American Economic Review 2017, 107(12): 3635–3689
https://doi.org/10.1257/aer.20151425

The Welfare Effects of Coordinated Assignment: Evidence from the New York City High School Match†

By Atila Abdulkadiroğlu, Nikhil Agarwal, and Parag A. Pathak*

Coordinated single-offer school assignment systems are a popular education reform. We show that uncoordinated offers in NYC's school assignment mechanism generated mismatches. One-third of applicants were unassigned after the main round and later administratively placed at less desirable schools. We evaluate the effects of the new coordinated mechanism based on deferred acceptance using estimated student preferences. The new mechanism achieves 80 percent of the possible gains from a no-choice neighborhood extreme to a utilitarian benchmark. Coordinating offers dominates the effects of further algorithm modifications. Students most likely to be previously administratively assigned experienced the largest gains in welfare and subsequent achievement. (*JEL* C78, D82, I21, I28)

In recent years, market design theory has inspired dramatic changes in how children are assigned to public schools across numerous American cities and around the world. The first new system adopted was for placing eighth graders into high schools in New York City (NYC). NYC's new system has not only received widespread scientific and popular acclaim (Economic Sciences Prize Committee 2012; Tullis 2014; Roth 2015), but also became a template for reforms in other cities.[1] Despite widespread adoption of and apparent consensus on the value of market-design-inspired

* Abdulkadiroğlu: Department of Economics, Duke University, 213 Social Sciences Building, Durham, NC 27708, and NBER (email: atila.abdulkadiroglu@duke.edu); Agarwal: Department of Economics, MIT, 50 Memorial Drive, Cambridge, MA 02142, and NBER (email: agarwaln@mit.edu); Pathak: Department of Economics, MIT, 50 Memorial Drive, E52-426, Cambridge, MA 02139, and NBER (email: ppathak@mit.edu). This paper was accepted to the *AER* under the guidance of Pinelopi Goldberg, Coeditor. We thank Neil Dorosin, Jesse Margolis, Sonali Murarka, and Elizabeth Sciabarra for their expertise and for facilitating access to the data used in this study. Special thanks to Alvin E. Roth for his collaboration, which made this study possible. Jesse Margolis provided extremely helpful comments. We also thank Josh Angrist, Steve Berry, Stefan Bonhomme, Eric Budish, Francesco Decarolis, Matt Gentzkow, Arda Gitmez, Jerry Hausman, Phil Haile, Hortaçsu, Peter Hull, Yusuke Narita, Derek Neal, Alex Olssen, Ariel Pakes, Amy Ellen Schwartz, Jesse Shapiro, and seminar participants at the University of Chicago, Harvard University, Federal Reserve Bank of NY, London School of Economics, Yale University, and the NBER Market Design conference for input. We received excellent research assistance from Alonso Bucarey, Red Davis, Weiwei Hu, Danielle Wedde, and a hard-working team of MIT undergraduates. Agarwal acknowledges support from the National Science Foundation (grant SES-1427231). Pathak acknowledges support from the Alfred P. Sloan Foundation (grant BR2012-068), National Science Foundation (grant SES-1056325), and the William T. Grant Foundation. Abdulkadiroğlu and Pathak are on the Scientific Advisory Board of the Institute for Innovation in Public School Choice. The authors declare that they have no relevant or material financial interests that relate to the research in this paper.

† Go to https://doi.org/10.1257/aer.20151425 to visit the article page for additional materials and author disclosure statement(s).

[1] Cities that adopted new coordinated matching systems following NYC include Camden, Denver, New Orleans, Newark, and Washington, DC.

centralized assignment schemes, there is remarkably little evidence on whether, why, and how much a coordinated assignment system affects pupil allocation to schools or the extent to which the new system created losers as well as winners. The empirical performance of alternatives to NYC's deferred acceptance-based scheme, the quantitative aspects of particular design trade-offs, and whether the new mechanism is associated with improvements in downstream educational outcomes also remain open questions.[2]

Characterizing the state of the market prior to the new mechanism is a major challenge because decentralized and uncoordinated systems do not usually generate systematic data. This paper surmounts this hurdle by exploiting new data on the system used in New York before 2003 to study the effect of moving from an uncoordinated assignment system to a coordinated single-offer system for allocating students to schools. Using our rich micro-data on applications, assignments, and enrollments, we describe high school student placement in both systems and assess whether students receive one of their preferred choices. Tracking students from application to assignment allows us to comprehensively describe the drawbacks of NYC's previous system, but still leaves unanswered questions. First, we do not know whether the reform realized most of the possible gains associated with a new assignment system or how those gains were distributed across applicants. Second, we know little about the magnitude of further algorithmic improvements, as considered by the market design literature compared to other design aspects. Our paper addresses these questions using an estimated model of student preferences exploiting the straightforward incentive feature of New York's new system.

Prior to 2003, rising NYC high school students applied to five out of more than 600 school programs; they could receive multiple offers and be placed on wait lists. Students were allowed to accept only one school and one wait list offer, and the cycle of offers and acceptances was repeated two more times. The vast majority of students not assigned in these rounds went through an administrative process that manually placed them at schools close to their residences. Since admissions offers were not coordinated across schools, we refer to this as the *uncoordinated mechanism.*[3] In fall 2003, this system was replaced by a single-offer assignment system, based on the student-proposing deferred acceptance algorithm (DA) for the main round. Applicants were allowed to rank up to 12 programs for enrollment in 2004–2005, and a supplementary round placed students unassigned in the main round. Since the central office coordinated all schools into a single offer, we refer to this new system as the *coordinated* mechanism. The mechanisms could produce different allocations for three main reasons: (i) the new mechanism allows students to rank up to 12 choices, whereas the old mechanism only allowed for 5; (ii) the old mechanism's limited number of offer and acceptance rounds led to congestion, as students held on to less-preferred choices while waiting to be offered seats at more preferred schools after others decline; and (iii) unlike the new mechanism, the old

[2] There is an active scholarly and policy debate about alternative designs. The OneApp process used in the Recovery School District in New Orleans is based on an entirely different assignment algorithm that relaxes the stability constraint in New York's system (Abdulkadiroğlu et al. 2017). Abdulkadiroğlu, Che, and Yasuda (2015) argue that ordinal strategy-proof mechanisms may not lead to improvements in cardinal utility.

[3] Although the offer timetable and number of rounds was uniform across schools, there was no coordination of admissions decisions among schools.

mechanism invited strategic considerations on student ranking, as schools were able to see the entire rank ordering of applicants in the old mechanism, and some advertised they would only consider those who ranked them first.

Offer processing and matriculation patterns provide rich details on how and why the new mechanism improves on the old one. In the old mechanism, 18.6 percent of students matriculate at schools different from their assignment at the end of the match compared to 11.4 percent under the new mechanism. Multiple offers and short rank-order lists in the old mechanism advantage few students but leave many without offers. Roughly one-fifth of students obtain multiple offers, while half of applicants receive no first round offer, and 59 percent of these applicants are administratively assigned. The take-up rates for administratively-assigned students are similar across mechanisms, but the number of students assigned in that round is three times larger in the old mechanism. In addition, 8.5 percent of applicants left the district after assignment in the old mechanism while only 6.4 percent left under the new mechanism. While these observations suggest welfare improvements, it is nevertheless necessary to estimate a model of school demand to quantify the distribution of student welfare effects and the relevance of design issues that are central to the school matching market design literature.

The new mechanism is based on DA, which is strategy-proof. This fact motivates treating stated preferences as true preferences, thereby sidestepping challenges associated with inferring preferences from highly manipulable systems.[4] As far as we know, our paper is the first to fit econometric models of school demand using data generated by DA, though many have argued for strategy-proof mechanisms because they generate credible preference data for guiding policy.[5] Our preference estimates characterize the heterogeneous nature of student preferences and allow us to quantify which aspects of school choice market design are most important for allocative efficiency. That our estimates are robust to variations on our assumptions about ranking behavior and evidence of in-sample and out-of-sample fit reassures us that using stated preferences is suitable for welfare analysis.

We use these estimates to evaluate the allocative and distributional aspect of various assignment mechanisms, an exercise that provides a quantitative counterpart to the theoretical literature on matching market design. To scale the magnitude of welfare effects, we first measure aggregate welfare from two benchmark cases: a neighborhood assignment allocation, wherein each student was assigned to the closest school subject to capacity constraints, and the utilitarian optimal assignment, which maximizes the equally weighted average of distance-equivalent utility. The coordinated scheme achieves 80 percent of this idealized benchmark. Next, we find relatively modest gains from relaxing the mechanism design constraints emphasized by a large theoretical market design literature (Erdil and Ergin 2008; Abdulkadiroğlu, Pathak, and Roth 2009; Kesten 2010; Kesten and Kurino 2012). Had the mechanism produced a student-optimal stable matching, the average student welfare would improve by another 0.6 percent of this range. An ex post Pareto

[4] Section VIIIB probes this assumption in greater detail.

