Signals

(a)

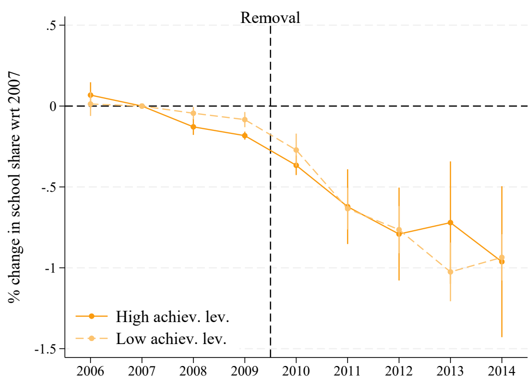


(b)

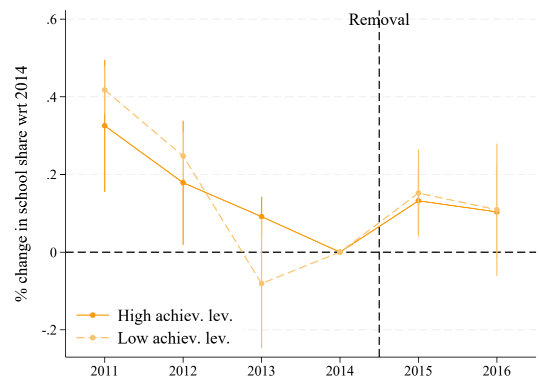


Introduction and Removal of Negative Quality Signals

(c)



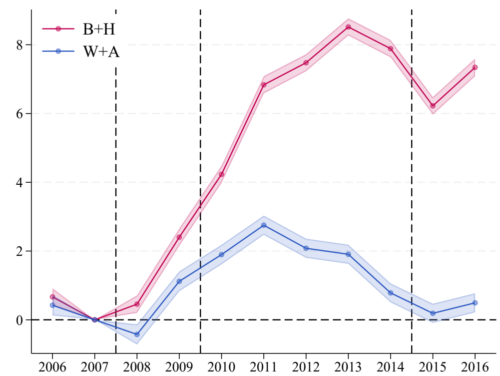
(d)



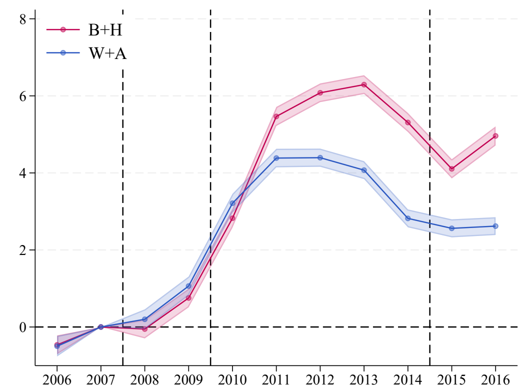
Notes: The figure plots event study estimates of the coefficient β_L^t of equation (3), from separate regressions for schools of different types and different peer quality. Panels (a) and (c) consider share changes relative to 2007, respectively for Type A and Type Low schools, using applicant cohorts of 2006-2014. Panels (b) and (d) share changes around the removal of letters, normalizing shares to 0 in 2014, and using cohorts of 2011-2016. Blue lines are for changes in shares of Type A schools, orange ones for shares of Type low schools. Dashed lighter lines are for changes in choice shares of schools enrolling lower achieving students, solid ones for schools enrolling higher achieving students.

Figure A8: Evolution of Ranked School Characteristics by Race

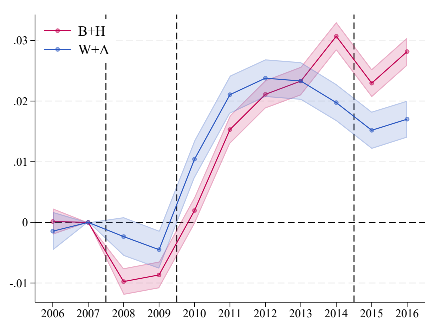
(a) Math VA (pct)



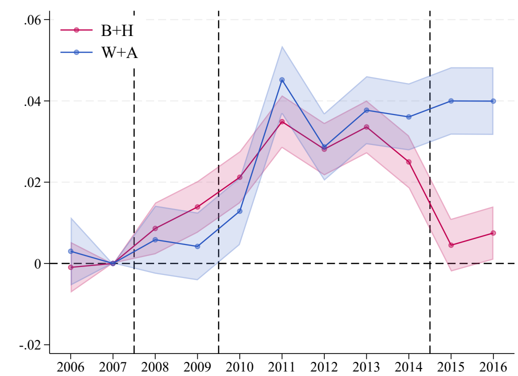
(b) Peer Quality (pct)



(c) White share

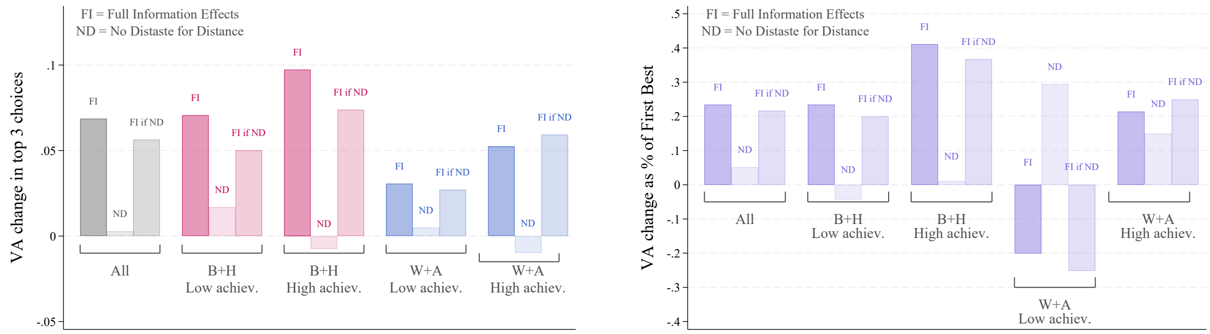


(d) I(applies out of borough)



Notes: The figure plots regression estimates of changes in average characteristics of applicants' first three high school choices over time with respect to 2007, by applicant race. Blue lines are for choices of white and Asian students, pink lines for choices of Black and Hispanic students. The lines plot coefficients of year dummies, normalizing the 2007 coefficient to 0. Controls include gender, ell status, subsidized lunch status and fixed effects for combinations of student borough and baseline test score terciles. Panel (a) shows trends in Regents value-added and panel (b) in peer quality of a student's first three choices, panel (c) in the share of white students enrolled in the student's first three school choices and panel (d) in the probability of applying to a school outside one's borough.

Figure A9: Role of Distance in Explaining Choice Gaps and Effects of Information

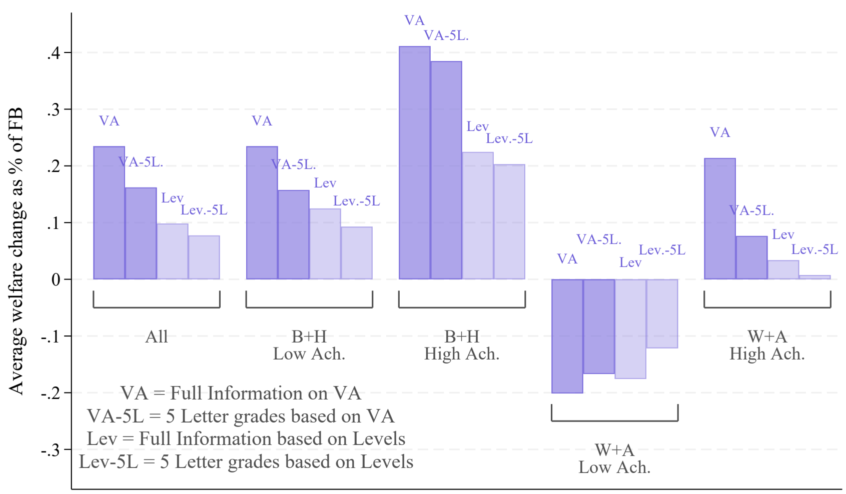


(a) Average ΔVA in Choices

(b) Welfare Gains

Notes: Panel (a) shows changes in VA of top 3 choices across three different simulations, taking averages within student groups defined by race and baseline achievement (above or below median). Panel (b) does the same thing for the resulting change in offered VA, expressed as a percentage of the average first-best achievement gains. Within each students subgroup, the first bar corresponds to the full-information benchmark that uses the real model estimates. The second bar corresponds to differences between the status-quo and a simulation in which students do not have distaste for commuting but the information environment is as in the status-quo. The third bar simulates changes with respect to the status-quo of providing full information if students do not have a distaste for commuting.

Figure A10: Information About VA vs. Information About Achievement Levels

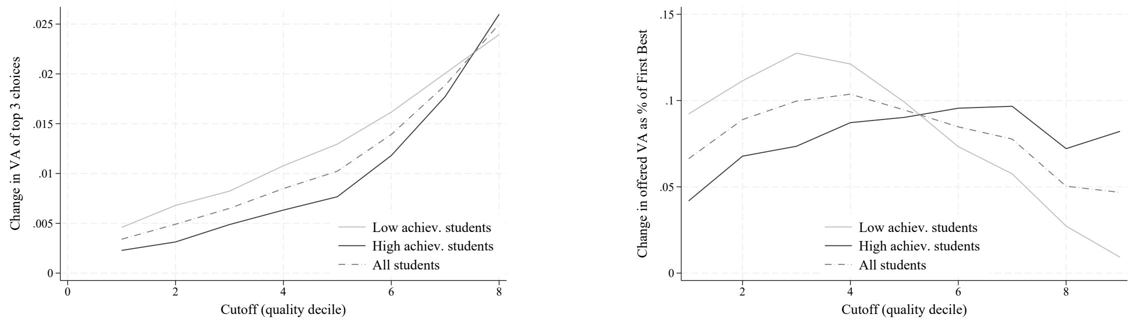


Notes: This figure plots the average welfare change, as defined by the average change in student test scores with respect to the status quo, by student subgroups for four different counterfactual simulations of student assignment. Welfare gains are expressed as a percentage of the average first-best achievement gains. Student subgroups are defined by combinations of race and baseline achievement. "VA" denotes the simulated student assignment under full information about school value added, "VA-5L" a counterfactual in which schools are rated from 1 to 5 based on their VA quintile, "Lev" a counterfactual in which students are told about differences in school achievement levels and these are presented as differences in VA, while "VA-5L" the counterfactual in which schools are rated from 1 to 5 based on their achievement level quintile.

Figure A11: School Quality of Choices and Offers as Signal Precision Increases at the Top

(a) Δ VA in Choices

(b) Welfare Gains



Notes: This figure plots how value added of the top three school choices and of school offers changes as a function of the cutoff used to assign schools to a low or a high quality rating. Panel (a) plots changes in the average value-added of students top three choices, while Panel (b) plot achievement gains as a share of first-best gains. The dotted lines are for averages across all students, dark gray lines are for students with above median baseline test scores, light gray lines for students with below median baseline test scores.

Table A1: Changes in Information Provided by the DOE

Year (fall 9th grade)	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
Letter grades			○	○	✓	✓	✓	✓	✓		
Letter grade subcategories			○	○	○	○	○	○	✓		
Graduation %	✓	✓	○	○	○	✓	✓	✓	✓	✓	✓
College %	○	○	○	○	○	○	○	✓	✓	✓	✓
Regents performance	✓	✓	○	○	○	○	○	○	○	○	○
State quality review			○	○	✓	✓	✓	✓	✓	✓	✓
Quality measures - de Blasio											○
Survey-based measures										✓	✓
(feel safe, satisfaction, variety of classes)											

[1] ✓- information provided on the school directory (and online)
[2] ○- information provided online only

Notes: This table summarizes which type of information about school performance was shown on the printed high school directory and online (denoted with ✓) and which was only shown online (denoted with ○). Years denote applicant cohorts and refer to the fall of their enrollment in 9th grade. Information is distributed (and applicants apply) in the preceding year.

