Essays on the Economics of Education and Market Design

Thi Hoang Lan Nguyen

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2020

Abstract

Essays on the Economics of Education and Market Design

Thi Hoang Lan Nguyen

This dissertation consists of three essays on the economics of education and market design. The first two chapters are united in their attention on school choice issues. Chapter 1 considers a specific application, whereas chapter 2 focuses on a matching mechanism widely used in multiple applications. Both chapters 1 and 3 explore equity concerns in education but through very different lenses (affirmative action vs. educational investment) and very different settings (the United States vs. Vietnam).

Chapter 1 addresses the diversity issue that is especially prevalent in elite schools that select students based on exams. Whereas previous studies only consider the direct impact on elite schools, I quantify the effects of two widely-discussed affirmative action plans on both elite and regular schools in New York City. I find that the two plans have quite different effects. First, there is a trade-off between improving diversity and maintaining student quality in elite schools as measured by state test scores in middle school. Despite taking into account the socioeconomic status of students' neighborhoods, the Chicago plan gives rise mostly to reshuffling within elite schools. Thus, both the overall racial composition and quality of incoming students are largely preserved as in the status quo. In contrast, the Top 7% plan, which would accept into the elite sector students in the top 7% by academic performance of each public middle school, causes considerable flows of students between the elite and regular sectors. The elite sector experiences a substantial increase in the proportions of Black

and Hispanic students, along with a decrease in average student quality. Analyzing the difference between the outcomes of these two policies provides some insight into how the two objectives—diversity and peer quality in elite schools—might be better balanced in general. The second difference between the plans arises because they transform the distribution of diversity across schools in different ways. The Chicago plan reduces the differences among schools within the elite sector, while the Top 7% plan reduces the gap in diversity between the two sectors even as it increases within-sector dispersion. Both plans result in considerable changes in school assignments in the regular school sector, thus affecting the average student quality in these schools.

Chapter 2, joint work with Guillaume Haeringer and Silvio Ravaioli, uses a lab experiment to study learning dynamics when participants receive feedback in centralized matching mechanisms. Our design allows for two types of learning: to coordinate within the same environment as well as to understand the underlying mechanisms. We provide additional evidence to previous work that the majority of the deviations from truth-telling, the dominant strategy in the Deferred Acceptance mechanism, are those that do not affect payoffs. Furthermore, by explicitly analyzing learning, we can confirm that at least some of the participants learn about the optimality of truth-telling, and their departures from it happen primarily when they face the same environment being repeated. Finally, we find that when learning to coordinate, agents tend to retain their previous strategy when the payoff from this strategy is high. This is suggestive evidence of reinforcement learning.

Chapter 3 documents the pattern of educational investments for high school students across different demographics and their effects on performance on the college entrance exam and in college. Survey data from Vietnam shows that high school students from higher-income households have higher education expenditure and participation in extra classes (both at the

extensive and intensive margin). Minority and rural students invest less than their non-minority and urban counterparts even after controlling for income. Out of these investments, only extra classes during the school year education expenditure other than that on extra classes are effective in increasing college entrance exam scores. In terms of college performance, a higher entrance exam score leads to a slightly higher grade point average at graduation, controlling for academic department fixed effects and investments in high school. Neither education expenditure or participation in extra classes in high school show any significant effects on college performance, except that already captured in the entrance exam scores. I record multiple gender differences. Female high school students tend to receive more investments. Even though they perform slightly worse on the entrance exam than their male peers with the same investments, they perform better in college, given the same entrance exam scores.

Contents

List of Tables

List of Figures

Acknowledgments

I would like to thank Columbia University, Trudy and Paul Woodruff Fellowship, and Columbia Experimental Laboratory in the Social Sciences for their funding and support. I am also grateful to the New York City Department of Education for their generous permission to allow me access to their data.

I would like to express my deepest gratitude to my advisors, Yeon-Koo Che and Miguel Urquiola, for their dedication and invaluable guidance. Without their tremendous help and unwavering support throughout the years, especially during an unexpectedly difficult job market, I would not have been where I am today. I am enormously grateful for Bernard Salanie, who has also been an excellent advisor to me, in all aspects but the official record. I thank Cailin Slattery and Alex Eble for their counsel and feedback as members of my dissertation committee.

I would like to thank Mike Riordan, Kate Ho, Tobias Salz, Adam Kapor, and John Asker. Even though I am sad that you could not accompany me towards the end of my Ph.D. journey, whether because of a move or retirement, I am happy to have received your mentorship along the way. I am also thankful for my co-authors, Guillaume Haeringer, Heiko Gerlach, and Silvio Ravaioli, who have been more helpful and understanding than I could ever ask for.

I am tremendously grateful to our program coordinators Shane Bordeau and Amy Devine, who have gone above and beyond for all of the students in our graduate program, myself included. I thank my upper-and under-classmen, especially those who also work in Industrial Organization and/or Education—Dong-Woo Hahm, Nate Mark, Anh Nguyen, Ilton Soares, to name a few—for their advice and feedback.

To Claudia Allende, Evan Friedman, Han Huynh, RC Lim, Janet Martin, Xavier Moncasi, Suanna Oh, and Jon Zytnick, thank you for sticking with me through thick and thin throughout this journey. I am blessed to have been part of such a close-knitted group. May our friendship endure and flourish for years to come.

To my Mum and Dad, to my Grandma and little sister, thank you so much for always being there for me. I love you, always.

CHAPTER 1: Predicting the Effect of Affirmative Action Plans in New York City Elite Public High Schools

1.1 Introduction

In recent years, the lack of diversity in schools has been a concern for policymakers. This problem is particularly intense for selective "elite" public schools. For example, this academic year, Stuyvesant, arguably the most prestigious high school in New York City, initially made 895 offers of admission, only 7 of which went to Black students (Shapiro, 2019). This situation has prompted urgent calls for affirmative action policies that might allow more minority and low socioeconomic status students to enroll.

The effects of such policies could be quite complex as they would likely propagate throughout the entire public high school system. Specifically, some seats in regular public high schools might be freed up when policy beneficiaries leave them in favor of elite schools, and they might be taken up by students displaced from elite schools. Previous studies have abstracted from this issue due to the technical challenge of accounting for students' choices in two connected school systems. In short, they have overlooked the impacts of affirmative action on those who would neither gain nor lose offers to elite high schools.

This paper addresses this challenge and quantifies the effects of affirmative action on both elite and regular schools. I consider two widely-discussed affirmative action plans in New

York City. The first has been put in place by the Chicago Public School District and takes into account the socioeconomic status of students' neighborhoods. The second was proposed by New York City Mayor Bill de Blasio and would accept into the elite sector students in the top 7% by academic performance of each public middle school. The latter plan is reminiscent of others in the country, such as those implemented at the college level in Texas and California.

I find that the two plans have quite different effects. First, there is a trade-off between improving diversity and maintaining student quality in elite schools as measured by state test scores in middle school. The Chicago plan mostly gives rise to reshuffling within elite schools, which does little to increase the racial diversity of this sector as a whole but, at the same time, preserves the quality of incoming students. In contrast, the Top 7% plan causes considerable flows of students between the elite and regular sectors. The elite sector experiences a substantial increase in the proportions of Black and Hispanic students, along with a decrease in average student quality. Analyzing the difference between the outcomes of these two policies provides some insight into how the two objectives—diversity and student quality in elite schools—might be better balanced in general. The second difference between the plans arises because they transform the distribution of diversity across schools in different ways. The Chicago plan reduces the differences among schools within the elite sector. In contrast, while the Top 7% plan reduces the gap in diversity between the two sectors, it increases the dispersion within each sector.

One effect that both plans have in common is that they change the average student quality in regular high school programs: more than half of these schools have higher student quality than under the status quo, whereas the rest undergo decreases in quality. This is due to the movements of students between the elite and regular sectors, as well as reshuffling within the regular sector. This result confirms that not considering regular schools is a damaging omission in evaluating affirmative action policies.

In New York City, the elite sector includes eight specialized high schools, which are public high schools that select their students based on the Specialized High School Admission Test (SHSAT). In the 2011-2012 school year, White and Asian students made up about 85% of specialized high schools, whereas Blacks and Hispanics made up only about 15% of these schools. These percentages are out of proportion to those in the broader population.[1]

The key feature that facilitates spillover effects from specialized high schools to regular schools is that the Department of Education (DoE) determines most school assignments in two rounds: a "Specialized Round" and then a "Main Round." This assignment procedure begins after all students have already submitted their rank-ordered lists over both types of schools (when applicable), and both rounds use these lists as inputs. The Specialized Round selects students who get offers to specialized high schools and also provides each of them with a regular school offer. These students are then allowed to choose between these two offers. Based on their decisions, the availability of seats at regular schools is adjusted before the Main Round is run to determine the rest of the regular school assignments.

I exploit properties of Gale and Shapley (1962)'s student-proposing Deferred Acceptance (DA) mechanism—defined broadly to include cases both with and without restrictions on the length of the rank-ordered lists—to inform my treatment of the link between the school rankings that students submit and their true preferences. Following Abdulkadiroglu, Agarwal and Pathak (2017), my empirical strategy relies on weak truth-telling.[2] Unlike strict truth-telling, this assumption does not presume any particular preference relation between

[1] For instance, public middle schools comprised of 28% White and Asian and 71% Black and Hispanic students.

[2] There is an alternative approach where this weak-truth-telling assumption is relaxed to allow for payoff-irrelevant "mistakes" (?Fack, Grenet and He, 2019). However, this method is not applicable in my setting due to the higher level of uncertainty students face, which I will discuss in more details in Section 1.5.

unranked schools and the outside option of not attending any New York City public high school (i.e., attending a private school or a school outside of the city). Specifically, weak truth-telling includes two requirements. First, the order of schools in the submitted rank-ordered list is according to true preferences, which is likely to hold for both specialized high schools and regular schools as this is a dominant strategy in the student-proposing DA mechanism. Second, ranked schools are preferred to all unranked schools. The latter may be violated for regular schools, where students can only rank twelve out of nearly 700 programs. When students are constrained to ranking only a strict subset of the schools available, it may be optimal for them to drop some schools to which they are unlikely to be admitted to make space for less-preferred schools where they have a better chance (Haeringer and Klijn, 2009). However, the assumption may still be reasonable in this particular setting, given that only 14.6% of students ranked all twelve programs.

I devise a two-step estimation procedure to infer preferences between the two connected school systems: specialized and regular. Within each system, given truth-telling, students' rankings over schools give rise to a generalization of the standard discrete choice model in which instead of choosing one single alternative, agents express preferences over multiple alternatives. Such a model, especially one with logit errors, is well-studied.[3] With two systems, the lack of a universal ranking poses a problem in ensuring that the cardinalizations of preferences for the two sets of schools are comparable. To address this, I make use of the choices between elite and regular schools, which are observed for students with elite school offers. In each step of the procedure, I estimate a rank-ordered logit model for preferences over one set of schools via maximum likelihood. Estimation for regular schools is done first so that the results can be used to normalize utilities for specialized high schools both in terms

[3] See, for instance, Beggs, Cardell and Hausman (1981) and Hausman and Ruud (1987).

of location and scale. This normalization involves a modification of the standard likelihood function to account for the regular vs. elite choice mentioned above.

The Chicago affirmative action plan assigns 40% of total specialized high school capacity purely on merit. The rest are divided equally among each of four neighborhood tiers. In my implementation of this policy in New York City, each student's tier is the tier of his or her zip code (assigned based on its median family income). I assume that those who wish to gain admission to specialized high schools would still have to take the SHSAT, which would be used as the criterion for merit seats as well as a tie-breaker among students of the same tier. Meanwhile, the Top 7% plan involves abolishing the current admission test and guaranteeing specialized high school seats for students in the top 7% of each public middle school in the city, determined by their combined English Language Arts (ELA) and mathematics scores on the Grade 7 state tests. I assume that these students would be asked for their rank-ordered lists over specialized high schools, and schools would rank students based on the aforementioned scores.

For the analysis of both counterfactual policies, estimates of the students' underlying preferences are used to simulate rankings that were not submitted in reality. First, students who did not apply to specialized high schools did not submit rankings for them, but in the Top 7% plan, some of these students would be eligible for seats there, and thus, such rankings are required to determine their school assignments. Second, we can only compare students' preferences between specialized high schools and regular schools for those who received specialized high school offers, and for each of these students, only between a specific pair of specialized high school and regular school. In both counterfactuals, new comparisons arise either when students receive specialized high school offers while they previously would not have or when they obtain a different specialized high school offer than the one they would have received in the current system. These "missing" rankings are constructed using the systematic part of the utility function that I estimated from the aforementioned random utility model; that is, a student is said to prefer one school over another in the constructed rankings if they are expected do so on average across all possible realizations of the idiosyncratic preference shocks.

Using the estimated parameters, I simulate both policies in the short run when preferences remain unchanged and find considerable differences in their effects. The Chicago plan causes very little exchange of students between the two school sectors (approximately 9% of total specialized high school capacity) but instead mostly gives rise to reshuffling within specialized high schools. As a result, both the racial composition and student quality[4] of specialized high schools as a whole remain relatively stable. Meanwhile, different specialized high schools become more similar to one another. Concretely, for Stuyvesant and Brooklyn Latin (the two schools with the lowest and highest percentages of Blacks and Hispanics, respectively), these percentages become closer, go from approximately 3% and 38% to 8% and 34%. The Top 7% plan results in considerable inflows from regular public high schools and vice versa (amounting to approximately 49% of the total specialized high school capacity), radically increasing the percentages of Black and Hispanic students in specialized high schools while lowering the overall quality. Under this plan, the elite and regular sectors become more similar in terms of racial composition, but within the elite sector, schools become more different from one another (Black and Hispanic students making up about 19% of Stuyvesant compared to 80% of Brooklyn Latin).

Despite differences in the volumes of flow between specialized high schools and regu-

[4]Quality is measured by combined scores on Grade 8 state tests, which are taken after the admission process but before students enter high school. This measure is therefore distinct from the selection criterion of the Top 7% plan.

lar schools, under both the Chicago and Top 7% plans, many students who are not directly affected—that is, they neither gain nor lose specialized high school offers—still experience a change in their school assignments. Within regular schools, there is slightly more reshuffling for Black and Hispanic students than for White and Asian students. Both plans induce increases in average student quality in more than half of the regular high school programs and decreases or no change in the rest. However, only some of these changes are statistically significant. The Top 7% plan additionally changes the distribution of racial compositions in such a way that there is a notable increase in the number of programs with very low percentages (0-10%) of Black or Hispanic students, a sign of increased segregation within the regular school sector.

On the student level, changes in peer quality are heterogeneous across the students' own scores. Under both plans, those belonging to the lowest two score deciles in the population of all public middle school students lose out in terms of peer quality, while those belonging to higher deciles (fifth to ninth) gain better peers. The sizes of the increases for the latter are either similar across plans or higher for the Top 7% plan. The two plans differ in the direction of the changes when it comes to students in the highest-scoring decile, who experience an increase in peer quality in the Chicago plan but a decrease in the Top 7% plan. The same type of assignment changes (e.g., from specialized to regular high schools) under the two plans can result in different experiences in terms of peer quality.[5] For instance, under the same label "always specialized high schools," students who stay in specialized high schools under both the status quo and the Chicago plan can be either better off or worse of in terms of peer quality depending on their own scores, while those who stay in specialized high schools

[5]The full list of types that I consider includes: always specialized high schools, specialized to regular, regular to specialized (offer gained), regular to specialized (better offer), different regular schools, same regular schools (main round), and same regular schools (specialized round).

under both the status quo and the Top 7% plan are always worse off.

Three main factors drive the differences in the outcomes of the two policies. First, the pool from which students are selected into specialized high schools is much more restrictive in the Chicago plan, where the students must have taken the SHSAT to be considered. On the contrary, the Top 7% plan considers all aspiring high school students, and in fact, nearly half of the students who are reassigned from regular schools to specialized high schools under this plan (1280 out of 2666) did not take the SHSAT. Second, the two plans currently use two different criteria to define "merit," with the Chicago plan using the SHSAT and the Top 7% plan using Grade 7 state test scores. When I restrict the selection pool to specialized high school applicants (for whom both types of scores are available) and compare a policy that selects the top 7% of each public middle school based on state test scores versus one that selects the top 7% based on SHSAT scores, the overlap in offers comprises of only 29% of the total specialized high school capacity. Third, the policy outcomes depend on how students are assigned into subgroups, either neighborhood tiers in the Chicago plan or public middle schools in the Top 7% plan. An affirmative action plan is more likely to promote diversity if the subgroups being used have sufficiently different racial compositions, which is the case for public middle schools. Meanwhile, a plan is more likely to maintain student quality in the elite sector if student quality is sufficiently similar across subgroups, such as neighborhood tiers.

The rest of the paper is organized as follows. Section 2 reviews the related literature and highlights my contributions. Section 3 provides relevant background on the New York City high school admission process. Section 4 introduces the data sources and my descriptive analysis. Section 5 describes the structural model, estimation procedure, and estimated parameters. Section 6 presents the effects of the two affirmative action plans and the policy implications. Section 7 concludes by discussing the paper's key findings and avenues for future research.

1.2 Related Literature

This paper makes several contributions to the literature on affirmative action in centralized school choice, particularly that in selective public high schools. In doing so, it also contributes in terms of empirical methodology to the emerging literature on divided enrollment systems.

Among papers that evaluate affirmative action in selective public high schools, this is the first that considers the indirect impact on regular public high schools in addition to the direct impact on selective ones. In the same setting of New York City specialized high schools, Treschan (2015) and Corcoran and Baker-Smith (2018) both simulate alternative admission rules keeping the total number of offers in specialized high schools fixed. Due to data limitations, neither study takes into account the students' preference rankings over specialized high schools, whether with or without the alternative policy. They treat all specialized high schools as a single entity and only provide results in terms of offers, not actual enrollment. In contrast, my paper accounts for students' preferences over specialized high schools as well as regular schools and finds two types of effects on specialized high schools that these studies miss. First, a non-trivial share of students with specialized high school offers decline them in favor of regular high school programs in all scenarios (status quo and alternative plans). The affirmative action plans that I consider both reduce the fraction of such students compared to the status quo by giving them offers to a specialized high school that they prefer over the specialized high school they would have been assigned before. Thus, even for specialized high schools as a whole, looking at offers instead of enrollment presents a consequential oversight. Second, by not considering assignments to specific specialized high schools, previous studies also overlook the possible reshuffling within specialized high schools, an important effect of the Chicago plan in my paper. In addition to neglecting some direct impacts on the elite sector, these studies do not consider the impact on the regular sector, which I find to be substantial in my analysis.

In terms of outcomes that are considered by both these authors and myself, Corcoran and Baker-Smith (2018) offer a similar result for their Top 10% rule (based on grade 7 state test scores and course grades) as my result for the Top 7% plan (based on grade 7 state test scores). Specifically, both show substantial changes in specialized high school offers that bring the racial composition of specialized high schools closer towards that of the population of public middle school students but at the same time reduces average student quality. It must be noted that Corcoran and Baker-Smith (2018) most likely understate the magnitude of the decrease as the authors themselves admit that they measure quality by achievement in grade 7 state tests, which is part of the selection criteria. To assess student quality, my paper uses grade 8 state tests, taken after the admission process but before students enter high schools. This is a better measure because it is both independent of the admission process and more recent. In contrast to the aforementioned results, Treschan (2015) finds that his proposed policy slightly improves both diversity and student quality in specialized high schools. The opposite result regarding student quality should be taken with a grain of salt because the policy assigns all except for about 9% of specialized high school offers based on city-wide rank in grade 7 state test scores, and student quality is then measured by the same tests.[6] The

[6]The paper is unclear about which state tests are used for student quality, but given the lack of expressed distinction between them and those used for selection, it is likely that they are the same.

increase in "quality" thus follows by construction.[7]

Another paper on affirmative action in elite high schools is Ellison and Pathak (2016), which considers the Chicago exam schools. However, their focus is on the efficiency of policies based on proxies for membership of the underrepresented group compared to policies based directly on membership status, for instance, race-blind affirmative action policies compared to race-based ones. To define efficiency, the authors assume that social welfare is the sum of students' expected educational outcomes, which, for each student, depends on the student's own type (ability and/or preparation), how well the school curriculum matches the student's type, and how well the composition of students in the school reflects the composition in the population. Using data from Chicago for the year 2013-2014, the authors evaluate the efficiency of the current Chicago tier-based policy as well as some alternative policies (including a Top 10% rule). The authors focus only on the two most selective schools. In all policies, only students who did apply to elite schools are considered, which means submitted school rankings are enough for counterfactual simulations, precluding the need for preference estimation as in my paper.

My second contribution is to provide further empirical evidence to illustrate previous theoretical results regarding the importance of precedence order, that is, the order in which the school seats are filled in the Deferred Acceptance (DA) mechanism in the presence of reserved seats such as the tier seats in the Chicago plan. In fact, this plan directly inspires Dur, Pathak and Sonmez (2016)'s theoretical model, which is simplified to a single school setting. The authors show that even with a tier-blind precedence order—where tiers are treated as anonymous, and thus, none is explicitly favored—one can still exploit the statistical prop-

[7]This problem is of greater concern here than in Corcoran and Baker-Smith (2018), where top students from each middle school are selected instead of city-wide.

erties of the scores to generate (dis)advantageous allocations to certain tier(s). Specifically, assuming that the school is oversubscribed, the precedence order where all merit seats are filled before all tier seats is the best tier-blind precedence for the tier with the worst distribution of scores, and the worst tier-blind precedence for the tier with the best distribution of scores.[8] The opposite is true for the precedence order where all tier seats are filled before all merit seats. My empirical analysis in the multiple-school setting of New York City specialized high schools gives similar results to that by Dur, Pathak and Sonmez (2016) of the Chicago exam schools and supports their theoretical results. Dur et al. (2017) instead focus on the effect of the change in the size of the reserve compared to the change in the precedence order. Theoretically, increasing the size of the reserve and moving non-reserve seats up in the precedence order both weakly increase the number of reserve-eligible students admitted. In their empirical analysis of the Boston Public School system, the effects of these two different adjustments are quantitatively similar. In this paper, I also find that moving (non-reserve) merit seats up in the precedence order can be as effective as increasing the number of tier seats (the reserve).

My third contribution is to show empirical evidence of the effect of affirmative action in a real-life setting. This complements the theoretical literature on affirmative action under different centralized school choice mechanisms, which is mostly agnostic when it comes to the effect of affirmative action under specific conditions. Both Kojima (2012)'s and Dogan (2016)'s theoretical results show that there exist students' preference profiles under which more affirmative action can hurt the minority that it aims to help. Dogan (2016) points out that this problem is because for some minority students, being treated as a minority does not

[8]"Worst" ("best") in the sense that at every possible score, they have the lowest (highest) representation among all tiers.

affect their own outcomes, but it can negatively affect the outcomes for other minority students. Thus, the author proposes an alternative mechanism where this problem is alleviated. Although the type of affirmative action he considers is different, a similar insight applies to the Chicago plan. Specifically, having merit seats filled last is bad for the disadvantaged tier as mentioned above because high-achieving students in this tier, who would have been admitted purely on merit, end up filling tier seats, leaving their lower-achieving peers to compete for merit seats. It must be noted, however, that both Kojima (2012) and Dogan (2016) are silent on what would happen under specific students' preference profiles (such as those occurring in practice). My empirical results demonstrate that in New York City, affirmative action policies do help more students in disadvantaged groups gain access to specialized high schools.