[5] The fact that strategy-proof mechanisms generate reliable demand data is a common argument in their favor (see, e.g., Abdulkadiroğlu et al. 2006; Abdulkadiroğlu, Pathak, and Roth 2009; Sönmez 2013). Following our paper, Pathak and Shi (2014) examine the out-of-sample performance of school demand forecasts using data from Boston's DA-based system.

efficient matching, which abandons the stability constraints in the current mechanism, results in a further improvement of about 2.7 percent of the range. While these alternatives are infeasible without sacrificing some appealing features of the mechanism, this exercise shows that the magnitude of student welfare gains from any potential algorithmic improvements are swamped by the effect of simply having choice in a coordinated system implemented by deferred acceptance.

We then deploy several approaches to use our demand estimates to evaluate the transition from an uncoordinated to a coordinated mechanism. The first approach evaluates welfare changes by treating student behavior in the two mechanisms symmetrically, and the second approach approximates the best-case scenario for the uncoordinated mechanism in which students behave optimally. We find the new mechanism has made it easier for students to obtain a choice they want, and this result is robust to how we interpret data from the uncoordinated mechanism. Even though the coordinated mechanism produces school assignments that are geographically further away, our estimated distribution of preferences indicates the degree to which students prefer these assignments more than compensates for this difference. Under our estimate of an approximate best-case for the uncoordinated mechanism, we estimate that admissions coordination represents about a 45 percent improvement in the range from no choice to the utilitarian optimum. Students across all demographic groups, boroughs, and baseline achievement levels receive a more preferred assignment on average from the new mechanism. The largest gains are for student groups that were more likely to be unassigned after the old mechanism's main round, a result that suggests that the old mechanism's congestion and ad-hoc placement of unassigned students are primarily responsible for misallocation. This comparison also shows that eliminating congestion through offer coordination via DA dominates the allocative effect of further modifications to the matching algorithm within the coordinated system. Finally, we show both test scores and graduation rates increase for students who were most likely to be administratively assigned in the old mechanism.

This paper relates to two distinct literatures on school choice and matching mechanisms. Our focus aligns with research on how choice affects student assignment and sorting (Epple and Romano 1998; Urquiola 2005) rather than the competitive effects of choice on student achievement (Hoxby 2003; Rothstein 2006). Further, we concentrate on allocative efficiency, and only briefly examine subsequent achievement, a topic of several other studies (Abdulkadiroğlu et al. 2011; Deming et al. 2014; Walters 2014; Neilson 2013). The allocative issues on which we focus are likely important for understanding potential long-term effects on residential choices and school productivity. A number of recent papers use micro data from assignment mechanisms to understand school demand (Hastings, Kane, and Staiger 2009; He 2012; Ajayi 2013; Agarwal and Somaini 2014; Calsamiglia, Fu, and Güell 2014; Hwang 2014; Burgess et al. 2015), typically using data from manipulable mechanisms like the Boston mechanism based on specific models of student information and sophistication. While some of these papers have compared the Boston mechanism and DA, ours is the first to examine congestion in an uncoordinated school assignment system. An approach based on estimated preferences complements survey data for comparing mechanisms. For instance, Budish and Cantillon (2012) use survey data on a multi-unit course allocation mechanism and

find that more students are assigned to preferable choices in a strategy-proof mechanism than in the draft mechanism used at the Harvard Business School. Finally, our work relates to comparisons of decentralized and centralized medical labor markets by Niederle and Roth (2003).

I. High School Choice in NYC

A. *School Options*

Aspiring high school students may apply to any school or program throughout New York City. A single school may host several programs that have curricula ranging from the arts to sciences to vocational training. The 2002–2003 High School directory describes program types. Specialized high schools, such as Stuyvesant and Bronx Science, admit students by admissions test performance on the Specialized High Schools Admissions Test (SHSAT).[6] There are three ways in which student screening differs for non-specialized high schools. *Unscreened* programs admit students by random lottery, in some cases giving priority to students from specific residential zones or to students who attend the school's open house. *Screened* programs evaluate students individually using an assortment of criteria, including grades, standardized/diagnostic test scores; attendance and punctuality; interviews; and essays. Such programs might also evaluate students for proficiency in specific performing or visual arts, music, or dance. *Education Option* programs also evaluate students individually, but only for half of their seats. The other half is allocated by lottery. Seat allocation in each half targets the following student ability distribution as measured by citywide seventh grade English language arts scores: 16 percent of seats should be allocated to high performing readers, 68 percent to middle performers, and 16 percent to low performers.

Throughout the last decade, the NYC Department of Education (DOE) opened new small high schools throughout the city, each with roughly 400 students. A big push for these small high schools came as part of the New Century High Schools Initiative launched by Mayor Bloomberg and Chancellor Klein. Eleven new small high schools were opened in 2002, 23 new small schools were opened in 2003, and small high school openings peaked in 2004 (Abulkadiroğlu, Hu, and Pathak 2013). Most of these schools are small and have about 100 students per entering class. As a result, the new small high schools have a relatively minor effect on overall enrollment patterns during our study period, which focuses on school options available in 2002–2003 and 2003–2004.

B. *Uncoordinated Admissions in 2002–2003*

Forms of high school choice have existed in New York City for decades. Before 2002, high school assignment in New York City featured many choice options, mostly controlled by borough-wide superintendents. Significant admissions power resided with school administrators, who could directly enroll students. Admissions

[6] Abdulkadiroğlu, Angrist, and Pathak (2014) describe the exam school admissions process in more detail.

to the specialized high schools and the LaGuardia High School of Music & Art and Performing Arts, however, have been traditionally administered as a separate process, which did not change with the new mechanism.[7] Our study therefore focuses on admissions to non-specialized public schools.

About 80,000 students interested in regular high schools visit schools and attend city-wide high school open houses before submitting their preferences in the fall. In the *main round* in 2002–2003, students could apply to at most five regular programs in addition to the specialized high schools. Programs receiving a student's application were able to see the applicant's entire preference list, including where their program was ranked. Programs then decided whom to admit, place on a waiting list, or reject. Applicants were sent a decision letter from each program to which they had applied, and some obtained more than one offer. Students were allowed to hold on to at most one admission and one wait list offer. After receiving responses to the first letters, programs with vacant seats could make new offers to students from waiting lists. After the second round, students who did not have a zoned high school were allowed to participate in a *supplementary round* known as the variable assignment process. In the supplementary round, students could rank up to eight choices and were assigned based on seat availability negotiations between the enrollment office and high school principals. After replies to the second letter were received, a third round of letters were mailed. New offers did not necessarily go to rejected or wait listed students in a predetermined order. Unassigned students were either placed at their zoned programs or administratively placed as close to home as possible by the central office. We refer to this final stage as the *administrative round*.[8]

Three features of this assignment scheme motivated the NYC DOE to abandon it in favor of a new mechanism. First, there was inadequate time for offers, wait list decisions, and acceptances to clear the market for school seats. DOE officials reported that in some cases, high-achieving students received acceptances from all of the schools to which they applied, while many other students received none (Herszenhorn 2004). Comments by the Deputy Schools Chancellor summarized the frustration: "Parents are told a school is full, then in two months, miracles of miracles, seats open up, but other kids get them. Something is wrong" (Gendar 2000).

Second, some schools awarded admissions priority to students who ranked them first on their application forms. The high school directory advises that when ranking schools, students should "determine what [their] competition is for a seat in this program" (DOE 2002, p. x). This recommendation reflects strategic incentives for ranking decisions. Students have to consider both the limited number of potential applications and whether the school only considers first-choice applicants.

[7] The 1972 Hecht-Calandra Act is a New York State law that governs admissions to the original four specialized high schools: Stuyvesant, Bronx High School of Science, Brooklyn Technical, and Fiorello H. LaGuardia High School of Music & Art and Performing Arts. City officials indicated that this law prohibits including these schools within the common application system without action by the state legislature.

[8] Students who are new to New York City or did not submit an application participate in an "over-the-counter" round over the summer. Our analysis follows applicants through to assignment and therefore does not consider students who joined the process after the high school match. Arvidsson, Fruchter, and Mokhtar (2013) provide further details on the over-the-counter round.

Third, a number of schools managed to conceal capacity to fill seats later on with more preferred students. For example, the Deputy Chancellor stated, "before you might have a situation where a school was going to take 100 new children for ninth grade, they might have declared only 40 seats, and then placed the other 60 outside the process" (Herszenhorn 2004, p. 3). Overall, critics alleged that the old mechanism disadvantaged low-achieving students and those without sophisticated parents (Hemphill and Nauer 2009).

C. *Coordinated Admissions in 2003–2004*

The new mechanism was designed with input from economists (see Abdulkadiroğlu, Pathak, and Roth 2005 and Abdulkadiroğlu, Pathak, and Roth 2009). When publicizing the new mechanism, the DOE explained that its goals were to utilize school places more efficiently and reduce the gaming involved in obtaining school seats (Kerr 2003). As in previous years, in the first round, students apply to specialized high schools when they take the SHSAT. Offers are produced according to a serial dictatorship with priority given by SHSAT scores.[9]

In the *main round*, students can rank up to 12 regular school programs in their applications, which are due in November. The DOE advises parents: "You must now rank schools very carefully, to reflect your true preferences (DOE 2003, p. ii)," because this round is built on Gale and Shapley's (1962) student-proposing deferred acceptance algorithm. Schools with programs that prioritize applicants based on auditions, test scores, or other criteria are sent lists of students who ranked the school, but these lists do not reveal where on the preference lists the schools were ranked. Schools return orderings of applicants to the central enrollment office. Schools that prioritize applicants using geographic or other criteria have those criteria supplied by the central office. That office uses a single lottery to break ties among students with the same priority, thus generating a strict student order at each school.