Table A2: School Progress Report Score Components

Component	Description	2006	2007	2008	2009	2010	2011	2012	Average
Progress	*Students on track for graduation (credits), Students in school lowest 3rd on track for graduation, Regents pass rate*	51%	56%	56%	57%	56%	50%	47%	53%
Performance	*Graduation rate, Regents Diploma rate*	31%	25%	24%	24%	25%	20%	19%	24%
Environment	*Attendance rate, answers from school environment survey*	13%	14%	14%	14%	14%	14%	16%	14%
College and Career readiness	*College readiness index, college enrollment rate*	0%	0%	0%	0%	0%	10%	10%	3%
Extra points	*ELL diploma rate, city lowest 3rd diploma rate, sped regents pass rate*	5%	6%	6%	5%	5%	7%	8%	6%
Total		100%	100%	100%	100%	100%	101%	101%	100%

Notes: This table describes the components of the quality score used to assign letter grades, their year-specific weight and the outcomes used to create them. The last column reports the average weight of each component across years. The year refers to the fall of the school year of the progress report. For example, 2006 refers to the 2006-2007 school progress report, which graded schools existing in the 2006-2007 school year. This progress report was made available to the public during the 2007-2008 school year, and therefore would have been used by the 2008 high school enrollment cohort to decide where to apply.

Table A3: Demand Responses to Introduction and Removal of School Quality Signals - Pooled Pre-Post Estimates

	School share			School log share		
	minority	white	difference	minority	white	difference
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: effect of introduction of information						
Type A $\cdot Post2010$	0.25***	0.17***	0.09**	0.26***	0.16**	0.10**
	(0.05)	(0.04)	(0.04)	(0.04)	(0.05)	(0.03)
Average $\cdot Post2010$	-0.13**	-0.08*	-0.05**	-0.10	-0.18**	0.07**
	(0.05)	(0.04)	(0.02)	(0.07)	(0.06)	(0.02)
Type Low $\cdot Post2010$	-0.47***	-0.28**	-0.20***	-0.66***	-0.71***	0.05
	(0.12)	(0.09)	(0.05)	(0.15)	(0.13)	(0.05)
Never graded $\cdot Post2010$	0.08	0.07	0.02	0.39	0.17	0.22
	(0.07)	(0.05)	(0.02)	(0.34)	(0.22)	(0.19)
Graduation % (SD) $\cdot$ Visible	-0.03	-0.02	-0.01	0.00	-0.04	0.04
	(0.03)	(0.02)	(0.01)	(0.04)	(0.04)	(0.02)
College % (SD) $\cdot$ Visible	0.02*	0.01*	0.01	0.02	0.02*	-0.01
	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)
N	54855	54855	109710	27896	15267	43163
N schools	463	463	463	446	432	446
Panel B: effect of removal of letters						
Type A $\cdot Post2015$	-0.13**	-0.10*	-0.03	-0.15***	-0.11**	-0.04
	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)	(0.03)
Type Average $\cdot Post2015$	-0.00	0.01	-0.01	-0.03	0.01	-0.04
	(0.01)	(0.01)	(0.01)	(0.03)	(0.02)	(0.02)
Type Low $\cdot Post2015$	0.02	0.00	0.02	0.02	0.06	-0.04
	(0.03)	(0.02)	(0.02)	(0.10)	(0.09)	(0.07)
Never graded $\cdot Post2015$	0.15**	0.10**	0.05**	0.20**	0.15**	0.05
	(0.04)	(0.03)	(0.02)	(0.05)	(0.04)	(0.05)
N	38,370	38,370	76,740	18,596	10,411	29,007
N schools	453	453	453	432	427	432

Notes: This table presents regression estimates of changes in demand for schools after the introduction (panel A) or after the removal (panel B) of letter grades for different categories of schools. The dependent variable is the share of students (or log share in columns (4)-(6)) in demographic cell c and application cohort t ranking the school among their first three choices. Demographic cells are defined by the interaction of student race, residential borough and baseline test score tercile. Schools are divided into mutually exclusive categories, fixed over time: *Type A* indicates schools receiving a grade of A in most years, *Type Low* indicates schools receiving a grade of C, D or F in most years, *Never graded* indicates schools that were never graded, while *Type Average* is a residual category for the remaining schools. Columns (1), (2), (4) and (5) report changes in the school shares over time separately by applicant race, pooling the event study coefficients β_L^t of equation (3) into pre-post differences. Columns (3) and (6) report estimates of the race-difference in changes in demand over time, pooling the event study coefficients δ_L^t of equation (5) into pre-post differences. Panel A uses application cohorts of 2006-2014, while panel B uses the 2011-2016 cohorts.

Table A4: Demand Responses to Quality Signals - Heterogeneity by Applicant Race

	School share				School log share			
	Black+Hispanic students		White+Asian students		Black+Hispanic students		White+Asian students	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A	0.19**	0.19***	0.08*	0.04	0.18**	0.29***	0.22	0.25***
	(0.04)	(0.04)	(0.03)	(0.02)	(0.07)	(0.06)	(0.11)	(0.05)
B	0.10*	0.09**	0.05	0.01	0.05	0.17***	0.07	0.13**
	(0.04)	(0.02)	(0.03)	(0.01)	(0.05)	(0.03)	(0.09)	(0.04)
C	-0.03		0.01		-0.19**		-0.14*	
	(0.04)		(0.02)		(0.05)		(0.06)	
D	-0.12	-0.07	-0.01	-0.01	-0.37***	-0.12	-0.34**	-0.18*
	(0.07)	(0.04)	(0.02)	(0.02)	(0.05)	(0.06)	(0.08)	(0.07)
F	-0.30**	-0.23*	-0.01	0.01	-0.44**	-0.26**	0.12	0.07
	(0.10)	(0.10)	(0.03)	(0.02)	(0.12)	(0.09)	(0.14)	(0.06)
Graduation % (SD)	0.00	-0.09**	-0.01	-0.04**	0.06	0.01	0.05	0.01
	(0.06)	(0.03)	(0.02)	(0.01)	(0.08)	(0.04)	(0.08)	(0.04)
College % (SD)	0.03	-0.00	-0.00	-0.02	0.01	-0.02	-0.01	-0.05
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)	(0.05)
Graduation % (SD) · Visible	0.02**	0.21***	0.01**	0.09***	0.05***	0.26***	0.05***	0.26***
	(0.01)	(0.03)	(0.00)	(0.02)	(0.01)	(0.04)	(0.01)	(0.04)
College % (SD) · Visible	0.00	0.03*	-0.00	0.01	0.02	0.06**	0.04	0.06**
	(0.02)	(0.01)	(0.01)	(0.02)	(0.03)	(0.02)	(0.03)	(0.01)
Only graded schools		X		X		X		X
N	32,190	22,815	32,190	22,815	15,213	11,936	8,266	6,597
N schools	458	338	458	338	429	334	409	319
Average school share	0.625	0.782	0.579	0.745	0.625	0.782	0.579	0.745

Notes: This table presents regression estimates of letter grade effects on demand for schools, separately measuring effects on the school choices of Black and Hispanic students and of white and Asian students. The dependent variable is the share (or log share in columns (4)-(6)) of students in demographic cell c and application cohort t ranking the school among their first three choices. Demographic cells are defined by the interaction of student race, residential borough and baseline test score tercile. The first 5 rows report estimates of the coefficients β_g in equation (4) for each letter grade. The other rows the coefficients of a school graduation or college rates in the year prior to when cohort t applies and of their interaction with an indicator (*Visible*) for years when these statistics were printed on the school directories. Other controls include school-cell fixed effects, year-cell fixed effects, a school average Regents performance and the share of white and Asian students enrolled at the school in the year prior cohort t applies to school. Standard errors are clustered at the school-year level. Estimates use applicant cohorts from 2010 to 2014 included. Even columns restrict the observations in the preceding columns to schools receiving a grade, so that the omitted category is receiving a grade of C.

Table A5: Demand Responses to Quality Signals - Robustness to Using Both Letters

	All (1)	Minority (2)	White (3)	All (4)	Minority (5)	White (6)	All (7)	Minority (8)	White (9)
A	0.14***	0.19***	0.03	0.15***	0.20***	0.04			
	(0.02)	(0.04)	(0.02)	(0.02)	(0.03)	(0.02)			
B	0.07**	0.10**	0.01	0.07**	0.10**	0.02			
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)			
D	-0.06	-0.08	-0.01	-0.05	-0.07	-0.00			
	(0.04)	(0.04)	(0.02)	(0.04)	(0.04)	(0.02)			
F	-0.18*	-0.27*	0.01	-0.15	-0.24*	0.02			
	(0.07)	(0.10)	(0.03)	(0.08)	(0.11)	(0.04)			
A - 2				0.09**	0.12*	0.02			
				(0.03)	(0.04)	(0.02)			
B - 2				0.05	0.07*	0.02			
				(0.02)	(0.03)	(0.02)			
D - 2				-0.01	-0.03	0.01			
				(0.03)	(0.04)	(0.01)			
F - 2				-0.20	-0.25	-0.08			
				(0.11)	(0.14)	(0.05)			
Two As							0.13**	0.17**	0.02
							(0.03)	(0.04)	(0.04)
One A							0.07**	0.09**	0.01
							(0.02)	(0.02)	(0.03)
Graduation %	-0.07**	-0.08**	-0.04*	-0.10**	-0.12**	-0.05**	-0.07**	-0.08**	-0.04*
	(0.02)	(0.02)	(0.01)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)
College %	-0.01	0.00	-0.01	-0.00	0.00	-0.01	0.00	0.01	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Graduation % · Visible	0.20***	0.24***	0.10***	0.20***	0.24***	0.11***	0.20***	0.24***	0.10***
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.03)	(0.01)
College % · Visible	0.02	0.03	0.01	0.03	0.03	0.01	0.03	0.04	0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
N	20,685	20,685	20,685	20,685	20,685	20,685	20,685	20,685	20,685
N schools	316	316	316	316	316	316	316	316	316
Average school share	0.766	0.782	0.745	0.766	0.782	0.745	0.766	0.782	0.745

Notes: This table presents robustness checks on estimates of letter grade effects presented in table 5 and A4 by separately estimating the effect of the two letter grades (one for each of the two preceding years) printed on the directory received by cohort t. The dependent variable is the share of students from a given demographic group listing the school among their first three choices. The sample includes applicant cohorts of 2010 - 2014. Columns (1) - (3) report for comparison the benchmark estimates of the effects of the most recent letter grade printed on the directory in equation 4. The equation estimated in columns (4) - (6) extends equation 4 by adding letter grade dummies for the additional grade printed on the directory, corresponding to that received two years prior to when cohort t applies to high school. Columns (7) - (9) substitute letter grade indicators in equation 4 with indicators for receiving two consecutive As or only one A (in one out of the two years), leaving as omitted category the event of not receiving an A in any of the two years considered in the directory of cohort t. Controls in columns (1), (4) and (7) are the same as in table 5 and those in the remaining columns are the same as in table A4. These always include school-cell and year-cell fixed effects.

Table A6: Demand Responses to Quality Signals - Robustness to Using Additive or Interactive Models

	All (1)	Minority (2)	White (3)	All (4)	Minority (5)	White (6)
A	0.07**	0.10***	0.02			
	(0.02)	(0.02)	(0.03)			
Low	-0.08**	-0.11**	-0.02			
	(0.02)	(0.03)	(0.02)			
A - 2	0.03	0.05	-0.00			
	(0.02)	(0.02)	(0.02)			
Low - 2	-0.07*	-0.09*	-0.03			
	(0.03)	(0.04)	(0.02)			
A-A				0.10**	0.14**	0.01
				(0.03)	(0.04)	(0.05)
A-B				0.06*	0.09**	0.01
				(0.03)	(0.03)	(0.04)
A-Low				-0.06	-0.08	-0.02
				(0.06)	(0.07)	(0.05)
B-A				0.02	0.04	-0.01
				(0.02)	(0.02)	(0.03)
B-Low				-0.09*	-0.11*	-0.05
				(0.04)	(0.05)	(0.03)
Low-A				-0.09	-0.12*	-0.03
				(0.05)	(0.05)	(0.04)
Low-B				-0.10**	-0.14**	-0.04
				(0.03)	(0.04)	(0.03)
Low-Low				-0.15**	-0.20**	-0.04
				(0.05)	(0.06)	(0.03)
Graduation %	-0.09**	-0.11**	-0.04*	-0.09**	-0.11**	-0.04*
	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)
College %	-0.00	0.00	-0.01	-0.00	0.00	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Graduation % · Visible	0.20***	0.24***	0.10***	0.20***	0.24***	0.10***
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)
College % · Visible	0.03	0.04	0.00	0.03	0.04	0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Constant	-0.06	-0.44	0.56	-0.07	-0.46	0.55
	(0.43)	(0.51)	(0.53)	(0.42)	(0.51)	(0.53)
N	20,685	20,685	20,685	20,685	20,685	20,685
N schools	316	316	316	316	316	316
Average school share	0.766	0.782	0.745	0.766	0.782	0.745

Notes: This table presents robustness checks on benchmark estimates of letter grade effects by considering models that estimate the effect of the two letter grades (one for each of the two preceding years) printed on the directory received by cohort t and that allow the two grade effects to be either additive (columns (1)-(3)) or interactive (columns (4)-(6)). The dependent variable is the share of students from a given demographic group listing the school among their first three choices. The sample includes applicant cohorts of 2010 - 2014. Grades of C, D, and F are pooled in one "Low grade" category. Controls in columns (1) and (4) are the same as in table 5 and those in the remaining columns are the same as in table A4. These always include school-cell and year-cell fixed effects.