In terms of methodology, I devise an estimation procedure to handle the two connected school systems in New York City, thus contributing to the literature on divided enrollment systems that has been mostly theoretical until now.[9] Divided enrollment systems refer to school systems in which students may apply to and receive offers of admission from multiple groups of schools, which conduct their admission processes separately. Manjunatha and Turhan (2016) show that this can lead to wastefulness in the sense that many students may not receive offers despite seats still being available in some schools. They propose an iterative mechanism in which these different groups can each independently match and re-match students to its schools to alleviate this problem while accounting for schools' priorities. This process is demonstrated to Pareto-improve the outcome in every iteration and arrive at a non-

[9] Hahm (2019) also estimates students' preferences for these two connected school systems, but in order to answer a different research question. As such, the author focuses his analysis only on specialized high school applicants. In terms of estimation strategy, he uses the choice between specialized high schools and regular school programs to create a global ranking for each student, and thus, preferences for both types of schools are estimated jointly.

wasteful outcome in a finite number of iterations. Turhan (2019) studies the effects of the partition structure—in other words, how schools are divided into groups—on the properties of the aforementioned mechanism. He finds that students' welfare increases and the mechanism becomes less manipulable as the partition becomes coarser. Dogan and Yenmez (2018) prove that given substitutable school priorities, a unified system, where each student submits one ranking over all schools and receives at most one school offer, achieves an outcome that (weakly) Pareto dominates that from a divided system. They also characterize three different sources of inefficiency in the divided system. My setting involves the coarsest possible partition short of unification, and while it does not use Manjunatha and Turhan (2016)'s mechanism, some waste is reduced when the DA mechanism is re-run for regular schools. The first source of inefficiency identified by Dogan and Yenmez (2018) is evidenced in the fact that declined specialized high school offers are not re-distributed. I show that under affirmative action plans, this waste is lessened when more students choose specialized high schools over regular schools.

Finally, this research is also related to the literature on affirmative action in the decentralized higher education markets, particularly Kapor (2016)'s paper on Texas Top Ten, which guarantees admission to Texas public universities for students in the top ten percent of their high school class. The paper finds that Texas Top Ten has a substantial effect on admission for all students due to displacement effects, like in my setting. The author considers a counterfactual where there is no affirmative action. In such a scenario, attendance at flagship universities decreased by 10% for Black and Hispanics students compared to under Texas Top Ten, but increases by 17% for students from affluent high schools. Entering students under Texas Top Ten or in its absence perform similarly in their first- and second- year GPA in college. In summary, Texas Top Ten improves diversity without compromising the educational outcomes of admitted students.[10]

1.3 New York City Public High School Admission

Middle-school students in NYC have two public high school options: specialized high schools and regular schools. A key feature that facilitates spillover effects from specialized high schools to regular school admissions is that admission results for the former are finalized first and thus affect admission to the latter. Private schools and schools outside of the New York City school districts are also among possible options.

1.3.1 Application Process

Grade 8 students can be divided into two groups: specialized high school applicants and specialized high schools non-applicants, that is, those who plan to apply for specialized high schools and those who do not. The sets of actions that needed to be taken differ for these two groups.

Specialized high school applicants need to go through a multi-step process. First, in October, they take the specialized high school Admission Test (SHSAT). The test includes two sections (verbal and mathematics) and can only be taken once.[11] Testing locations are assigned based on the geographic district of each student's middle school. In 2010-2011, Queens borough had two test locations, and the other boroughs each had one. During the test, students must express their preferences over the specialized high schools by filling out a rank-ordered

[10]For details on other papers in this strand of literature, Arcidiacono, Lovenheim and Zhu (2015) and Arcidiacono and Lovenheim (2016) provide excellent literature surveys.

[11]Strictly specking, students who fail to gain admission into specialized high schools can take the test again the following year when they are in grade 9. However, the test result will be used for admission into grade 10 in specialized high schools, which is separate from the admission into grade 9 that I focus on.

list (ROL) on the answer booklets. Each student must select one specialized high school as the first choice, and then can additionally specify the second choice, third choice, and so on, up to the eighth choice (i.e., ranking all eight specialized high schools) if he/she so chooses. These lists will be used as an input to determine specialized high school offers. Second, in December, they must submit their applications for regular schools, in which they can rank up to 12 programs out of approximately 700. At this point, students know about the admission methods of the regular school programs—which I will elaborate on in the next subsection—but not their actual chance of getting into these programs. They also have not yet received their scores for the SHSAT. Even if a student is not interested in any regular schools, this step is necessary for receiving results from specialized high schools. Finally, in February, those who qualify for offers from specialized high schools receive the aforementioned offer as well as an offer from a regular school program. They will then need to choose among these two options and the outside option of exiting the NYC public school system.

Specialized high school non-applicants go through a more straightforward process. They only need to submit one application for regular schools in December, the same as in the second step for specialized high school applicants. The ROLs from both specialized high school applicants and non-applicants will be used as an input to determine regular school assignments. Specialized high school non-applicants and specialized high school applicants without specialized high school offers will receive one regular school offer after the admissions for specialized high school applicants with specialized high school offers have been set.

1.3.2 School Assignment Procedures

The NYC Department of Education determines most school assignments centrally in two rounds: a "Specialized Round" and then a "Main Round."[12] In both rounds, Gale and Shapley (1962)'s student-proposing Deferred Acceptance (DA) mechanism, as defined below, is used.

In addition to students' submitted ROLs described in the previous subsection, the DA mechanism also takes schools' capacities and schools' strict priorities over students as inputs. All specialized high schools have a common priority rule: students are ranked according to their total scores on the SHSAT. For regular schools, different programs have different admission methods ranging from generating priority based on a random lottery in unscreened programs to evaluating each individual student based on criteria such as grades from the prior school year, standardized test scores, attendance and punctuality, and interviews or essays in screened programs. Programs that evaluate students individually are given a list of students who include them on their ROLs and asked to return a priority ranking over these students to the central enrollment office. Crucially, schools do not observe where they are ranked on the students' ROLs, so students do not have any incentive to influence schools' priorities by strategically changing their rankings.

Given the aforementioned inputs, students are matched to school seats using the following algorithm:

Step 1:

- Each student proposes to her first choice school according to her submitted ROL.
- Each school tentatively accepts the applicants who have applied to it, one at a time,

[12] There are also supplementary rounds afterwards for those unassigned or unhappy with their assignment by the end of the Main Round. However, these student only make up 14% of the total in application year 2010-2011.

using the school's priority list and starting with the applicant with the highest priority. It does so until it has admitted as many students as its capacity or runs out of applicants. The remaining applicants (if any) are rejected.

Step $k > 1$:

- Students who are rejected in the previous step apply to the next school on their submitted ROLs. If a student has already applied to all the schools in her ROL, then that student remains unassigned (and the algorithm ends for that student).

- Each school considers the set of students it accepted at the previous step together with the set of new applicants. From this larger set, the school tentatively accepts students, one at a time, using the school's priority list and starting with the applicant with the highest priority. It does so until it has admitted as many students as its capacity or runs out of applicants. The remaining applicants (if any) are rejected.

End: The algorithm ends when there are no new proposals; that is, either no one was rejected in the previous step or all rejected students have exhausted their ROLs.

When students are allowed to rank all available schools, student-proposing DA is strategy-proof for students; that is, it is a weakly dominant strategy for students to report their true preferences (Dubins and Freeman, 1981; Roth, 1982).[13] When they are only allowed to rank a subset of schools, it is still a weakly dominant strategy to submit a truthful order of the schools one does rank, but it may be optimal to drop schools (for instance, at the top of the

[13] Strictly speaking, these papers state the result for the marriage market, where one agent on each side is matched to each other. In the school choice context where each school can be matched with multiple students, this result holds if schools' preferences over students are responsive—informally, schools' preferences over sets of students can be completely described by their preferences over individual students—which is the case in New York City high school admission.

true preference ranking) to make space for worse schools where the student have a higher chance of getting in. A more detailed discussion on how this affects my empirical strategy can be found in Section 1.5.

In the Specialized Round, the DA mechanism is run separately for specialized high schools and regular schools. Due to the common priority among specialized high schools, the mechanism is essentially reduced to a serial dictatorship where the serial order is from the highest- to the lowest-scoring student. Specifically, the student with the highest score chooses her first-choice school, then the student with the second-highest score chooses her highest-ranked school among those with available seats, and so on. For regular school programs, priorities vary depending on the admission methods described above.

The resulting assignments are sent out to those who received specialized high school offers. Once these students have made their choices between their specialized high school offer and regular school offer, their assignments are finalized, they are removed from the system, and the numbers of remaining seats in regular school programs are adjusted accordingly. Then, the DOE proceeds to the Main Round, where DA is run again for the remaining students and regular school seats.

1.4 Data and Descriptive Analysis

1.4.1 Students

I link together three data sets from the New York City Department of Education containing the universe of students who applied to any public high schools during the 2010-2011 application year. The New York City High School Admission Process data contains all high school applicants with information on their zip codes, middle schools, scores on the Grade 7 English Language Arts (ELA) and mathematics state tests, submitted rank-ordered lists, priority ranks, and admission decisions. The SHSAT data set contains a subset of the above who applied to specialized high schools with additional information on the specialized high school admission process. Finally, the June biographic data supplements the aforementioned data sets with demographic information such as ethnicity, home language, eligibility for subsidized lunch, etc.

I make two restrictions for the main analysis sample, which is described in this section and used for the first step of structural estimation in Section 1.5. First, only students in New York City's public middle schools during the application year are included because demographic information is not available for students in private middle schools. Second, I only consider only those who appear in the Specialized or Main Round of the admission process. The sample used for the second step of estimation, henceforth referred to as the specialized high school sample, is a subset of the main analysis sample, restricted to only specialized high school applicants.

Table 1.1 summarizes the characteristics of students in the main analysis sample as well as the specialized high school sample. There is selection into applying to specialized high schools in terms of all observable characteristics. On average, specialized high school applicants score higher on the grade 7 state tests, come from more affluent neighborhoods, and are less likely to be eligible for subsidized lunch. In terms of scores, specialized high school applicants are doing better not only on average: Figure 1.4.1 shows that for the specialized high school sample, there is a higher density of students at every score above 1330. The three spikes on the right tail of each distribution are likely attributable to a property of the scoring

Table 1.1: Students' Characteristics for Application Year 2010-2011

	All High School Applicants	specialized high schools Applicants Only
Grade 7 State Math	672.7	695.03
	(32.7)	(28.7)
Grade 7 State ELA	661.5	677.7
	(28.9)	(33.0)
Zip code income ($)	54940	60211
	(25705)	(29001)
Subsidized lunch eligibility (%)	66.49	56.03
White	13.41	17.73
Asian	14.78	29.74
Black	31.34	26.70
Hispanic	39.88	25.11
N	66068	23255

Notes: Statistics is based on the analysis sample, consisting of 66068 students. Zip code income is median family income for each zip code from the American Community Survey 2007-2011.

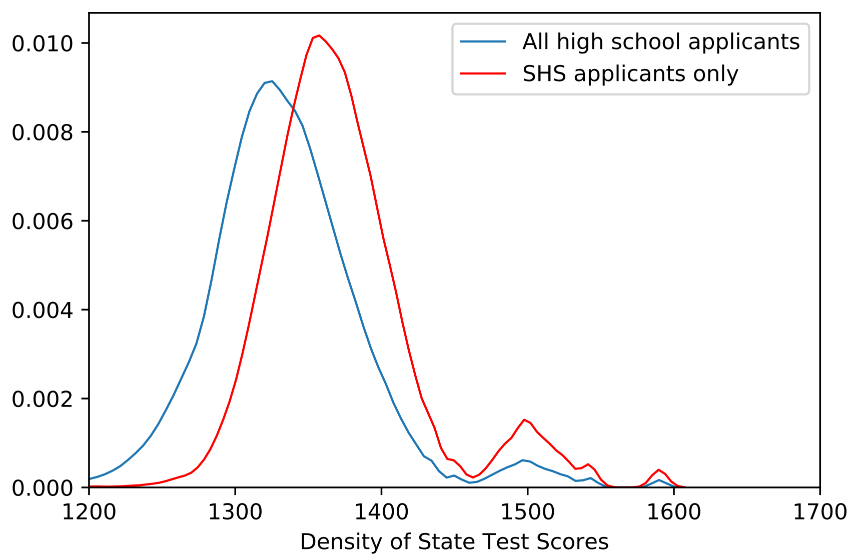


Figure 1.4.1: Distribution of State Test Scores - All High School Applicants vs. specialized high schools Applicants Only

scale of the test.[14] The percentages of White and Asian students among specialized high school applicants are higher than those in the full sample, and the percentage of Black and Hispanic students are lower. The differentials in compositions are especially pronounced for Asians and Hispanics.

In terms of students' choice, the amount of information I observe differs by student type as defined by their actions during the application process. For specialized high schools non-applicants (i.e., those who appear in the main analysis sample but not the specialized high school sample), only their rank-ordered lists over regular school programs and regular school

[14]For both tests, the score increments are much smaller for the lower range of scores compared to that between the top two scores (698 and 790 for ELA; 752 and 800 for mathematics), and there are bunchings at these top scores. Therefore, from left to right, the three spikes in the distribution of combined scores correspond to those who obtain the second-highest score in both tests, those who obtain the second-highest score on one test and highest score on the other, and those who obtain the highest score in both tests. This is unlikely due to the specific student population as test score data for other school years exhibit the same pattern.

offers are available. For specialized high school applicants, I also observe rank-ordered lists over specialized high schools. Among these students, some receive specialized high school offers, which I observe along with their choice between their specialized high school and regular school offers.

Table 1.2 describes the rank-ordered lists over regular school programs for all students in the analysis sample. Regular school programs that are ranked higher tend to be closer to the students' home, have a bigger grade 9 cohort, higher percentages of White students, lower percentages of students eligible for subsidized lunch, and higher percentages of admitted students who performed well in the Regent exams. The type of schools (based on their characteristics) that are included in the students' ranking differ by their own race, their baseline mathematics achievement as measured by their score on the grade 7 state test, and the median family income of their neighborhoods. Details for this heterogeneity can be seen in Table 1.7 of the appendix, where average school characteristics by the rank of student choice are shown for different subgroups.

From the observed responses between specialized high school and regular school offers summarized in table 1.3, we can see that it is not an obvious choice. A considerable fraction (19.08%) of students with specialized high school offers opted for a regular public school instead. Out of these students, 70% got their first-choice regular school programs, suggesting that the reason they applied to specialized high schools but did not accept their offers may be due to uncertainty in their admission chance at their preferred regular schools.

Table 1.2: Regular School Application - (Average) School Characteristics by Rank of Student Choice (All Students)

	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th
Students ranking choice (%)	100.00	91.01	86.39	79.01	69.65	59.08	48.21	39.10	30.83	24.69	19.51	14.61
Distance in miles: Mean	4.76	5.06	5.19	5.31	5.49	5.60	5.71	5.79	5.89	5.87	5.91	5.67
Distance in miles: Median	3.69	3.92	4.11	4.25	4.42	4.54	4.63	4.72	4.79	4.77	4.85	4.53
Size of grade 9	465.53	488.26	469.82	447.98	426.87	404.53	384.38	368.35	348.39	340.86	329.43	331.33
Percent white	17.96	16.64	15.67	14.47	13.38	12.53	11.43	10.98	10.15	9.72	9.08	8.43
Percent subsidized lunch	62.77	63.84	65.45	66.88	68.03	68.95	70.01	70.56	71.45	72.17	72.84	73.44
High math achievement (%)	11.78	9.98	9.65	9.25	8.84	8.64	8.08	7.78	7.40	6.96	6.48	5.82

Notes: Statistics is based on the analysis sample, consisting of 66068 students. The high math achievement measure for each high school is the percentage of admitted students in said school who score higher than 85% in the Regents mathematical exams in the 2009-2010 school year.

Table 1.3: Responses to Specialized High School Offers

	Number	Percent
Accept	3319	76.11
Regular Public Schools	832	19.08
Other Options	210	4.81
Total	4361	100

Notes: This table only accounts for offers to public middle school students, whom the estimation and counterfactual simulations focus on. There are additionally about 1000 other offers to private middle school students.

1.4.2 Schools

School characteristics for 2009-2010, the school year before the application year, come from the New York State Report Card Database publicly accessible on the New York State Education Department website. Program descriptions are taken from the official New York City High School Directory. Both types of information would have also been available for parents and students at the time of application.[15]A few schools and, therefore, their associated programs have their characteristics censored due to the Department's restriction on publishing data for groups with fewer than five students. Table 1.4 summarizes the characteristics of all high school programs for the 2009-2010 school year. On average, specialized high schools have a higher percentage of White students, lower percentage of students eligible for subsidized lunch, and markedly outperform regular high school programs in terms of their current students' achievement in the Regent mathematics exam.

[15]For school characteristics from the Report Card Database, there may be some slight difference between the later version that I use compared to the initial version available at the time.

Table 1.4: Program Characteristics for School Year 2009-2010

	Regular HS Programs	specialized high schools
Size of grade 9	430	464
Percent white	10.82	28.63
Percent subsidized lunch	71.21	42.00
High math achievement (%)	4.98	62.04
N	697	8

Notes: The high math achievement measure for each high school is the percentage of admitted students in said school who score higher than 85% in the Regents mathematical exams in the 2009-2010 school year.

1.5 Estimating Student Preferences

To quantify students' preferences in both the elite and regular high school sectors, I devise a two-step estimation procedure, exploiting the choice between one elite and one regular school that is observed for a subset of students.

1.5.1 Model and Estimation

Preferences over specialized high schools and regular schools both take the form of a random utility model in which students with different backgrounds have heterogeneous tastes over multiple dimensions of school characteristics. Although student preferences over the two types of schools have the same functional form, I allow the parameters to be different to reflect the fact that students may evaluate the characteristics of specialized high schools and regular schools differently.

Estimation step 1

The first step of the estimation procedure recovers the utility function for attending regular school programs from the rank-ordered lists over these programs submitted by all students in the analysis sample. Suppose $s = 1, ..., S$ is a regular public high school program. Student i's utility of attending program s is:

$$u_{is} = \underbrace{\delta_s + \sum_m \sum_n \alpha_{mn} x_{sm} z_{in} - d_{is}}_{V_{is}} + \lambda \epsilon_{is} \tag{1.5.1}$$

where x_s denotes program characteristics, z_i denotes students characteristics, and d_{is} denotes the distance between the student's home and the school. Following the convention in the school choice literature, I normalize the parameter for distance to -1 so that the unit can be interpreted in terms of the disutility of traveling. The program fixed effect δ_s represents how much an "average" student in the analysis sample[16] values program s (disregarding traveling distance). Parameters α_{mn} capture how students of different backgrounds evaluate program characteristics differently from the average. Peer preferences are captured by including racial and socioeconomic compositions as program characteristics. I assume that these compositions are uncorrelated to ϵ_{is} because the changes in preferences and, hence, behaviors of a single student in a large market should not make a difference in the aggregate compositions.

Each student always has the outside option to exit the NYC public school system (e.g.,

[16]This student belongs to the omitted category for all dummies (White and not eligible for subsidized lunch) and has characteristics that are equal to the average of the analysis sample for all continuous variables, which means all z_i are 0 in the standardized data. For a full description of such a student, see table 1.1.

attend a private high school) with utility:

$$u_{i0} = \lambda \epsilon_{i0}.$$

In the above notation, V_{is} is the systematic part of the utility function. The idiosyncratic preference shocks ϵ_{is} and ϵ_{i0} are independently and identically distributed according to type I extreme value distribution.

As in Abdulkadiroglu, Agarwal and Pathak (2017), the identification assumption I use is weak truth-telling, which consists of two parts. First, the order of schools in the submitted rank-ordered list is according to true preference. Second, ranked schools are preferred to all unranked schools. Unlike strict truth-telling, this assumption does not presume any relationship between unranked schools and the outside option of not attending any New York City public high school (i.e., attending a private school or a school outside of the city).

Given this assumption, the validity of which will be discussed later in this section, one can derive the likelihood of the observed data as follows. For student i who submits a rank-ordered list $r_i = (r_{i1}, .., r_{ik})$, the individual likelihood is the probability of observing said ranking:

$$\begin{aligned}
L_i(\boldsymbol{\theta_{reg}}) &= Pr(r_i) \\
&= Pr(u_{i1} > u_{i2} > ... > u_{ik} > u_{ij} \forall j \notin \{1, 2, ...k\}) \\
&= Pr(u_{i1} > u_{ij} \forall j \neq 1) \times Pr(u_{i2} > u_{ij} \forall j \notin \{1, 2\}) \times ... \\
&\quad ... \times Pr(u_{i1} > u_{i2} > ... > u_{ik} > u_{ij} \forall j \notin \{1, 2, ...k\}) \\
&= \prod_{s=1}^{k} \frac{exp\left(\frac{V_{is}}{\lambda}\right)}{\sum_{t=s}^{S+1} exp\left(\frac{V_{it}}{\lambda}\right)}
\end{aligned}$$

where $\theta_{\boldsymbol{reg}} = (\boldsymbol{\delta_{reg}}, \alpha, \lambda)$ is the vector of preference parameters. From the first part of the weak truth-telling assumption, observing r_i implies that the true utility ranking is $u_{i1} > u_{i2} > ... > u_{ik}$. From the second part of the assumption, the lowest-ranked regular high school program k has utility $u_{ik} > u_{ij}$, where j represents indices of unranked schools and the outside option. The third line follows from the independence of irrelevant alternatives (IIA) property of the logit model, and the last line utilizes the analytical form of the logit probability.

Thus, the log-likelihood for a sample of I students, each faced with $S+1$ alternatives (S programs and one outside option), is:

$$\mathcal{L}(\theta_{\boldsymbol{reg}}) = \sum_{i=1}^{I}\sum_{s=1}^{k}\frac{V_{is}}{\lambda} - \sum_{i=1}^{I}\sum_{s=1}^{k}\log\left[\sum_{t=s}^{S+1} exp\left(\frac{V_{it}}{\lambda}\right)\right]$$

I estimate the parameters $\theta_{\boldsymbol{reg}}$via maximum likelihood and use the standard asymptotic theory for the estimation of the variance-covariance matrix of the estimates.

Estimation Step 2

The second step of the estimation procedure recovers the utility function for attending specialized high schools from the rank-ordered lists over them submitted by specialized high school applicants in the specialized high school sample. Suppose $j = 1, ..., J$ is a specialized high school. Student i's utility of attending specialized high schools j is:

$$u_{ij} = \delta_j + \sum_m \sum_n \beta_{mn} x_{jm} z_{in} - \gamma d_{ij} + \hat{\lambda}\epsilon_{ij}, \tag{1.5.2}$$

with notations being similar to equation (1.5.1).

For utilities for the two types of schools to be comparable, their normalizations should be compatible in both scale and location.[17] To ensure that utilities are of the same scale, the coefficient of the idiosyncratic preference shock in equation (1.5.2) is fixed at $\hat{\lambda}$, the estimated value of λ in the utility function for attending regular school (equation (1.5.1)). Given this normalization, the coefficient of distance to specialized high schools $-\gamma$ potentially differs from -1 (that for regular schools). To ensure that utilities are comparable in terms of location, I link the two sets of utilities by the comparison of utilities between one specialized high school and one regular school. For students who are allowed to choose between a specialized high school offer and a regular school match, these two schools are used, and the ranking is based on the student's enrollment decision. For students who do not have the opportunity to choose, I assume that they prefer their highest-ranked specialized high schools to their lowest-ranked regular school. Otherwise, for this student, all ranked regular schools must have been better than all specialized high schools, which means they would not have applied to specialized high schools in the first place given any non-negative cost of application.

For student i who submits a rank-ordered list $r_i = (r_{i1}, .., r_{ik})$ over specialized high schools and chooses specialized high schools m over regular program s, the individual likelihood is now the probability of simultaneously observing said rank-order list and choice:

$$\begin{aligned} L_i(\boldsymbol{\theta_{SHS}}, \boldsymbol{\hat{\theta}_{reg}}) &= Pr(r_i \text{ and } u_{im} > u_{is}) \\ &= Pr(r_i) \times Pr(u_{im} > u_{is}) \\ &= \prod_{j=1}^{k} \frac{exp\left(\frac{V_{ij}}{\hat{\lambda}}\right)}{\sum_{t=j}^{J} exp\left(\frac{V_{it}}{\hat{\lambda}}\right)} \times \frac{exp\left(\frac{V_{im}}{\hat{\lambda}}\right)}{exp\left(\frac{\hat{V}_{is}}{\hat{\lambda}}\right) + exp\left(\frac{V_{im}}{\hat{\lambda}}\right)}, \end{aligned}$$

where $\hat{V_{is}}$ is the systematic part of the utility of attending regular program s, evaluated at the

[17]This is because given an ordinal ranking, cardinal utilities are only identified up to an affine transformation.

estimated $\hat{\theta}_{reg}$. We can take the product of unconditional probabilities as in the second line due to the IIA property of the logit model. The first part of the product is derived, under the assumption of weak truth-telling, in a similar fashion to that in the regular school case, and the second part is the logit probability of choosing m over s.