The centralized clearinghouse assigned schools using DA with inputs that include student preferences, school capacities, and schools' strict ordering. After lottery numbers are drawn, DA works as follows:

Step 1: Each student proposes to her first choice. Each school tentatively assigns seats to its proposers one at a time, following their priority order. The student is rejected if no seats are available at the time of consideration.

Step $k > 1$: Each student who was rejected in the previous step proposes to her next best choice. Each school considers the students it has tentatively assigned together with its new proposers and tentatively assigns its seats to these students one at a time following the school's priority order. The student is rejected if no seats are available when she is considered.

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

[9] There is very limited overlap between the specialized round and subsequent rounds. In 2003–2004, 4,175 out of 4,553 of students offered a specialized high school placement accepted that offer.

DA is run for all students in February. In this first round, only students who receive a specialized high school offer receive an offer letter from a regular school if admitted, and they are asked to choose. After they respond, students who accept are removed from the pool, school capacities are adjusted, and the algorithm is re-run with the remaining students. After the main round, all students receive a letter notifying them of their assignment or whether they are unassigned after the main round.

Students unassigned after the main round receive a list of programs with vacancies and are asked to rank up to 12 of these programs. In 2003–2004, the admissions criteria at the remaining school seats were ignored in this *supplementary round*. Students are ordered by their random number, and DA is run with this ordering in place at each school. Students who remain unassigned in the supplementary round are assigned administratively. These students and any appealing students are processed on a case-by-case basis in the *administrative round*.

II. Data and Sample Restrictions

A. *Students*

For this study, the NYC DOE provided us with several datasets: student choices and assignments, student demographics, and October student enrollment. Each student has a unique identification number. For 2002–2003, the assignment files record students' main round rank-order list, their offers and rejections for each round, whether they participate in the supplementary round, and their final assignment at the conclusion of the assignment process as of July 2003. For 2003–2004, the assignment files contain students' choice schools in order of preference, priority information for each school, assignments at the end of each round, and final assignment as of early August 2004. The student demographic files for both years contain information on home address, gender, race, limited English proficiency status, special education status, and performance on seventh grade citywide tests. We use addresses to compute the road distance between each student and her school and to place each student in a census block group.[10] We also have access to similar files for 2004–2005. Further details are in the online Data Appendix.

Our analysis sample makes two restrictions. First, since we do not have demographic information for private school applicants, we restrict the analysis to students in NYC's public middle schools in the year prior to application. Second, we focus on students who are not assigned to specialized high schools because that part of the assignment process did not change in the new mechanism. Given these restrictions, we have two main analysis files: the *mechanism comparison* sample and the *demand estimation* sample.

The mechanism comparison sample is used to compare the assignment across the two mechanisms. This sample is the largest set of students assigned through the high school assignment mechanism to a school that exists when the high school directory

[10] Though we use road distance, we also computed subway distance using the Metropolitan Transportation Authority GIS files; the correlation between driving and subway commuting distance for all student-program pairs is 0.96.

Table 1—Characteristics of Student Sample

	Mechanism comparison		Demand estimation
	Uncoordinated mechanism (1)	Coordinated mechanism (2)	Coordinated mechanism (3)
Number of students	70,358	66,921	69,907
Female	49.4	49.0	49.0
Bronx	23.7	23.3	23.7
Brooklyn	31.9	34.1	33.3
Manhattan	12.5	11.8	12.0
Queens	25.0	24.8	24.7
Staten Island	6.9	6.0	6.3
Asian	10.6	10.9	10.6
Black	35.4	35.7	35.7
Hispanic	38.9	40.4	40.3
White	14.7	12.6	13.0
Other	0.4	0.4	0.4
Subsidized lunch	68.0	67.4	67.8
Neighborhood income ($)	38,360	37,855	37,920
Limited English proficient	13.1	12.6	12.6
Special education	8.2	7.9	7.5
SHSAT test-taker	22.4	24.3	23.9

Notes: Percents unless otherwise noted. Uncoordinated mechanism refers to the 2002–2003 mechanism and coordinated mechanism refers to the 2003–2004 mechanism based on deferred acceptance. Neighborhood income is the median census block group family income from the 2000 census. SHSAT stands for Specialized High School Achievement Test.

is printed. A key property of the mechanism comparison sample is that every student has an assignment. Columns 1 and 2 of Table 1 summarize student characteristics in the mechanism comparison samples across years; 3,500 fewer students are involved in the mechanism comparison sample in the coordinated mechanism, a difference mainly due to the students assigned either to schools created after the high school directory was printed or to closed schools (as shown in online Appendix Table C2).

New York City is the nation's largest school district, and like many urban districts, the majority of students are low-income and nonwhite. Nearly three-quarters of students are black or Hispanic, and about 10 percent of students are Asian. Brooklyn has the largest number of applicants, followed by Queens and the Bronx, both of which account for roughly one quarter of students. Manhattan and Staten Island account for considerably smaller shares of students at about 13 percent and 7 percent, respectively. Consistent with the sudden announcement of the new mechanism, applicant characteristics are similar across years.

The demand sample, drawn from the assignment files, contains participants in the main round in the new mechanism. These students' school choices represent the overwhelming majority of students. From the set of main round participants, we exclude a small fraction of students who are classified as the top 2 percent because

Table 2—Descriptive Statistics for Schools and Programs

	Uncoordinated mechanism (1)	Coordinated mechanism (2)
Panel A. Schools		
Number	215	235
High math achievement (%)	10.2	10.0
High English achievement (%)	19.1	19.3
Percent attending four-year college	47.8	47.2
Percent inexperienced teachers	54.7	55.6
Attendance rate (%)	85.5	85.7
Percent subsidized lunch	62.5	62.6
Size of ninth grade	465.7	451.3
Percent white	10.9	10.9
Percent Asian	8.7	8.6
Percent black	38.5	38.4
Percent Hispanic	41.9	42.1
Panel B. Programs		
Number	612	558
Screened	233	208
Unscreened	63	119
Education option	316	119
Spanish language	27	24
Asian language	10	9
Other language	6	7
Arts	80	80
Humanities	89	93
Math and science	53	60
Vocational	55	59
Other specialties	163	162

Notes: Panel A reports means and panel B reports counts, unless otherwise noted. Uncoordinated mechanism refers to the 2002–2003 mechanism and coordinated mechanism refers to the 2003–2004 mechanism based on deferred acceptance. The data Appendix presents information on the availability of school characteristics. High math and High English Achievement is the fraction of students who scored more than 85 percent on the Math A and English Regents tests, respectively. Inexperienced teachers are those who have taught for less than two years.

these students are guaranteed a school only if they rank it first, and this may distort their incentives to rank schools truthfully. Additional details on the sample restrictions are in the online Data Appendix.

B. *Schools*

School data were taken from New York State report card files provided by NYC DOE. Program information comes from the official NYC High School Directory made available to students before they submit an application. Table 2 summarizes school and program characteristics across years. The number of schools increases from 215 to 235, and the number of students enrolled in ninth grade per school decreases by about 14 students. This decrease is driven by the replacement of some large schools with smaller schools that took place concurrently in 2003–2004, as described above. Despite these shifts, there is little change in either average school achievement level or school demographic composition as measured by the report

card data. We are not aware of other significant changes in school inputs, recruitment campaigns, or materials, including the format of the high school directory.[11]

Students can choose from among roughly 600 programs throughout the city. Programs vary substantially in focus, post-graduate orientation, and educational philosophy. For instance, the Heritage School in Manhattan is an Educational Option program where the arts play a substantial role in the curriculum, while Townsend Harris High School in Queens is a Screened program with a rigorous humanities department, making it among the most competitive in the city. Using information from high school directories, we identify each program's type, language orientation, and specialty. With the new mechanism, there are more Unscreened programs and fewer Educational Option programs, a change driven by the conversion of many Educational Option programs to Unscreened programs. This change in labeling was due to overlapping admissions criteria and similarity of educational programming. We code language-focused programs as Spanish, Asian, or Other, and we categorize program specialties as Arts, Humanities, Math and Science, Vocational, or Other. Not all programs have specialties, though about 70 percent fall into one of these categories. (Details on our classification scheme are in the online Data Appendix.) The menu of language program offerings and program specialties changes little across years.

III. Congestion and Changes in Assignments

The similarity of student and school attributes in Tables 1 and 2 suggest that there were few major systematic changes in either participant attributes or school supply across years. Moreover, there is no large-scale change in student location across years, as shown in Figure 1, which maps both student and school locations. Taking these facts into consideration, we attribute differences in allocations between 2002–2003 and 2003–2004 primarily to the assignment mechanism rather than changes in student participation or range of school options.