Table A7: Demand Responses to Quality Score and Its Subcomponents

	School share in top 3 choices					
	(1)	(2)	(3)	(4)	(5)	(6)
Progress score (SD)	0.03**	-0.00	-0.00			
	(0.01)	(0.01)	(0.01)			
Performance score (SD)	0.04***	0.02*	0.02**			
	(0.01)	(0.01)	(0.01)			
Environment score (SD)	0.02	0.01	0.01			
	(0.01)	(0.01)	(0.01)			
Progress score^2 (SD)			0.00			
			(0.00)			
Performance score^2 (SD)			0.00			
			(0.01)			
Environment score^2 (SD)			-0.01			
			(0.01)			
Quality score (SD)				0.06**	-0.01	-0.18*
				(0.01)	(0.02)	(0.08)
Quality score^2 (SD)						0.02*
						(0.01)
A		0.12***	0.12***		0.15***	0.18***
		(0.02)	(0.02)		(0.03)	(0.04)
B		0.06**	0.06**		0.07**	0.10**
		(0.02)	(0.02)		(0.02)	(0.02)
D		-0.04	-0.04		-0.06	-0.09**
		(0.03)	(0.02)		(0.03)	(0.03)
F		-0.13	-0.12*		-0.15	-0.22*
		(0.06)	(0.06)		(0.07)	(0.08)
N	22,815	22,815	22,815	22,815	22,815	22,815
N schools	338	338	338	338	338	338
Average school share	0.766	0.766	0.766	0.766	0.766	0.766

Notes: This table presents regression estimates of the effect of the quality score components on demand for schools, with and without controlling for letter grade fixed effects. It shows that demand does not respond to changes in the quality score and its components, beyond the variation controlled for by letter grade fixed effects in columns (2)-(4) and (5)-(6). The dependent variable is the share of students (or log share in columns (4)-(6)) in demographic cell c and application cohort t ranking the school among their first three choices. Demographic cells are defined by the interaction of a student residential borough and baseline test score tercile. The set of controls is the same as for regressions in table 5. The estimation sample includes cohorts from 2010 to 2014 included.

Table A8: Applicants Descriptive Statistics by Information Exposure Treatment Status

	All		Minority		Non-Minority	
	Control	Treated	Control	Treated	Control	Treated
	(1)	(2)	(3)	(4)	(5)	(6)
Black	0.15	0.40	0.31	0.48	0.00	0.00
Hispanic	0.34	0.42	0.69	0.52	0.00	0.00
white	0.24	0.07	0.00	0.00	0.46	0.41
Asian	0.27	0.10	0.00	0.00	0.53	0.54
Subsidized lunch	0.69	0.86	0.78	0.88	0.60	0.77
Ell	0.07	0.12	0.08	0.12	0.07	0.14
7th grade Math	0.47	-0.10	0.09	-0.19	0.83	0.34
7th grade English	0.38	-0.13	0.12	-0.18	0.63	0.11
Bronx	0.05	0.38	0.08	0.42	0.02	0.18
Brooklyn	0.30	0.37	0.27	0.38	0.32	0.31
Manhattan	0.14	0.03	0.17	0.04	0.12	0.01
Queens	0.41	0.19	0.40	0.15	0.43	0.33
Staten Island	0.10	0.03	0.08	0.00	0.12	0.17
Share of mostly grade A seats in neighborhood	0.17	0.10	0.16	0.11	0.18	0.07
Share of mostly low grade seats in neighborhood	0.16	0.23	0.16	0.21	0.15	0.31
Average 7th grade math in neighborhood	0.00	-0.15	-0.06	-0.17	0.06	-0.07
Minimum distance to A (minutes)	24.09	23.73	22.92	22.52	25.24	29.90
Minimum distance to Log grade (minutes)	29.98	23.28	28.52	22.33	31.42	28.08
% of 2006 MS students applying to mostly A	0.67	0.42	0.66	0.42	0.69	0.39
% of 2006 MS students applying to mostly low	0.12	0.39	0.16	0.41	0.09	0.30
N	333,938	254,810	162,358	209,571	171,580	45,239

Notes: This table provides student descriptive statistics across values of the indicator $Treated_i$ defined in section 3.2. Columns (1)-(2) report mean statistics considering all students, while columns (3)-(6) split students by race. The term "Minority" refers to Black and Hispanic students, while "Non-Minority" includes both white and Asian students.

Table A9: Consequences of Letter Grade Introduction on Simulated Offers

	(1) Grade A	(2) Low grade	(3) Regents VA σ	(4) Regents VA pct	(5) Peer quality pct	(6) White and Asian %	(7) Screened	(8) P(matched) or P(enrolls)
Panel A: simulated offers under no screening								
$Post2010 \cdot M_i$	0.034***	-0.041***	0.026***	3.716***	0.615***	-0.011***	-0.002	-0.012***
	(0.003)	(0.002)	(0.001)	(0.158)	(0.135)	(0.001)	(0.003)	(0.002)
$Post2010$	0.017***	-0.049***	0.028***	2.464***	4.893***	0.028***	0.033***	-0.025***
	(0.003)	(0.002)	(0.001)	(0.126)	(0.102)	(0.001)	(0.002)	(0.002)
N	431,526	431,526	431,443	431,443	431,526	431,373	422,654	503,150
Black+Hispanic mean	0.154	0.198	-0.0119	48.94	56.67	0.190	0.157	0.914
White+Asian mean	0.263	0.0979	0.111	68.15	77.68	0.496	0.214	0.763
Panel B: offers								
$Post2010 \cdot M_i$	0.026***	-0.052***	0.029***	4.492***	1.582***	-0.006***	-0.005*	-0.000
	(0.003)	(0.002)	(0.001)	(0.142)	(0.118)	(0.001)	(0.003)	(0.002)
$Post2010$	0.017***	-0.042***	0.025***	1.967***	4.108***	0.021***	0.036***	-0.031***
	(0.002)	(0.001)	(0.001)	(0.108)	(0.084)	(0.001)	(0.002)	(0.001)
N	459,617	459,617	459,617	459,617	459,617	459,617	459,617	502,923
Black+Hispanic mean	0.144	0.211	-0.0264	46.75	53.93	0.167	0.224	0.929
White+Asian mean	0.276	0.0857	0.123	69.78	79.49	0.503	0.443	0.919

Notes: This table presents pooled differences in differences estimates of the differential changes in the attributes of school offers (panel B) and of school offers simulated using admission rules that remove all priorities based on residential address and academic screening (panel A). The sample includes students applying to enroll in 9th grade between 2006 and 2014. Controls include gender, ell status, subsidized lunch status and fixed effects for combinations of student borough and baseline test score terciles.

Table A10: Consequences of Letter Grade Introduction on Achievement Inequality

	Regents Math (1)	SAT Math (2)	Graduates in time (3)	College in time (4)
Panel A: pooled diff-in-diff estimates by race				
$Post2010 \cdot M_i$	0.06***	0.01	0.05***	0.07***
	(0.00)	(0.00)	(0.00)	(0.00)
N	339,182	292,828	428,789	426,937
Black+Hispanic mean	-0.189	-0.283	0.660	0.452
White+Asian mean	0.628	0.656	0.874	0.768
Panel B: pooled diff-in-diff estimates by exposure to new information (Black and Hispanic students)				
$Post2010 \cdot Treated_i$	0.02***	-0.02***	0.01***	0.02***
	(0.01)	(0.01)	(0.00)	(0.00)
N	210,817	143,634	244,386	243,302
Treated Black+Hispanic mean	-0.304	-0.395	0.617	0.394
Control Black+Hispanic mean	-0.0170	-0.147	0.720	0.534
Panel C: pooled diff-in-diff estimates by exposure to new information (White and Asian students)				
$Post2010 \cdot Treated_i$	0.07***	0.01	0.03***	0.05***
	(0.01)	(0.01)	(0.00)	(0.01)
N	85,391	108,516	129,746	129,551
Treated White+Asian mean	0.294	0.264	0.802	0.645
Control White+Asian mean	0.728	0.750	0.894	0.802

Notes: This table presents pooled differences in differences estimates of the differential changes in the achievement outcomes by student race (panel A), and by values of the variable $Treated_i$ defined in section 3.2 within race (panels B and C) after the introduction of letter grades. The sample includes students from cohorts between 2006 and 2014, who enroll in the district and have non-missing achievement outcomes. Controls include gender, ell status, subsidized lunch status and fixed effects for combinations of student borough and baseline test score terciles.

Table A11: Model Estimates

	Student Demographic Cell (Race x Baseline Tercile)								
Race:		Black			Hispanic			White	
Baseline tercile:	Low	Median	High	Low	Median	High	Low	Median	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
γ_c	4.24	6.80	6.79	3.51	4.61	6.25	2.56	4.34	6.53
	(0.033)	(0.049)	(0.047)	(0.027)	(0.027)	(0.039)	(0.036)	(0.037)	(0.034)
β_c^{white}	1.71	-1.02	2.73	-0.26	0.46	2.25	-0.48	-2.02	-0.62
	(0.049)	(0.081)	(0.086)	(0.043)	(0.053)	(0.078)	(0.063)	(0.066)	(0.073)
$\beta_c^{peerquality}$	3.97	5.56	5.84	4.06	3.84	5.74	1.80	4.96	4.66
	(0.027)	(0.049)	(0.055)	(0.026)	(0.031)	(0.045)	(0.04)	(0.043)	(0.049)
μ_{c0}	-0.15	-0.06	-0.02	-0.09	-0.09	0.00	-0.07	0.09	0.10
	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.011)	(0.006)	(0.004)
μ_{c1}	-0.09	0.11	-0.03	0.06	-0.06	0.18	0.25	-0.01	0.17
	(0.012)	(0.017)	(0.016)	(0.016)	(0.015)	(0.016)	(0.024)	(0.018)	(0.015)
μ_{c2}	0.13	0.16	0.17	0.08	0.25	0.12	0.42	0.01	0.13
	(0.006)	(0.007)	(0.011)	(0.008)	(0.009)	(0.01)	(0.007)	(0.007)	(0.012)
σ_{c0}	2.56	3.64	2.17	2.86	2.33	2.14	3.15	3.69	2.94
	(0.054)	(0.057)	(0.046)	(0.05)	(0.035)	(0.045)	(0.094)	(0.058)	(0.041)
σ_{c1}	-0.31	0.15	-0.15	-0.14	-0.02	-0.17	0.12	0.47	0.17
	(0.055)	(0.069)	(0.067)	(0.037)	(0.025)	(0.036)	(0.06)	(0.09)	(0.076)
σ_{c2}	0.05	0.13	0.09	0.12	0.09	0.10	0.12	0.11	0.09
	(0.043)	(0.067)	(0.039)	(0.069)	(0.054)	(0.028)	(0.102)	(0.074)	(0.045)
$\tilde{\xi}_{cj}$									
mean	480	522	423	451	444	411	410	371	233
within-cell SD	16	18	20	13	14	17	16	18	23
Corr($\tilde{\xi}_{cj}$, VA)	0.015	0.240	0.377	0.034	0.304	0.417	0.367	0.485	0.600
p-value	[0.76]	[0]	[0]	[0.49]	[0]	[0]	[0]	[0]	[0]
Corr($\tilde{\xi}_{cj}$, Peer quality)	0.137	0.407	0.519	0.067	0.429	0.555	0.502	0.623	0.737
p-value	[0]	[0]	[0]	[0.17]	[0]	[0]	[0]	[0]	[0]
Corr($\tilde{\xi}_{jc}$, % white)	0.050	0.202	0.251	0.145	0.337	0.381	0.635	0.651	0.700
p-value	[0.31]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]
N school	423	423	422	423	423	422	423	423	420
N students	36,433	27,521	13,279	44,676	38,949	21,485	13,120	25,938	53,816

Notes: This table presents the model estimates by student demographic cells defined by the interaction of student race and baseline test score tercile. Asymptotic standard errors in parenthesis take into account the first stage sampling error and rely on numerical approximations when necessary. Square brackets report the p value of a test of the significance of the correlation coefficient in the row above.