Symmetrically, if student i chooses regular program s over specialized high schools m, her individual likelihood would be:

$$L_i(\theta_{SHS}, \hat{\theta}_{reg}) = \prod_{j=1}^{k} \frac{exp\left(\frac{V_{ij}}{\hat{\lambda}}\right)}{\sum_{t=j}^{J} exp\left(\frac{V_{it}}{\hat{\lambda}}\right)} \times \frac{exp\left(\frac{\hat{V}_{is}}{\hat{\lambda}}\right)}{exp\left(\frac{\hat{V}_{is}}{\hat{\lambda}}\right) + exp\left(\frac{V_{im}}{\hat{\lambda}}\right)}.$$

The log-likelihood for a sample of I students, where i chooses specialized high schools m_i for $i = 1, .., n$ and h choose regular program s_h for $h = n+1, ..., I$, is:

$$\begin{aligned}\mathcal{L}(\theta_{SHS}, \hat{\theta}_{reg}) = & \sum_{i=1}^{I}\sum_{j=1}^{k} \frac{V_{ij}}{\hat{\lambda}} - \sum_{i=1}^{I}\sum_{j=1}^{k} \log\left[\sum_{t=j}^{J} exp\left(\frac{V_{it}}{\hat{\lambda}}\right)\right] \\ & + \sum_{i=1}^{n} \frac{V_{im_i}}{\hat{\lambda}} + \sum_{i=n+1}^{I} \frac{\hat{V}_{is_i}}{\hat{\lambda}} - \sum_{i=1}^{I} \log\left[exp\left(\frac{\hat{V}_{is_i}}{\hat{\lambda}}\right) + exp\left(\frac{V_{im_i}}{\hat{\lambda}}\right)\right]\end{aligned}$$

I also estimate the parameters θ_{SHS} via maximum likelihood. In computing the standard errors for the estimated parameters $\hat{\theta}_{SHS}$, I need to account for the fact that the estimates from the first step are used as an input. Thus, I take independent and identically distributed draws from the distribution of $\hat{\theta}_{reg}$ and re-estimate the second step for each draw. The resulting distribution of $\hat{\theta}_{SHS}$ is used to calculate its standard errors.

The Validity of Weak Truth-telling Assumption

As previously stated in the description of the student-proposing Deferred Acceptance (DA) mechanism, when students are allowed to rank all available schools as is the case for specialized high schools, it is a weakly dominant strategy for students to report their true preferences (strict truth-telling). Consequently, both parts of the weak truth-telling assumption are likely to hold for specialized high schools. Reporting the correct order among those that are ranked (part 1 of weak truth-telling) is still a dominant strategy when the length of the rank-ordered list is restricted as in the case for regular school programs. However, because students can only rank twelve out of nearly 700 programs, the second part may be violated because it may be optimal for students to drop some schools that they are unlikely to get in to make space for less-preferred schools where they have a better chance (Haeringer and Klijn, 2009). Only 14.61 % of students ranked all twelve choices, which somewhat alleviates concerns over students dropping true top choice(s) due to restricted rank-ordered lists.

The possibility of payoff-insignificant mistake as introduced by **?**, that is, students dropping schools they are unlikely to get in even when they still have space for them in the rank-ordered list, is still present for both types of schools. It must be noted, however, that there is more uncertainty in admission chances in my setting compared to the empirical setting of their paper, where priority is based on a score and students already know their own score before submitting rank-ordered lists. Although specialized high schools also use one single score to determine priority, specialized high school applicants have to submit their rank-ordered lists before learning their own scores. Yet more uncertainty is present for regular school programs, some of which based priority on a lottery. Even for those which have deterministic criteria for assigning priority like screened programs, how these different criteria are weighed is somewhat opaque from the students' perspective.

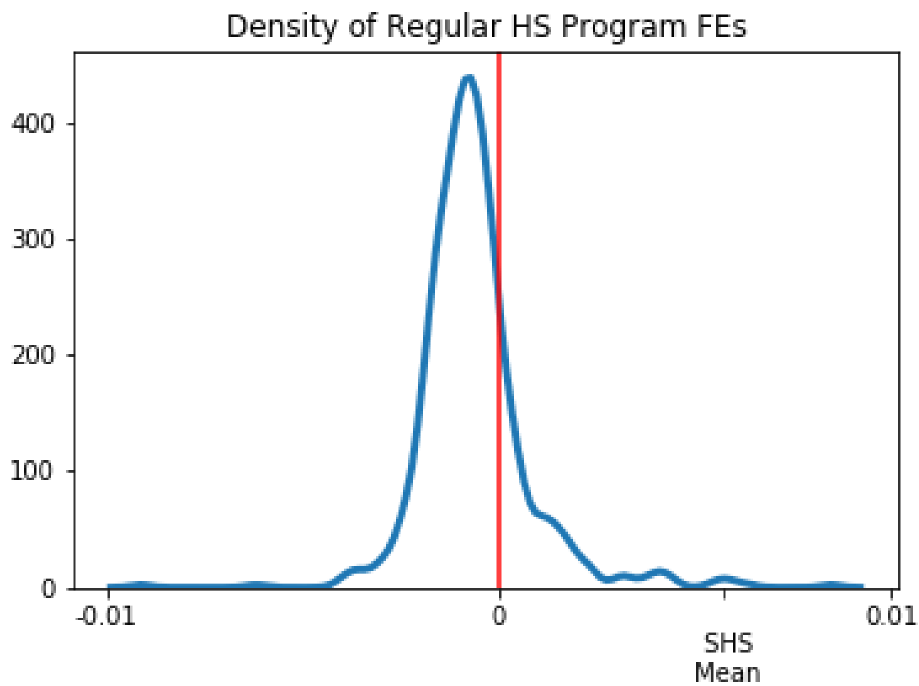


Figure 1.5.1: Distribution of Estimated Fixed Effects for Regular High School Programs

1.5.2 Preference Estimates

The distribution of estimated program fixed effects shown in figure 1.5.1 looks approximately normal. As expected, the mean fixed effect for specialized high schools falls within the right tail of this distribution, demonstrating the fact that these elite schools are highly valued. There is some small mass of programs around as well as above this point in accordance with the observation in Section 1.4 that there exist regular high school programs that are preferred to some specialized high schools.

The estimates for the coefficients of the utility functions for both regular high school programs and specialized high schools are reported in Table 1.8. In both estimations, schools' characteristics are centered around the average characteristics of the population of regular high school programs as seen in the first column of Table 1.4. Students' continuous characteristics are standardized based on the mean and standard deviation in the main analysis sample. There are some notable patterns of differences between the coefficients for regular school programs and specialized high schools. First, while students of all races tend to prefer regular programs with higher percentages of White students, the opposite is true for specialized high schools, although it is only significant for Hispanics and students in the other race category.[18] Second, the effects of the size of grade 9 are reversed for students of all races except Black. Third, although regular school programs tend to be more attractive to students with higher baseline achievement in both ELA and mathematics, the corresponding effect is only significant for students with higher baseline mathematics when it comes to specialized high schools.

[18] This category includes multiracial and Native American students.

1.6 The Effects of Affirmative Action Plans

There are three main findings from my simulation of the affirmative action plans currently considered. First, there is a trade-off between improving diversity and maintaining student quality in specialized high schools. The Chicago plan does poorly in terms of increasing overall racial diversity but preserves the quality of incoming students. In contrast, the Top 7% plan increases diversity substantially at the cost of student quality. Second, the Chicago plan reduces the differences among schools within the elite sector, whereas the Top 7% plan increases these differences. Third, both plans also change regular school assignments and thereby the student quality in regular high school programs.

1.6.1 Implementing Counterfactual Policies

I simulate the effects of each affirmative action plan in the first year that it is implemented, using the estimated parameters presented in the previous section. The analysis is based on two main assumptions. First, in this short-run scenario, I assumed that parents and students are myopic with respect to the possible changes in the school characteristics, including the composition of incoming students; that is, their preferences over schools do not change. Subsequent versions of the paper will address the effects in medium- and long-run when parents and students also adjust their preferences in response to the policy. Second, to recover preferences over specialized high schools of students who did not apply, I assume that these can be extrapolated from students who did apply with the same observables. Rankings are constructed using the systematic part of the utility function; that is, a student is said to prefer one school over another in the constructed rankings if they are expected do so on average across all possible realizations of the error terms.

In both counterfactuals, I run two rounds: Specialized and Main, as in the status quo. In the Specialized Round, the selection rules are different for specialized high schools, leading for different offers. For regular school programs, the Specialized Round proceeds in the same way as in the status quo. Given the new specialized high school offers, students who receive them are allowed to choose between one specialized high school and one regular school program. If a student would have chosen between these exact same schools in the status quo, his or her actual choice is used in the counterfactuals. If not, then I predict this choice by comparing the systematic parts of the utilities of going to these two schools. Given the difference in the outcomes of the Specialized Round between the counterfactuals and the status quo, the inputs for the Main Round are different. However, the procedure remains the same.

I assume that the capacities of specialized high schools used in the Specialized Round under the status quo are maintained in both counterfactuals, but the total numbers of incoming students, given their responses to specialized high school offers, can change. In particular, although specialized high school offers fill up their capacities, some students decline these offers in favor of other options. In the current system, there are no waiting lists dependent on the number of accepted offers.[19] I maintain this feature under the affirmative action policies, which means if more or fewer students decide to accept their specialized high school offers, the total numbers of admitted students under the counterfactual policies may indeed differ from those in reality.

When considering the effects on regular schools, I assume that the specialized high school

[19] The current analysis does not consider the Discovery Program, which allows students belonging to disadvantaged who barely miss the cutoff to apply for conditional offers at specialized high schools. However, the numbers of seats for this program are committed to before the entire application process and thus do not depend on how many students accept their specialized high school offers.

offers and choices between specialized high schools and regular schools (if applicable) do not change for private middle school students. This is because there are no identification numbers that allow matching private school students between the specialized high school and regular school admission data sets. In both counterfactuals, since these students do experience changes in specialized high school offers (or lack thereof), thereby influencing the school assignments in regular schools, we can consider the current results with respect to the changes in regular school admission as a lower bound for the actual changes. Because the number of private school students being affected under the Chicago plan is much lower than that in the Top 7% Plan, the extent to which this assumption influences the results is different for the two plans. As such, we should view within-regular-school outcomes for the two plans as two examples of the types of spillover effects that are possible instead of comparing them head-to-head.

Chicago Plan

In the Chicago plan, some seats are assigned purely on merit, and some are reserved for each of four tiers of socioeconomic status determined by students' neighborhood. An implementation in NYC would still require those who wish to gain admission to specialized high schools to take the SHSAT, which would be used as the criterion for merit seats as well as a tie-breaker among students of the same tier.

I assume that 40% of total capacity at each specialized high school is set aside as merit seats, and the remaining seats are divided equally among four tiers. Tier 1 student has the highest priority for tier 1 seats, and so on. Similar to the current policy in Chicago, when running the Deferred Acceptance mechanism to determine specialized high school offers, I fill all merit seats first, then tier seats.[20] In theory, if the applicants in one tier run out before the seats in that tier are filled, students from other tiers would be admitted, which means the precedence order in which different tiers are filled is also important. In practice, all tier seats are oversubscribed, so any precedence order among the four tiers gives the same result.

Tiers are determined based on students' zip codes. For public middle school students, actual zip codes are used. For private middle school students, individual zip codes do not appear in the specialized high school admission data, so the tier for each is proxied by assigning his or her middle school the average tier of all specialized high school applicants from that school. Each zip code is given a tier depending on its median family income. Tier 1 has income in the lowest quantile compared to the rest of New York City; tier 4 has income in the highest quantile. To avoid dependency on the population of students who submitted public high school applications, income quantiles are calculated based on the population of families in New York City in the American Community Survey 2007-2011.

Top 7 % Plan

In the Top 7% plan, I assume that academic performance is measured by the combined English Language Arts (ELA) and mathematics scores on the Grade 7 state test. The top 7% of students in each public middle school based on this measure obtain automatic eligibility for one of the specialized high schools. The specific school assignments among these students are determined by running the Deferred Acceptance mechanism, using Grade 7 state test scores as the criterion of the common priority of all specialized high schools. For the 3474 students who applied before, I use their submitted rank-ordered lists over specialized high

[20] In Appendix 1.9, I explore different precedence orders as well as different proportions of total capacity set aside as merit seats. I find that filling merit seats first, as in the main specification, does in fact favor the more disadvantaged tier(s).

schools. For the 1280 students who never applied, I simulate their rank-ordered lists from the estimated preferences of those who applied. After the top 7% public middle school student have taken their seats, there are 652 seats (approximately 12% of the total capacity) left for private school students. I do not consider the allocation of these seats due to two reasons. First, the demographic information of this group is unavailable, thereby prevent me from any conclusion regarding their compositions. Second, even if I establish their assignments in specialized high school admission, they cannot be linked to regular school admission.

1.6.2 Students' Movements

The Chicago plan causes reshuffling within specialized high schools but little flow of students between specialized high schools and regular schools, whereas the Top 7% plan causes considerable flows of students between specialized high schools and regular schools. In both cases, the exchange of students between the two sectors also results in extensive changes in regular school assignments for those who neither gain nor lose specialized high school offers.

As shown in table 1.5, there are considerably more exchanges of students between the elite and regular sectors under the Top 7% plan. The second row for each plan corresponds to the proportions of students out of all those from public middle schools who move from specialized high schools to regular schools. Under both policies, only small fractions of Black and Hispanic students lose access to specialized high schools (0.08% and 0.05% for Chicago, and 0.56% and 0.49% for Top 7%) contrasted with considerably more movement of White and Asian students. However, the plans differ starkly in terms of the magnitude of the flows of these students from specialized high schools to regular schools: approximately 11 times more White students and 22 times more Asian students are displaced from specialized high schools

Table 1.5: Comparison of School Assignments between the status quo and Affirmative Action Plans

	White	Asian	Black	Hispanic
School Assignments - status quo vs. Chicago Plan (% of total)				
Always specialized high schools	8.43	20.23	0.90	0.84
Specialized to regular	0.57	0.64	0.08	0.05
Regular to specialized (offer gained)	0.65	1.59	0.46	0.48
Regular to specialized (better offer)	0.60	0.92	0.10	0.10
Different regular schools	28.89	30.32	49.89	49.04
Same regular schools (Main Round)	55.64	43.24	48.12	48.98
Same regular schools (Specialized Round)	2.06	2.02	0.16	0.17
School Assignments - status quo vs. Top 7% Plan (% of total)				
Always specialized high schools	2.62	6.67	0.42	0.39
Specialized to regular	6.38	14.20	0.56	0.49
Regular to specialized (offer gained)	4.68	4.83	5.28	4.77
Regular to specialized (better offer)	0.72	0.87	0.07	0.10
Different regular schools	28.34	29.90	47.74	46.72
Same regular schools (Main Round)	53.86	42.10	45.61	47.14
Same regular schools (Specialized Round)	0.70	0.50	0.06	0.06
Total Number of Students	9011	9813	21306	26893

Notes: The last row shows the total numbers of public middle school students in four racial categories. For each column, the corresponding number represents 100%.
In terms of school assignments, there is one category omitted from the table for students who would always go to La Guardia (a performing art school) or exit the public school system.
Theoretically, those who move from specialized high schools to regular schools can be divided into those who lose offers and those who choose regular school because they receive worse specialized high school offers. I do not make the distinction in this table because the second category is very small (0 and 7 students, across all races, for Chicago and Top 7%, respectively).
Students who move from regular schools to specialized high schools are divided into 1) those who gain specialized high school offers under the affirmative action plan when they would not have had any offers under the status quo and 2) those who would get a better offer under the affirmative action plan and thus choose specialized high schools even though they would have chosen regular schools before. Students who attend the same regular school are divided into 1) those who are assigned in the Specialized Round (i.e. they choose regular schools over specialized high school offers) under both status quo and affirmative action plan and 2) those who are assigned in the Main Round under both scenarios.

under the Top 7% plan compared to Chicago plan. As a result, the total inflow of students from regular schools to specialized high schools is also much larger under Top 7%. There are two types of inflow: 1) those who gain specialized high school offers under the affirmative action plan when they would not have had any offers under the status quo and 2) those who would get a better offer under the affirmative action plan and thus choose specialized high schools even though they would have chosen regular schools before. Most of the disparity between the two plans comes from the first type. First, there are substantial differences in magnitude between the two plans across all the races considered (0.48-1.59% and 4.68-5.28% for Chicago and Top 7%, respectively). Second, a markedly higher percentage of Asians gain access compared to other races under the Chicago plan, while the percentages for all four races are similar under the Top 7% plan. Note that in absolute numbers, there are still many more Black and Hispanic than White and Asian students who gain access in the Top 7% plan.

Substantial portions of students are also reshuffled among regular school programs under both plans. Specifically, Black and Hispanic students are more affected both in percentage and absolute terms. Surprisingly, even a small exchange between the two sectors in the Chicago plan causes substantial spillover effects in regular schools. Although it is not appropriate to compare the outcomes of the two plans here as a way of evaluating their merits against one another,[21] the results here show that exchanges of different sizes can cause similar disturbances in regular school admission, and the identities of the individuals being exchanged (i.e., both their submitted preferences and priorities at various schools) most likely play an important role.

[21] This is due to an assumption on private middle school students. For more details, please refer to Subsection 1.6.1.

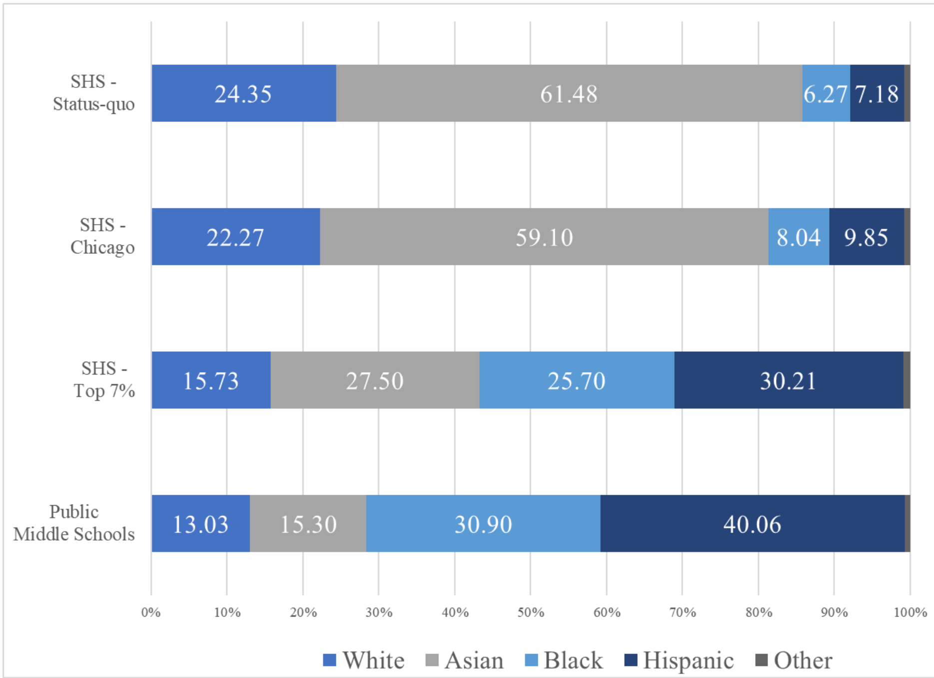


Figure 1.6.1: Racial Composition of specialized high schools under Different Policies

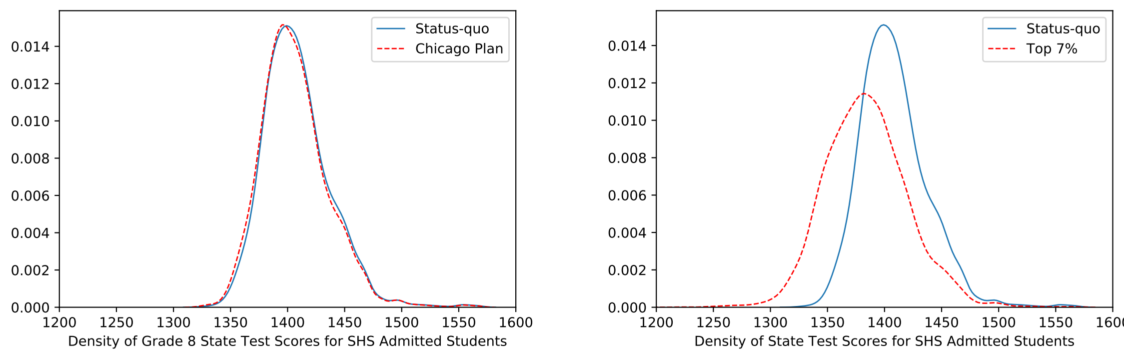
Notes: This figure only takes into account public middle school students for whom racial identities are known, so 100% for each bar corresponds to the total number of such students who would enroll in specialized high schools in each scenario. For comparison purpose, I include the last bar that represents the population of all public middle school students, the main source of high school applicants.

1.6.3 Effects on specialized high schools

The overall diversity and student quality in specialized high schools remain relatively unchanged under the Chicago plan, whereas there is a substantial increase in the proportions of Black and Hispanic students and a decrease in overall student quality under the Top 7% plan.

In terms of racial composition, both plans increase the percentages of Black and Hispanic students in specialized high schools but with remarkably different orders of magnitude. Figure 1.6.1 shows that the changes in the Chicago plan are small enough (all less than 5 percentage points) that the resulting composition still look very similar to that under the status quo. Meanwhile, the changes in the Top 7% plan are all very drastic: the percentage of Black students increases from 6.27% to 25.70% and that of Hispanic students from 7.18% to 30.21%. While the proportions of White and Asian both decrease, Asian students, who are most over-represented in specialized high schools under the status quo, experience the most substantial change that more than halves their percentage. The resulting racial composition in specialized high schools under the Top 7% plan thus becomes much closer to that of the public middle school population, the main source of high school applicants, although Blacks and Hispanics are still under-represented.

The effects on compositions are heterogeneous across individual specialized high schools, as seen in Table 1.9. Under the Chicago plan, this is mostly due to shuffling within specialized high schools in the direction of reducing the dispersion in racial compositions across specialized high schools. First, although all specialized high schools experience decreases in percentages of Whites, the greatest changes in terms of percentage points are at the two most predominantly White schools originally. Second, all but the three schools with the lowest fractions of Asian originally experience decreases in percentages of Asians. Finally, percentages of Black students increase in all schools except for one where Blacks would have been most well-represented among specialized high schools under the status quo. Under the Top 7% plan, even though percentages of Black and Hispanic students increase across the board, the high variance in the magnitude of the increases cause more dissimilarities among specialized high schools. On one end of the spectrum, Stuyvesant still has a severe under-representation of Blacks and Hispanics, which only account for 9.98% and 9.70% of its population, respectively. In contrast, more than 50% of Bronx Science's and Brooklyn Latin's students under the Top 7% plan are Black, which would be an over-representation if the population of public middle school students is used as a benchmark.