A. *Congestion in the Main Round*

Table 3 reports the number of students assigned across rounds of the uncoordinated and coordinated mechanisms. The most noteworthy pattern is that, in the uncoordinated mechanism, more students obtain their final assignments in the administrative round than in the first round. Panel A of the table shows that 37 percent of students are assigned administratively, compared to 34 percent in the first round. Panel B shows that only 33,894 students obtained one or more first-round offers. This is consistent with application patterns that are concentrated at relatively few schools and conservative yield management practices. As a consequence, only 23,867 students received their final assignment in the first round; 10,027 students with a first round offer were finalized with offers made in subsequent rounds. These students were processed as schools revised offers based on first-round rejections and made new offers in the second and third rounds. However, the relatively small number of students placed in the second and third round indicates that three rounds

[11] Appendix Figure A3 shows that the market share of most programs, except for about 20 rarely-ranked programs in the uncoordinated mechanism, is similar across years.

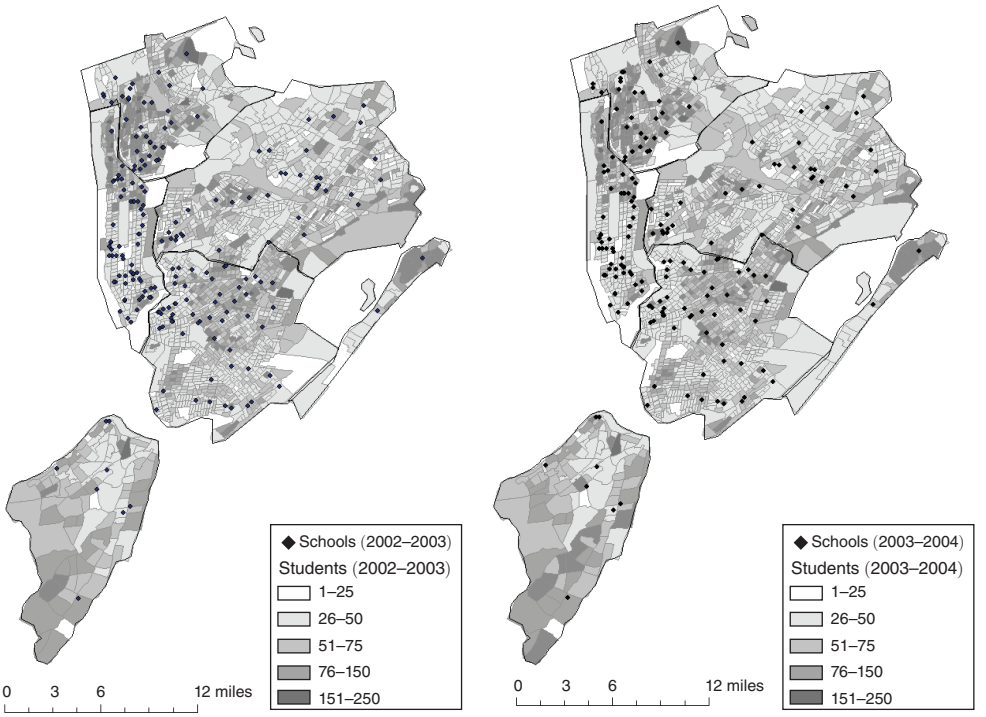

Figure 1. School Locations and Students by New York City Census Tract in 2002–2003 and 2003–2004

Table 3—Offer Processing across Mechanisms

	Number of students (1)	Distance to school (miles) Assignment (2)	Distance to school (miles) Enrollment (3)	Exit from NYC public schools (%) (4)	Enrolled in school other than assigned (%) (5)
Panel A. Uncoordinated mechanism, by final assignment round					
Overall	70,358	3.36	3.50	8.5	18.6
First round	23,867	4.23	4.11	5.2	9.6
Second round	5,780	4.55	4.44	4.8	11.4
Third round	4,443	4.35	4.26	4.9	14.2
Supplementary round	10,170	4.61	4.37	7.8	25.4
Administrative round	26,098	1.64	2.11	13.3	26.8
Panel B. Uncoordinated mechanism, by number of first round offers					
No offers	36,464	2.80	3.12	10.4	24.4
One offer	21,328	3.89	3.85	7.1	13.8
Two or more offers	12,566	4.07	4.03	5.7	9.8
Panel C. Coordinated mechanism, by final assignment round					
Overall	66,921	4.05	3.91	6.4	11.4
Main round	54,577	4.02	3.86	6.1	9.9
Supplementary round	5,201	5.10	4.90	4.8	10.4
Administrative round	7,143	3.50	3.52	9.6	23.6

Notes: Columns 2–5 report means. Uncoordinated mechanism refers to the 2002–2003 mechanism and coordinated mechanism refers to the 2003–2004 mechanism based on deferred acceptance. Student distance is calculated as road distance using ArcGIS. Assignment is the school assigned at the conclusion of the high school assignment process. Enrollment is the school in which a student enrolls in October following application. Assigned students exit New York City if they are not enrolled in any NYC public high school in October following application. Enrolled in school other than assigned means student is in a NYC public high school, but in a school other than that assigned at end of match. Final assignment round is the round during which an offer to the final assigned school was first made.

were insufficient to process all students. That only half of the students were placed in the main round of the old mechanism contrasts sharply with the new mechanism, wherein 82 percent of students were placed in the main round.[12]

These observations about the old mechanism are characteristic of congestion, as described in Roth and Xing's (1997) study of the labor market for entry-level clinical psychologists. In that market, training position offers were made in an uncoordinated fashion during a seven-hour window, and Roth and Xing (1997) argue that uncoordinated processing and a small market-clearing window led to mismatch. In NYC, holding few rounds and serially processing batches of offers, so that programs waited for previous offers to be rejected before making new offers, combined to have similar effects. In addition to insufficient offer processing, the small number of applications allowed in the old mechanism also led to situations where students who applied to oversubscribed schools fell through the cracks. Since rank-order lists were short, the mechanism considered a smaller number of alternate choices for these students. Had more applications been allowed, schools where these students were ultimately placed may have been assigned in the main round.

The new mechanism relieved congestion by increasing both the number of choices students can rank and the number of rounds of offer processing. To investigate the role of these two forces—short rank order lists and limited offer processing—in producing administrative assignments, we used data from the coordinated mechanism to simulate two variations: (i) the main round, with only the top 5 choices considered and no restriction in the number of rounds, and (ii) the main round with 12 choices, but only 3 sets of proposals from the deferred acceptance algorithm. The first simulation is intended to illustrate the role of 5 choices, while the second illustrates the role of few offer-processing rounds. Since we do not model student behavioral responses, we only intend this exercise to shed light on mechanical features generating administrative assignments in the uncoordinated mechanism. The five-choice constraint with an unlimited number of rounds leaves about one quarter of applicants unassigned. The unconstrained mechanism with three proposal rounds leaves roughly half of applicants unassigned. Relative to the uncoordinated mechanism, then, the new coordinated mechanism appears to reduce administrative assignments by computerizing offer processing and avoiding the need for active student and school participation once preferences are submitted. Short rank-order lists also generate administrative assignments, but less so than few offer processing rounds.

B. *Distance*, *Exit*, *and Matriculation*

Across mechanisms, there are stark differences in distance to assigned school and offer take-up. Figure 2 reports the distribution of distance between students' residences and their assigned schools in both mechanisms. New York City spans a large geographic range, with nearly 45 miles separating the southern tip of Staten Island from the northernmost areas of the Bronx, and 25 miles separating the western edge of Manhattan near Washington Heights from Far Rockaway at the easternmost

[12] The marked shift in the number assigned in the main round also appears in the second year of the coordinated mechanism, where even more students, 87.3 percent, were placed in the main round (shown in online Appendix Table B1).

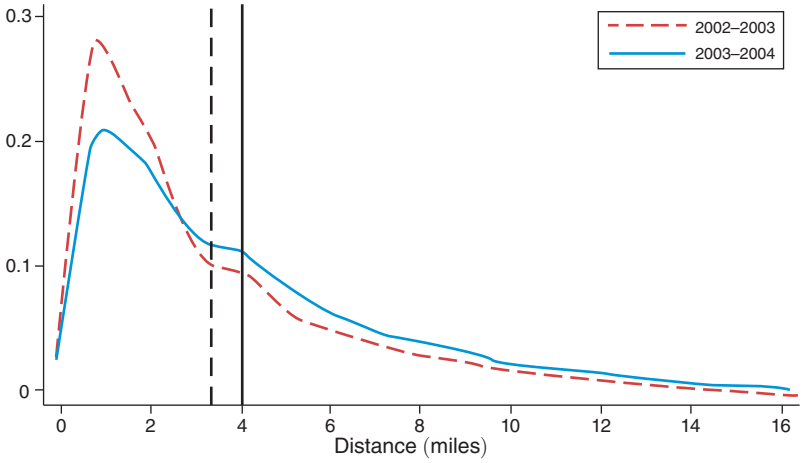

Figure 2. Distribution of Distance to Assigned School in Uncoordinated (2002–2003) and Coordinated (2003–2004) Mechanism

Notes: Mean (median) travel distance is 3.36 (2.25) miles in 2002–2003 and 4.05 (3.04) miles in 2003–2004. Top and bottom 1 percent are not shown in figure. Line fit from Gaussian kernel with bandwidth chosen to minimize mean integrated squared error.

tip of Brooklyn.[13] The closest school for a typical student is 0.82 miles from home, and students in the uncoordinated mechanism on average traveled 3.36 miles to their assigned school. In the coordinated mechanism, the average distance is 4.05 miles. Panels A and C of Table 3 show that average distances were lower in the uncoordinated system because a large number of students were administratively assigned to schools close to their homes.