Table A12: Model Fit

	All students		Minority students		Non-Minority students		Below median Math		Above median Math	
	Real	Simulated	Real	Simulated	Real	Simulated	Real	Simulated	Real	Simulated
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Average in top 3 choices										
Regents VA (σ)	0.66	0.67	0.54	0.52	0.88	0.93	0.43	0.43	0.88	0.89
Regents VA (percentile)	67.5	67.4	63.3	62.4	75.0	76.3	60.6	60.4	74.0	74.0
Regents VA (σ) left unexploited	1.74	1.74	1.97	1.98	1.35	1.29	1.96	1.96	1.54	1.53
SAT VA (σ)	0.89	0.89	0.60	0.59	1.40	1.40	0.42	0.46	1.33	1.29
SAT VA (percentile)	72.9	72.9	66.6	66.3	84.0	84.7	63.3	64.0	81.9	81.4
SAT VA (σ) left unexploited	1.55	1.55	1.78	1.78	1.13	1.12	1.70	1.66	1.40	1.44
Peer quality	0.26	0.25	0.12	0.11	0.52	0.50	0.03	0.05	0.48	0.44
White+Asian %	0.38	0.38	0.27	0.28	0.57	0.57	0.27	0.28	0.48	0.48
Commuting time	39.87	37.40	40.47	37.97	38.79	36.39	38.92	36.53	40.76	38.22
Panel B: 2016-2014 changes in application behavior										
P(applies to A) as 1st	-0.038	-0.060	-0.049	-0.066	-0.023	-0.054	-0.050	-0.056	-0.036	-0.072
P(applies to A) in top3	-0.028	-0.051	-0.036	-0.064	-0.014	-0.029	-0.039	-0.064	-0.023	-0.044
P(applies to A) ever	-0.007	-0.018	-0.010	-0.022	-0.002	-0.010	-0.010	-0.023	-0.007	-0.015
P(applies to C/D/F) as 1st	0.006	0.022	0.010	0.031	0.000	0.009	0.013	0.028	0.002	0.020
P(applies to C/D/F) in top3	0.023	0.048	0.037	0.065	0.003	0.024	0.041	0.058	0.014	0.048
P(applies to C/D/F) ever	0.064	0.078	0.086	0.098	0.033	0.053	0.087	0.081	0.055	0.090
Panel C: simulated school offers										
Regents VA (σ)	0.39	0.29	0.20	0.11	0.73	0.62	0.08	-0.01	0.68	0.56
Regents VA (percentile)	59.7	56.1	53.4	50.2	71.0	66.9	50.1	46.5	68.8	64.9
SAT VA (σ)	0.51	0.40	0.17	0.07	1.12	1.01	0.03	-0.09	0.96	0.85
SAT VA (percentile)	64.9	61.7	56.8	53.5	79.4	76.7	53.6	49.5	75.6	72.8
Peer quality	0.07	0.01	-0.09	-0.15	0.38	0.32	-0.18	-0.24	0.31	0.25
White+Asian %	0.31	0.27	0.20	0.16	0.52	0.47	0.20	0.15	0.42	0.38
Commuting time	38.00	33.70	38.90	33.50	36.38	34.08	37.68	32.29	38.30	35.00
Share matched	0.94	0.96	0.94	0.98	0.93	0.94	0.94	0.95	0.93	0.98
N	53,014		33,896		19,118		25,698		27,316	

Notes: This table assesses the model fit. It compares summary statistics of the characteristics of students' first three school choices (Panel A), school offers (Panel C) and changes in the probability of applying to high or low letter grade schools in the real data with those simulated using the model estimates (Panel B). The sample is the 2016 applicant cohort in panels A and C, and applicants in 2014 and 2016 for panel B. Simulations of the school match are based on priorities that are reconstructed on the basis of the admission rules used in the 2016 general education high school match and real school capacities.

Table A13: Robustness of Model Estimates - non-parametric prior mean and variance

	By Race			By 7th Grade Math Tercile		
	Black	Hispanic	White	Low	Median	High
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: second step - preferences						
γ_c	5.2	4.2	5.4	3.2	4.8	6.7
β_c^{white}	2.0	1.5	1.1	1.2	1.3	2.0
$\beta_c^{peerquality}$	4.4	4.4	3.7	3.7	4.1	4.7
$\tilde{\xi}_{cj}$ SD	17.1	14.2	20.3	14.3	16.4	20.8
$\tilde{\xi}_{cj}$ range	104.9	86.9	107.2	92.4	97.7	106.8
$\tilde{\xi}_{cj}$ skewness	0.06	0.23	0.71	0.15	0.31	0.58
Corr($\tilde{\xi}_{cj}$, VA)	0.16	0.21	0.54	0.07	0.34	0.52
Corr($\tilde{\xi}_{cj}$, Peer quality)	0.30	0.30	0.67	0.15	0.48	0.66
Corr($\tilde{\xi}_{cj}$, % white)	0.14	0.26	0.68	0.18	0.38	0.56
Panel B: second step - beliefs						
μ_{cj} below med. R_j, below med. Q_j	-0.06	-0.02	-0.05	0.03	-0.07	-0.09
μ_{cj} above med. R_j, below med. Q_j	0.03	0.09	0.04	0.02	-0.09	0.23
μ_{cj} below med. R_j, above med. Q_j	-0.10	-0.02	0.41	0.04	0.21	0.05
μ_{cj} above med. R_j, above med. Q_j	0.18	0.29	0.28	0.05	0.33	0.40
σ_{cj}^{-1} below med. R_j, below med. Q_j	1.84	1.02	1.65	1.15	1.48	1.77
σ_{cj}^{-1} above med. R_j, below med. Q_j	1.44	1.11	2.15	1.00	1.58	2.12
σ_{cj}^{-1} below med. R_j, above med. Q_j	3.25	1.64	4.26	0.97	3.04	5.04
σ_{cj}^{-1} above med. R_j, above med. Q_j	3.47	1.74	4.61	1.03	3.14	5.55

Notes: This table summarizes the second step model estimates when prior means and precision are a non parametric function of four discrete school types which combine whether a school has above or below median achievement levels, and above or below median quality. Panel A reports estimates of the preference parameters $\gamma_c, \beta_c, \xi_{cj}$ and panel B of the first and second moments of priors for each school type, taking a weighted average of cell-specific estimates across cells sharing the same covariate (race or baseline test score), using weights proportional to cell size.

Table A14: Robustness of Model Estimates to Different Functional Forms

	By Race						By 7th Grade Math Tercile					
	Location and scale shift			Log Normal			Location and scale shift			Log Normal		
	Black	Hispanic	White	Black	Hispanic	White	Low	Median	High	Low	Median	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: second step - preferences												
γ_c	5.0	5.0	4.9	4.7	4.6	4.7	4.9	5.0	5.1	4.5	4.7	4.8
β_c^{white}	1.1	1.0	1.0	1.1	1.0	1.0	1.0	1.0	1.1	1.0	1.0	1.0
$\beta_c^{peerquality}$	2.3	2.2	2.1	2.1	2.1	2.0	2.1	2.1	2.3	2.1	2.1	2.1
$\tilde{\xi}_{cj}$ SD	18.1	15.1	21.7	18.1	15.1	21.5	14.5	17.4	22.9	14.3	17.3	23.0
$\tilde{\xi}_{cj}$ range	111.6	90.5	110.5	112.6	89.9	108.5	93.9	101.7	114.5	92.9	101.1	114.3
$\tilde{\xi}_{cj}$ skewness	0.2	0.3	0.7	0.2	0.3	0.7	0.3	0.4	0.7	0.3	0.3	0.7
Corr($\tilde{\xi}_{cj}$, VA)	0.12	0.17	0.33	0.13	0.17	0.33	0.07	0.22	0.34	0.07	0.23	0.35
Corr($\tilde{\xi}_{cj}$, Peer quality)	0.41	0.41	0.70	0.43	0.42	0.70	0.25	0.56	0.73	0.26	0.57	0.74
Corr($\tilde{\xi}_{cj}$, % white)	0.25	0.36	0.69	0.26	0.36	0.69	0.26	0.45	0.62	0.26	0.46	0.63
Panel B: second step - beliefs												
μ_{cL}	-0.10	-0.03	-0.02	-0.46	-0.47	-0.48	0.04	-0.06	-0.12	-0.44	-0.45	-0.53
μ_{cH}	0.24	0.15	0.21	-0.41	-0.38	-0.34	0.08	0.20	0.29	-0.44	-0.37	-0.31
σ_{cL}^{-1}	1.63	1.92	1.52	1.67	1.73	1.31	1.96	1.66	1.47	1.98	1.54	1.16
σ_{cH}^{-1}	1.74	1.81	1.74	1.69	1.78	1.45	1.91	1.69	1.71	2.00	1.63	1.28

Notes: This table summarizes the second step model estimates for models using alternative functional forms for prior distributions. “Location and scale shift” refers to a model in which priors have the same distribution as quality in the city up to a location and a scale shift that varies across school types. “Log Normal” refers to a model in which priors have a log normal distribution. Parameters of the log normal distribution are re-scaled to make them comparable to other estimates, so that value-added is always rescaled to have mean zero and standard deviation 1 across schools.

Table A15: Robustness of Model Estimates - Strategic Reporting

	By race			By 7th grade Math tercile		
	Black	Hispanic	White	Low	Median	High
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: first step						
δ_{cjt} SD	53	31	33	37	38	39
δ_{cjt} range	302	182	169	217	211	206
Corr(δ_{cjt}, VA)	0.29	0.41	0.59	0.35	0.42	0.55
Corr(δ_{cjt}, Peer quality)	0.41	0.55	0.74	0.48	0.56	0.69
Corr(δ_{cjt}, % white)	0.28	0.44	0.68	0.38	0.46	0.59
Panel B: second step - preferences						
γ_c	5.6	4.3	5.9	2.7	5.7	7.4
β_c^{white}	3.2	3.7	2.5	2.5	4.3	2.7
$\beta_c^{peerquality}$	7.3	4.6	4.3	3.7	4.8	7.4
$\tilde{\xi}_{cj}$ SD	26	19	25	22	22	25
$\tilde{\xi}_{cj}$ range	159	118	122	134	127	131
$\tilde{\xi}_{cj}$ skewness	0.27	0.24	0.54	0.47	0.13	0.45
Corr($\tilde{\xi}_{cj}$, VA)	0.25	0.32	0.54	0.33	0.32	0.47
Corr($\tilde{\xi}_{cj}$, Peer quality)	0.41	0.45	0.69	0.47	0.47	0.62
Corr($\tilde{\xi}_{jc}$, % white)	0.21	0.30	0.64	0.36	0.31	0.52
Panel C: second step - beliefs						
μ_{cL}	-0.56	-0.01	-0.10	-0.03	-0.25	-0.31
μ_{cH}	0.64	0.95	0.41	1.20	0.36	0.46
σ_{cL}^{-1}	3.45	1.70	2.17	2.78	2.25	2.00
σ_{cH}^{-1}	2.45	1.49	2.52	1.08	2.80	2.47
Absolute Bias	0.41	0.41	0.37	0.42	0.31	0.46

Notes: This table summarizes the model estimates when students are allowed to report preferences strategically. Specifically, students are assumed to only consider schools where they have a non-zero probability in admission and to rank schools truthfully within this set. Panel A reports summary statistics for the estimates of the mean school utility δ_{cjt} obtained in the first step. Panel B reports the second step estimates of the preference parameters $\gamma_c, \beta_c, \xi_{cj}$ and panel C of the prior moments $\boldsymbol{\mu_c}, \boldsymbol{\sigma_c^{-1}}$ taking a weighted average of cell-specific estimates across cells sharing the same covariate (race or baseline test score), using weights proportional to cell size.