(a) Chicago Plan

(b) Top 7% Plan

Figure 1.6.2: Distribution of Grade 8 State Test Scores of Admitted Students to specialized high schools

Table 1.6: Grade 8 State Test Scores of Admitted Students to Specialized High Schools

	status quo	Chicago Plan		Top 7% Plan	
	Mean	Mean	Difference	Mean	Difference
Stuyvesant	1427.5	1419.1	-8.4***	1412.2	-15.3***
Bronx Science	1410.8	1407.7	-3.2**	1386.1	-24.8***
Queens Sciences	1405.4	1404.3	-1.1	1392.4	-13.0***
HS of American Studies	1400.5	1404.8	4.3	1377.6	-22.9***
Staten Island Tech	1398.1	1398.9	0.8	1373.0	-25.1***
HS of Maths, Science & Engineering	1404.1	1400.3	-3.8*	1395.0	-9.2***
Brooklyn Tech	1400.7	1403.7	2.9***	1368.9	-31.8***
Brooklyn Latin	1384.8	1394.3	9.4***	1362.7	-22.1***
All	1408.3	1406.4	-1.8***	1384.1	-24.2***
Standard Deviation (specialized high schools Admitted Students under status quo): 29.1					

Student quality, as measured by Grade 8 state test scores, remains almost the same under the Chicago plan but decreases considerably under the Top 7% plan. Figure 1.6.2 shows that the distribution of test scores of admitted students under the Chicago plan tracks that under the status quo quite closely, which is expected given that the overall specialized high school student population changes very little. In contrast, the distribution clearly shifts left under the Top 7% plan compared to the status quo. As seen in Table 1.6, the decreases in mean score of specialized high schools are significant at the 1% level for both plans, but the magnitude for Chicago plan is negligible whereas the decrease for the Top 7% plan is nearly one standard deviation of scores for the status quo specialized high school student population. In terms of school-specific average student quality, some increase and some decrease under the Chicago plan, leaving the specialized high schools more similar to each other than before, while average qualities decrease for all specialized high schools under the Top 7% plan.

1.6.4 Effects on Regular High School Programs

Due to considerable reshuffling within regular school programs, both the distributions of racial compositions and student qualities across these programs change.

The top two histograms in Figure 1.6.3 show that concerning the percentages of White and Asian students across programs, the three policies (status quo and two affirmative action plans) look quite similar. However, concerning the percentages of Black and Hispanic students, as shown in the bottom histograms, there are noticeable changes compared to the status quo, especially for the Top 7% plan in which there are more schools with very low percentages (0-10%) of these two races.

In terms of student quality as measured by mean scores on Grade 8 state test of admitted

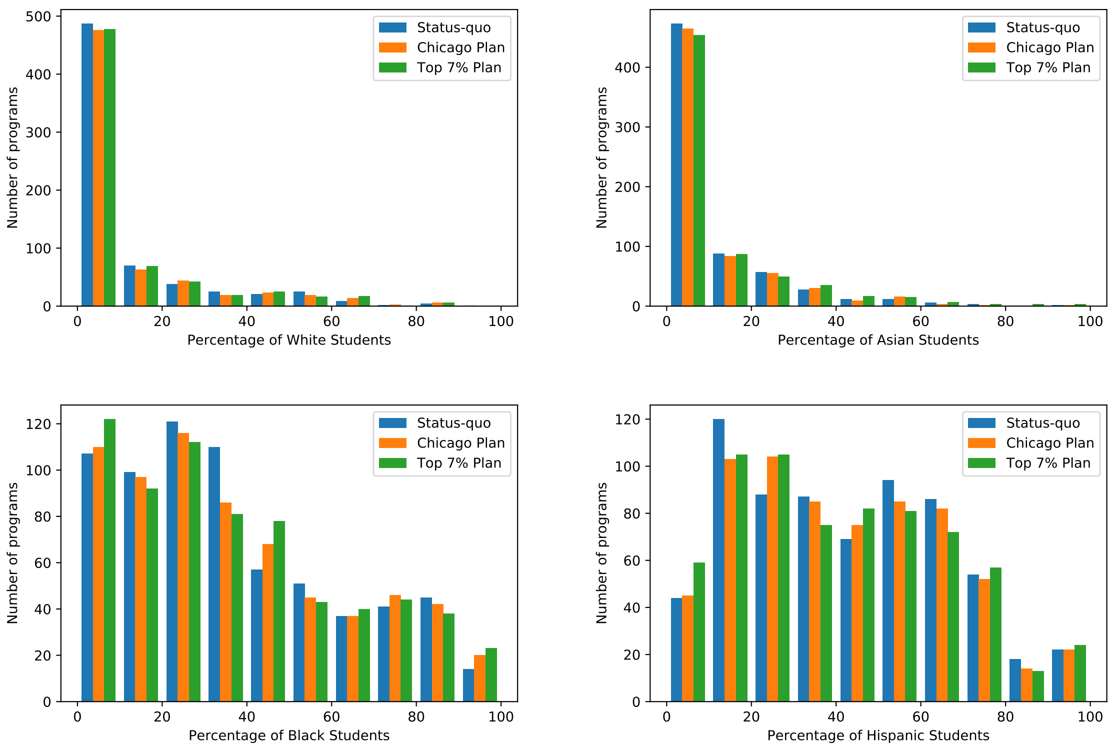
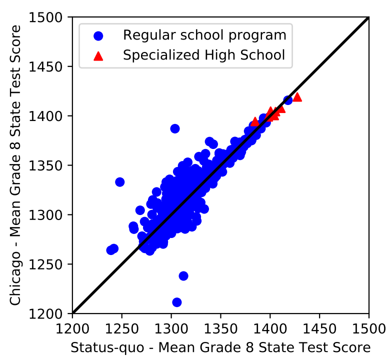
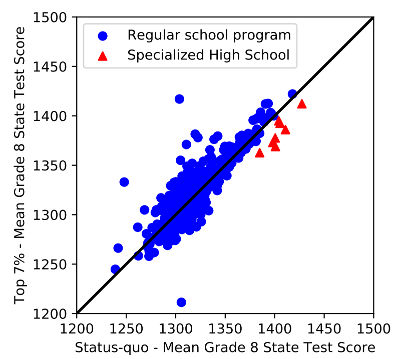


Figure 1.6.3: Distribution of Racial Compositions across Regular High School Programs



(a) Chicago Plan



(b) Top 7% Plan

Figure 1.6.4: Mean Grade 8 State Test Scores of Admitted Students - status quo v.s. Affirmative Action Plans

students, the effects of the two plans are similar: about 52% of regular high school programs experience increases while the rest experience decreases or no change. Figure 1.6.4 visualizes this result. When one considers only changes that are statistically significant at the 5% level, the mean scores increase for about 9% and 10% of the programs under Chicago and Top 7% plans, respectively, and decrease for about 3% and 4% of the programs.

1.6.5 Heterogeneous Changes in Peer Quality

In the previous two subsections, I have looked at changes in student quality at the school level. In this subsection, I will consider peer quality at the student level. For each student, the peer quality is measured by the mean student quality at his or her assigned school, excluding the student him-or-herself.

Figure 1.6.5 shows the average Grade 8 state test scores of peers for students in the ten different deciles of scores in the entire population of public middle school students. The changes in peer quality under both affirmative action plans (compared to the status quo) are in the same direction for all deciles except for the highest one. In particular, students in the two lowest deciles experience lower peer quality under the counterfactuals (although their peers still outperform them on average). Peer quality for students in the third and fourth deciles remains relatively stable. Meanwhile, peer quality increases for students in the fifth to ninth deciles. The magnitudes of the increases are similar across plans for the fifth and ninth deciles, whereas they are higher for the Top 7% plan for the sixth to eighth deciles. Students in the highest decile, in other words, the best-performing students in the population, experience an increase in peer quality under the Chicago plan but a decrease under the Top 7% plan.

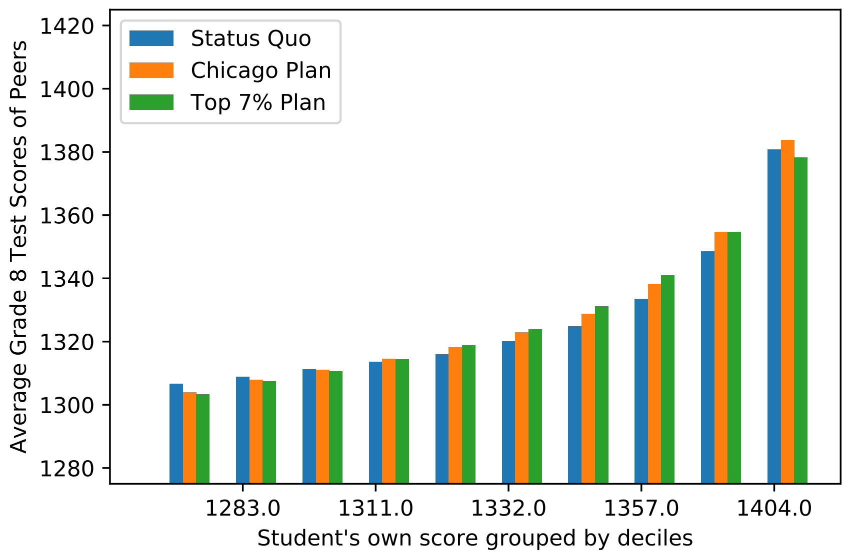


Figure 1.6.5: Changes in Peer Quality in High Schools - All students

Notes: Peer quality is measured by average Grade 8 state test scores of peers. Deciles are based on the entire population of public middle school students. Deciles are based on the entire population of public middle school students.

To further investigate the heterogeneity of the effects, I divide students into groups based on the changes (or lack thereof) in their high school assignments. Figures 1.8.4 and 1.8.5 show the changes in peer quality of these different groups, compared to in the status quo, for the Chicago and Top 7% plans, respectively. Deciles are still based on the entire population of public middle school students, so if certain deciles are missing for a particular group, it means there are no students from said group who belong to those deciles. Although there may be overlaps, groups with the same label are generally different across affirmative action plans. For instance, "Always Specialized High Schools" for each plan refers to students who are assigned to specialized high schools in both the status quo and said plan (regardless of their assignments in the other plan). Despite this disparity, it is helpful to look at them side-by-side to highlight the fact that the same type of assignment changes under the two plans can mean different experiences in terms of peer quality.

As seen in Panel (a) of each figure, students who stay in specialized high schools under both the status quo and the Chicago plan all belong to one of the top six score deciles in the broader population. Changes in peer quality are heterogeneous across the students' own score deciles: students in the lowest decile (out of the six) experience a sizable increase in peer quality, those in the next four deciles experience moderate increases, and those in the highest decile experience a small decrease. This is consistent with the within-specialized-high-school reshuffling documented in Subsection 1.6.2. Meanwhile, students who stay in specialized high schools under both the status quo and the Top 7% plan, who are part of the top four score deciles in the population, all experience considerable decreases in peer quality due to the drop in overall student quality in specialized high schools under this plan.

Panel (b) shows, for each plan, that across all relevant deciles, students who move from specialized high schools (under the status quo) to regular schools (under the plan) experience considerable decreases in peer quality, which is expected given the marked difference in student quality between specialized high schools and most regular schools. The effects are worse for lower-scoring students in the groups. Notably, under Top 7%, those scoring between 1326 and 1327 points (i.e., in the fifth decile of the large population and lowest relevant decile for the "Specialized to Regular" group under this plan) go from peers who perform much better than them on average (at 1385 points) to having peers who perform worse than them on average (at 1299 points).

Panels (c) and (d) demonstrate that students who move from regular schools to specialized high schools, whether due to gaining offers or receiving better offers, enjoy higher peer quality under both plans compared to under the status quo, regardless of their own scores. The sizes of the gains are generally greater for lower-scoring deciles. In two cases, the increases in peer quality for the lowest relevant deciles in the groups are so large that the students in these deciles end up with better-performing peers than those in the next deciles. First, those gaining offers to specialized high schools under the Chicago plan who are in the fifth score decile[22] would have had worse peers compared to the other deciles under the status quo but have the best-performing peers under the plan. Second, those gaining better offers under the Top 7% plan who are in the seventh decile would have had only slightly better peers than those in the eighth decile under the status quo but now have much better peers that are more similar to the peers of the two highest deciles.

From Panels (e) and (f), we can see a similar trend under both plans for students who end up in a different regular schools as well as those who are assigned the same regular school during the the Main Round. Specifically, those in the lowest deciles are exposed to lower peer quality under affirmative action compared to the status quo, whereas higher deciles

[22] The graph shows a lower score decile for this group, but there is only one student in that decile.

either experience little changes or increases. This pattern agrees with that stated above for the broader population of all public middle school students under both plans for the first nine deciles. For the highest decile, it agrees with the overall effect under the Chicago plan. We have the opposite overall effect for this decile under the Top 7% plan due to the effect for these groups being canceled out by that for those moving from specialized to regular schools, who are much more numerous here than under the Chicago plan.

1.6.6 Factors Driving the Results and Policy Implications

Three main factors drive the differences in the outcomes of the two policies: the pools of students to select from, the "merit" criteria, and the division of students into subgroups.

First, the pool from which students are selected into specialized high schools is much more restrictive in the Chicago plan, where the students must have taken the SHSAT to be considered. On the contrary, the Top 7% plan considers all aspiring high school students. In fact, the 1280 students who are eligible for specialized high schools under this plan but never took the SHSAT account for 26.9% of the total offers and 48.0% of the policy beneficiaries. Given that it is not feasible to incentivize disadvantaged students into taking the SHSAT by explicitly favoring them with bonus scores or similar measures, this result demonstrates that it is hard to get far in improving specialized high school diversity without abolishing the SHSAT barrier.[23]

Second, the two plans currently use two different criteria to define "merit," with the Chicago plan using the SHSAT and Top 7% plan using grade 7 state test scores. To investigate this, I restrict attention to only specialized high school applicants, for whom both

[23] At the same time, it must be noted that if state tests are used, the current specialized high school applicants will most likely increase their effort in preparing for the state tests, which may still give them some advantage.

types of scores are available. Given this population as the selection pool, a policy that selects the top 7% of each public middle school based on state test scores versus one that selects the top 7% based on SHSAT scores result in an overlap in specialized high school eligibility that comprises only 29% of the total number of specialized high schools seats. This is because the quantile rank to which a student belongs within this population can be quite different depending on the score being used. Figure 1.8.1 plots the quantile rank of all specialized high school applicants based on SHSAT scores against that based on grade 7 state test scores. Although there is some concentration of students along the 45-degree line, especially for lowest and highest quantiles, there is also significant dispersion. There exist students who may rank very high on one test but very low on the other.

Third, the policy outcomes depend on the dispersion of race as well as student quality across subgroups into which students are assigned. On one hand, a policy is more likely to increase diversity if the racial compositions are very different across subgroups. On the other, too much dispersion in student quality across subgroups means that even the best students selected from lower-performing subgroups are likely to be low-performing compared to the overall population, thereby reducing the average quality of those admitted to the elite schools.

The first of the two aforementioned points explains why the Top 7% plan is so much more effective than the Chicago plan at improving overall specialized high school diversity. Figure 1.8.2 illustrates severe racial segregation within New York City public middle schools, which are the relevant subgroups for Top 7%. In the first two panels, we can see that there is a large number of schools where there are no White or Asian students. Thus, these schools contribute greatly to the number of Black and Hispanic students selected for specialized high school eligibility. In contrast, the differences in racial compositions across neighborhood tiers, the relevant subgroups for the Chicago plan, are more subdued, as seen in Figure 1.8.3.

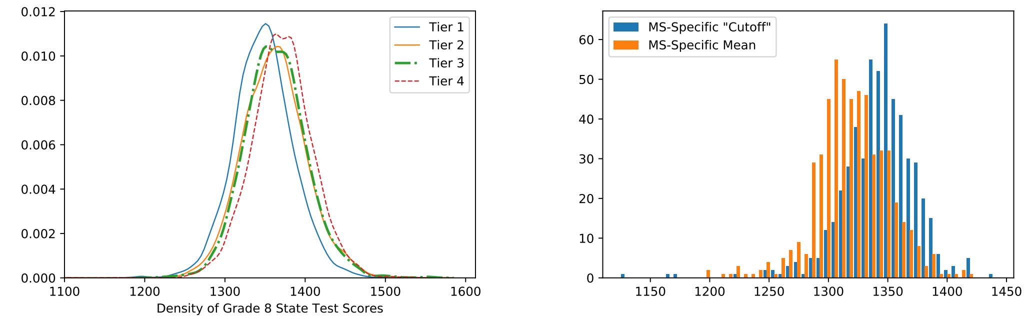
The more similar the compositions within subgroups are to one another, and thus to the overall composition, the more the selection of top students from each subgroup resembles the selection of top students from the entire population. As a result, the set of specialized high school offers in Chicago plan is close to that in the status quo.

Likewise, the more similar the student qualities are across subgroups, the more likely that the top students within each subgroup are also top students in the population, and their selection into elite schools would lead to higher student quality in this sector. This is true for the Chicago plan, where the distributions of student qualities across tiers as measured by Grade 8 state test scores are quite alike, as shown in Panel (a) of Figure 1.6.6. By comparison, a lot of dispersion is evidenced in the histograms of two middle-school-specific quality measures in Panel (b) of Figure 1.6.6. The first measure, labeled as the middle-school-specific "cutoff," is the lowest Grade 8 state test score among students eligible for specialized high schools at each middle school;[24] and the second measure is the mean Grade 8 state test score of all students in the school. By plotting them together, one can see that for most schools, except very high-performing ones, the "cutoff" is lower than the mean at some other schools. This means some students from the middle school currently being considered receive specialized high school offers even though they are not as qualified (solely on academic performance) as an average student at another school, who would not have received offers.

1.7 Conclusions

In this paper, I estimate a model of students' preferences during the New York City high school admission process and simulate the effect of two counterfactual affirmative action

[24] This is not the actual cutoff used during the selection process, which is based on Grade 7 state test scores.



(a) Chicago Plan (b) Top 7% Plan

Figure 1.6.6: Dispersion of Student Qualities (Grade 8 State Test Scores) across Subgroups

Notes: The two panels are not exact counterparts of each other, due to the difference in the numbers of subgroups: 4 tiers (Chicago) v.s. 569 public middle schools (Top 7%). Panel (a) shows the full distribution of student qualities for each tier. Panel (b) shows the histograms of two middle-school-specific statistics: the "cutoff" for being eligible for specialized high schools and the mean student quality. Here, the "cutoff" at each school refers the lowest Grade 8 state test score among students eligible for specialized high schools under Top 7% plan; it is not the actual cutoff used during the selection process, which is based on Grade 7 state test scores.

plans on both elite and regular high schools in the city.

There are three key findings from this analysis. First, there is a trade-off between improving diversity and maintaining student quality in elite schools. A tier-based plan similar to the one implemented by the Chicago public schools (Chicago plan) barely increases overall racial diversity but preserves the quality of incoming students. In contrast, a plan that admits top students in each public middle school based on their academic performance (Top 7% plan) increases diversity substantially while lowering student quality. Second, the Chicago plan reduces the differences among schools within the elite sector, whereas the Top 7% plan increases these differences despite making the elite and regular sectors more similar. Third, both plans also change regular school assignments and thereby the student quality in regular high school programs.

The factors driving the results give further insights into how to design affirmative action plans in general. Most noticeably, the division of the overall population into subgroups that are treated separated during the assignment process plays an important role. To increase racial diversity among those selected, there must be sufficient dispersion in racial compositions across groups, but to maintain student quality, there should not be too much dispersion in student quality across groups. I am currently working on fine-tuning a hybrid policy that aims to achieve the same overall diversity in specialized high schools as the Top 7% plan but do better than that plan in terms of specialized high school student quality.

Another direction for further research is to consider the medium- and long-term effects of these affirmative action plans. Unlike the short-run where I assume that preferences over schools remain unchanged, in the long-run, parents and students will have learned what to expect of the policy, how it changes school characteristics, and thus their "new" preferences over schools. It is reasonable to assume that they have rational expectations over these characteristics, and the new equilibrium is found in the simulation by iterating until a fixed point in the distribution of school characteristics is reached.

Both topics of the hybrid policy and longer-term effects of affirmative action will hopefully be available in subsequent versions of this paper in the near future.

1.8 Appendix A: Additional Tables and Figures

Table 1.7: Regular School Application - School Characteristics by Rank of Student Choice (by Subgroups)

	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th
Panel A. Split by high math achievement												
Students with low baseline math												
Percent white	10.27	9.71	9.13	8.63	7.87	7.42	7.06	7.11	6.49	6.58	6.54	6.28
Percent subsidized lunch	72.24	72.67	73.25	74.05	74.53	75.28	75.68	75.45	76.17	76.52	76.28	76.70
High math achievement (%)	5.31	5.24	4.98	4.90	4.74	4.67	4.58	4.51	4.41	4.31	4.24	3.97
Students with high baseline math												
Percent white	28.90	26.42	25.18	23.32	21.86	20.45	18.20	17.16	16.30	15.03	13.80	12.52
Percent subsidized lunch	48.62	51.00	53.96	56.18	58.28	59.62	62.01	62.58	64.15	65.29	66.83	67.73
High math achievement (%)	23.32	16.95	16.26	15.69	15.14	14.79	13.49	13.21	12.50	11.66	10.76	9.31
Panel B. Split by zip code income												
Students from bottom zip code income quantile												
Percent white	9.16	9.25	8.87	8.55	8.21	7.78	7.55	7.52	7.18	7.07	6.61	6.17
Percent subsidized lunch	73.10	73.03	73.74	74.32	74.84	75.38	75.52	75.72	75.89	76.49	76.61	77.34
High math achievement (%)	6.85	6.74	6.44	6.34	5.98	5.86	5.63	5.48	5.37	5.31	5.02	4.45
Students from top zip code income quantile												
Percent white	36.26	32.10	29.35	26.00	23.59	21.74	19.22	18.00	16.88	15.96	14.41	14.63
Percent subsidized lunch	45.03	48.56	51.29	53.89	55.55	56.90	59.12	59.86	61.47	62.46	63.91	63.80
High math achievement (%)	17.76	14.30	14.49	13.93	13.64	13.41	12.66	12.17	11.84	10.45	9.58	9.76
Panel C. Split by race												
White students												
Percent white	41.66	36.66	34.47	31.76	29.77	27.67	25.63	24.00	23.17	21.23	19.43	19.56
High math achievement (%)	16.19	14.30	14.62	14.58	14.23	14.31	13.56	12.62	12.89	11.84	11.28	9.69
Asian students												
Percent white	24.08	23.61	23.34	22.33	21.01	20.25	18.38	17.35	15.95	15.88	14.97	13.35
High math achievement (%)	25.44	17.62	16.63	16.04	15.47	15.39	14.44	13.94	13.25	12.93	11.95	10.55
Black students												
Percent white	11.61	11.34	10.73	10.11	9.74	9.42	8.91	9.02	8.41	8.25	8.03	7.85
High math achievement (%)	7.93	7.71	7.42	6.99	6.88	6.71	6.41	6.49	6.24	5.76	5.41	5.31
Hispanic students												
Percent white	13.46	12.76	12.07	11.48	10.68	10.25	9.57	9.26	8.92	8.63	7.83	7.10
High math achievement (%)	8.64	8.00	7.79	7.69	7.32	7.28	6.90	6.67	6.46	6.37	5.80	5.02

Notes: Statistics is based on the analysis sample, consisting of 66068 students. High baseline math students score above the seventy-fifth percentile for Grade 7 math state test; low baseline math students score below the twenty-fifth percentile. Zip code income is median family income for each zip code from the American Community Survey 2007-2011; bottom and top income quantiles correspond to Tiers 1 and 4, respectively, in the Chicago counterfactual.
The high math achievement measure for each high school is the percentage of admitted students in said school who score higher than 85% in the Regents mathematical exams in the 2009-2010 school year.