The increased distance between home and assigned school parallels Niederle and Roth's (2003) study of the gastroenterology labor market, wherein physician mobility increased following a centralized match. While these observations may suggest that coordinated mechanisms expand market scope, in the school choice context daily travel to school imposes a cost on students. It is therefore essential to measure how students value proximity relative to other aspects of their school choices to assess whether improved assignments compensate for the distance increase.

Student enrollment patterns documented in Table 3 indicate that student assignments in the uncoordinated mechanism, particularly those made in the administrative round, are less desirable than assignments made in the coordinated mechanism. After receiving an assignment, a student may opt for a private school, leave New York, or even drop out. Families may switch schools after their final assignments are announced, but before the school year starts.[14] In the uncoordinated mechanism, principals had greater discretion to enroll students, and the DOE officials quoted above alleged that students with sophisticated parents might just show up at a school in the fall and ask for a seat at the school. The exit rate from city public schools is higher in the uncoordinated mechanism (8.5 percent compared to 6.4 percent), and

[13] Our analysis focuses on road distance, which correlates highly with subway distance. Online Appendix B presents a detailed comparison of both measures.

[14] Narita (2016) theoretically and empirically studies mechanisms to handle post-match reassignments.

the fraction of students who enroll at a school other than their assignment is higher (18.6 percent compared to 11.4 percent).

In the uncoordinated mechanism, students assigned in earlier rounds appear more satisfied with their assignments than those assigned in later rounds. The fraction of students who exit NYC public schools is 13.3 percent among administrative placements, compared to 5.2 percent among those assigned in the first round. More than a quarter of students assigned in the administrative round who are still in NYC public schools matriculate at schools other than those to which they were assigned. By comparison, the take-up of offered assignments is much higher for those assigned in the first three rounds. Based on exit and matriculation, students with multiple offers in the first round are more satisfied with their assignment than students with zero or one offers. Students with multiple offers also travel further to their final assignments or enrolled schools. In contrast, the majority of students with no offers are assigned through the administrative round, which likely accounts for their higher rates of exit and enrollment at a school other than their assignment. Even though the coordinated mechanism has substantially fewer administratively assigned students, the exit rates are highest and the matriculation rates are lowest for the participants in that round.[15]

C. *Mismatch in Administrative Round*

To further evaluate student assignments in the administrative round, we compare the attributes of schools that students wanted (i.e., ranked) to the attributes of schools to which they were assigned. Students processed in earlier rounds are assigned to schools with attributes more similar to the schools they ranked than students processed in later rounds. Table 4 shows that, in both mechanisms, students assigned in the main round ranked schools with similar or better attributes, with the exception of distance, than the schools they received. For instance, ranked schools have higher math and English performance, more students attending four-year colleges, and higher attendance rates. Ranked and assigned schools are similar in terms of teacher experience, poverty (as measured by the percent of students receiving subsidized lunch), and racial make-up.

For students placed in the supplementary round, assigned schools are also less desirable than ranked schools, and many of the gaps are wider than in the main round. In the uncoordinated mechanism, for instance, the math performance gap between ranked and assigned schools is 0.7 percentage points for those assigned in the main round and 2.5 points in the supplementary round. The gap between ranked and assigned alternatives for ninth grade size is quite pronounced under both mechanisms. For example, in the uncoordinated mechanism, ranked schools have about 200 fewer ninth graders than the schools where students are assigned. Since students participate in the supplementary round when they do not obtain a main round assignment, it is not surprising that the difference between what students wanted and what they received widens.

[15] In the second year of the mechanism, the average distance to the assigned school is 4.07 miles and the average exit rate is 6.6 percent (shown in online Appendix Table B1). The take-up rate is higher than the first year and the fraction in the administrative round decreases to 5.4 percent.

Table 4—Ranked versus Assigned Schools by Student Assignment Round

	Uncoordinated mechanism		Coordinated mechanism	
	Ranked schools (1)	Assigned (2)	Ranked schools (3)	Assigned (4)
Panel A. Main round				
Distance (miles)	4.82	4.30	5.10	4.00
High math achievement (%)	12.4	11.7	13.0	10.7
High English achievement (%)	20.9	20.2	22.1	19.1
Percent attending four-year college	49.1	47.1	50.6	48.3
Percent inexperienced teachers	45.3	45.6	46.6	43.8
Attendance rate (%)	85.1	84.6	85.7	83.8
Percent subsidized lunch	60.0	60.5	57.6	56.7
Size of ninth grade	694.3	698.8	675.0	819.2
Percent white	15.1	14.7	16.7	17.8
Panel B. Supplementary round				
Distance (miles)	4.87	4.59	5.87	5.17
High math achievement (%)	11.8	9.3	16.6	14.2
High English achievement (%)	19.9	15.8	26.5	20.0
Percent attending four-year college	48.6	44.9	54.1	50.1
Percent inexperienced teachers	46.0	41.5	45.3	36.9
Attendance rate (%)	85.1	82.2	87.4	83.2
Percent subsidized lunch	62.0	61.8	53.5	51.0
Size of ninth grade	685.3	908.0	638.5	1,129.7
Percent white	13.8	13.3	17.4	15.3
Panel C. Administrative round				
Distance (miles)	5.11	1.62	5.33	3.43
High math achievement (%)	14.9	10.5	14.3	10.7
High English achievement (%)	24.3	17.5	24.2	19.2
Percent attending four-year college	52.0	46.7	51.7	47.9
Percent inexperienced teachers	41.9	39.4	47.8	42.1
Attendance rate (%)	85.8	80.8	86.7	82.9
Percent subsidized lunch	53.8	50.4	57.2	53.1
Size of ninth grade	760.6	1,181.9	607.6	984.0
Percent white	18.5	19.1	17.6	17.9

Notes: Means unless otherwise noted. Analysis restricts the sample to students from the welfare sample. Uncoordinated mechanism refers to the 2002–2003 mechanism and coordinated mechanism refers to the 2003–2004 mechanism based on deferred acceptance. Main round in the uncoordinated mechanism corresponds to the first round. Rankings are from the main round. Student distance is calculated as road distance using ArcGIS. See Table 2 notes for details on school characteristics.

The most striking pattern in Table 4, however, is for students who are administratively assigned. As noted above, the uncoordinated mechanism produces three times more students assigned in this round in the uncoordinated mechanism. Panel C shows that, in the uncoordinated mechanism, administratively-assigned students ranked schools, on average, 5.1 miles away from home and were assigned to schools only 1.6 miles away, a much larger gap than in either the main or supplementary round. For other school characteristics, the difference between what students wanted and what they were assigned widens relative to the supplementary round, suggesting that mismatch is greatest for students in the administrative round. For instance, the 2.5 point spread in math achievement in the supplementary round is 4.4 points in the administrative round, and the fraction of students attending four-year colleges similarly widens. The difference in ninth grade size is also considerable for the

TABLE 5—OFFER PROCESSING BY STUDENT TYPE

	Uncoordinated mechanism			Coordinated mechanism		
	Main round (1)	Supplementary (2)	Administrative (3)	Main (4)	Supplementary (5)	Administrative (6)
Students	48.5	14.5	37.1	81.6	7.8	10.7
Female	51.0	14.3	34.7	82.1	7.7	10.2
Bronx	53.3	20.2	26.5	81.7	6.7	11.6
Brooklyn	49.8	16.2	33.9	82.9	8.0	9.1
Manhattan	66.8	19.2	14.0	78.9	7.4	13.7
Queens	43.1	8.3	48.6	79.2	10.0	10.8
Staten Island	11.9	0.0	88.1	88.3	2.4	9.3
Asian	46.1	5.4	48.5	82.3	7.3	10.3
Black	53.2	18.4	28.4	81.3	8.7	10.0
Hispanic	51.2	17.3	31.5	81.8	7.9	10.3
White	31.5	3.8	64.7	81.4	5.0	13.6
High baseline math	57.3	7.4	35.3	85.2	5.1	9.7
Low baseline math	46.8	19.8	33.4	79.9	7.2	12.9
Subsidized lunch	51.8	15.9	32.3	82.7	7.7	9.6
Bottom neighborhood income quartile	55.4	23.3	21.3	81.8	7.2	11.0
Top neighborhood income quartile	41.3	8.1	50.6	80.8	7.4	11.8
Limited English proficient	46.9	16.3	36.8	81.8	7.6	10.7
Special education	38.9	18.8	42.3	71.8	0.0	28.2
SHSAT test-taker	61.9	10.3	27.8	82.6	7.3	10.0

Notes: Table reports mean percentages. Uncoordinated mechanism refers to the 2002–2003 mechanism and coordinated mechanism refers to the 2003–2004 mechanism based on deferred acceptance. Table reports on final assignment round, which is the round during which an offer to the final assigned school was accepted. Neighborhood income is median census block group family income from the 2000 census.

administratively assigned: ranked schools have more than 400 fewer students than assigned schools.