Table A16: Full-Information Benchmark

	All students		Black + Hispanic		White+Asian		Below median Math		Above median Math	
	No info	Full info	No info	Full info	No info	Full info	No info	Full info	No info	Full info
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Top 3 choices										
ΔW		0.069		0.081		0.047		0.064		0.073
VA - pct	68	74	62	71	76	81	60	68	74	81
Peer math - pct	76	79	70	74	88	89	67	71	85	87
White+Asian %	0.383	0.403	0.276	0.298	0.572	0.588	0.278	0.297	0.481	0.502
Panel B: Offers										
$\Delta W\|$ offered		0.011		0.013		0.008		0.010		0.013
ΔW		0.009		0.012		0.005		0.008		0.011
ΔW as % of first best		24%		30%		12%		19%		28%
VA - pct	55	57	50	52	65	66	46	47	64	66
Peer math - pct	63	63	55	55	77	77	49	49	75	75
White+Asian %	0.263	0.265	0.159	0.161	0.449	0.449	0.146	0.147	0.374	0.375
Offered	0.964	0.961	0.975	0.971	0.944	0.943	0.952	0.949	0.975	0.973
N	52,997		33,901		19,096		25,706		27,291	

Notes: This table compares summary statistics of the characteristics of students' first three school choices and school offers in the simulated status-quo ("No info") and in the full-information benchmark ("Full info"). In the status quo students receive no additional information about school quality from the policy maker, and form beliefs about quality only based on their priors. In the full information counterfactual, students are perfectly informed about the VA of each school. Welfare is measured by the student average Regents Math test scores. The first two columns report averages for the entire set of applicants, while the remaining columns split applicants by race or by baseline achievement (above and below the median 7th grade math test score).

Table A17: Targeted Outreach

	Targeted students		Non- targeted students	
	Outreach	Full info	Outreach	Full info
	(1)	(2)	(3)	(4)
Panel A: Targeted = students from lowest performing middle schools				
ΔW - choices	0.074	0.074	0.000	0.065
ΔW - offers	0.033	0.011	-0.008	0.008
ΔW - offers under no screening	0.037	0.016	-0.012	0.005
% B+H in top 20% schools	0.027	0.026	0.027	0.026
N	17197		35800	
Panel B: Targeted = top performing Black and Hispanic students				
ΔW - choices	0.105	0.105	0.000	0.064
ΔW - offers	0.053	0.018	-0.006	0.008
ΔW - offers under no screening	0.054	0.012	-0.006	0.008
% B+H in top 20% schools	0.023	0.026	0.023	0.026
N	6089		46908	

Notes: This table compares changes in average value-added (ΔW) of students' first three school choices and school offers in the targeted outreach counterfactual ("Outreach") and in the full-information benchmark ("Full info"), relative to the status-quo. In the outreach counterfactuals only students denoted with "targeted" receive perfect information about school quality, while everyone is perfectly informed in the full information benchmark. Panel A considers an outreach intervention that provides information only to students in the bottom half of middle school performance, while panel B one providing information to Black and Hispanic students with test scores in the top tercile of the 7th grade test score distribution. The last row of each panel also reports the change in the share of Black and Hispanic students receiving offers to the best 20% of schools with respect to the status quo.

Table A18: Best and Worst 5 Letter Rules

	Cutoff percentile				Offered VA change				
	D	C	B	A	All	Low achiev.	High achiev.	Black+Hispanic	White+Asian
					(1)	(2)	(3)	(4)	(5)
Full information					23.4%	15.9%	30.6%	30.0%	11.8%
Naïve	20^{th}	40^{th}	60^{th}	80^{th}	17.2%	11.4%	22.6%	25.1%	3.3%
Best on average	10^{th}	30^{th}	70^{th}	90^{th}	20.1%	13.5%	26.2%	28.4%	5.0%
Worst on average	70^{th}	80^{th}	90^{th}	95^{th}	11.3%	2.2%	20.0%	15.3%	4.5%
Best for low achieving	5^{th}	10^{th}	30^{th}	70^{th}	18.5%	17.0%	20.0%	27.6%	2.0%
Worst for low achieving	70^{th}	80^{th}	90^{th}	95^{th}	11.3%	2.2%	20.0%	15.3%	4.5%
Best for high achieving	10^{th}	40^{th}	70^{th}	90^{th}	20.1%	13.2%	26.4%	28.4%	5.0%
Worst for high achieving	5^{th}	10^{th}	20^{th}	30^{th}	14.2%	15.2%	13.3%	21.0%	2.2%
Best for Black and Hispanic	10^{th}	35^{th}	70^{th}	95^{th}	20.1%	16.1%	23.9%	29.4%	3.5%
Worst for Black and Hispanic	70^{th}	80^{th}	90^{th}	95^{th}	11.3%	2.2%	20.0%	15.3%	4.5%
Best for white and Asian	5^{th}	30^{th}	90^{th}	95^{th}	17.2%	11.4%	22.8%	22.0%	8.9%
Worst for white and Asian	30^{th}	40^{th}	50^{th}	60^{th}	14.2%	11.7%	16.7%	21.5%	1.2%

Notes: This table compares changes in welfare relative to the status-quo in the full-information benchmark (top row) with those induced by information disclosure policies that rate school quality with five letters, from A to F, varying the position of the cutoffs along the quality distribution. Welfare gains are expressed as a percentage of the average first-best achievement gains. The naïve intervention places the cutoffs evenly apart, while the other rating policies reported are those that maximize or minimize the test scores of a given subgroup of students. The unnumbered columns describe the position of the letter cutoffs in terms of value added percentile ranking. The remaining columns report changes in test scores by student subgroup.

B Robustness to using alternative measures of value added

Here I consider alternative ways of measuring school value-added than those used in the main analysis and provide evidence that it makes a little difference for the results of this paper.

B.1 Robustness to using race-specific measures of value-added

I relax the assumption that school value-added is constant across students embedded in the model in equation (1) and allow school effectiveness to vary by student race as captured by OLS estimates of α_{jr} in the following regression:

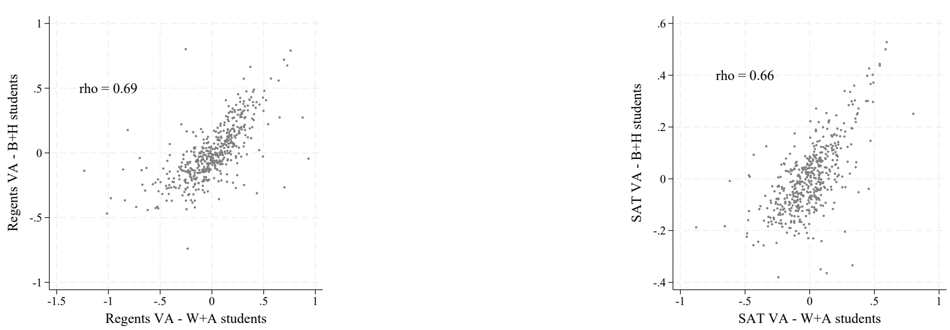
$$Y_i = \alpha_0 + \sum_{j=1}^{J} \alpha_{jr(i)} D_{ij} + X_i' \Gamma_{t(i)} + \epsilon_i \tag{10}$$

where $r(i)$ indicates student i's binary race (pooling Black and Hispanic student into the "minority" category and white and Asian into the "white" category). I provide evidence that the constant-effect model used in the main analysis is already a good approximation of reality because value-added changes little across student race, therefore using a race-specific measure of value-added would not change the main results.

Figure B1 shows that race-specific measures of VA are highly correlated (correlation coefficient of about 0.7) for both Regents and SAT test scores. Table B1, instead, compares lottery-based tests of bias for these two VA models. The idea behind these tests is to use random variation in school offers embedded in the centralized school match to test whether the VA estimates in (1) predict student outcomes (Angrist et al., 2016, 2021, 2022).

Figure B1: Correlation of race-specific measures of school quality

(a) Regents VA (b) SAT VA



Notes: This figure shows that race-specific estimates of school value-added are strongly correlated within schools by presenting scatter plots of value-added estimates for white and Asian students (x-axis) against estimates of value-added for Black and Hispanic students (y-axis) for both SAT and Regents test scores.

The forecast test captures the extent to which the estimated value added α_j predicts causal school effectiveness on average. In practice it is conducted by instrumenting estimates of the value added $\alpha_{d(i)} = \sum_j \alpha_j D_{ij}$ of the school where i enrolls (D_{ij} denotes enrollment indicators), with random school lottery offers. That is, it test the null hypothesis that the IV estimate $\hat{\psi}$ of the following second stage equation is equal to 1

$$Y_i = \tau_0 + \psi \alpha_{d(i)} + X_i' \tau + \nu_i \tag{11}$$

, meaning that a one-unit increase in α_j translates into a one-unit increase in Y_i. The second stage parameter ψ is often referred to as a forecast coefficient and deviations from the null that $\psi = 1$ are called forecast bias. The omnibus test provides another way of testing the CIA by testing that the regression residuals ϵ_i are unrelated to any randomness in school offers. This is a joint test of l orthogonality restrictions $E[(Z_{il} - p_{il})\epsilon_i] = 0$, one for each of L school lotteries available to the econometrician. Z_{il} indicate offers in lottery l, while p_{il} is an assignment propensity score measuring student i's probability of receiving an offer in lottery l. In practice, these restrictions are tested by asking whether $\tau_1 = ... = \tau_L = 0$ in the residual regression equation:

$$\hat{\epsilon}_i = \tau_0 + \sum_{l=1}^{L} \tau_l Z_{il} + \sum_{l=1}^{L} \mu_l p_{il} + X_i' \Delta + \nu_i \tag{12}$$

. Angrist et al. (2016) show how this test can be decomposed into two separate test statistics. The first is equivalent to the one used in the forecast test, while the second is the Sargan LM statistic for a test of 2SLS overidentifying restrictions, which checks whether VAM estimates are equally predictive within every lottery. In practice, in all the tests reported in this appendix, schools are classified into 10 bins defined by deciles of the distribution of the estimated conventional value-added in equation (1). The testing equation (12) is estimated using bin-level (rather than single-school) offers and propensity scores. Propensity scores at the school level are computed using the method derived in (Abdulkadiroğlu et al., 2022) and then aggregated at the bin level taking a sum over the propensity scores of schools in the bin. Offers are random conditional on propensity scores and running variable controls defined and constructed in Abdulkadiroğlu et al. (2022).

Table B1: Pooled and Race-Specific VAM Bias Tests for Regents and SAT scores

	Pooled VAM			Race-specific VAM		
	All	Black+Hispanic	White+Asian	All	Black+Hispanic	White+Asian
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Regents math VA						
Forecast coefficient	0.966	1.02	0.837	0.968	0.983	0.970
	(0.032)	(0.037)	(0.072)	(0.032)	(0.036)	(0.082)
First stage F statistic	1771	1369	309	1620	1374	223
Bias tests						
Forecast	1.11	0.217	5.19	0.968	0.229	0.136
	[0.292]	[0.641]	[0.023]	[0.325]	[0.632]	[0.712]
Overidentification (9 d.f.)	12.1	13.1	5.32	12.0	13.2	4.90
	[0.208]	[0.158]	[0.805]	[0.213]	[0.154]	[0.843]
Omnibus (10 d.f.)	13.2	13.3	10.5	13.0	13.4	5.03
	[0.213]	[0.207]	[0.396]	[0.225]	[0.201]	[0.889]
N (testing)	49322	35739	13583	49322	35739	13583
N (estimation)	179978	130281	49697	179978	130281	49697
Panel B: SAT math VA						
Forecast coefficient	0.756	0.736	0.636	0.794	0.821	0.564
	(0.050)	(0.055)	(0.131)	(0.052)	(0.061)	(0.113)
First stage F statistic	1203	1267	121	901	1020	117
Bias tests						
Forecast	23.8	23.3	7.78	15.5	8.63	14.8
	[0.000]	[0.000]	[0.005]	[0.000]	[0.003]	[0.000]
Overidentification (9 d.f.)	6.28	9.83	10.4	6.58	8.70	9.44
	[0.712]	[0.364]	[0.317]	[0.681]	[0.466]	[0.398]
Omnibus (10 d.f.)	30.1	33.1	18.2	22.1	17.3	24.2
	[0.001]	[0.000]	[0.052]	[0.015]	[0.067]	[0.007]
N (testing)	46679	34693	11986	46679	34693	11986
N (estimation)	179978	130281	49697	179978	130281	49697

Notes: This table reports tests for bias in OLS value-added models (VAMs). The pooled VAM uses all students to estimate school value-added as measured by the coefficient α_j in equation (1), regardless of their race or ethnicity. The Race-specific VAM instead estimates school value-added on separate sub-samples of students, dividing students according to their race or ethnicity. Both VAMs control for cubic functions of baseline math and ELA scores and indicators for sex, race, subsidized lunch, special education, limited English proficiency, each interacted with application year. Forecast coefficients are from instrumental variables regressions of test scores on VAM fitted values, instrumenting fitted values with binned assignment indicators. Assignments are binned by decile of the estimated conventional VAM. IV models control for propensity scores, running variable controls, and baseline demographics and achievement. Test scores for outcomes and VAMs are standardized to be mean zero and standard deviation one in the student-level test score distribution, separately by year. The forecast bias test checks whether the forecast coefficient equals 1; the overidentification test checks overidentifying restrictions implicit in the procedure used to estimate the forecast coefficient. The omnibus test combines tests for forecast bias and overidentification. Standard errors are reported in parentheses; test p-values are reported in brackets. Different columns use different samples of students for testing: columns (1) and (4) pool all students together, while the remaining columns split students by race.