Table 1.8: Estimated Coefficients

	Regular high schools programs		specialized high schools	
	Estimate	Std. Error	Estimate	Std. Error
Percent subsidized lunch				
× ELA State test score	-0.0033***	9.73E-06	-0.0031	0.0034
× Mathematics State test score	-0.0020***	9.65E-06	-0.0060	0.0052
× Zipcode income	0.0118***	1.11E-05	0.0046	0.0038
× Eligible for subsidized lunch	-0.0070***	1.78E-05	-0.0008	0.0217
× Asian	-0.0169***	3.95E-05	-0.0101	0.0279
× Black	0.0041***	3.77E-05	0.0000	0.0287
× Hispanic	-0.0100***	3.90E-05	-0.0546***	0.0200
× Other	-0.0183***	9.98E-05	0.0040	0.0026
Percent white				
× ELA State test score	0.0052***	1.76E-05	-0.0044	0.0035
× Mathematics State test score	0.0033***	1.92E-05	-0.0088*	0.0048
× Zipcode income	-0.0010***	2.35E-05	0.0034	0.0033
× Eligible for subsidized lunch	-0.0130***	3.05E-05	0.0048	0.0202
× Asian	0.0117***	6.84E-05	-0.0294	0.0257
× Black	0.0600***	7.80E-05	-0.0227	0.0269
× Hispanic	0.0438***	8.21E-05	-0.0626***	0.0177
× Other	0.0274***	2.68E-04	-0.0039**	0.0512
Size of grade 9				
× ELA State test score	-3.95E-05***	4.15E-07	0.0001***	0.0000
× Mathematics State test score	-0.0002***	4.37E-07	0.0004***	0.0001
× Zipcode income	-0.0006***	5.93E-07	0.0001	0.0000
× Eligible for subsidized lunch	-3.75E-05***	7.16E-07	0.0004*	0.0002
× Asian	-0.0007***	1.35E-06	0.0013***	0.0003
× Black	0.0001***	1.59E-06	0.0010***	0.0003
× Hispanic	-0.0001***	1.38E-06	0.0009***	0.0002
× Other	-0.0002***	5.35E-06	0.0013	0.0006
Percent high math achievement				
× ELA State test score	0.0042***	1.50E-05	0.0021	0.0018
× Mathematics State test score	0.0130***	1.96E-05	0.0095***	0.0030
× Zipcode income	0.0098***	1.30E-05	0.0032*	0.0017
× Eligible for subsidized lunch	0.0018***	2.77E-05	0.0005	0.0091
× Asian	0.0340***	6.37E-05	0.0271***	0.0101
× Black	0.0275***	6.64E-05	0.0142	0.0097
× Hispanic	0.0192***	6.54E-05	0.0020	0.0075
× Other	0.0121***	1.74E-04	-0.0059	0.0701
Distance ($-\gamma$)	-1	–	-0.0024	0.0072
Scale (λ)	1.5700***	1.20E-03	–	–

Notes: *** p<0.01; ** p<0.05; * p<0.1.

Table 1.9: Racial Composition of Individual specialized high schools (%)

	status quo				Chicago Plan				Top 7% Plan			
	White	Asian	Black	Hispanic	White	Asian	Black	Hispanic	White	Asian	Black	Hispanic
Stuyvesant	22.05	74.38	1.51	1.51	21.59	69.64	2.97	5.13	28.19	52.20	9.98	8.70
					-0.46	-4.75	1.46	3.62	6.14	-22.18	8.47	7.19
Bronx Science	21.63	67.63	3.34	6.68	19.90	62.40	4.78	12.02	9.62	26.46	52.51	10.88
					-1.73	-5.23	1.44	5.34	-12.00	-41.17	49.17	4.20
Queens Sciences	5.68	78.41	11.36	4.55	3.16	75.79	12.63	8.42	8.06	39.52	20.97	30.65
					-2.52	-2.62	1.27	3.88	2.38	-38.89	9.60	26.10
HS of American Studies	60.53	17.11	7.89	14.47	47.92	21.88	15.63	14.58	20.75	4.40	42.14	2.38
					-12.61	4.77	7.73	0.11	-39.77	-12.70	34.24	18.23
Staten Island Tech	19.61	47.06	9.80	22.55	20.33	42.28	11.38	24.39	3.92	6.37	57.84	31.86
					0.72	-4.78	1.58	1.84	-15.69	-40.69	48.04	9.31
HS of Maths, Science & Engineering	52.67	42.39	0.41	3.70	43.30	46.74	4.60	4.98	55.29	27.40	11.54	5.29
					-9.38	4.36	4.19	1.28	2.61	-14.98	11.13	1.58
Brooklyn Tech	21.46	59.70	8.90	9.14	20.84	59.26	9.91	9.19	9.42	19.93	25.63	43.76
					-0.62	-0.44	1.01	0.05	-12.04	-39.77	16.73	34.62
Brooklyn Latin	22.07	37.93	24.83	13.79	16.89	47.97	18.24	16.22	14.81	5.56	53.70	25.93
					-5.18	10.04	-6.58	2.42	-7.25	-32.38	28.88	12.13

Notes: The second row for each school represents the changes in composition in percentage points.

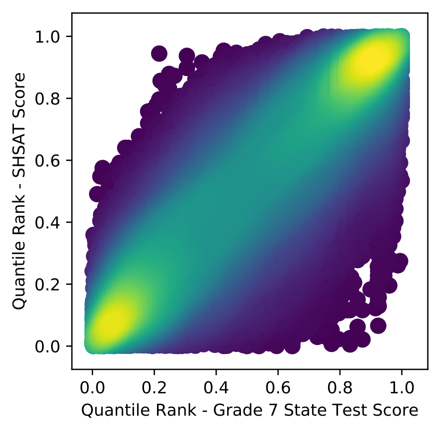
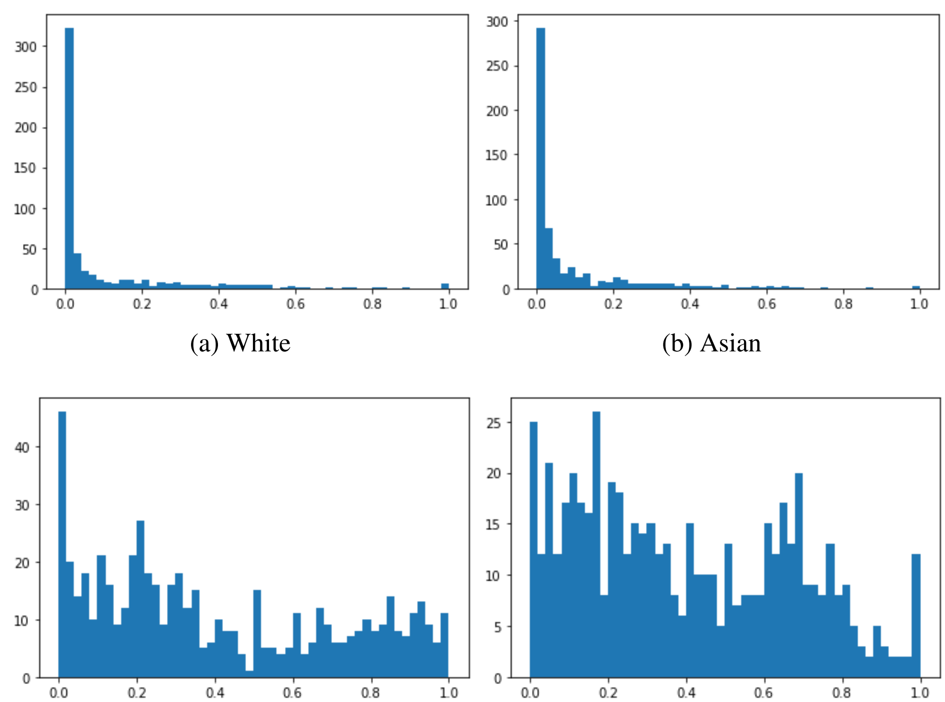


Figure 1.8.1: Quantile Rank of Students within specialized high schools Applicants Based on Grade 7 State Test v.s. SHSAT

Notes: From dark blue to solid yellow, the density of students increases.



(a) White (b) Asian

(c) Black (d) Hispanic

Figure 1.8.2: Distribution of Racial Compositions across Public Middle Schools

Notes: Each panel is a histogram of the school-specific percentages of one race. The population includes all public middle schools.

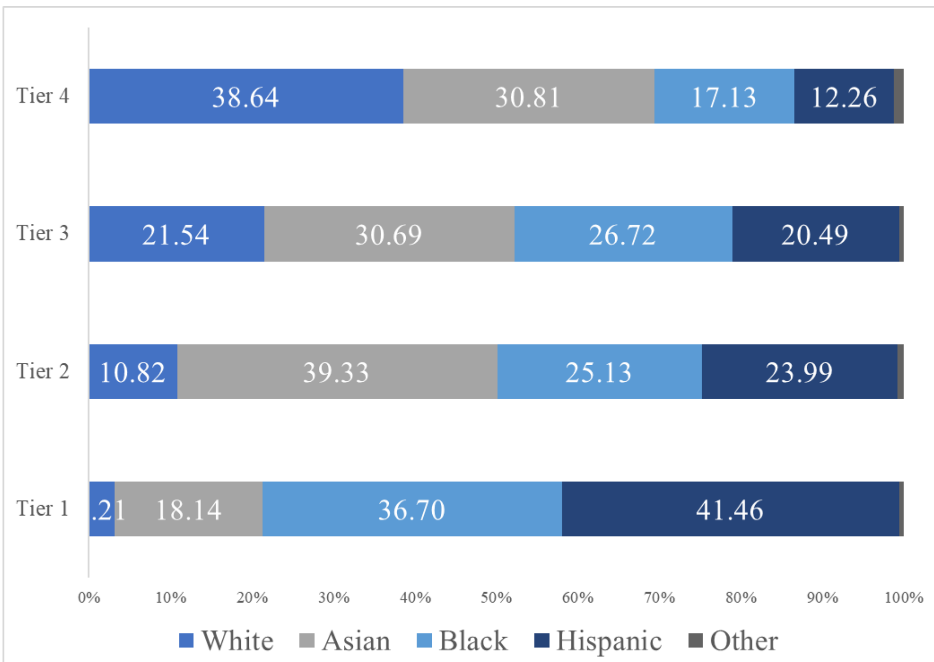
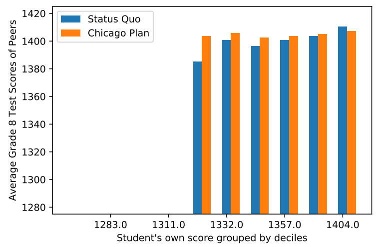
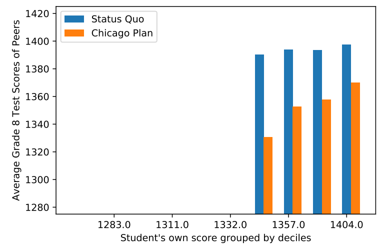


Figure 1.8.3: Racial Compositions by Neighborhood Tiers

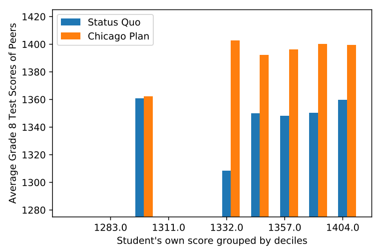
Notes: Tiers are assigned based on median family income of each zip code according to the American Community Survey 2007-2011. Tier 1 has income in lowest quantile compared to the rest of New York City; tier 4 has income in the highest quantile.



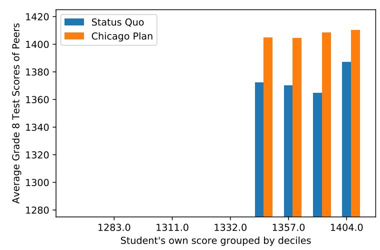
(a) Always Specialized High Schools



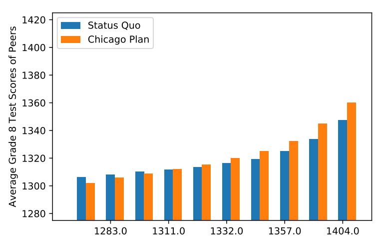
(b) Specialized to Regular



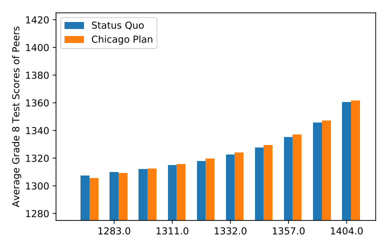
(c) Regular to Specialized (offer gained)



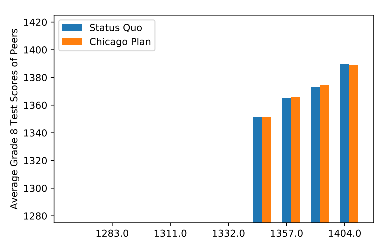
(d) Regular to Specialized (better offer)



(e) Different Regular Schools



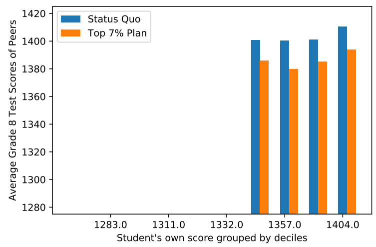
(f) Same Regular Schools (Main Round)



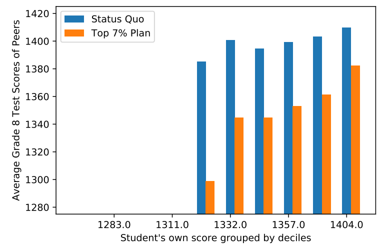
(g) Same Regular Schools (Specialized Round)

Figure 1.8.4: Changes in Peer Quality in High Schools - Status Quo v.s. Chicago Plan

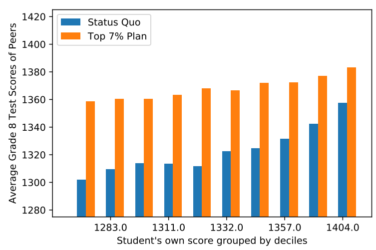
Notes: Peer quality is measured by average Grade 8 state test scores of peers. Deciles are based on the entire population of public middle school students.



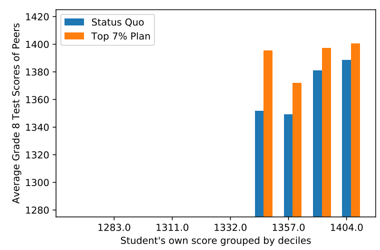
(a) Always Specialized High Schools



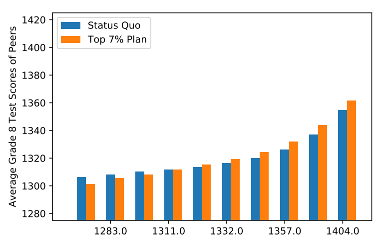
(b) Specialized to Regular



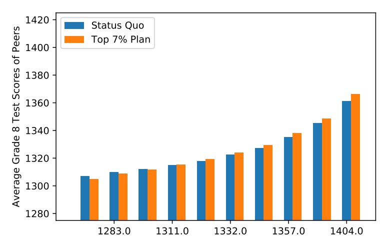
(c) Regular to Specialized (offer gained)



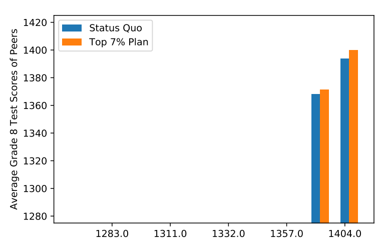
(d) Regular to Specialized (better offer)



(e) Different Regular Schools



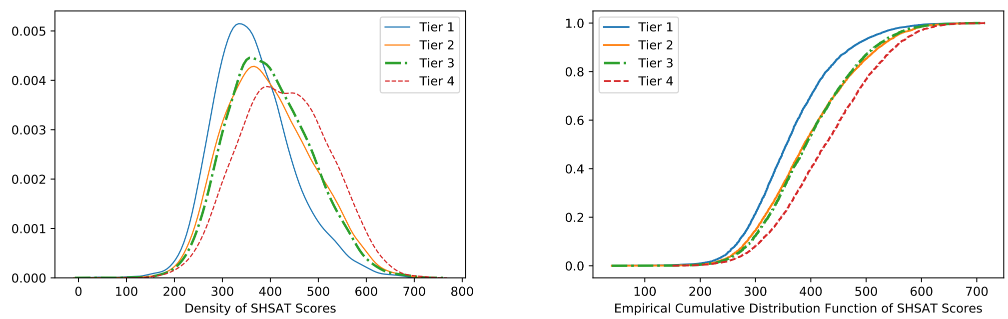
(f) Same Regular Schools (Main Round)



(g) Same Regular Schools (Specialized Round)

Figure 1.8.5: Changes in Peer Quality in High Schools - Status Quo v.s. Top 7% Plan

Notes: Peer quality is measured by average Grade 8 state test scores of peers. Deciles are based on the entire population of public middle school students.



(a) Empirical Probability Density Function (b) Empirical Cumulative Distribution Function

Figure 1.9.1: Distributions of SHSAT Scores by Tiers

1.9 Appendix B: Different Precedence Orders and Merit Reserves for the Chicago Plan

I examine four variations of the Chicago plan which combine two different tier-blind precedence orders (merit first and merit last) with two different percentages of merit seats (40% merit/60% tier and 30% merit/70% tier). My empirical result in the multiple-school setting of New York City specialized high schools supports the theoretical result in Dur, Pathak and Sonmez (2016). Since their result is only in terms of offers, not enrollment, I also restrict my attention to offers in this appendix.

For there to be the best and worst tier-blind precedence orders for different tiers, the score distributions need to satisfy certain assumptions, which is approximately the case in my current setting. Specifically, their Assumption 2 formally requires the "worst-scoring" tier to have the lowest density among all tiers at every possible score, but the authors also state that this property of the density function needs only hold for sufficiently high scores to obtain the result. In my setting, Tier 1 does have the lowest density among all tiers for all SHSAT

Table 1.10: Percentages of Offers by Tiers for Different Variations of the Chicago Plan

	Tier 1	Tier 2	Tier 3	Tier 4
Merit First - 40% Merit/60% Tier	18.78	25.95	25.18	26.38
Merit First - 30% Merit/70% Tier	20.33	25.62	25.21	25.23
Merit Last - 40% Merit/60% Tier	15.04	25.58	25.82	29.43
Merit Last - 30% Merit/70% Tier	17.54	24.81	24.68	28.95

scores above 420 as seen in Panel (a) of Figure 1.9.1. Similarly, their Assumption 3 regarding the "best-scoring" tier is not satisfied for every possible score, but the required property does hold for the range of scores where students may receive offers. The lowest score that allows admission into at least one of the specialized high schools is 472 under the status quo, where all offers are based on merit, and Tier 4 does have the highest density for all scores above this point as seen in Panel (a) of Figure 1.9.1. Although the theoretical result is not based on first-order stochastic dominance, it is interesting to also note that the empirical score distribution of Tier 4 first-order stochastically dominates those of Tiers 2 and 3, which are very similar and both dominate the score distribution of Tier 1 as shown in Panel (b) of Figure 1.9.1. This order agrees with the order from the most to the least disadvantaged tiers in terms of neighborhood income.

Table 1.10 shows the percentages of students receiving offers from specialized high schools by their tiers. Given the same proportions of merit v.s. tier seats, having merit seats filled first is the best tier-blind precedence for Tier 1 and the worst for Tier 4, and having merit seats filled last is the worst tier-blind precedence for Tier 1 and the best for Tier 4, as per the theoretical prediction. To see this, we need to compare the first and third rows, which both have 40% merit seats but different precedence, and compare the second and last rows. In both cases, higher percentages of admitted students belong to Tier 1 and lower percentages belong to Tier 4 when merit seats are filled first. Note that when merit seats are filled last, for both variations considered here, the percentages of admitted students that belong to Tier 1 are almost exactly the same as the percentages set aside for Tier 1 seats (15% and 17.5% respectively), that is, almost no Tier 1 students get in through merit seats.

Given the same precedence order, when the proportion of merit seats is reduced, Tier 1 students, who have systematically lower scores and hence less likely to get in through merit seats, gain more access whereas some Tier 4 students lose access. Note that by decreasing the percentage of merit seats from 40% to 30% (concurrently increasing the percentage of Tier 1 seats from 15% to 17.5%), the percentages of Tier 1 students among those admitted only increase by around 1.5 to 2.5 percentage points depending on the precedence. Meanwhile, changing the precedence order form merit last to merit first has larger effects, resulting in increases of 2.7 to 3.7 percentage points depending on the proportion of merit seats. A similar comparison of the effects of change in the proportion of merit seats v.s. change in precedence order can be made for Tier 4 students.

CHAPTER 2: Matching and Learning – An Experimental Study

Joint with Guillaume Haeringer and Silvio Ravaioli

2.1 Introduction

Gale and Shapley (1962)'s student-proposing Deferred Acceptance mechanism has become increasingly ubiquitous in school choice to assign students to schools. This is, in part, thanks to its strategyproofness, an attractive property, which, informally in this context, means that it is in the students' best interest to report their true preferences. However, a growing body of experimental literature has recorded large deviations from truth-telling when subjects participate in the Deferred Acceptance mechanism in laboratory settings. At the same time, there has also been empirical evidence that some of these "mistakes" have little to no effects on payoff (Artemov, Che and He, 2020).

This paper introduces learning into a lab experiment of centralized matching mechanisms, Deferred Acceptance in particular. Our contribution is two-fold. First, we shed light on the nature of "mistakes" that occur in the lab. We find that the majority of these mistakes are payoff-irrelevant as in Artemov, Che and He (2020). We show that payoff-irrelevant mistakes are more common when subjects have had the opportunities to learn to coordinate within the same environment, and they occur despite evidence that subjects have gained an understanding of the strategyproofness of the mechanism. Second, our results provide insights into the learning process of students who receive feedback in centralized matching.

In our laboratory experiment, we adopt a 2x2 design, with two matching mechanisms (between-subjects) and two treatments (within-subjects). The two mechanisms are Gale-Shapley Deferred Acceptance (DA) and Boston Immediate Acceptance (IA) algorithms. While DA is our focus, IA serves as a control. In each experimental session, we use only one of the mechanisms. At the beginning of the session, the rules of the relevant mechanism are clearly explained, accompanied by an example. Each session is divided into two parts for the two different treatments. Part one involves a random environment without school zones, and part two, a designed environment with school zones (where students are informed of their higher chance of being admitted in two out of five schools). Within each part, there are 8 rounds, each consisting of 5 periods. In every period, each participant plays the role of a student who applies to multiple schools. There are 20 students and 5 schools. Participants are informed of their own payoffs for being admitted into the schools and asked for an order in which they would apply to those schools. Within a round, the environment (schools' priorities and students' preferences) is unchanged, and periods differ only in the rankings submitted by participants. Across the rounds, the environment changes, but the mechanism does not.

Focusing on DA, we observe two opposite patterns in truthfulness within a round and across rounds. Across periods within the same round, when the environment remains constant, truthful reporting decreases over time. However, payoffs do not change much throughout the round, and the vast majority of participants obtain the same payoff as if they were fully truthful. In contrast, truthful reporting increases over time across rounds when the environment changes but the same matching mechanism is used. We do not observe the same trend for IA, for which truth-telling is not necessarily a good strategy.

Given the patterns above, we argue that participants learn about the strategyproofness of the DA mechanism throughout the experimental session. This is supported both by a further comparison with the results in the IA sessions and by an examination of the first period of every round when a new environment is generated. Within each round, when subjects are allowed to learn to coordinate within the same environment, the best possible match each of them can get is often easy to predict in later periods. Subsequently, their untruthfulness is mostly payoff-irrelevant and relates to the concepts of robust equilibrium and stable-response strategy introduced by Artemov, Che and He (2020). We also find that a higher payoff leads to a subject being more likely to keep the same strategy in the next period, consistent with reinforcement learning.

This paper contributes to the experimental literature on matching markets, particularly school choice, by exploring the learning process of students who receive feedback in interactive environments. Hakimov and Kubler (2019) offer a systematic review of the relevant literature. Starting from the earliest lab experiments on school choice (Chen and Sonmez, 2006; Pais and Pinter, 2008; Calsamiglia, Haeringer and Klijn, 2010, 2011), the designs put significant emphasis on the comparison of different mechanisms (DA, IA, and TTC), from the perspective of efficiency and stability. The main results on strategyproofness of the mechanisms are systematically aligned with the theoretical predictions, with the lowest proportion of truthful reporting under IA, but efficiency results are mixed. Surprisingly, there is unequivocal evidence of departure from the complete truth-telling, with fractions of truthful reporting under DA ranging between 40% and 80% (on average, larger fractions were registered when the design included fewer schools to rank). Other common results include systematic biases in favor of schools with many seats and schools that belong to the same "zone" as the student (in both cases, perceived with a higher probability of admission). More recent experiments compared traditional and modified mechanisms, such as the parallel mechanism (Chen and Kesten, 2019) and the secure Boston mechanism (Dur, Hammond and Morrill, 2019).

A small number of experimental papers on school choice provide more detailed results on the learning process in this domain, either by testing it directly or by adopting designs in which participants play multiple rounds of matching with the same mechanism. In the vast majority of the studies, we observe a positive trend of reported truthful over time under DA (up to a 30% increase), but with a ceiling around 80% regardless of the characteristics of the environment. Furthermore, the starting level and slope of truthfulness trends over time are negatively correlated with the number of schools in the task, with slower increases (or even decreases) being observed when the environment contains more than four schools (Chen and Kesten, 2019; Bo and Hakimov, 2020).