In the coordinated mechanism, the difference between ranked and assigned schools also widen for later rounds. Rank and assigned schools differences in math and English achievement and four-year college going are narrower for the administrative round of the coordinated mechanism than for the uncoordinated one. On the other hand, assigned schools are not as close to home in the coordinated mechanism. Therefore, it is not possible to assert which mechanism's administrative round generates better matches. What is clear, however, is that being processed in the administrative round is undesirable for students in both mechanisms. As a result, it is reasonable to expect that significant changes in student welfare produced by the coordinated mechanism will be driven by reducing the number of administratively assigned.

D. *Offer Processing*

Table 5 reports student attributes across rounds compared to the overall applicant population. Students from Manhattan, those with high math baseline scores, and those who have applied to exam schools (as indicated by taking the SHSAT) tend to obtain offers earlier in the uncoordinated mechanism. These applicants are also overrepresented among students who receive multiple first round offers (not shown). Students from Staten Island, students who are white, and those from high-income

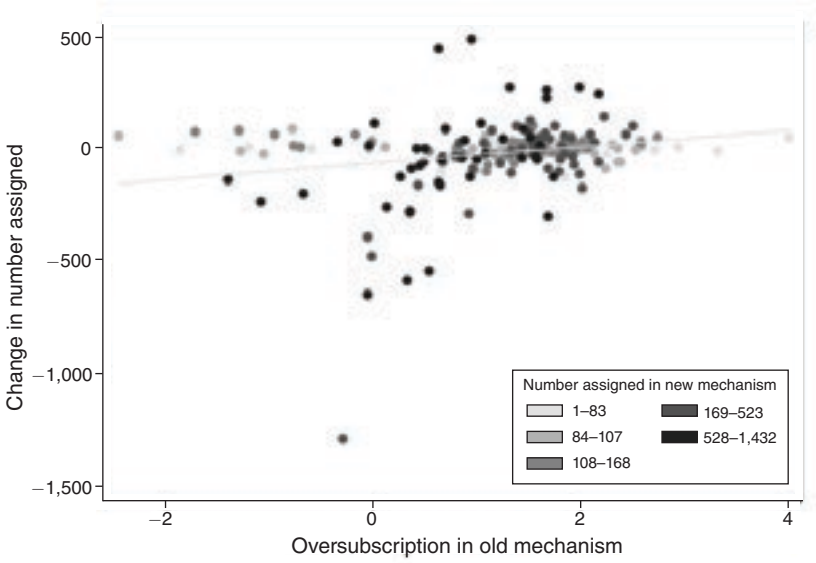

FIGURE 3. CHANGE IN NUMBER ASSIGNED BY OVERSUBSCRIPTION IN UNCOORDINATED MECHANISM

Notes: The figure shows the change in the number of students assigned to the school in the new mechanism minus the old mechanism (on the vertical axis) compared to oversubscription in the uncoordinated mechanism (on the horizontal axis). Oversubscription is measured as the log of the number of applications divided by the number assigned to the program.

neighborhoods tend to obtain offers later in the process. Compared to the overall population, these groups are overrepresented in the administrative round.

The coordinated mechanism distributes school access more evenly across rounds. That is, the differences in the types of students assigned in each round are not as pronounced under the coordinated mechanism. This can be seen by comparing students across boroughs or racial groups in column 4 of Table 5. The fraction of students assigned in the main round is similar across all five boroughs, as is the racial composition of students. Higher baseline applicants are more likely to be assigned in the main round in the new mechanism than low baseline applicants, but are not as overrepresented as in the old mechanism.

The coordinated mechanism also assigns fewer students to schools that were undersubscribed in the uncoordinated mechanism. Figure 3 reports the change in the number of students assigned to a school compared to a measure of how oversubscribed the school was in the uncoordinated mechanism. For example, in 2002–2003, 1,455 students were assigned to the Louis Brandeis High School, a struggling Manhattan high school with four-year graduation rates among the lowest in the city, but in 2003–2004 only 911 students were assigned there.[16] The upward sloping line indicates that if a school is more oversubscribed in the old mechanism, the new mechanism tends to assign more students to that school. This phenomenon suggests that the coordinated mechanism was able to use the submitted preferences to more

[16] The NYC DOE announced the closure of this school in 2009. The largest size reduction is the Evander Childs High School in the Bronx, which went from 1,739 to 453 ninth graders. This high school had a longstanding reputation for violence and disorder, and it was eventually closed in 2008.

effectively place children into their desired schools. The extent to which this change represents an improvement in student welfare depends on the heterogeneity of student preferences, an issue we turn to next.

IV. Estimating Student Preferences

A. *Student Choices*

Families in NYC obtain information about high school programs from many sources including guidance counselors, teachers, and other families. Each year the DOE publishes the *Directory of New York City Public High Schools*, a booklet with information about school size, course offerings, Regents and graduation performance, the school's address, the closest bus and subway, and a description of each program, including its extracurricular activities and sports teams. Families can also learn about schools at high school fairs and open houses and from local newspapers, online guides, and books (e.g., Hemphill 2007).

While a family may rank a school for reasons that we do not observe, the observable dimensions of their choices display consistent regularities: students prefer closer and higher quality schools as measured by student achievement levels, shown in Table 6.[17] The first row of the table shows that only 20 percent of applicants rank 12 school choices; the majority rank 9 or fewer choices, and nearly 90 percent rank at least 3 choices. A student's top choice is on average 4.43 miles away from home. Since the closest school is on average 0.82 miles away, the school closest to home is not most students' first choice. The typical student's first choice is 0.38 miles closer to home than her second choice, and her second choice is 0.24 miles closer than her third choice. Distance increases monotonically until the ninth choice, which is 5.65 miles away.

Lower-ranked schools are also less desirable on other measures of school quality. Math performance decreases going down rank-order lists. (English performance exhibits the same trends as math and is therefore not reported.) Other measures of performance (also not reported) such as the percent of students attending a four-year college and the fraction of teachers classified as inexperienced change monotonically going down rank-order lists. Schools enrolling lower shares of poor students or higher proportions of white students tend to be ranked higher.

Using requests for individual teachers, Jacob and Lefgren (2007) find that parents in low-income and minority schools value a teacher's ability to raise student achievement more than parents in high-income and nonminority schools. In contrast, Hastings, Kane, and Staiger (2009) report that higher-SES families are more likely to choose higher-performing schools than lower-SES ones based on stated reports under Charlotte's school choice plan. This difference across groups motivates our investigation of how baseline ability and neighborhood income influence ranking behavior. High-achieving students tend to rank schools with higher math

[17] Online Appendix Table B3 provides additional information on school assignments. 31.9 percent of students receive their top choice, 15.0 percent receive their second choice, and 2.4 percent receive a choice ranked tenth, eleventh, or twelfth. 17.5 percent of students are asked to participate in the supplementary round because they are unassigned in the main round.

Table 6—School Characteristics by Rank of Student Choice

	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th
Panel A. All students												
Students ranking choice (%)												
Coordinated	100.0	93.4	88.7	82.8	76.2	69.1	62.3	56.0	49.7	43.2	34.4	20.3
Uncoordinated	100.0	93.5	85.8	71.7	46.7							
Distance in miles: mean												
Coordinated	4.43	4.81	5.05	5.21	5.38	5.49	5.59	5.63	5.65	5.58	5.43	5.12
Uncoordinated	4.80	4.91	4.94	4.88	4.79							
Distance in miles: median												
Coordinated	3.51	3.95	4.20	4.37	4.57	4.63	4.72	4.77	4.78	4.71	4.59	4.24
Uncoordinated	3.87	4.00	4.05	4.05	4.02							
High math achievement (%)												
Coordinated	16.7	15.3	14.7	13.9	13.4	12.8	12.4	11.9	11.5	11.1	10.8	10.4
Uncoordinated	14.1	13.3	12.8	12.1	11.7							
Percent subsidized lunch												
Coordinated	51.4	53.4	54.5	56.2	57.4	58.7	59.8	60.6	61.3	61.9	62.7	63.1
Uncoordinated	56.6	58.0	59.1	60.7	62.0							
Size of ninth grade												
Coordinated	713.4	708.1	689.3	668.0	655.3	635.9	620.7	609.5	608.8	612.1	623.8	649.2
Uncoordinated	720.7	720.7	709.3	696.5	686.6							
Percent white												
Coordinated	19.1	16.7	15.7	14.4	13.3	12.2	11.2	10.8	10.4	9.8	9.4	9.0
Uncoordinated	14.6	13.4	12.5	11.4	10.8							
Panel B. Student subgroups												
Split by high math achievement												
Students with low baseline math (%)												
Coordinated	10.9	10.9	10.5	10.1	10.0	9.7	9.6	9.4	9.4	9.1	8.8	8.8
Uncoordinated	9.5	9.5	9.4	8.9	8.7							
Students with high baseline math (%)												
Coordinated	26.0	21.4	20.5	19.1	18.2	17.3	16.4	15.6	15.2	14.2	13.9	12.8
Uncoordinated	21.5	19.0	17.8	16.9	16.1							
Split by neighborhood income												
Students from bottom neighborhood income quartile (%)												
Coordinated	11.4	10.9	10.5	10.4	10.1	9.9	10.0	9.9	9.6	9.3	9.1	8.7
Uncoordinated	9.5	9.6	9.5	9.1	8.7							
Students from top neighborhood income quartile (%)												
Coordinated	23.3	20.7	19.6	18.7	17.7	16.8	16.0	15.1	15.0	14.1	13.2	12.7
Uncoordinated	21.4	18.5	17.6	16.5	16.1							

Notes: Statistics based on 69,907 students for the coordinated mechanism and 59,277 students for the uncoordinated mechanism. Uncoordinated mechanism refers to the 2002–2003 mechanism and coordinated mechanism refers to the 2003–2004 mechanism based on deferred acceptance. Student distance is calculated as road distance using ArcGIS. High math achievement is the fraction of students who scored over 85 percent on the Math A Regents in New York State Report Cards. High baseline math students score above the seventy-fifth percentile for seventh grade middle school math; low baseline math students score below the twenty-fifth percentile. Neighborhood income is median family income from the 2000 census.

achievement relative to low achievers, though both groups place less emphasis on achievement as they move further down their preference list. Similarly, students from low-income neighborhoods tend to put less weight on math achievement than do students from high-income neighborhoods, but both groups rank higher achieving schools higher. These differences suggest the importance of allowing for preferred school achievement levels to differ by baseline achievement and income groups in the demand model.