The lottery-based tests of bias show that measures of VA that do not vary by race ("Pooled VA") have a good predictive validity for student Regents scores of both races. Black and Hispanic students Regents test scores are equally well predicted by the pooled and by the race-specific measures of VA. White and Asian student outcomes are instead better predicted by the race-specific VA (the forecast bias test of the pooled VA rejects the null, unlike the one of the race-specific VA) although the forecast coefficient of the pooled VA is relatively high even for this student subgroup.Moreover, overidentification test results of race-specific VA are similar to those for the pooled VA, which further supports the existence of little heterogeneity in school effects across student races (Angrist et al., 2017). As noted above, SAT OLS VA is instead more biased. Forecast bias tests always reject the null, but this is true regardless of whether VA is estimated by race or on the pooled sample, suggesting that bias in not related to heterogeneity in treatment effects by race.

Finally, I directly show that Black and Hispanic students choose worst schools even when considering measures of race-specific value added, indicating that the reason behind cross-race gaps in choices is not that students are choosing schools that are best for their own demographic group while constant VAM models are failing to capture race-specific school match effects As shown by comparisons of cross-race gaps in the table below with those reported in the main text, if anything, cross-race gaps are larger when considering measures of value-added that vary with student race.

Table B2: Gap in Choice of School Quality - Robustness to Using Race-Specific Value-Added

		Race gap							
	N	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable: school race-specific value-added (test score σ)									
Regents VA in top 3 choices	734,854	-0.11***	-0.10***	-0.07***	-0.09***	-0.08***	-0.07***	-0.05***	-0.06***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
SAT VA in top 3 choices	734,853	-0.15***	-0.14***	-0.11***	-0.11***	-0.11***	-0.11***	-0.09***	-0.09***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
borough FE			X				X		
zipcode FE				X				X	
test score controls						X	X	X	X
mean and max in choice-set					X				X

Notes: This table consider whether estimates in table 3 are robust to using race-specific estimates of VA. It reports race differences in the quality of school choices as estimated by the coefficient β in equation (2), using measures of VA that vary by applicant race. The regressions in the first column correspond to raw race gaps, while columns (2)-(8) progressively add controls for residential location, test scores and quality available in the students' feasible set.

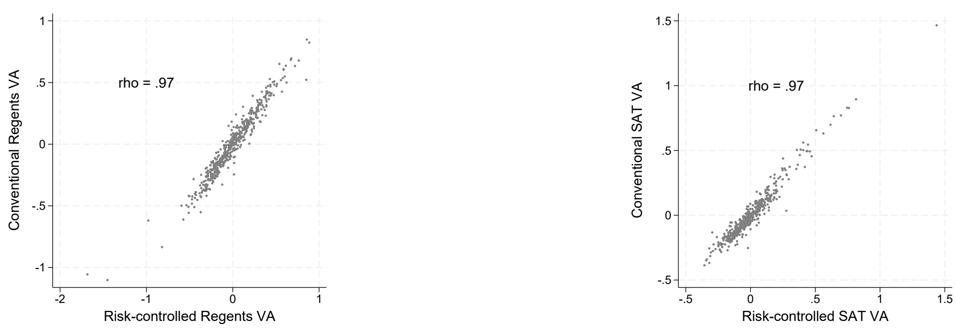
B.2 Robustness to using risk-controlled value added

Next, I relax the CIA by estimating risk-controlled (RC) VAM, as introduced by Angrist et al. (2021). RC VAM supplements the vector of controls with applicant characteristics integral to school matching, such as where they apply and the priority status that a school assigns them. This requires restricting the sample to the subset of applicants cohorts for which I have the necessary information to replicate the high school match, that is, starting from 2012 applicants. Because student typically take these SAT tests and Regents exams after their sophomore year, these measures rely on tests taken between 2014 and 2019. As a consequence, RC VAM estimates are not available for a subset of schools in my sample that were phased out before these dates, and rely on a much shorter time span. For these reasons, in the main analysis I rely on OLS VAM estimates of school quality and I provide evidence that conventional and risk-controlled VAM measures in this setting are largely equivalent.

Figure B2 shows that conventional and risk-controlled measures of value added (VA) are highly correlated, with correlation coefficients very close to 1. Table B3 compares lottery-based tests of bias for these two VA models, which confirm that the predictive validity of conventional estimates of Regents VA is incredibly similar to that of risk-controlled measures, and that both are very good. Adding risk-related controls, however, substantially improves the predictive validity of conventional SAT VA measures, that are otherwise substantially more biased. For this reason, I use Regents VA as the primary measure of school quality unless otherwise noted.

Figure B2: Correlation of conventional and risk-controlled measures of school quality

(a) Regents VA (b) SAT VA



Notes: This figure shows that conventional and risk-controlled estimates of school value-added are strongly correlated within schools by presenting scatter plots of risk controlled value-added estimates (x-axis) against conventional estimates of value-added (y-axis) for both SAT and Regents test scores.

Table B3: Conventional and Risk-Controlled VAM Bias Tests for Regents and SAT scores

	Coventional VAM			Risk-controlled VAM		
	All	Black+Hispanic	White+Asian	All	Black+Hispanic	White+Asian
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Regents math VA					
Forecast coefficient	0.966	1.02	0.837	0.927	0.986	0.817
	(0.032)	(0.037)	(0.072)	(0.031)	(0.037)	(0.067)
First stage F statistic	1771	1369	309	1839	1391	351
Bias tests						
Forecast	1.11	0.217	5.19	5.53	0.154	7.48
	[0.292]	[0.641]	[0.023]	[0.019]	[0.694]	[0.006]
Overidentification (9 d.f.)	12.1	13.1	5.32	9.78	5.84	7.75
	[0.208]	[0.158]	[0.805]	[0.369]	[0.756]	[0.560]
Omnibus (10 d.f.)	13.2	13.3	10.5	15.3	5.99	15.2
	[0.213]	[0.207]	[0.396]	[0.121]	[0.816]	[0.124]
N (testing)	49322	35739	13583	49291	35595	13696
N (estimation)	179978	130281	49697	179978	130281	49697
	Panel B: SAT math VA					
Forecast coefficient	0.756	0.736	0.636	0.960	0.925	0.939
	(0.050)	(0.055)	(0.131)	(0.061)	(0.066)	(0.169)
First stage F statistic	1203	1267	121	1007	1069	90.6
Bias tests						
Forecast	23.8	23.3	7.78	0.430	1.30	0.130
	[0.000]	[0.000]	[0.005]	[0.512]	[0.255]	[0.718]
Overidentification (9 d.f.)	6.28	9.83	10.4	9.68	9.82	7.35
	[0.712]	[0.364]	[0.317]	[0.377]	[0.365]	[0.601]
Omnibus (10 d.f.)	30.1	33.1	18.2	10.1	11.1	7.48
	[0.001]	[0.000]	[0.052]	[0.431]	[0.348]	[0.680]
N (testing)	46679	34693	11986	47008	34763	12245
N (estimation)	179978	130281	49697	179978	130281	49697

Notes: This table reports tests for bias in OLS value-added models (VAMs). The conventional VAM controls for cubic functions of baseline math and ELA scores and indicators for sex, race, subsidized lunch, special education, limited English proficiency, each interacted with application year. Risk-only VAM adds propensity score and running variable controls to the uncontrolled specification. RC VAM adds propensity score and running variable controls to the controls in the conventional VAM. Forecast coefficients are from instrumental variables regressions of test scores on VAM fitted values, instrumenting fitted values with binned assignment indicators. Assignments are binned by decile of the estimated conventional VAM. IV models control for propensity scores, running variable controls, and baseline demographics and achievement. Test scores for outcomes and VAMs are standardized to be mean zero and standard deviation one in the student-level test score distribution, separately by year. The forecast bias test checks whether the forecast coefficient equals 1; the overidentification test checks overidentifying restrictions implicit in the procedure used to estimate the forecast coefficient. The omnibus test combines tests for forecast bias and overidentification. Standard errors are reported in parentheses; test p-values are reported in brackets. Different columns use different samples of students for testing: columns (1) and (4) pool all students together, while the remaining columns split students by race.

C Model and counterfactuals appendix

C.1 Model identification

Separating preferences for quality from priors This proof is analogous to the argument used in the proof of proposition 1 in Vatter (2022), modified to the case in which quality is scalar and identification comes from changes in letter grades or their absence within schools. In what follows, for simplicity, I focus on variation within a school over time and thus I drop the school subscript j to write: $\delta_t = X_t'\beta + \gamma E[q|s_t = r] + \xi$. The argument developed here can be directly applied to the demand of different demographic cells and for schools (or school types) that receive at least 3 different quality ratings, or 2 ratings and no rating. For simplicity, I also assume no variation over time in X_t to focus only on identification of beliefs from preferences for quality but the argument is easily extended to consider preferences for other time-varying school attributes X_t as long as these are not perfectly collinear with letter grades. Throughout, I assume the identification of $\delta_t = \gamma\,\mathbb{E}[q|s_t = r] + \xi$ up to a constant. In this simplifying case in which I have dropped the dependence on X_t, these are effectively only letter grade fixed effects δ_r (including the case of lack of grades) for each school type $h(j)$.

Lemma 1. *Let f and g be two distinct, strictly positive. densities, supported over $Q = [\underline{q}, \bar{q}]$. Then there exists $\underline{x} < \tilde{x} < \bar{x}$ s.t. $\mathbb{E}_f[x|x \in (\underline{x}, \tilde{x})] \leq \mathbb{E}_g[x|x \in (\underline{x}, \tilde{x})]$ and $\mathbb{E}_f[x|x \in (\tilde{x}, \bar{x})] \geq \mathbb{E}_g[x|x \in (\tilde{x}, \bar{x})]$ with one of the inequalities strict.*

Also, there exists $x' < x'' \in [\underline{q}, \bar{q}]$ such that $\mathbb{E}_f[x|x \in (x', x'')] = \mathbb{E}_g[x|x \in (x', x'')]$.

Proof. Because f and g are distinct and continuous over a common support, they must cross at an interior point $\tilde{x} \in (\underline{\mathrm{q}}, \bar{q})$. By continuity, $\exists\epsilon > 0$ such that $f(x) > g(x)\forall x \in (\tilde{x}, \tilde{x}+\epsilon)$ and $f(x) \leq g(x)\forall x \in (\tilde{x} - \epsilon, \tilde{x})$ where there role of f and g is wlog. Define $h_f(x, \epsilon) = \frac{f(x)}{F(\tilde{x}+\epsilon)-F(\tilde{x})}$ and analogously for g, where F and G denote the cumulative distribution functions of f ad g respectively. Both $h_f(\cdot, \epsilon)$ and $h_g(\cdot, \epsilon)$ are continuous and integrate to 1 over $(0, \epsilon)$ and therefore intersect at an interior point in $(0, \epsilon)$. Moreover, note that $\forall\tilde{\epsilon} \in (0, \epsilon)$, $h_f(\cdot, \tilde{\epsilon}) < h_g(\cdot, \tilde{\epsilon})$. Pick $\tilde{\epsilon} \in (0, \epsilon)$ such that $h_f(\cdot, \tilde{\epsilon})$ and $h_g(\cdot, \tilde{\epsilon})$ intersect only once at a point $\hat{x}$ and denote $\bar{x} = \tilde{x} + \tilde{\epsilon}$. Then we have that $E_f[x|x \in (\tilde{x}, \bar{x})] - E_g[x|x \in (\tilde{x}, \bar{x})]$:

$$\begin{aligned}\int_{\tilde{x}}^{\bar{x}} (h_f(q, \tilde{\epsilon}) - h_g(q, \tilde{\epsilon})qdq) &= \int_{\tilde{x}}^{\hat{x}} \underbrace{(h_f(q, \tilde{\epsilon}) - h_g(q, \tilde{\epsilon})}_{<0} qdq) + \int_{\hat{x}}^{\bar{x}} \underbrace{(h_f(q, \tilde{\epsilon}) - h_g(q, \tilde{\epsilon})}_{>0} qdq) \\ &> \hat{x}\Big[\int_{\tilde{x}}^{\hat{x}} (h_f(q, \tilde{\epsilon}) - h_g(q, \tilde{\epsilon})dq) + \int_{\hat{x}}^{\bar{x}} (h_f(q, \tilde{\epsilon}) - h_g(q, \tilde{\epsilon})dq)\Big] = 0\end{aligned}$$

which proves the first inequality. The proof for the second part of the statement is analogously applied to $(\tilde{x} - \epsilon, \tilde{x})$. The proof of the equality follows from defining $w(\lambda) = \mathbb{E}_f[x|x \in (\underline{\mathrm{x}}, \tilde{x} + \lambda\tilde{\epsilon})] - \mathbb{E}_g[x|x \in (\underline{\mathrm{x}}, \tilde{x} + \lambda\tilde{\epsilon})$. Because $w(0) \leq 0$, $w(1) > 0$ and $w(\cdot)$ is a continuous function, by the intermediate value theorem, $\exists\lambda^*$ such that $w(\lambda^*) = 0$ and thus the interval cutoffs for the second part are $x' = \underline{\mathrm{x}}$ and $x'' = \tilde{x} + \lambda^*\tilde{\epsilon}$. □

Assumption 1 - (Variation in signals) Let r_t be the quality rating received by the school in year t, which is defined by a compact interval of possible quality values $[\underline{\mathrm{c}}_t^r, \bar{c}_t^r]$. In the absence of letter grades, this is degenerate and coincides with the entire quality space $Q = [\underline{q}, \bar{q}]$. The possible quality partitions r_t are drawn from a distribution over intervals of quality and at least $N \geq 3$ are observed in the data.