In contrast with all the previously listed experiments, we focus more explicitly on learning in an interactive environment. As in the previous cases, we let the participants play under the same mechanism for several rounds, allowing them to reach a clearer understanding of how the algorithm operates. In addition, we have an explicit form of feedback, that is represented by the realized match given the (unobservable) schools' priorities, and the combination of own and others' application lists. Each participant submits a list for multiple periods within each environment (experimental round), with the possibility to revise the list at the beginning of each period. This procedure allows the participant to learn about the elements of the environment that are not directly observable. At the same time, the comparison of the effect of own change across periods is confounded by the presence of other human participants, who also can revise their own lists.

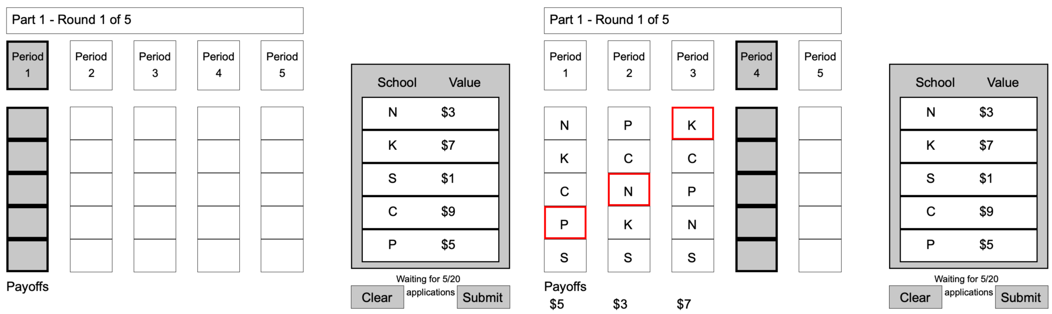
Ding and Schotter (2019) also consider learning in an interactive environment, and they address similar research questions, but with a different paradigm. In their experiments, participants repeatedly played DA or IA in a market comprised of five students and three schools (one seat each). Participants were informed about their own preferences and schools' priorities, and groups were randomly re-matched in every round. They observe a significant increase in truthful reporting under DA (from 64% to 77% over 20 rounds) and no increase in IA. Our design differs in many aspects, such as the size of the environment (20 students and 5 schools), the fact that all students can be successfully assigned to a school, the lack of explicit information about school priority, and the absence of re-matching across rounds.

A different group of experimental papers focus on learning in the form of advice, either from a previous participant to the experiment (Ding and Schotter, 2017) or from a third party (Guillen and Hing, 2014). Most of the studies conclude that providing correct advice increases the rates of truthful reporting in strategyproof mechanisms, but the opposite effect is even stronger when wrong advice is provided, with a large number of participants manipulating their preference reports.

The rest of the paper is organized as follows. Section 2 describes the experimental design. Section 3 presents our analysis of the results, and section 4 concludes.

2.2 Experimental Design

We design our experiment to compare the learning behavior under different assignment mechanisms. We implement a two-by-two design: for each of the two mechanisms, Deferred Acceptance (DA) and Immediate Acceptance (IA), we examine a *random* environment (without school zones) and a *designed* environment (with school zones). The first manipulation occurs between subjects, with each participant being assigned to one mechanism only. The second manipulation occurs within subjects, with the random environment being presented first, fol-



(a) Beginning of period 1 (b) Beginning of period 4 (same round)

Figure 2.2.1: Experimental interface (Part 1)

Notes: The experimental interface contains the list of the schools, each with an associated monetary value, and an empty application list that the participant can fill freely by clicking on the schools' names. The screen also contains a shrinking timer that indicates the time available to complete the choice, and two buttons, to submit the application in advance or restart the application list.

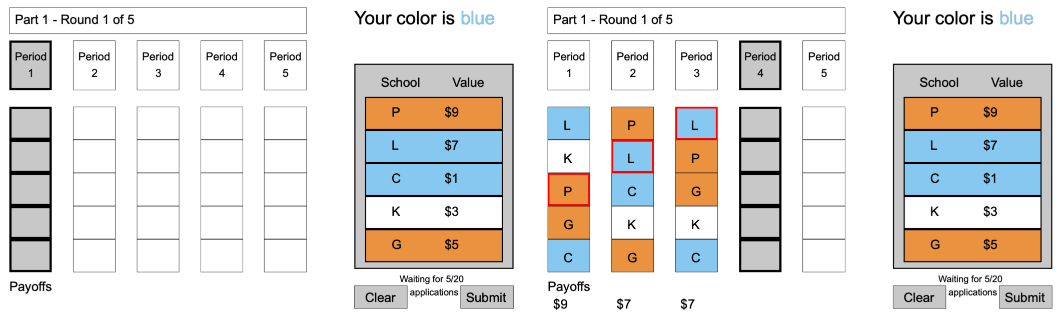
lowed by the designed environment. Despite the possible concern for the order effect, we prioritize clarity of the task for the participants and frame the second environment as a variation on the first part. Each environment condition is used for eight subsequent rounds, with each round comprised of five periods. In each period, twenty students compete for twenty seats (five schools, four seats each). We now describe the details of the experimental design, which apply to all the conditions unless otherwise specified.

Students' Preferences and Schools' Priorities. Each student receives a random ranking of the schools, with values ranging from \$1 to \$9, with \$2 increments. Subjects are informed that different participants may have different payoff values, and receive detailed information about how the values are generated, with particular emphasis on the fact that the ranking is not independent across students, and on average participants like the schools located at the top of the list more than the schools at the bottom. The ranking of each student is generated by summing, for each school, a common score and a random score. The common score (0 to 12 points, equally spaced) is the same for all students and determined by the rank used to display the schools on the screen. The random score (0 to 8 points) is independent across students and across schools, which allows students to have different preferences. The sum of the two scores determines the subjective ranking of the schools, and this ranking is associated with the monetary amounts displayed as values for the seat. Participants can observe their own preferences, but not other students' preferences. They also know that all preferences stay constant within each round but change across rounds. In the designed environment condition, students' zones and schools' zones have no effect on the ranking.

Schools' priorities are generated differently in the *random* environment (without school zones) and in a *designed* environment (with school zones). In both environments, students are not directly informed about schools' priorities, but they are aware that priorities astay constant within each round but change across rounds. In the random environment, each school independently generates a random ranking of the twenty students. In the designed environment, each student receives a color (orange or blue, indicating her own zone, ten students for each color), and the five schools are partitioned into two schools for each color, plus a white (neutral) school. Colors are randomly assigned at the beginning of each round and remain unchanged within the round. For each school, the probability of a student being among the top 4 positions of the priority list depends on her color. If the student's color matches [does not match] the school, she has a 30% probability [10% probability] of being among the top four positions.[1]

School Application Task. Each period is comprised of an *application* phase and a subsequent assignment phase. During the application phase, each student submits an application

[1] A separate random draw determines which students are among the top 4 positions. The relative ranking of those four students is random. Similarly, the relative ranking of the students ranked 5th and below is randomly generated.



(a) Beginning of period 1

(b) Beginning of period 4 (same round)

Figure 2.2.2: Experimental interface (Part 2)

Notes: The experimental interface differs from the previous one only because of the “student’s color" (top right of the screen) and the colors associated with the various schools.

list, which indicates in which order they want to apply to schools. Students are not forced to apply to all five schools, although not doing so means that they may end up unassigned (an outcome with value $0, strictly lower than a seat in any school). The application phase has a limited time (60 seconds in period 1, 30 seconds in all the subsequent periods), but participants can submit their own application lists before the timeout.

In the assignment phase, each participant observes to which school she would be assigned, given other students’ current application lists as well as her own. The assignment is salient, and it is displayed both with the school’s name in the application list and its value below the application list (see Figures 2.2.1 and 2.2.2). No other information (schools’ priorities or other students’ assignments) is communicated to the participants. The application phase has a fixed time of 10 seconds, after which all the participants move to the next period.

Each round is comprised of five periods, and within a round, there is no change in students’ preferences, schools’ priorities, nor zones (in the designed environment condition). From period 2 to 5, participants can revise the application list, and submit either the previous one or a new one, with no direct cost or benefit from the revision. The task is repeated for eight different rounds for each environment (random and designed).

Procedure. The experiment was run in CELSS (Columbia Experimental Laboratory for Social Sciences, Columbia University, New York, USA) between March and April 2019. The experiment was coded in MATLAB (Release 2018b) using Psychotoolbox 3 (Psychophysics Toolbox Version 3). Seventy-six volunteers (38 in the Boston treatment, and 38 in the Gale-Shapley treatment) were recruited from the students registered in the Marked Design undergraduate course at Columbia University as optional assignment for the class' requirements.[2] On average, the whole experiment took 60 minutes, including instructions. Participants did not receive any monetary incentive for their participation in the study, but they were motivated by the fact that their scores in the task would be posted on the website of the course. Anecdotal evidence suggests that participants were motivated to get a high score and put effort into the task as if they were facing a monetary incentive.

We ran four sessions in total (two sessions for each mechanism). Each session is designed for a group of twenty subjects, and the actual number of participants ranged between eighteen and twenty. When the number was not exactly twenty, the missing participants were replaced by automatic computer programs. Human participants were informed that each program was submitting a list of schools in the first period of each round, but not revising it in the following periods. Instructions were read aloud by the experimenter, and were also provided to each participant as paper printout (see Appendix 2.6). There are two versions of the instructions (one for each matching mechanism), containing the same information verbatim with the exception of the tailored information on the mechanism being used. The instructions included

[2] Crucially, the experiment was run before the instructor discussed the mechanisms for school assignment in class.

detailed information about the task, the data generating process used to generate priorities and payoffs, as well as the matching mechanism (described formally first, and illustrated with an example after). In addition to the written instructions, the main projectors of the laboratory were used to illustrate the matching mechanism in the example and to display screenshots of the experimental interface. Clarifying questions, if any, were answered publicly.

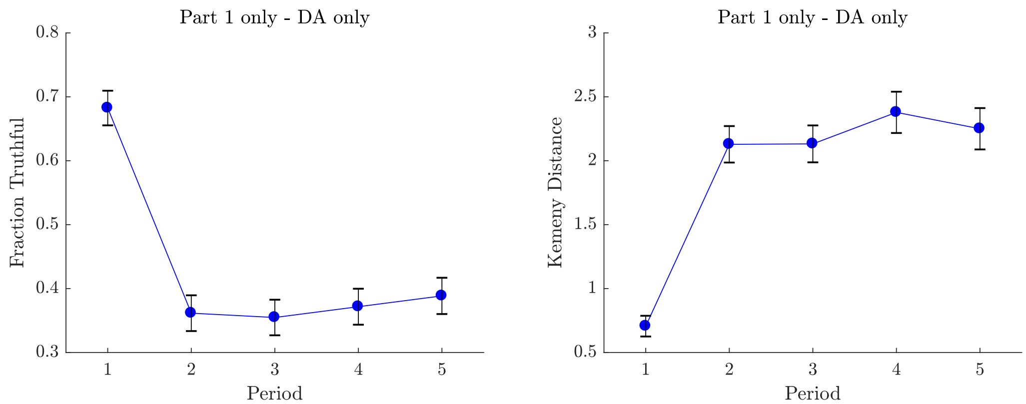
2.3 Results

We focus our analysis on the results of the Deferred Acceptance (DA) sessions, and contrast it with Immediate Acceptance (IA) when appropriate. We also restrict our attention to the first part of the experiment to avoid the potentially confounding effect of time and school zones in part 2.

2.3.1 Across Periods within A Round

We find that for all ex-ante measures of (un)truthfulness, that is, measures that are independent of the resulting match, truthful reporting decreases over time within a round. However, payoffs remain relatively stable throughout the round because the majority of untruthful participants behave in such a way that they still obtain the same payoff as under truthful reporting.

Figure 2.3.1 includes the plots for two ex-ante (un)truthfulness measures for the five periods, averaged over rounds. Both show that truthful reporting decreases further into the round. As seen in Panel (a), the fraction of participants who are fully truthful (i.e., reporting their exact true preferences) drops sharply from the first to the second period before mostly leveling out. Panel (b) shows are similar story when we move from the aforementioned binary



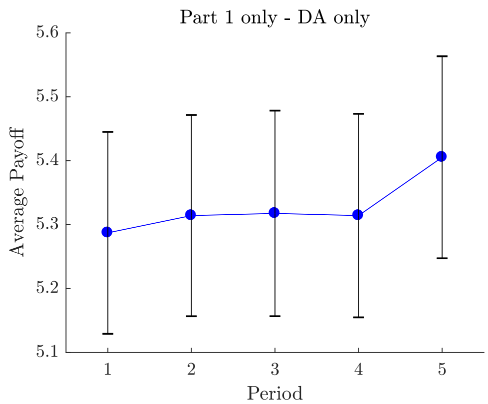
(a) Fraction of fully truthful participants

(b) Kemeny distance from submitted ranking to true preference

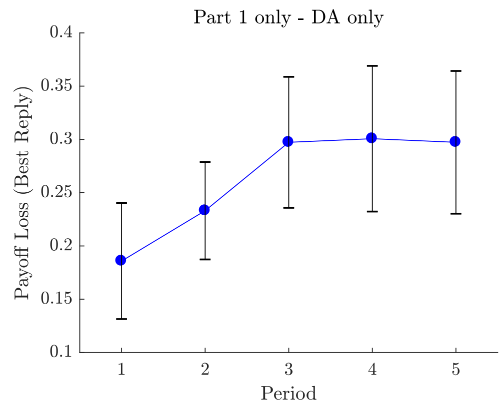
Figure 2.3.1: Across periods - Measures of ex-ante (un)truthfulness

measure to one that also takes into account the intensity of truthful reporting for each participant. The Kemeny distance from the submitted list to the true preference is the number of pairwise comparisons in the submitted list that deviate from the true preference: the higher this measure is, the less truthful the participants are. It increases considerably from close to 0 in the first period to more than 2 in period 2 and then remains between 2 and 2.5 until the end of the round. With five schools, there are a total of ten pairwise comparisons, so during periods 2 to 5, an average participant is untruthful for 20-25% of the pairwise rankings.

Figure 2.3.2 plots two different measures of payoff for the five periods, averaged over rounds. Panel (a) demonstrates that average payoffs remain relatively stable across the five periods. Meanwhile, Panel (b) shows the payoff loss for each participant compared to what they would have gotten by best-replying to the actual strategy profile of their opponents. For DA, as is the case here, full truthfulness is a best-reply to any strategy profile of the opponents, so this measure can also be considered as payoff loss compared to truthful reporting. There is a slight increase in payoff loss at the beginning of the round before it levels out.



(a) Average Payoff



(b) Payoff loss (compared to best reply)

Figure 2.3.2: Across periods - Measures of Payoff

The magnitude of both the average payoff loss (ranging from 0.15 to 0.3) and the increase in payoff loss are small compared to the average payoff (5.3-5.4). This suggests that each participant was very close to best-replying to their opponents' strategies even though they might not be using the dominant strategy (truth-telling).

To further investigate this phenomenon, we divide the participants into three categories within each period: truthful, untruthful with no effect on payoff (i.e. the payoff loss outlined above is zero), and untruthful with effect (i,e, the payoff loss is positive). We find that the vast majority of the 38 participants (between 90-94%) belong to the first two groups for all periods, which explains why although the fraction of full truthfulness decreases, this has little effect on payoff. In line with the pattern in truthful reporting, the most drastic change in composition of participants occur in the second period, where 49% of those who were truthful in the first period switch to untruthful with no effect on payoff (see the transition matrix on the left in Figure 2.3.3). While there exist movements between groups for all the later periods, the composition always stay close to that attained in period 2. As seen in the

Period 1 \ Period 2	Truthful (36%)	N-Truthful No effect (55%)	N-Truthful W/ effect (9%)
Truthful (68%)	45%	49%	6%
N-Truthful No effect (26%)	18%	73%	9%
N-Truthful W/ effect (5%)	18%	38%	44%

Period 4 \ Period 5	Truthful (39%)	N-Truthful No effect (53%)	N-Truthful W/ effect (8%)
Truthful (37%)	75%	24%	1%
N-Truthful No effect (54%)	16%	75%	9%
N-Truthful W/ effect (9%)	23%	46%	31%

Figure 2.3.3: Transition matrices: periods 1 to 2 (left) and periods 4 to 5 (right).

transition matrix on the right in Figure 2.3.3, by the time we reach the transition from period 4 to period 5, a large fraction (75%) of those in the truthful and untruthful unaffected groups stay within the same group. The proportion of participants in the untruthful affected group rises from 5% in period 1 to 8-9% in later periods, which accounts for the slight increase in payoff loss documented above.

2.3.2 Across Rounds

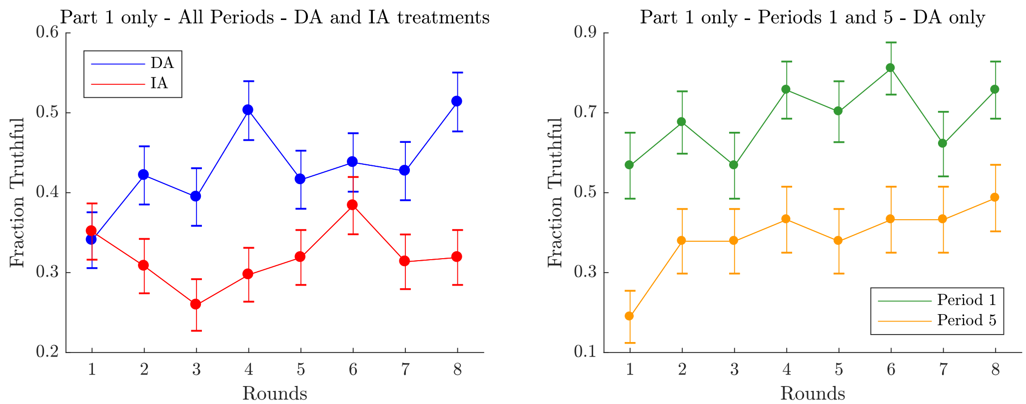
When we make the comparison across rounds, truthful reporting increases over time during the session for DA. The same trend is not observed for IA (the control), suggesting that (some) participants eventually learn that DA is a strategy-proof mechanism while IA is not.

While the fractions of truthful participants start out at comparable levels for the first round of both mechanisms, they diverge in later rounds. Panel (a) of Figure 2.3.4 shows an upward trend in the fraction of full truthfulness under the DA mechanism, from about 34% in the first round to 51% in the eighth round (within the range of the previous literature). In contrast, the equivalent fraction under the IA mechanism fluctuates between 25 and 40% throughout

Table 2.1: Truthfulness and Payoff - DA vs. IA

	(1) Full Truthfulness	(2) Truthful or Untruthful Unaffected	(3) Payoff	(4) Payoff Loss (vs best reply)
Constant	0.542***	0.905***	5.291***	0.284***
	(0.029)	(0.016)	(0.163)	(0.048)
IA Dummy	-0.057		0.089	1.712***
	(0.042)		(0.236)	(0.138)
Round	0.019***	0.007**	-0.003	-0.023**
	(0.006)	(0.003)	(0.031)	(0.011)
Period	-0.047***	-0.005	0.024	0.029
	(0.009)	(0.005)	(0.050)	(0.019)
IA * Round	-0.013*		0.055	0.010
	(0.008)		(0.044)	(0.027)
IA * Period	0.004		-0.040	-0.134***
	(0.013)		(0.072)	(0.043)
R-squared	0.029	0.004	0.0023	0.1562
N	2960	1480	2960	2960

Notes: Observations are at the session-round-period-participant level. For columns (1), (3), and (4), the sample includes observations from part I (rounds 1-8) of all four experimental sessions. For columns (2), the sample is restricted to the two DA sessions.
The round and period numbers are normalized to 0-7 and 0-4, respectively, so that the constant indicates the average level of truthfulness in the first period of the first round under DA.



(a) DA vs. IA - All periods

(b) DA only - Periods 1 and 5

Figure 2.3.4: Across rounds - Fraction of truthful participants

the session. The same pattern can be seen in the first column of Table 2.1, which presents the results for the regression of an indicator for full truthfulness on a dummy for the IA mechanism, round number (1 to 8), period number (1 to 5), and their interactions. In the first period of the first round, there is no significant difference in the level of truthfulness between the two mechanisms. As the experiment progresses, the probability of truthfulness improves by 1.9 percentage point per round on average (significant at the 1% level) under DA as opposed to the 1.3 percentage point decrease (significant at the 10% level) under IA.

Under DA, there is no significant change in payoff, but payoff loss decreases marginally across rounds (according to columns (3) and (4) of Table 2.1). The decrease in payoff loss is attributable to the modest increase in the fraction of participants who are either truthful or untruthful but unaffected in terms of payoffs. Although both of the aforementioned changes are significant (at the 5% level), the magnitudes are negligible.

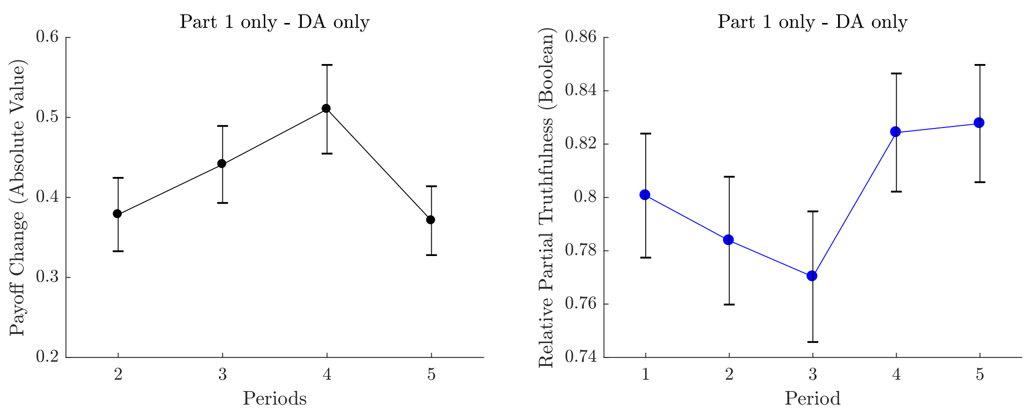
2.3.3 Explaining the Results

2.3.3.1 Learning about the Mechanism and Payoff-Irrelevant Mistakes

Although participants learn about the optimality of truth-telling under the DA mechanism, when they are allowed to learn to coordinate within the same environment, a large fraction deviate from full truthfulness in a way that is payoff-irrelevant. Two main arguments support the existence of learning about the mechanism over time: when we compare DA to the control IA and when we concentrate our attention on periods in which participants were faced with a new environment. We find that subjects were presumably able to predict their best possible match in later periods within each round, and their untruthfulness relates to the concepts of robust equilibrium and stable-response strategy introduced by Artemov, Che and He (2020).

The comparison between DA and IA points to not only the fact that subjects had some understanding of the strategyproofness of DA but also that at least some of them acquired this over time during the experimental session. As we described in the previous subsection, there is a significant increase in truthful reporting across rounds under DA that is not replicated under IA. Given the random assignment of subjects into the two mechanisms, it is likely that this result is driven by subjects' understanding of the mechanism's underlying property. In theory, this understanding can be obtained immediately after reading the description and example of the mechanism in the instructions. However, we can attribute it to learning during actual participation (with feedback) in the mechanism for at least some participants because there is no significant difference in truthfulness between IA and DA in the first round (and the first period of that round, in particular).

Now, let us focus on the first period of each round, when participants had not yet learned to coordinate with each other in the new environment. Panel (a) of Figure 2.3.4 demonstrates



(a) Across Periods - Absolute changes in payoff (b) Across Periods - Relative partial truthfulness

Figure 2.3.5: Payoff Predictability and Observed Strategy

Notes: Relative partial truthfulness is defined as reporting the truthful relative rankings among schools above or equal to the realized match.

that the increase in truth-telling across rounds, previously reported for the average over all periods, persists under this restriction. Moreover, we can see that although truthfulness is lower in period 5 than period 1 of every round, this low level in the last period of a round does not carry over the period immediately after it (i.e. period 1 of the next round) when the environment is new. This further confirms that the decrease in truthful reporting we observe across periods within a round can be ascribed to learning to coordinate and does not discount the learning-about-the-mechanism hypothesis. We also confirm that participants who became consistently truthful in the first periods towards the end of the experimental session account for a considerable part of the increase in truthful reporting across rounds. Figure 2.5.1 shows the fractions of participants (compared to the total) who were truthful in the first period for all 8 rounds, all of the last 7 rounds, all of the last 6 rounds, and so on. Half of of the participants who were truthful in the last round (76% of the total) had been truthful for the last 5 rounds (inclusive of round 8).