Schools with lower math achievement or higher fraction subsidized lunch are also ranked lower in the uncoordinated mechanism. For distance, however, the gradient is lower relative to the coordinated mechanism. In the uncoordinated

mechanism, the distance to a students' first choice school is 4.80 miles, while it is 4.79 miles for their fifth choice. In the coordinated mechanism, the fifth choice is about one mile further than the first choice. Such a pattern is consistent with students being more expressive with their choices in the new mechanism, which would be expected given that more choices can be ranked.

All else being equal, based on their submitted preferences, students prefer attending a school closer to home. The fact that students in the new mechanism are assigned to schools further from home might suggest that school assignments are worse on average than in the old mechanism. On the other hand, students may prefer schools outside their neighborhood because these schools are a better fit. To assess the new mechanism, we must therefore weigh the greater travel distance in the new mechanism against changes in other aspects of the assigned school. Our next task is to quantify how students evaluate distance relative to school attributes, including average achievement levels, demographic composition, and size, based on their submitted preferences in the coordinated mechanism.

B. *Model and Estimation*

Comparing the characteristics of schools ranked higher, lower, or not at all reveals students' tastes over these characteristics. To quantify these trade-offs, we estimate a random utility model, which represents the ordinal rankings in cardinal terms. Let $i \in \mathcal{I}$ index students and $j \in \mathcal{J}$ index programs. The indirect utility of student i from program j is

$$u_{ij} = v(\mathbf{x}_j, \mathbf{z}_i, \xi_j, \gamma_i, \varepsilon_{ij}) - d_{ij}, \tag{1}$$

where $\mathbf{x}_j$ is a vector of program j's observed characteristics, $\mathbf{z}_i$ is a vector of observed student characteristics, ξ_j is a program-specific unobserved vertical characteristic, γ_i is a vector capturing idiosyncratic tastes for program characteristics, ε_{ij} captures idiosyncratic tastes for programs, and d_{ij} is distance between student i's address and program j, measured in miles. Our main assumption is

$$(\gamma_i, \varepsilon_{ij}) \perp d_{ij} \,|\, \mathbf{z}_i, \mathbf{x}_j, \xi_j,$$

which states that, conditional on observed student and school characteristics, and the vertical school characteristics ξ_j, unobserved tastes for programs are independent of distance. This assumption may be violated if students systematically reside near schools they prefer, given our controls for student characteristics. In that case, we're likely to underestimate the value of being assigned to a nearby school. On the other hand, the assumption is plausible if the set of observed characteristics is sufficiently rich. The specification also treats distance as the numeraire. All else being equal, if students dislike traveling to school, then the coefficient of -1 on distance is a scale normalization, which allows us to measure utility in distance units, expressed as a "willingness to travel." We also normalize $v(\cdot) = 0$ if all of its arguments are zero.[18]

[18] These scale and location normalizations are without loss of generality and equivalent to other normalizations since the additively separable form of utilities is well defined only up to positive affine transformations.

To assist estimation in finite samples while allowing for a computationally tractable procedure, we parametrize u_{ij} as follows:

$$u_{ij} = \delta_j + \sum_l \alpha^l z_i^l x_j^l + \sum_k \gamma_i^k x_j^k - d_{ij} + \varepsilon_{ij}, \tag{2}$$

$$\text{with } \delta_j = \mathbf{x}_j \beta + \xi_j.$$

We further assume that

$$\gamma_i \sim \mathcal{N}(0, \Sigma_\gamma), \qquad \xi_j \sim \mathcal{N}(0, \sigma_\xi^2), \qquad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma_\varepsilon^2).$$

The vector of random coefficients, γ_i, capture idiosyncratic tastes for program characteristics. It is worth noting that for our welfare calculations, we do not interpret the coefficients on school characteristics as measuring the causal effect of different school characteristics. Throughout the welfare comparisons, we examine the effects of different assignments on utility holding school attributes fixed. This parametrization is an ordered choice version of the model in Rossi, McCulloch, and Allenby (1996), who show that these distributional assumptions allow for estimation via Gibbs' sampling.[19] To this end, we specify conjugate priors for $\theta = (\alpha, \beta, \Sigma_\gamma, \sigma_\xi^2, \sigma_\varepsilon^2)$. For additional details on the procedure and the specification of priors, see the online Computational Appendix.

Intuitively, the degree to which otherwise identical students' propensity to rank a school changes with distance reveals how important distance is relative to other factors. Since our model is nonparametrically identified, these assumptions are made for tractability.[20]

Our specification allows for students to have idiosyncratic tastes for schools that are not captured by the student characteristics in our dataset. To exploit the richness of rank-ordered data, we do not restrict the correlation across the dimensions of γ_i. Berry, Levinsohn, and Pakes (2004) show that data on top and second choices improves on estimates that only use first choice by revealing common characteristics between subsequent rankings for a given student. Rank-ordered data also allow us to relax the common assumption that random coefficients on choice characteristics are independently distributed.

We do not explicitly model an outside option because our primary interest is studying the allocation within inside options rather than substitution outside of the NYC public school system. Moreover, the commonly-used model of the outside option, which infers that a school is unacceptable if not ranked, would require us to

[19] We use Gibbs' sampling rather than simulated maximum likelihood because of biases in datasets with a large number of choices (Train 2009). The posterior means we report have the same asymptotic distribution as maximum likelihood estimates (see van der Vaart 2000, chapter 10.1).

[20] In particular, under our assumption that there is one characteristic (distance) that is additively separable and independent of the unobservable (γ_i, ε_{ij}), we can use variation in distance to trace out the distribution of utility. Identification for binary and multinomial models is studied by Ichimura and Thompson (1998); Lewbel (2000); and Briesch, Chintagunta, and Matzkin (2002). Ordered choice data contains strictly more information than in these settings. Agarwal and Somaini (2014) study identification in the school choice context with a potentially manipulable mechanism. Nonparametric identification results in these settings apply to our case.

assume that students who have not ranked all 12 choices prefer their outside option to a NYC high school.[21] We set the mean of u_{ij} to zero if all observables are zero.

The demand sample for 2003–2004 contains the rankings of 69,907 participants across 497 programs in 235 schools, for a total of 542,666 school choices. Our specifications follow other school demand models and include average school test scores and racial attributes as characteristics (see, e.g., Hastings, Kane, and Staiger 2009). We focus on four main school characteristics: high math achievement, percentage of students receiving subsidized lunch, percent white, and ninth grade size.

We start by assuming that rankings reflect true preferences. For rank-order list $\mathbf{r}_i = (r_{i1}, \ldots, r_{iK_i})$, let r_{ik} be the kth-choice program and K_i is the number of choices i ranks. In terms of indirect utilities, we assume that

$$u_{ir_{ik}} > u_{ir_{ik+1}} \qquad \text{for all } k < K_i,$$

$$u_{ir_{iK_i}} > u_{ij} \qquad \text{for all } j \neq r_{ik} \text{ and } k \leq K_i.$$

This benchmark is motivated by the mechanism's straightforward incentive properties and by the advice the NYC DOE provided in the 2003–2004 High School Directory and their information and outreach campaign.[22] For instance, the DOE guide advises participants to "rank your twelve selections in order of your true preferences" (DOE 2003, p. v) with the knowledge that "schools will no longer know your rankings." (DOE 2003, p. ii)

This model assumes that all programs are available to applicants, though in practice, 193 programs have an eligibility restriction whereby applicants are only eligible for a program if they reside in particular boroughs. We therefore also consider a model in which each applicant's choice set only comprises the programs for which she is eligible. About 5 percent of students rank a program for which they are ineligible and this corresponds to about 1 percent of all submitted choices. In those cases, we include the ineligible program as part of the choice set.[23] Section VIII reports on variations on these assumptions.

C. *Preference Estimates*

We report select estimates for four different specifications in Table 7 and a larger set of specifications in Table A3. The first specification includes school characteristics (high math achievement, percent subsidized lunch, percent white, and ninth grade size), but does not incorporate interactions between school and

[21] Online Appendix Table B3 shows that only about 14,000 students submit all 12 ranks and that students with incomplete lists are also more likely to remain unassigned. Online Appendix Table B4 shows that the majority of these students enter the supplementary round and are likely to accept an assignment in this round. This suggests that the choice of the number of schools to rank is not directly related to preferences.