Assumption 2 - (Knowledge of the rating design) Consumers know the quality cutoffs $\underline{c}_t^r, \bar{c}_t^r$ that define the quality rating s_t. Because s_t may also include a degenerate signal, i.e. its absence, this means they also know the boundaries of the quality space Q. Moreover, they use the ratings and Bayes' rule to update a continuous prior density $f : Q \to R+$.

Proposition 1. *Let assumptions 3 and 4 hold and let student preferences over quality be linear, i.e., $v(q) = \gamma q$. Then $(\gamma, f(\cdot))$ are identified.*

Proof. By contradicion, suppose there exist two distinct elements in the identified set $(\gamma_0, f_0), (\gamma_1, f_1)$. By the above lemma and assumption 2, there exists three possible quality ratings r, r' and r'' which may be drawn with positive probability, such that: $\mathbb{E}_{f_0}[q|r] = \mathbb{E}_{f_1}[q|r]$, $\mathbb{E}_{f_0}[q|r'] < \mathbb{E}_{f_1}[q|r']$ and $\mathbb{E}_{f_0}[q|r''] \geq \mathbb{E}_{f_1}[q|r'']$. Therefore, we have that:

$$\gamma_0(\mathbb{E}_{f_0}[q|r'] - \mathbb{E}_{f_0}[q|r]) = \delta_{r'} - \delta_r = \gamma_1(\mathbb{E}_{f_1}[q|r'] - \mathbb{E}_{f_1}[q|r]) \implies \gamma_0 > \gamma_1$$

$$\gamma_0(\mathbb{E}_{f_0}[q|r''] - \mathbb{E}_{f_0}[q|r]) = \delta_{r''} - \delta_r = \gamma_1(\mathbb{E}_{f_1}[q|r''] - \mathbb{E}_{f_1}[q|r]) \implies \gamma_0 \leq \gamma_1$$

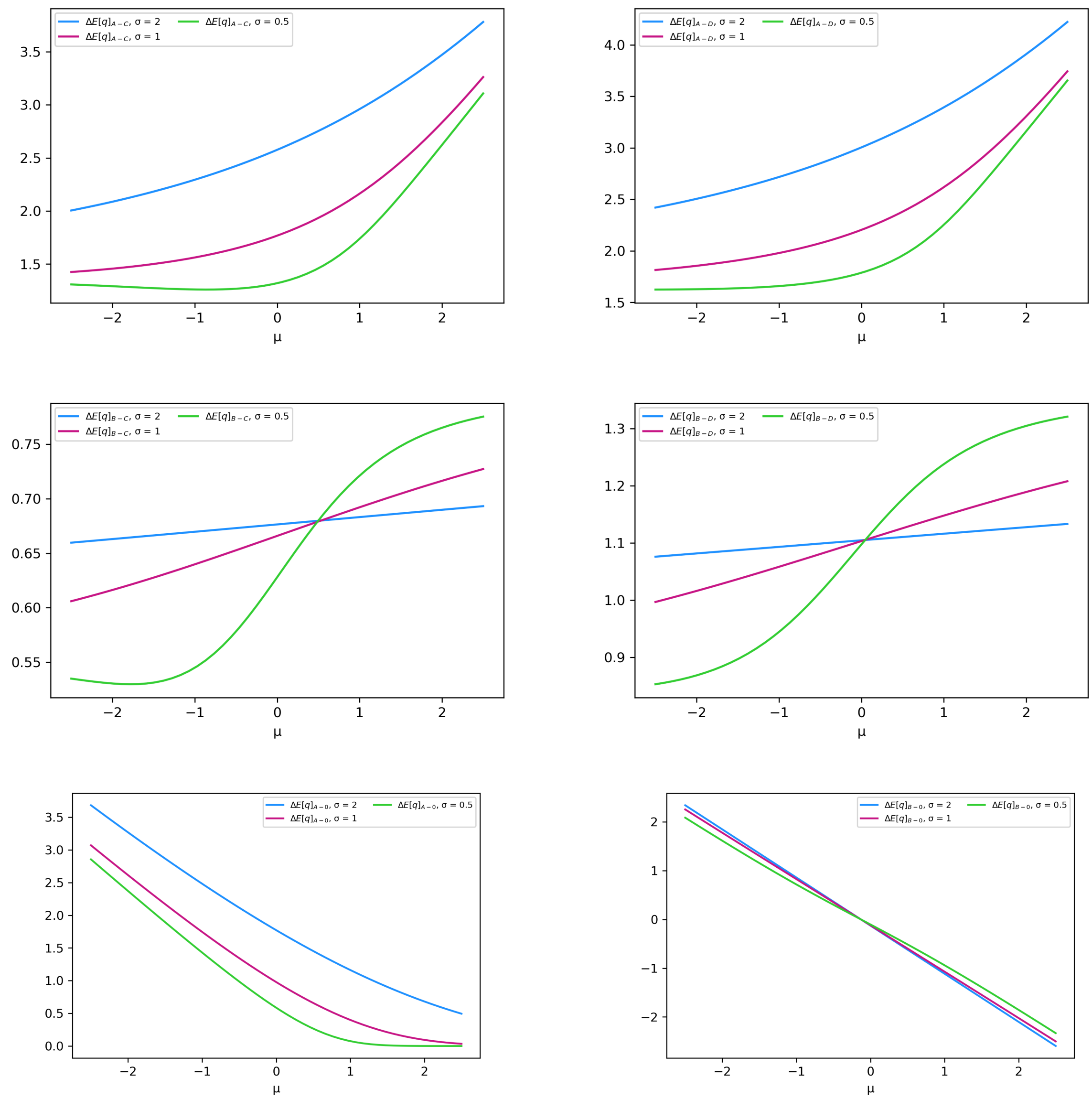
, a contradiction.

□

In theory, point identification without a functional form restriction for priors requires a lot of variation in partitioning cutoffs of ratings r_t over the years. In practice, in my case, quantiles used to assign letter grades vary little from year to year, but I still have variation from the presence or absence of letter grades, which is equivalent to a degenerate partitional design. Moreover, the fact that the same school may receive different grades in different years and the normal functional form assumption together allow to identify the model parameters.

To build intuition for how this argument is used in practice given the observed letter grades cutoffs, figure C1 below plots the change in expected quality for pairs of different letter grades as a function of the prior parameters μ and σ. While the positive updating in beliefs between receiving a D and and A is always larger the larger the variance of the prior, σ, this is no longer the case when looking at the belief updating between receiving a D and a B. This variation is what allows to separate the prior precision from preferences for quality. Changes in demand between years with and without letter grades help pin down the prior mean μ , as they vary most strongly in this parameter as compared to the precision of the prior. Intuitively, if following the removal of positive (negative) signals the fall (increase) in utility is larger, this means that the mean prior of quality was lower (higher), as people were more surprised by the positive quality signals. Changes in demand when the school receives different letter grades, help identifying preferences for quality and prior uncertainty.

Figure C1: Belief updating as a function of prior parameters and changes in letter grades



Notes: This figure plots the change in expected school quality $\Delta E[q]_{s'-s''}$ due to a change in quality signal from s'' to s' as a function of the prior mean μ and variance σ. $s = 0$ denotes the absence of letter grades.

School quality distribution I assume that each school in the city has quality q_j that is fixed over time and that is not observed by students. For the empirical estimation I re-center value added around the city-wide average without loss of generality, since value added is always measured relative to the value-added of some arbitrarily picked school. Moreover, I rescale it by its across-school standard deviation which simply changes the interpretation of the preference parameter γ to measure willingness to travel for 1 additional standard deviation of quality in the cross-school distribution. Denoting with $[\underline{q}, \bar{q}]$ the space of possible quality values observed in the city, the empirical distribution of quality in NYC is well approximated by a truncated normal distribution over $[\underline{q}, \bar{q}]$.[60] Given the normalization, I have that the mean and standard deviation of the quality distribution are $\mu_q = 0$ and $\sigma_q = 1$.

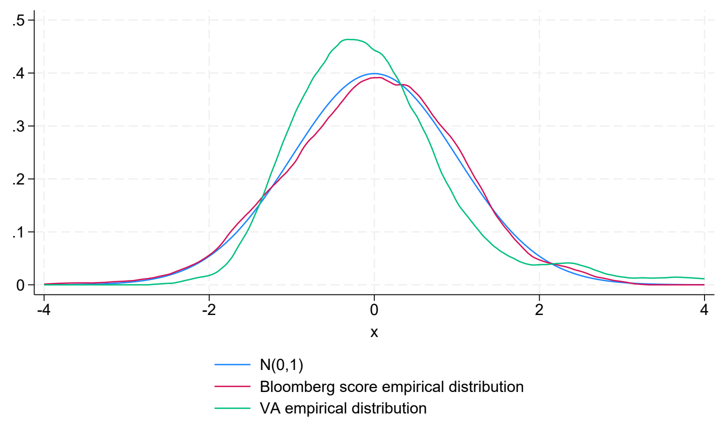
Figure C2 shows that the normal functional form is a very good approximation of reality. The figure compares the distribution of value added in the city (standardized across schools) and the distribution of the standardized quality scores underlying letter grades used in one random year against the probability density function of a standard normal. The three densities are similarly bell-shaped and approximately symmetric.

The point of the figure is to show that, while the underlying quality score was only imperfectly correlated with value added, quantiles of the two measures correspond roughly everywhere. This is important in light of my assumption that students observed quantile cutoffs of the quality score distribution used to assign letter grades and that they believed these were the cutoffs of the underlying value-added distribution. Because the two coincide in practice, this makes it easier to believe the assumption and also easier to argue that changes in demand following letter signals based on quantiles of the quality score can be more easily generalized to counterfactual scenarios in which signals of quality are based on value added. The fact that quantiles of the two distributions coincide (and that their shape is also similar to the normal form that I picked for student priors) is nice but not necessary for my estimation argument to be valid. Under my assumption, I would still be measuring preferences for value-added even if the quantiles of the two distributions were different.

Additionally, the fact that the empirical distribution of quality is approximately normal motivates the adoption of a truncated normal functional form for student priors. Priors are assumed to be normally distributed over the possibly quality space but are not constrained in any additional way, that is, they may be incorrect on average. This allows students to believe that certain schools are of better or worse quality than the average school and be more or less certain about that, but the shape of their beliefs is still realistic, in the sense that it is similar to what is observed in the city.

[60] While the number of schools and the distribution of quality in the city are naturally discrete, this continuous approximation is a convenient simplifying assumption because it allows to work with continuous probability densities. This assumption is more innocuous in a setting like NYC, in which the number of high schools families can choose from is large ($\sim$400) compared to the average public school district in the U.S.