The observed pattern in the data provides suggestive evidence of the ease of predicting the best possible match during later periods of each round. We remarked in subsection 2.3.1 that average payoffs remain stable across periods, but this could be masking large positive and negative changes occurring at the same time for different participants. Panel (a) of Figure 2.3.5 shows that this is not the case: the absolute changes in payoff, ranging on average between 0.3 and 0.5 through out the round, are small compared to the increment of $2 between consecutive prizes and the difference of $1 between the unmatched outcome and the lowest prize. Note that there is a one-to-one correspondence between payoffs and schools. Therefore, given little change in payoffs and 90-94% of participants receiving as-if-truth-telling payoffs, participants could assume (often correctly) that the best match they could obtain was the same as their previous match.

To examine the subjects' ranking strategies, we define an "ex-post" notion of truthfulness that relies on knowledge of the match. From a participant's point of view when choosing their strategy, it is their best guess of the match, which is unobservable to the experimenter. Hence, we use the realized match, which as argued above was easy to predict. Relative partial truthfulness is, then, a binary measure indicating that a subject reported the truthful relative rankings among schools of higher or equal positions to the match in the submitted list. In other words, all schools ranked above the match are in the correct order and preferred to the match. Panel (b) of Figure 2.3.5 shows that the majority of paticipants (between 76 and 84%, depending on the period) satisfy relative partial truthfulness with respect to the realized match. In period 1, the large fraction of relative partial truthfulness is mostly driven by fully truthful participants (68 out of the 80 percentage points). Since this measure uses realized instead of predicted match, the decrease in periods 2 and 3 can be plausibly explained by the fact that it may take time to notice a pattern and accurately predict the best possible match, after which it rises in periods 4 and 5. Notably, even though full truthness is low in these last two periods (37-38%), relative partial truthfulness is even higher than that in period 1.

The untruthful behavior of the participants in our experiment relates to the theory in Artemov, Che and He (2020). First, they formalize the robust equilibrium concept, which allows for mistakes along as they become payoff irrelevant when the market grows arbitrarily large and the set of feasible schools for each participant becomes close to deterministic. Even though the market is small in our experiment, we show above that empirically the most preferred feasible school is stable and easy to predict in the later periods of each round. Second, the relatively partially truthful strategies employed by the majority of our participants belong to a subset of Artemov, Che and He (2020)'s stable-response strategy, which requires the most preferred feasible school to be ranked above all other feasible schools in the limit continuum economy. Note that while a stable response strategy allows for a worse but infeasible school to be above the most preferred feasible school, relative partial truthfulness rules out this behavior. It is understandable that this is the case in a small economy, where there is more uncertainty regarding the set of feasible schools.

2.3.3.2 Reinforcement Learning within a Round

We examine subjects' tendency to revise their strategies when going from one period to another and find that the higher the payoff the more likely a subject would keep the same strategy in the next period, which is consistent with reinforcement learning.

Table 2.2 presents logistic regressions of the probability of no revision in four different specifications, considering the previous period (specifications 1 and 2) and the previous two periods (specifications 3 and 4). In all specifications, the coefficient on lagged payoff is positive and significant at the 1% level although the magnitude are smaller in specifications 3

Table 2.2: Logistic Regression - Probability of No Revision

	Pr(No revision in period t)			
	(1)	(2)	(3)	(4)
Constant	-3.848***	-4.938***	-1.602***	-3.438***
	(0.354)	(0.780)	(0.475)	(1.007)
Full Truthfulness $(t-1)$	2.049***	1.631***	1.767***	1.706***
	(0.299)	(0.320)	(0.385)	(0.428)
Relative Partial Truthfulness Only $(t-1)$	0.764**	0.644*	0.925***	0.834**
	(0.311)	(0.335)	(0.354)	(0.392)
Payoff $(t-1)$	0.365***	0.481***	0.210***	0.351***
	(0.033)	(0.044)	(0.046)	(0.059)
ΔPayoff $(t-2$ to $t-1)$			0.314*	0.301
			(0.161)	(0.193)
Revision $(t-1)$			-2.326***	-1.891***
			(0.226)	(0.243)
ΔPayoff *Revision $(t-1)$			-0.080	-0.084
			(0.172)	(0.206)
Round FEs	Yes	Yes	Yes	Yes
Participant FEs	No	Yes	No	Yes
Periods	2-5	2-5	3-5	3-5
Pseudo R-squared	0.265	0.390	0.410	0.476
N	1184	1184	888	888

Notes: “Relative partial truthfulness only” refers to strategies that satisfy relative partial trutfulness without being fully truthful.

and 4 when we look at a sample of period 3 onwards and control for revision in the previous period (or lack thereof) and the associated payoff change. In other words, the probability of keeping the current strategy is increasing in the payoff said strategy brings, consistent with reinforcement learning. This result is not driven by individual's intrinsic propensity to revise. In fact, the estimated impact of payoff grows when we include participant fixed effects (specification 2 as opposed to 1 and 4 as opposed to 3). The change in payoff also has a positive effect on not revising, but it becomes insignificant once participant fixed effects are controlled for. Note that having adopt a relatively partially truthful strategy (whether or not it is fully truthful) in the previous period is also less likely to lead to revision.

2.4 Conclusion

We use a lab experiment to study learning when participants receive feedbacks in centralized matching mechanisms widely used in school choice and other real-world applications. Our results provide a deeper understanding of participants' "mistakes" in strategyproof mechanisms like Deferred Acceptance as well as their learning process.

A novel aspect of our experimental design is to include the distinction between rounds and periods, allowing for two types of learning. Within a round, the environment (schools' priorities and students' preferences) is unchanged, and periods differ only in the rankings submitted by participants. Thus, agents learn to coordinate, which in our experiment induces them to become untruthful in a way that is payoff irrelevant. Across the rounds, the environment changes, but the mechanism does not. There is suggestive evidence that agents do have a better grasp of the underlying property of the mechanism through this experience.

We plan to run more experimental sessions with an improved design, which includes randomizing the order of the environments with and without school zones and allowing for more periods within a round. Another direction of future research would be to explore sequential admission procedures, as seen in German university admission, where the feedback and learning process (although very different from our current experiment) can be implemented in a real-world setting.

2.5 Appdendix A: Additional Figure

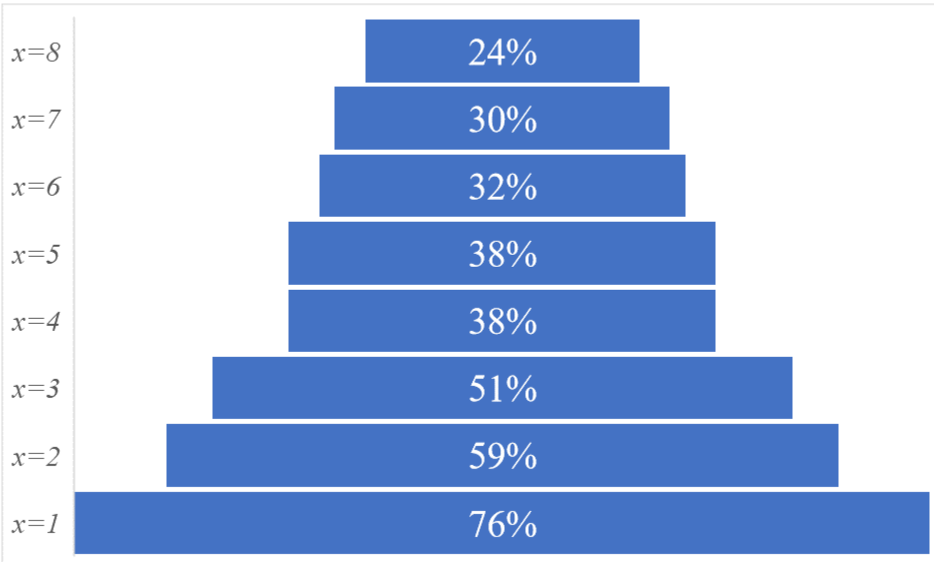


Figure 2.5.1: Fraction of participants who have been truthful in the first period of the last x round(s)

2.6 Appendix B: Instructions for the Experiment

2.6.1 Instructions for the Deferred Acceptance Mechanism (Mechanism G)

Thank you for participating in this experiment on decision making. From now until the end of the session any communication with other participants is forbidden. Foods and drinks are not allowed in the lab, and we ask you to **turn off your mobile phone**.

If you have any question, feel free to ask at any point of the experiment. Please do so by raising your hand and one of us will come to your desk and answer your question privately.

This session will last approximately 90 minutes and comprise of multiple rounds. Throughout the experiment the following rules apply.

You and every other participant play the role of a student who is applying to several schools sequentially. Different schools have different values for you (i.e. how much you like them), and you will choose in which order to apply to the schools. Your goal is to be accepted by a school with a high value. **Application** is done by submitting an *application list* of schools.

Each school has a **priority** list over students (i.e., a ranking of students) ,which determines in which order applicants are considered for acceptance.

Schools have different priorities over students; students have similar but not identical preferences over schools.

The computer will run a **matching** algorithm that utilizes the *application lists* submitted by the students and the schools' *priority lists* over students. This algorithm will determine who is assigned to which school.

The experiment involves N human participants and 20-N automated computer program(s), which will make decisions under the same conditions as yours.

Submitted application list

Each student (including the computer programs) submits an application list, which indicates in which order she wants to apply to schools. That is:

- The first school in the list is the first school to which the student will apply.
- The second school in the list is the the next school to which the student will apply if rejected by the first school.
- and so on...

Priority list over students

Each school has a priority list that will not be communicated to you. Schools decide to accept or reject students who applied to them using their own priority list. There are 5 schools with 4 seats at each school.

The priority list of each school is generated independently.

Student's payoff if accepted to a school

Your payoff depends on the value of the school where you are accepted. You will see on the screen a table outlining the value for each school.

Different participants may have different payoff tables.

Matching Algorithm G

The participants are assigned to schools by a multi-step algorithm.

Step 1:

- Each student applies to the school that is ranked first in her submitted application list.
- Each school tentatively accepts the applicants who have applied to it, one at a time, using the school's priority list and starting with the applicant with the highest priority. It does so until it has admitted as many students as its capacity or runs out of applicants. The remaining applicants (if any) are rejected by the school.

Step 2, 3, ...: This step is only for the students who have been rejected in the previous step.

- Each of those students apply to the next school on their submitted application list. If a student has already applied to all the schools in her application list, then that student remains unassigned (and the algorithm ends for that student).
- Each school considers the set of students it accepted at the previous step together with the set of new applicants. From this larger set, the school tentatively accepts students, one at a time, using the school's priority list and starting with the applicant with the highest priority. It does so until it has admitted as many students as its capacity or runs out of applicants. The remaining applicants (if any) are rejected by the school.

End: The algorithm stops when either no school receives any new applications, or no application is rejected.

Example (Mechanism G)

This is a small example that shows how the algorithm works. We consider

- 3 students: A, B and C
- 3 schools: 1, 2, and 3. Each school has 1 seat.

Schools' priority lists of students are described in Table 2.3.

Table 2.3: Schools' priority lists of students

Priority	School 1	School 2	School 3
1st	A	A	C
2nd	C	B	B
3rd	B	C	A

So, for instance, at school 3, student C is considered first, then B, then A.

The application lists submitted by the students are in Table 2.4.

Table 2.4: Students' submitted application lists

	Student A	Student B	Student C
1st application	1	1	2
2nd application	2	2	3
3rd application	3	3	1

So, for instance, student C will first apply to school 2. Then, if rejected, she will apply to school 3. Finally, if rejected again, she will apply to school 1. The other students' application lists are read similarly.

Let's run the algorithm.

Step 1 Students A and B apply to school 1, and C applies to school 2.

- School 1 starts accepting the students who applied to it following the order given by its priority. A is the top priority student, so she is accepted. School 1 cannot accept any more student (it reached its capacity of 1), so B is rejected.

- The other student, C, applies to school 2. She is the only applicant, so she is accepted.

At the end of the first step, A is tentatively accepted at school 1, C is tentatively accepted at school 2, and B is not accepted at any school.

Step 2 Student B applies to school 2.

- School 2 has now two applicants: C (from Step 1) and B. B has higher priority than C, and since school 2 has only one seat available, B is now tentatively accepted at school 2 and C is rejected.

At the end of the second step, A is tentatively accepted at school 1, B is tentatively accepted at school 2, and C is not accepted at any school.

Step 3 Student C applies to school 3.

- She is the only applicant, so she is accepted.

No student is rejected, so the algorithm stops.

The final assignment is

Student A	Student B	Student C
School 1	School 2	School 3

Part I

You can see on the main screen how the interface of the experiment looks like. The first half of the session consists of 8 Rounds, each divided into 5 Periods. We are in the first part of the session, first round, first period.

On the right-hand side, you can see the five schools (each with a different letter) and your value in case you are accepted there. Each school has 4 seats. If you are not accepted to any school, your payoff will be $0.

You create your application list by clicking on the school boxes on the right. After each click, the school is crossed out and it is added to your application list on the left, in the highest empty box still available. Keep clicking on the schools until you are ready to submit. Click the submit button to confirm your application list, or click on the clear button to reset the list and start from scratch. You can submit your list without applying to all 5 schools.

You have a limited amount of time to submit your application list. The available time is indicated by a timer below the school values. If you run out of time before clicking submit, the current application list on your screen is submitted automatically.

After a few seconds, a new period starts with the *same* school values for all students and *same* priority lists for all schools as in the previous period. You can observe the school you were accepted to in the previous period (highlighted in the red square) and the payoff associated with it (written below the application list). Now you can review your application list and submit the same or a different one. The experiment proceeds in this way for five periods.

After the end of the fifth period, a new round starts with *different* school priorities and *different* student payoffs. Remember that your objective is to get as many dollars as possible in each period.

The schools' priority lists and students' payoffs are generated as follows.

Priority list over students

Each school independently generates a random ranking of the 20 students.

Students' payoffs

Each student receives a random ranking of the schools, with values ranging from $1 to $9 with $2 increments. This ranking is not independent across students: on average students like the schools on the top more than the schools on the bottom. This means that the top school (N in the figure) is more likely to have a high payoff both for you and for the other students, and so on.

The ranking for each student is generated in this way. Schools in position 1 to 5 (ranked from top to bottom) start with a *common score* of 12, 9, 6, 3, and 0 points, that is the same for all the students. Then, each school receives a *random score*, uniformly distributed from 0 to 8. The random score is independent across students and across schools, and allows different students to have different preferences over schools. Common and random scores are summed, and the ranking is generated from the school with more points to the school with less points.

Computer program actions

At the beginning of each round, each computer program will choose an application list and submit the same list for all 5 periods. Only human participants can review their own list between periods.

If something is unclear, you can ask questions now. After answering your questions, we can proceed with the first part.

Part II

The second half of the session consists of 8 Rounds, each divided into 5 Periods. The same rules as Part I apply, with only a difference in how schools generate priority lists. You can see on the main screen the interface for the second part.

At the beginning of each round, half of the students will be randomly assigned the color orange, and the other half will be assigned the color blue. Your color is visible above the school values, and will remain the same during the round, **but your color might change between rounds**.

Schools are also randomly assigned to a color. There are three types of school colors:
2 orange schools, 2 blue schools, and 1 white school.
The color is used to fill the box with the school's name and value.

For each school, the probability of a student being among the top 4 positions of the priority list depends on her color:

- Each orange student has a 30% probability of being among the top 4 positions in the priority list of each orange school, 20% for the white school, and 10% for each blue school.
- Each blue student has a 30% probability of being among the top 4 positions in the priority list of each blue school, 20% for the white school, and 10% for each orange school.
- A separate random draw will determine which students are among the top 4 positions in each school's priority list. The relative ranking of those 4 students will be random. Similarly, the relative ranking of the students who will be ranked 5th, 6th, . . . will be randomly generated.

How you value schools does not depend on your color, and values are generated as in part 1. As before, in each period you create your application list by clicking the school boxes. Each round is comprised of 5 periods.

If something is unclear, you can ask questions now. After answering your questions, we can proceed with the second part.

2.6.2 Instructions for the Immediate Acceptance Mechanism (Mechanism B)

All instructions, except for the description of the mechanism and the illustrating example (as seen below), are identical to the instructions for the Deferred Acceptance mechanism.

Matching Algorithm B

The participants are assigned to schools by a multi-step algorithm.

Step 1:

- Each student applies to the school that is ranked first in her submitted application list.
- Each school accepts the applicants who have applied to it, one at a time, using the school's priority list and starting with the applicant with the highest priority. It does so until it has admitted as many students as its capacity or runs out of applicants. The remaining applicants (if any) are rejected by the school.

  The accepted applicants and their seats are removed from the system: these are their final acceptances.

Step 2, 3, ...: This step is only for the students who have been rejected in the previous step.

- Each of those students apply to the next school on their submitted application list. If a student has already applied to all the schools in her application list, then that student remains unassigned (and the algorithm ends for that student).
- If a school still has available seats remaining at the end of the previous step, then it accepts the applicants who have applied to it, one at a time, using the school's priority

list and starting with the applicant with the highest priority. It does so until it has admitted as many students as its capacity or runs out of applicants. The remaining applicants (if any) are rejected by the school.

The accepted applicants and their seats are removed from the system: these are their final acceptances.

End: The algorithm stops when either no school receives any applications, or no application is rejected.

Example (Mechanism B)

This is a small example that shows how the algorithm works. We consider

- 3 students: A, B and C
- 3 schools: 1, 2, and 3. Each school has 1 seat.

Schools' priority lists of students are described in Table 2.5.

Table 2.5: Schools' priority lists of students

Priority	School 1	School 2	School 3
1st	A	A	C
2nd	C	B	B
3rd	B	C	A

So, for instance, at school 3, student C is considered first, then B, then A.

The application lists submitted by the students are in Table 2.6.

Table 2.6: Students' submitted application lists

	Student A	Student B	Student C
1st application	1	1	2
2nd application	2	2	3
3rd application	3	3	1

So, for instance, student C will first apply to school 2. Then, if rejected, she will apply to school 3. Finally, if rejected again, she will apply to school 1. The other students' application lists are read similarly.

Let's run the algorithm.

Step 1 Students A and B apply to school 1, and C applies to school 2.

- School 1 starts accepting the students who applied to it following the order given by its priority. A is the top priority student, so she is accepted. School 1 cannot accept any more student (it reached its capacity of 1), so B is rejected.
- The other student, C, applies to school 2. She is the only applicant, so she is accepted.

At the end of the first step, students A and C have received their final school acceptances: A at school 1, and C at school 2. B is not accepted at any school. Both schools 1 and 2 have reached full capacity whereas school 3 still has 1 empty seat.

Step 2 Student B applies to school 2.

- School 2 is already full, so B is rejected.

At the end of the second step, B is still not accepted at any school, and school 3 still has 1 empty seat.

Step 3 Student B applies to school 3.

- She is the only applicant, so she is accepted.

No student is rejected, so the algorithm stops.

The final assignment is

Student A	Student B	Student C
School 1	School 3	School 2

CHAPTER 3: Participation in Extra Classes in High School and Its Effect on the College Entrance Examination and Performance: Evidence from Vietnam

3.1 Introduction

While preparations for high-stakes exams can be productive in improving students' knowledge and ability, there are concerns that such activities mostly involve "studying to the test" to game the system. Moreover, intense preparations can be costly and, therefore, not affordable for students of disadvantaged backgrounds, raising equity issues.

In the context of Vietnam, preparations for the national college entrance exam commonly take the form of extra classes organized by individual teachers or cram schools. This paper brings together survey data on the participation in extra classes of high school students and administrative data on college admission and performance to address some of the the matters presented above. First, I document the pattern of participation in extra classes as well as education expenditure of high school students across different demographics. Second, I use the average of these educational investment measures at the local level to proxy for the investments that college students originating from the corresponding locality made during high school. This allows me to quantify their effect on college entrance exam scores and performance in college.

The educational investments observed in the Vietnam Household Living Standards Survey in 2006 confirm the validity of the equity concerns. Students from higher-income households have higher education expenditure and participation in extra classes (both at the extensive and intensive margin). The effect is most notable in the differences between those from the highest income quintile nationwide and the rest of the country. Minority and rural students invest less than their non-minority and urban counterparts even after controlling for income.

I look at how these investments affect college admission and performance for the incoming class of 2007 in a major engineering university in Vietnam. Only extra classes during the school year and non-extra-class expenditure are productive in increasing college entrance exam scores while the rest have no significant effects. Although urban students tend to do better on the exam on average, this is fully captured through their higher investment in high school. In terms of college performance, a higher entrance exam score leads to a slightly higher grade point average at graduation, controlling for academic department fixed effects and investments in high school. Given the inclusion of entrance exam scores as an explanatory variable, neither education expenditure nor participation in extra classes in high school show any significant effects on college performance. However, this does not rule out the possibility that score inflation exists for the entrance exam but is canceled out due to correlations between extra classes and omitted factors that improve college performance.

Throughout the analysis, there are persistent differences between genders. Female high school students tend to receive more investments even though these investments are not significantly more productive for them than for their male counterparts in raising college entrance exam scores. Given the same investments/preparations, females perform slightly worse on the exam. However, conditional on the same entrance exam score, they outperform their male peers in college.

To my knowledge, data on extra classes or tutoring are generally unavailable, and I will be the first to document the pattern of participation in them across different demographics. Additionally, my paper contributes to the literature on education expenditure (Aslam and Kingdon, 2008; Azam and Kingdon, 2013; Acerenza and Gandelman, 2019). Among them, Acerenza and Gandelman (2019) also find more investment in females at the secondary education level in Latin American and Caribbean countries, similar to what I observe in Vietnam.

My paper is also related to the literature on the effect of admission test scores on college performance. In the US context, Rothstein (2004) examines SAT using data from the University of California administrative records for 1996 (for student race, gender, SAT score, and freshman GPA) and California public school records (for high school characteristics data). The author criticizes the previous literature for omitted-variable bias and overstating the predictive power of SAT on college performance. In particular, he shows that the portion of SAT score predicted by student- and high-school-level characteristics has more predictive power than the residual. Furthermore, when characteristics are included as controls, the coefficient of SAT score decreases compare to the more parsimonious model. In the context of the college entrance exam in China, which is more similar to Vietnam, Bai, Chi and Qian (2014) consider two universities with very different rankings in China and find that coefficients for total score are similar for the two universities, but the patterns for subject scores differ, which the authors suggest is due to differences in the major compositions. Meanwhile, Yang (2014) shows that at a national key university in China, the predictive power of subject tests differs by admission track (humanities vs. science) and college major. The pattern is robust when personal characteristics and high school achievements are controlled for.

The rest of the paper is organized as follows. Section 2 describes the setting and data. Section 3 records the pattern of education expenditure and participation in extra classes of high school students nationwide. Section 4 quantifies the effects of these educational investments on college entrance exam scores and college performance. Section 5 concludes.

3.2 Setting and Data

Grade 12 students in Vietnam need to complete a nationwide college entrance examination in order to be considered for college admission. During the years for which I have data, this examination was separate from and happened after the high school graduation examination.[1] The set of subjects students needed to take depend on their specialization track; those applying to engineering universities were typically in track A (Mathematics, Physics, and Chemistry). Every track has three subjects, each graded on a ten-point scale. Additionally, students may receive bonuses due to their disadvantaged background, such as living in remote areas or being an ethnic minority.[2] Admission is based solely on the total score, including bonuses. Although the content of the exam for each subject was the same nationwide, students took the exam at the university to which they were applying. Anecdotally, throughout high school, most students, especially those in urban areas, attended extra classes organized by individual teachers or cram schools in preparation for this exam. Private one-one-one tutoring was also present, but less common. Both extra classes and tutoring were on a voluntary basis.