Figure C2: Comparison of quality distribution relative to a standard normal



Notes: This figure plots the distribution of standardized value-added in NYC high schools (green line), of the standardized quality score used in the last year of the letter grade policy (pink line) and the distribution of a standard normal (blue line).

C.2 Counterfactual simulations details

In all simulations, I restrict the pool of general education 9th grade applicants to students for whom I know baseline test scores, residential address and race. These restrictions drop 20% of the applicant pool. I restrict capacities of all schools uniformly by this same factor. I simulate choices at the school rather than at the program, since all observable schools characteristics are measured at this level, and I use the admission rules (priority design and screening method) of the school largest program, in case the school has more than one, which is the case for approximately 30% of the schools. I am able to reconstruct student priorities and ranking at all schools based on student residential address, home language, ell status, middle school test scores, standardized state test scores and the middle school the student attended. Several schools give higher priority to students attending an information session at the school, which is a piece of information I do not have and that I cannot counterfactually estimate. Therefore the simulated priorities abstract from this.

Finally, I keep students' lottery number and preferences shocks constant across counterfactual simulations. The length of the rank ordered list is model-driven and coincides with all schools with utility above zero, the utility of the outside option. This typically results in longer lists, which explains why in my counterfactual simulations the probability of receiving an offer in the first round is slightly larger (by approximately 2 p.p.) than what is observed in reality.

D Data appendix

D.1 NYCDOE data

Lists of high school applicants, their rank ordered lists, priorities, lottery numbers, 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). Baseline middle school achievement in taken from the New York State Assessment. High school achievement outcomes come from Regents exams, graduation, SAT and college records that originate with different sources, and are collected by the NYCDOE. Geographic information on students comes from Zoned DBN data. All these data files were provided by NYCDOE. They 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 high school applications are maintained by the Student Enrollment Office of the NYCDOE. I received all applications for the 2006-07 through 2018-2019 school years. Application records include students' rank-order lists of academic programs submitted in each round of the application process, eligibility, priority group and rank at each program listed, the admission procedure used at the respective program, and the program to which the applicant was assigned. Lottery numbers and details on assignments at Educational Option (Ed-Opt) programs are provided in separate data sets only for the high school match of 2012 to 2017. The NYC high school match is conducted in three rounds, and separately for 9th grade and 10th grade seats and for general education and special education seats. I focus on first-time applicants to general education 9th grade seats using data from the first round.

OSPA Data I received registration and enrollment files for the 2005-06 through 2018-2019 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 Regents Exam Data Regents Examinations are statewide-standardized exams used to determine the type of New York State Diploma students are eligible for and more broadly to determine graduation eligibility. I received data on all Regents examinations conducted between the 2005-2006 and the 2020-2021 school years by students enrolled in a NYC public school. These years cover all high school applicant cohorts in my analysis sample, since most students take these tests before or during their junior year. I use the first test in each subject for multiple takers. I only consider tests taken in the subjects of ELA and Algebra 1 (or corresponding denominations that may vary slightly over the years). During my sample period, Regents in ELA and Math have been redesigned. In the 2013-24 SY NYC began administering the new test, designated as Common Core aligned Regents. During the first year the Common Core was rolled out, students were allowed to take the old and the new test and the higher of the two was counted for grading and other purposes. I adopt the same convention for students taking both tests during the transition period. Scores are then normalized to have mean zero and standard deviation one within a subject-year. A very small subset of students takes the ELA test during 8th grade. I only keep records for tests take after high school.

SAT Data I received data on SAT scores for tests conducted between the 2006-2007 and the 2020-2021 school years by students enrolled in a NYC public school. These years should cover all cohorts in my analysis sample, since most students take these tests before or during their junior year. These data originate with the College Board and but are provided by the NYCDOE. I use the first test for multiple takers. For applicants tested in the same month, I use the highest score. During my sample period, the SAT has been redesigned. I re-scale scores of SAT exams taken prior to the reform according to the official re-scaling scheme provided by College Board. Scores are normalized to have mean zero and standard deviation one within a subject-year among all students taking SAT in that year.

Graduation Data The DOE Graduation file records the discharge status for public school students enrolled from the 2005-2006 to 2020-2021 SY. I do not have graduation records for the last cohort of applicants in my sample, who were expected to graduate during the 2021-2022 SY. These records are used to construct a dummy indicating students graduating within 4 years of their enrollment in 9th grade.

College Data College enrollment data are generated by the DOE School Performance office based on data collected from the National Student Clearinghouse (NSC). They contain one record per student from the graduating cohort of that school-year, indicating whether a student enrolled in college in the fall that immediately followed their on-time graduation. I received data covering years from 2005 to 2020. These would cover students graduating on time up to the 2016 cohort included. Records before 2016 do not distinguish between 2 year and 4 years higher education institutions. For this reason, I use this data to construct an indicator for enrollment in any college in the fall that immediately followed their expected on-time graduation.

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 7th grade test scores from the 2003-04 to the 2017-18 SY 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 7th grade NYC public school students taking the test.

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 of school years from 2007-08 to 2018-19 to collect data on student residential districts and census tracts at the time of high school application.

D.2 Public data

Commuting time I collect commuting times between a student's residence and a school, estimated using the HERE Public Transit API. They are given by public transit travel time made of the shortest combination of walking, local and express bus, and subway modes, setting an arrival time of 8:00AM on September 9th, 2020. Residential addresses are approximate and given by the centroid of the census tract of residence.

High school directories I collect the pdfs of the printed high school directories used by applicant cohorts of 2006-2018 and the corresponding digitized versions. I use this data to understand the information displayed on each directory and as a basis for my school-year panel.

School progress reports I obtain publicly available data included in the NYC DOE School Progress Reports rating school performance in the school years from 2006-07 to 2013-14 from the NYC Open Data website.[61] These files include the overall grade and quality score a school received, as well as grades and sub-scores for each of the components of the overall quality score (e.g. school environment score, school progress score etc.) Grades referring to a school year (e.g. 2006-07) were typically made public during the following year (e.g. 2007-08).

D.3 Survey data

In February and March of 2023, we surveyed parents and guardians of students who had applied to 9th grade seats during the 2023-2024 NYC high school admission cycle, in partnership with the NYCDOE. A more extensive analysis of the survey data is presented in a companion paper (Corradini and Idoux, 2023). The survey was designed to be sent after families applied to high school but before the offers were sent out. The timeline allowed parents to have at least two weeks to complete it, and the survey had no time constraints beyond this deadline. Incomplete surveys were automatically submitted by the deadline. Participants who answered at least one question by the deadline received a $10 Amazon gift card. The survey was sent electronically using the email addresses of families used in the high school application process in the top three most spoken languages in NYC: English, Spanish, and Simplified Chinese. The survey was designed on Qualtrics and it was available in those same three languages. All questions were marked as optional, except the consent to participate which included declaring to be older than 18.

The survey was sent to 21,401 families. This sample consisted of a subset of parents or guardians of students applying to start 9th grade in fall 2023 who satisfied the following conditions: 1) they had been enrolled in a NYC middle school since 6th grade 2) they had non-missing demographic records within the NYCDOE database and 3) they had taken the

[61] https://opendata.cityofnewyork.us/

New York State Assessment standardized test in 4th grade and we observed that in our records. We received 3,628 responses (17% response rate).

Table D1 compares descriptive statistics of survey receivers (columns 2-3) and respondents (column 4-7) to those of NYC applicants (column 1). Respondents were slightly more likely to be white and Asian, less likely to be eligible for a subsidized lunch and had students with higher baseline achievement, compared to the sample of eligible families. We only use responses of students completing at least 50% of the survey (descriptives in column 5).

Table D1: Descriptive Statistics of Survey Receivers and Respondents

	Applicants with baselines			Respondents			
	NYC	Survey receivers	Receiving belief Q.	All	Answers > 50%	Ans. > 50% and gets belief Q	Ans. > 50% and ans. belief Q
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Asian	20%	25%	25%	29%	29%	28%	29%
Black	19%	16%	16%	14%	14%	13%	13%
Hispanic	42%	39%	39%	33%	32%	32%	30%
White	16%	17%	17%	21%	22%	23%	25%
Subsidized Lunch	76%	73%	73%	66%	64%	64%	62%
Brooklyn	29%	31%	31%	31%	31%	33%	32%
Queens	32%	35%	34%	37%	36%	35%	36%
Manhattan	10%	9%	9%	10%	11%	12%	12%
Bronx	21%	18%	18%	14%	14%	14%	12%
7th grade Math	0.11	0.34	0.33	0.55	0.58	0.59	0.61
7th grade ELA	0.11	0.35	0.35	0.51	0.54	0.55	0.58
N	47618	21401	7946	3628	3099	1142	849

Notes: This table provides descriptive statistics about the sample of households receiving (column (2)) and responding (column (3)) to our survey. Column (1) compares their characteristics to the entire sample of 9th grade applicants from which we drew the sample of survey receivers. Column (5) restricts the sample of respondents to those answering at least 50% of the survey questions, column (6) further conditions on respondents receiving the question about beliefs used in this paper and column (7) conditions on having responded to the question, which is the sample used in this paper.

To reduce the time it takes to complete the survey and increase participation, we devised eight different survey versions by creating different combinations of question subsets. Eligible participants were shown one version of the survey at random. In this paper , I only use the answers to one question asking respondents to situate a school in the distribution of school quality of their residential borough. The text of the question asked:"How well does *school name - (school code)* prepare students for their Regents exams compared to other schools in your borough?". Answers could vary from 1 to 4, with 1 corresponding to the worst 25% of schools and 4 to the best 25%. By design, 37.5% of the survey participants received this question in their questionnaire. Columns 6 and 7 restrict the sample of respondents to those receiving the question I study (N=1,142) and finally answering it (N=849). This last sub-sample is what I use in the analyses in section 2.3.

The schools that populated the question varied across respondents. Schools were assigned to respondents at random subject to a set of constraints. The set of high schools eligible for inclusion was determined on the basis of their proximity to the student's address and other criteria as follows:

1. we started from schools existing in the 2021-2022 high school directory, dropped specialized schools, special districts (75 and 79), and home schools and keep only schools participating in the 2023 high school match.

2. For each of the 32 residential districts, we took all schools that are located in the district borough

3. We then restricted to fairly popular schools in the district, as indicated by the fact that they were ranked by at least 5% of students in the district.

This returns, on average, 55 schools per district. A school in this subset is ranked, on average, by 11% of students residing in the district. For each district, we then selected at random 4 schools within this subset, subject to the constraint that each of the four school corresponded to a possible combination of two dummies. The first dummy indicated schools enrolling a high share of white and Asian students (above 26% of white and Asian students, corresponding to the 25% most white schools). These are schools typically also enrolling higher achieving students. The second indicated schools with Regents value-added above the median in the borough. For the purpose of designing the question, I classify a school as having above median Regents value-added if it is above the median for both Regents Math and Regents ELA.

If the intersection of high-white and above (below) median value-added returned an empty set, we selected the school with the highest share of white students, conditional on being above (below) the median value-added. If this also resulted in an empty set, we chose the school with the highest (lowest) value-added, conditional on being a high white school.

Similarly, if the intersection of non-high-white and above (below) median value-added returned an empty set, we selected the school with the lowest share of white students, conditional on being above (below) median value-added. If this also returned an empty set, we selected the school with the highest (lowest) value-added, conditional on being a non-high white school.

Each respondent was randomly assigned one of the four possible schools selected for her district. The balance table below confirms that the characteristics of schools assigned to the questionnaires did not differ by respondent race. It includes regression estimates of school characteristics on respondent's race, controlling for district fixed effects. The coefficient on respondent's race is never statistically different from zero.

Table D2: Balance of School Characteristics Across Respondent Race

	Regents VA (SD)		Average Regents (SD)	
	(1)	(2)	(3)	(4)
Respondent is white or Asian	-0.003	-0.039	0.040	0.042
	(0.109)	(0.053)	(0.094)	(0.045)
Average Regents (SD)		0.912***		
		(0.024)		
Regents VA (SD)				0.734***
				(0.020)
N	849	849	849	849
Dep. var. mean	0.141	0.141	0.467	0.467
Dep. var s.d.	1.147	1.147	1.162	1.162

Notes: This table shows that the quality and mean achievement levels of the schools populating the survey questions are balanced across respondent's race. The table reports estimates of a regression of school characteristics on a dummy indicating whether respondents are white or Asian. Columns (2) and (4) also control for the school attribute not used in the left hand side (e.g. school achievement levels in the regression with school quality as a dependent variable) to check balance also conditional on other school attributes.