I have two main data sources: survey data on education expenditure and extra

[1] Since 2015, these two examinations have been combined.
[2] Any ethnicity except for Kinh and Hoa is considered a minority.

classes/tutoring for high school students nationwide from from the General Statistics Office of Vietnam and administrative data on college admission and graduation from a major university in Hanoi, Vietnam. Although the two student populations are distinct, average expenditure and participation in extra classes/tutoring at the provincial level from the first source can be used as proxies for actual values for incoming college students from the same locality.

The data from the General Statistics Office of Vietnam comes from the education section of the Vietnam Household Living Standards Survey in 2006 (VHLSS 2006), which was conducted on a representative sample of households nationwide. The data I obtain consists of 10,877 students, which account for all high school students from the aforementioned households. For these students, I observe their individual answers to education survey questions as well as the characteristics of their households. Important variables at the individual student level includes total expenditure on education, expenditure on extra classes, and participation in extra classes (dummy and frequencies) in the last 12 months, as well as the number of years of participation since grade 1. Although there are similar questions on participation in tutoring, the data is missing for a lot of students, and the majority of the non-missing data indicates non-participation (which is consistent with anecdotal evidence). Therefore, for the rest of the paper, I will concentrate only on extra classes.

The data from the engineering university covers the incoming class of 2007, including two separate data sets: one on admission and one on graduation. The graduation data set is digitalized from a physical archive and matched with the admission data set based on students' name, gender, and date of birth. I was able to match 58% of the admission data to 68% of the graduation data, resulting in an analysis sample of 1477 students. In terms of admission, I observe the score for each of 3 subjects (Maths, Physics, Chemistry) and the bonus (if any). In term of college performance, I observe Grade Point Average (GPA) at the time of graduation, graduation grade (High Distinction/ Distinction/ Credit/ Strong Pass/ Pass), and the graduation date. Other variables include each students' province, district/town, high school during Grade 12, which allow me to match them with proxies for education expenditure and extra class participation from VHLSS 2006. Due to the fact that I only have access to admission and graduation data, there is selection into the sample. In particular, we neither observe students who took the entrance exam but failed to gain admission, nor the college performance of those who did not graduate.

3.3 Pattern of Education Expenditure and Participation in Extra Classes

In this section, I examine the education expenditure and participation in extra classes across demographics. Surprisingly, there is more investment in the education of female students across all measures. As expected, students from higher-income households tend to receive more educational investment, with those from the highest income quintile nationwide being markedly different from the rest. Minority and rural students invest less than their non-minority and urban counterparts even after controlling for income.

Table 3.1 shows the summary statistics for education expenditure and extra class participation for a subsample of 1487 high school students for which the information is available. Total education expenditure for a student in the last 12 months is, on average, 1449.28 thousand Vietnam dong (VND), which is approximately 1.8 times the average monthly total consumption expenditures per capita in urban area and 3.6 times that in rural areas. [3] Expenditure

[3]On average, the biggest component of total education expenditure is tuition. The average monthly total con-

on extra classes is 438.83 thousand VND and on average accounts for about 23% of total education expenditure. Note there is a lot of variation in this particular type of expenditure (the standard deviation is nearly twice the mean while its counterpart for total education expenditure is lower than the mean). In the last 12 months, a substantial majority (86%) of students participated in extra classes during the formal school year (from September to May, excluding 1-2 week(s) of winter holidays). On average, a student is in extra classes for more than half of the available weeks of the school year for nearly 8 hours per week. Participation in extra classes during school holidays (including winter and summer holidays) is more moderate on both the extensive and intensive margins.

I consider the differences in the aforementioned variables across demographics by regressing them on students' and households' characteristics, as seen in Table 3.2. Compared to her male countepart, an average female student receives 103.02 thousand VND more in total education expenditure, have attended extra classes for 0.29 more years since Grade 1, spends more hours per week in extra classes both during the formal school year and school holidays. The coefficients for female when expenditure on extra classes and on the number of weeks in extra class during the school year are the dependent variables are also positive but insignificant. I explore the possibility that this pattern is true conditional of enrolling in high school, but female may be less likely to enroll. However, among individuals at high school age (15-18) who have not yet completed Grade 12 in the full VHLSS 2006 sample, the fractions who have attended schools in the last 12 months are comparable for male and female (37.15% and 36.80%, respectively).

In terms of household income, several measures are included. Compared to a household in the first income quintile nationwide, one in the fifth quintile spends 291.91 VND more on the

sumption expenditures per capita are 811.8 and 401.7 thousand VND for urban and rural areas, respectively.

Table 3.1: Summary Statistics - Education expenditure and participation in extra classes in the last 12 months

	Mean	SD	25th Percentile	Median	75th Percentile
Total education expenditure ('000 VND)	1449.28	1211.23	802.50	1125.00	1633.50
expenditure on extra classes ('000 VND)	438.83	829.43	62.00	200.00	480.00
Extra classes as percentage of total expenditure (%)	23.43	19.95	7.71	20.31	34.72
Years of participation in extra classes since grade 1	4.96	3.00	3.00	4.00	7.00
During the formal school year:					
Participated in extra classes	0.86	0.34	1.00	1.00	1.00
Number of weeks in extra classes	22.48	13.67	10.00	27.00	36.00
Number of hours per week in extra classes	7.81	6.31	4.00	6.00	10.00
During school holidays:					
Participated in extra classes	0.53	0.50	0.00	1.00	1.00
Number of weeks in extra classes	4.04	4.41	0.00	3.00	8.00
Number of hours per week in extra classes	5.51	7.00	0.00	3.00	9.00
N			1487		

Notes: Although there are 10,877 observations in the full sample of high school students in VHLSS 2006, I have information on all of the variable listed above for a subsample of only 1487.
All expenditure are in 2006 currency.

student's education in total, most of which (214.95 VND) is attributable to expenditure on extra classes. Students from fifth-quintile households also spend more time on extra classes (1.35 more year since grade 1, 5.69 more weeks during the school year, 2.62 more weeks and 3.76 hours more per week during the holidays). Those in the middle quintiles do not differ significantly from the lowest one, except for their more frequent participation in extra classes during holidays. All else being equal (including being in the same quintile for *total* household income), one extra dollar in income per member increases total education expenditure on the high school student by 2 cents and expenditure on extra classes by 1 cent. The designation of "poor household", which is specific to the local level, offers additional information to the previous income measures. Specifically, a poor household spends 227.78 VND less on the student's education and have had them attending extra classes for nearly one year fewer on average than a similar household considered non-poor (which must be in a different location). This is true when province fixed effects are already controlled for.

Even after accounting for household income in multiple ways, ethnic minorities invest significantly less than non-minorities, and urban students invest significantly more than rural students across the board. Meanwhile, given the controls, the effect of belonging to a larger household is only significant for total expenditure and years of extra classes, with a relatively small magnitude.

Table 3.2: Linear Regression - Expenditure and Extra Classes (EC) on Student and Household Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Total education expenditure	Expenditure on EC	Years of EC	Weeks of EC (School year)	EC hours/week (School year)	Weeks of EC (School holidays)	EC hours/week (School holidays)
Female	103.02**	30.03	0.29**	0.10	0.73**	0.56**	1.01***
	(42.38)	(29.44)	(0.12)	(0.68)	(0.32)	(0.22)	(0.36)
Ethnic Minority	-290.59***	-46.16	-1.15***	-3.74**	-0.23	-1.91***	-1.52
	(89.21)	(61.79)	(0.26)	(1.76)	(0.84)	(0.58)	(0.93)
Urban	240.77***	157.74***	1.31***	3.15***	0.74*	0.53*	0.34
	(56.56)	(39.63)	(0.16)	(0.86)	(0.41)	(0.28)	(0.46)
Poor household	-227.78***	-65.16	-0.95***	-0.01	-0.06	-0.04	0.44
	(83.53)	(58.15)	(0.24)	(1.57)	(0.75)	(0.51)	(0.83)
Household size	-32.95**	-6.93	-0.08*	-0.10	-0.19	0.07	-0.15
	(16.55)	(11.55)	(0.05)	(0.27)	(0.13)	(0.09)	(0.14)
Income per household member	0.02***	0.01***	2.80E-06	-6.59E-05	-0.0000137	-4.70E-06	-2.49E-05
	(0.00)	(0.00)	(9.06E-06)	(5.43E-05)	(2.59E-05)	(1.78E-05)	(2.87E-05)
Income quintile 2	14.21	-16.26	0.23	1.86	-0.18	1.06**	2.16***
	(84.55)	(58.48)	(0.25)	(1.53)	(0.73)	(0.50)	(0.81)
Income quintile 3	10.91	-3.29	0.31	1.64	-0.62	1.06**	1.37**
	(88.00)	(60.89)	(0.26)	(1.53)	(0.73)	(0.50)	(0.81)
Income quintile 4	128.03	6.53	0.68	1.98	-0.06	1.54***	2.67***
	(91.79)	(63.70)	(0.27)	(1.58)	(0.75)	(0.52)	(0.83)
Income quintile 5	291.91***	214.95***	1.35***	5.69***	0.50	2.62***	3.76***
	(107.03)	(74.33)	(0.31)	(1.81)	(0.86)	(0.59)	(0.96)
Province FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.2927	0.200	0.3244	0.1217	0.0464	0.0946	0.0640
N	2159	2078	2159	1556	1556	1557	1556

Notes: All household information comes from the VHLSS 2006 data set. Ethnic minority refers to any ethnicity except Kinh and Hoa. Income quintile is based on total household income, in comparison to the entire country. All income and expenditure values are in thousand VND in 2006.

3.4 Effect of Extra Classes

3.4.1 On College Entrance Examination

I find that only extra classes during the school year and non-extra-class expenditure other than that on extra classes are productive in improving college entrance exam score. The advantage of urban students disappears when educational investment is controlled for, and female students consistently receive a lower score, conditional on the same amount of preparation.

Table 3.3 presents the results for four different linear regressions with total college entrance exam score as the dependent variables. The effect of expenditure on extra classes on exam score is a pretty precise zero while other education expenditure consistently raises score by 0.001 per thousand VND spent across specifications. Such expenditure includes tuition (and extra tuition for attending non-neighborhood schools); contributions to school, class, and parent funds; uniforms; textbooks and reference books; other stationery; and others. Extra classes during the school year in the last 12 months increase exam score. In contrast, having spent more years participating in extra classes or attending the during school holidays have no significant effects.

The first specification shows that students from urban areas have a higher score than those in rural areas on average, but the effect becomes insignificant once educational investments in high school are controlled for. Regardless of which other explanatory variables are present, female students score approximately 0.8 point lower on average compared to male students. Given gender differential in educational investment recorded in the previous section, I also test the hypothesis that these investments are more productive in increasing college entrance exam scores for female students. To this end, I run another series of regressions (not presented here) that additionally include interactions between gender and educational investment vari-

Table 3.3: Linear Regression - Total college entrance exam score on students' characteristics and proxies for expenditure and extra classes (EC) in high school

	Total College Entrance Exam Score (without bonus)			
	(1)	(2)	(3)	(4)
Constant	21.503***	20.691***	18.893***	20.424***
	(0.070)	(0.287)	(0.714)	(0.568)
Female	-0.801***	-0.796***	-0.794***	-0.787***
	(0.148)	(0.147)	(0.147)	(0.148)
Urban	0.459***	0.209	0.222	0.230
	(0.129)	(0.159)	(0.163)	(0.165)
Expenditure on EC		2.52E-04	-7.20E-05	-1.05E-05
		(4.33E-04)	(4.67E-04)	(5.53E-04)
Other Education expenditure		0.001**	0.001**	0.001**
		(0.000)	(0.000)	(0.000)
Years of EC			-0.019	-0.061
			(0.044)	(0.064)
EC during school year (dummy)			2.120***	
			(0.749)	
EC during school holidays (dummy)			0.197	
			(0.455)	
EC during school year (weeks)				0.047**
				(0.020)
EC during school year (hours/week)				-0.041
				(0.055)
EC during school holidays (weeks)				-0.108
				(0.137)
EC during school holidays (hours/week)				0.049
				(0.077)
R-squared	0.023	0.029	0.034	0.033
N	1477	1476	1476	1476

Notes: The analysis sample includes all students that were matched between the admission and graduation data for the incoming class of 2007 of a major engineering university in Hanoi, Vietnam. Expenditure and extra class proxies are based on province-level average including high school students of all grades from VHLSS 2006.

ables. However, none of these interactions have significant effects on exam score.

3.4.2 On College Performance

I study the effect of education expenditure and participation in extra classes in high school on college performance. The focus is on Grade Point Average (GPA) at graduation as graduation grade is simply a coarser version of the former variable.

Table 3.4 shows that a one-point increase in the total entrance exam score results in a slight increase of 0.077-0.078 point in the graduation GPA. Female students consistently outdo male students with similar observables by nearly 1 point (out of 10) across the different specifications. This is true when selection into different academic departments and the difficulty of the associated curriculum are both controlled for through department fixed effects. Given the controls, being from an urban area does not make any difference on performance.

Once college entrance exam scores, which are affected by extra classes in high school, are accounted for, there may still be two additional effects operating in opposite directions. On the one hand, extra classes may involve studying to the test that inflate exam scores without improving actual ability, resulting in a negative coefficient. On the other hand, extra classes, while not actually productive, may be correlated to omitted variables that do promote performance in college, resulting in a positive coefficent. In this case, neither the extra class nor expenditure variables have significant coefficients. Unfortunately, with the current data, I am unable to distinguish between two possible scenarios: 1) any influence of high school education expenditure and participation in extra classes only operates through what is captured in the entrance exam score, or 2) the two opposing effects outlined earlier are both present and cancel out each other.

Table 3.4: Linear Regression - Graduation GPA on students' characteristics and proxies for expenditure and extra class (EC) in high school

	Graduation Grade Point Average (out of 10)			
	(1)	(2)	(3)	(4)
Constant	6.651***	6.683***	6.800***	6.604***
	(0.027)	(0.068)	(0.221)	(0.170)
College Entrance Exam Score	0.077***	0.077***	0.078***	0.077***
(total without bonus)	(0.008)	(0.008)	(0.008)	(0.008)
Female	0.809***	0.808***	0.808***	0.810***
	(0.043)	(0.043)	(0.043)	(0.043)
Urban	0.012	-0.001	-0.011	-0.013
	(0.035)	(0.044)	(0.045)	(0.046)
Expenditure on EC		9.81E-05	1.41E-04	-2.12E-05
		(1.13E-04)	(1.23E-04)	(1.46E-04)
Other Education expenditure		-6.75E-05	-5.55E-05	-1.11E-04
		(8.35E-05)	(8.43E-05)	(8.51E-05)
Years of EC			-0.005	0.003
			(0.010)	(0.016)
EC during school year (dummy)			-0.069	
			(0.218)	
EC during school holidays (dummy)			-0.103	
			(0.117)	
EC during school year (weeks)				0.002
				(0.005)
EC during school year (hours/week)				0.012
				(0.018)
EC during school holidays (weeks)				0.057
				(0.037)
EC during school holidays (hours/week)				-0.036
				(0.020)
Department FEs	Yes	Yes	Yes	Yes
R-squared	0.312	0.312	0.313	0.314
N	1477	1476	1476	1476

Notes: The analysis sample includes all students that were matched between the admission and graduation data for the incoming class of 2007 of a major engineering university in Hanoi, Vietnam. Expenditure and extra class proxies are based on province-level average including high school students of all grades from VHLSS 2006. Total score on college entrance exam is centered at the sample average. Department fixed effects indicate which academic department the students belong too, which is coarser than major.

3.5 Conclusion

This paper documents the pattern in educational investments (education expenditure and participation in extra classes) of high school students and their effects on performance on the college entrance exam and in college. In terms of educational investments, I use survey data from the Vietnam Household Living Standards Survey in 2006. In terms of college data, I consider the incoming class of 2007 in a major engineering university in Vietnam. The investments these college students made in high school are proxied by the average investments of their peers from the same locality.

As expected, high school students from more advantageous backgrounds invest more. Those with larger household income, especially the highest income quintile nationwide, have higher education expenditure and participation in extra classes (both at the extensive and intensive margin). Minority and rural students invest less than their non-minority and urban counterparts even after controlling for income.

In terms of improving college entrance exam scores, extra classes are only effective during the school year but not during the school holiday. Expenditure on extra classes has no significant effects, while other education expenditure does promote scores. Although urban students tend to do better on the exam on average, this is fully captured through their higher investment in high school. In college, a higher entrance exam score leads to a slightly higher grade point average at graduation. Controlling for entrance exam scores, all measures of educational investment in high school show no significant effects on college performance, but there still exists the possibility of opposing effects offsetting one another.

An important limitation of the paper is due to data restrictions. There is selection into the sample of college students in two aspects. First, it only includes students who successfully gain admission to the college. Second, college performance is only observed for graduating students. Aside from selection, this also means that intermediate performance measures during college remain unobserved. Despite this, the data set can still be improved in other dimensions. The size of the matched sample between admission and graduation data can be increased (the selection into matched vs. unmatched is due to digitalization glitches and presumably independent of any variables of interest). Furthermore, information on the majors of individual students (finer categories than academic departments) can also be obtained.

Bibliography

Abdulkadiroglu, Atila, Nikhil Agarwal, and Parag A. Pathak. 2017. "The Welfare Effects of Coordinated Assignment: Evidence from the New York City High School Match." *American Economic Review*, 107(12): 3635–3689. Working Paper.

Acerenza, Santiago, and Nestor Gandelman. 2019. "Household education spending in latin america and the caribbean: Evidence from income and expenditure surveys." *Education Finance and Policy*, 14(1): 61–87. cited By 1.

Arcidiacono, Peter, and Michael Lovenheim. 2016. "Affirmative Action and the Quality-Fit Trade-off." *Journal of Economic Literature*, 54(1): 3–51.

Arcidiacono, Peter, Michael Lovenheim, and Maria Zhu. 2015. "Affirmative Action in Undergraduate Education." *Annual Review of Economics*, 7: 487–518.

Artemov, Georgy, Yeon-Koo Che, and Yinghua He. 2020. "Stragic 'Mistakes': Implications for Market Design Research." Working Paper.

Aslam, Monazza, and Geeta Gandhi Kingdon. 2008. "Gender and household education expenditure in Pakistan." *Applied Economics*, 40(20): 2573–2591.

Azam, Mehtabul, and Geeta Gandhi Kingdon. 2013. "Are Girls the Fairer Sex in India? Revisiting Intra-Household Allocation of Education Expenditure." *World Development*, 42: 143–164.

Bai, Chongen, Wei Chi, and Xiaoye Qian. 2014. "Do College Entrance Examination Scores Predict Undergraduate GPAs? A Tale Of Two Universities." *China Economic Review*, 30: 632 – 647.

Beggs, S., S. Cardell, and J. Hausman. 1981. "Assessing The Potential Demand For Electric Cars." *Journal of Econometrics*, 16: 1–19.

Bo, Inacio, and Rustamdjan Hakimov. 2020. "Iterative versus standard deferred acceptance: Experimental evidence." *The Economic Journal*, 130: 356–392.

Calsamiglia, Caterina, Guillaume Haeringer, and Flip Klijn. 2010. "Constrained school choice: An experimental study." *American Economic Review*, 100: 1860–1874.

Calsamiglia, Caterina, Guillaume Haeringer, and Flip Klijn. 2011. "A comment on âSchool choice: An experimental study." [J. Econ. Theory 127 (1) (2006) 202-231]." *Journal of Economic Theory*, 146: 392–396.

Chen, Yan, and Onur Kesten. 2019. "Chinese college admissions and school choice reforms: An experimental study." *Games and Economic Behavior*, 115: 83–100.

Chen, Yan, and Tayfun Sonmez. 2006. "School choice: An experimental study." *Journal of Economic Theory*, 127: 202–231.

Corcoran, Sean P., and Christy Baker-Smith. 2018. "Pathways to an Elite Education: Application, Admission, and Matriculation to New York City's Specialized High Schools." *Education Finance and Policy*, 13(2): 256–279.

Ding, Tingting, and Andrew Schotter. 2017. "Matching and chatting: An experimental study of the impact of network communication on school-matching mechanisms." *Games and Economic Behavior*, 103: 94–115.

Ding, Tingting, and Andrew Schotter. 2019. "Learning and mechanism design: An experimental test of school matching mechanisms with intergenerational advice." *The Economic Journal*, 129: 2779–2804.

Dogan, Battal. 2016. "Responsive affirmative action in school choice." *Journal of Economic Theory*, 165: 69–105.

Dogan, Battal, and M. Bumin Yenmez. 2018. "Unified Versus Divided Enrollment In School Choice: Improving Student Welfare In Chicago." Working Paper.

Dubins, Lester E., and David A. Freeman. 1981. "Machiavelli and the Gale-Shapley Algorithm." *American Mathematical Monthly*, 88(7): 485–494.

Dur, Umut, Parag A. Pathak, and Tayfun Sonmez. 2016. "Explicit vs. Statistical Targeting in Affirmative Action: Theory and Evidence from Chicago's Exam Schools." Working Paper.

Dur, Umut, Robert G. Hammond, and Thayer Morrill. 2019. "The secure Boston mechanism: Theory and experiments." *Experimental Economics*, 22: 918–953.

Dur, Umut, Scott Duke Kominers, Parag A. Pathak, and Tayfun Sonmez. 2017. "Reserve Design: Unintended Consequences and The Demise of Boston's Walk Zones." Working Paper.

Ellison, Glenn, and Parag A. Pathak. 2016. "The Efficiency of Race-Neutral Alternatives to Race-Based Affirmative Action: Evidence from Chicago's Exam Schools." Working Paper.

Fack, Gabrielle, Julien Grenet, and Yinghua He. 2019. "Beyond Truth-Telling: Preference Estimation with Centralized School Choice and College Admissions." *American Economic Review*, 109(4): 1486–1529.

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

Guillen, Pablo, and Alexander Hing. 2014. "Lying through their teeth: Third party advice and truth telling in a strategy proof mechanism." *European Economic Review*, 70: 178–185.

Haeringer, Guillaume, and Flip Klijn. 2009. "Constrained School Choice." *Journal of Economic Theory*, 144(5): 1921–1947.

Hahm, Dong Woo. 2019. "What Makes NYC Specialized High Schools So Special?: Relating School Effectiveness and Student Preferences." Working Paper.

Hakimov, Rustamdjan, and Dorothea Kubler. 2019. "Experiments on Matching Markets: a Survey." Working Paper.

Hausman, Jerry A., and Paul A. Ruud. 1987. "Specifying And Testing Econometric Models For Rank-ordered Data." *Journal of Econometrics*, 34: 83–104.

Kapor, Adam. 2016. "Distributional Effects of Race-Blind Affirmative Action." Working Paper.

Kojima, Fuhito. 2012. "School choice: Impossibilities for affirmative action." *Games and Economic Behavior*, 75: 685–693.

Manjunatha, Vikram, and Bertan Turhan. 2016. "Two School Systems, One District: What To Do When A Unified Admissions Process Is Impossible." *Games and Economic Behavior*, 95: 25–40.

Pais, Joana, and Agnes Pinter. 2008. "School choice and information: An experimental study on matching mechanisms." *Games and Economic Behavior*, 64: 303–328.

Roth, Alvin E. 1982. "The Economics of Matching: Stability and Incentives." *Mathematics of Operation Research*, 7(4): 617–628.

Rothstein, Jesse M. 2004. "College Performance Predictions and the SAT." *Journal of Econometrics*, 121: 297 – 317.

Shapiro, Eliza. 2019. "Only 7 Black Students Got Into Stuyvesant, N. Y.'s Most Selective High School, Out of 895 Spots."

Treschan, Lazar. 2015. "The Specialized High School Admissions Debate: Moving from Rhetoric to a Research-Based Solution." Community Service Society of New York.

Turhan, Bertan. 2019. "Welfare And Incentives In Partitioned School Choice Markets." *Games and Economic Behavior*, 113: 199–208.

Yang, Guangliang. 2014. "Are All Admission Sub-tests Created Equal? - Evidence From A National Key University In China." *China Economic Review*, 30: 600 – 617.