[22] As mentioned above, this formulation ignores potential information about the relative value of schools less preferable than the K_ith most preferred school. It is consistent with a model in which the student ranks all schools that are preferable to an outside option, but does not require this to be the case.

[23] This specification is similar in spirit to methods in Fack, Grenet, and He (2015) who advocate for restricting the student choice set based on the set of schools that may be achievable for that student. Their setting has the additional advantage that admission criteria for all schools are known, whereas we do not have direct information on admissions criteria used for all screened schools.

Table 7—Select Preference Estimates for Different Demand Specifications

		School characteristics × Student characteristics		
			Models with random coefficients	
	No student interactions (1)	Without random coefficients (2)	All choices (3)	Choice among eligible programs (4)
High math achievement				
Main effect	0.016 (0.016)	0.027 (0.014)	−0.029 (0.018)	−0.058 (0.039)
Baseline math		0.031 (0.001)	0.039 (0.001)	0.050 (0.001)
Percent subsidized lunch				
Main effect	−0.085 (0.007)	−0.057 (0.004)	−0.069 (0.009)	−0.113 (0.058)
Size of ninth grade (in 100s)				
Main effect	−0.164 (0.036)	−0.092 (0.032)	−0.113 (0.048)	−0.153 (0.178)
Percent white				
Main effect	−0.002 (0.014)	0.070 (0.012)	0.062 (0.016)	0.093 (0.062)
Asian		−0.054 (0.002)	−0.075 (0.003)	−0.100 (0.004)
Black		−0.084 (0.002)	−0.124 (0.002)	−0.189 (0.003)
Hispanic		−0.047 (0.002)	−0.084 (0.002)	−0.119 (0.003)
Standard deviation of ε	7.226 (0.010)	7.385 (0.011)	7.858 (0.013)	10.059 (0.022)
Standard deviation of ξ	3.519 (0.121)	2.954 (0.100)	3.676 (0.129)	5.151 (0.650)
Random coefficients (covariances)				
Size of ninth grade (in 100s)			1.584 (0.009)	1.837 (0.012)
Percent white			−0.006 (0.001)	−0.009 (0.001)
Percent subsidized lunch			−0.002 (0.000)	−0.002 (0.000)
High math achievement			−0.011 (0.001)	−0.015 (0.001)
Percent white			0.008 (0.000)	0.013 (0.000)
Percent subsidized lunch			−0.001 (0.000)	−0.002 (0.000)
High math achievement			0.005 (0.000)	0.007 (0.000)
Percent subsidized lunch			0.002 (0.000)	0.003 (0.000)
High math achievement			0.000 (0.000)	−0.001 (0.000)
High math achievement			0.016 (0.000)	0.022 (0.000)
		X	X	X
		X	X	X
	69,907	69,907	69,907	69,907
	542,666	542,666	542,666	542,666

Notes: Select estimates of demand system with submitted ranks over 497 program choices in 235 schools. Distance is calculated using ArcGIS. Dummies for missing school attributes are estimated with separate coefficients. Column 1 contains no interactions between student and school characteristics. Column 2 contains interactions among school characteristics and baseline achievement, gender, race, special education, limited English proficiency, subsidized lunch, and median 2000 census block group family income. Columns 3–4 include random coefficients on school size, percent white, percent subsidized lunch, and math achievement, with unrestricted covariance across characteristics. High math achievement is the fraction of students who score more than 85 on the Math A Regents in New York State Report Cards. Models estimate the utility differences among inside options only. Column 4 restricts each applicant's choice set to include eligible programs. If an applicant ranked an ineligible program, that program is included in the choice set. A total of 193 programs have eligibility restrictions and 3,854 students rank an ineligible program. Standard errors in parentheses.

student characteristics (i.e., $\alpha^l = 0$ for all l). Further, we do not include additional achievement characteristics examined in Table 4, such as high English achievement and percentage of students later attending a four-year college, because these both closely relate to high math achievement.

The next three specifications include student-school interactions. Each specification includes dummies for Spanish, Asian, and Other Language Programs, which are interacted with students' English proficiency status and whether they are Hispanic

or Asian. The model in column 2 assumes that there are no random coefficients ($\gamma_i = 0$), while the model in column 3 places no restriction on γ_i. Column 4 reports estimates wherein the choice set for each applicant is restricted to the set of schools for which she is eligible.

There are three main patterns in Table 7. First, student-school interactions are often estimated precisely; for instance, high baseline math students tend to prefer higher-achieving schools and minority students tend not to prefer schools with high white percentages. Second, the estimates are similar between the model with no restriction on the choice set and the model where the choice set only contains programs for which the student is eligible. Third, many of the random coefficients are significant, thereby suggesting the importance of a flexible specification in accounting for the underlying heterogeneity in student preferences. We further report on model fit in Section VIII. We use the estimates in column 3 for our primary calculations, because these fully exploit all choice data in the most flexible model, and discuss other specifications in Section VIII.

V. Comparing Alternative Mechanisms

A. *Measuring Welfare*

Our estimates allow us to compute measures of welfare for an assignment using the distribution of student preferences. Let $\mu : \mathcal{I} \to \mathcal{J} \cup \mathcal{I}$ denote a matching, such that each student is assigned either to only one program or is unassigned (here, assigned to herself), and no program's matches exceed its capacity. Note that $\mu(i)$ denotes student i's assignment. The average student welfare under μ is

$$W(\mu) = \frac{1}{|\mathcal{I}|}\sum_{i\in\mathcal{I}} u_{i\mu(i)},$$

where u_{ij} is the (indirect) utility student i associated with assignment to program j. Given the normalization in equation (1), welfare is measured in the same units as distance (in miles).

Since individual student utility is not directly observed, we compute the expectation of the utility for each student, $E[\,u_{i\mu(i)} \,|\, \mathbf{r}_i, \mathbf{z}_i, \mathbf{x}_j, \mathbf{d}_i; \theta]$, given the rank-order list $\mathbf{r}_i$, the observed characteristics $(\mathbf{z}_i, \mathbf{x}_j, \mathbf{d}_i)$, and the estimated distribution of preferences parameterized by θ. In what follows, we drop the conditioning on $(\mathbf{z}_i, \mathbf{x}_j, \mathbf{d}_i)$ and θ to simplify the notation. With this convention, we estimate the average student welfare $\bar{W}(\mu)$ as the expectation of $W(\mu)$ defined above:

$$\bar{W}(\mu) = \frac{1}{|\mathcal{I}|}\sum_{i\in\mathcal{I}} E[u_{i\mu(i)} \,|\, \mathbf{r}_i].$$

Evaluating the expectation of each student's utility requires an assumption about the relationship between utilities and observed rank-order lists. If preferences are truthfully submitted to coordinated mechanism, the kth ranked program yields the

kth highest utility. This assumption restricts values of unobserved tastes γ_i and ε_{ij}. The expectation

$$(3)\quad E[u_{ij} \mid \mathbf{r}_i] = \begin{cases} E\left[u_{ij} \mid u_{ir_{i1}} > \cdots > u_{ir_{ik-1}} > u_{ij} > u_{ir_{ik+1}} > \cdots > u_{ir_{iK_i}}\right] & \text{if program } j \text{ ranked } k\text{th} \\ E\left[u_{ij} \mid u_{ir_{i1}} > \cdots > u_{ir_{ik}} > \cdots > u_{ir_{iK_i}} > u_{ij}\right] & \text{if program } j \text{ not ranked,} \end{cases}$$

where K_i is the number of programs listed on rank-order list $\mathbf{r}_i$. The conditioning events in equation (3) translate observed rank-order lists $\mathbf{r}_i$ into restrictions on unobserved tastes γ_i and ε_{ij}. Intuitively, if program j is ranked among the top 12, it must yield higher utility than all unranked programs. Since there are many unranked programs, this means that idiosyncratic tastes for ranked programs are likely to be high. We compute the expression in equation (3) by fixing the estimated parameters at their posterior means, and using a Gibbs sampler to draw from the distribution of utilities given the rank order list.[24] For a student who did not submit a rank-order list in the main round, we compute the mean utility conditional only on $(\mathbf{z}_i, \mathbf{x}_j, \mathbf{d}_i, \theta)$. This simulated value of the expectation of utility is fixed for welfare comparisons across assignments.

The difference in average student welfare between two matchings, μ and μ', is given by

$$(4)\quad \bar{W}(\mu) - \bar{W}(\mu') = \frac{1}{|\mathcal{I}|} \sum_{i \in \mathcal{I}} \left(E[u_{i\mu(i)} \mid \mathbf{r}_i] - E[u_{i\mu'(i)} \mid \mathbf{r}_i]\right).$$

To report differences in average welfare for particular demographic groups, we simply restrict the sample to the set of students with that demographic characteristic when computing equation (4).

B. *Evaluating Mechanism Design Features*

There are several coordinated mechanisms that can be used to organize school admissions. They differ in how they deal with strategic considerations and school priorities. We summarize their theoretical properties before examining these alternatives empirically.

Theoretical Mechanism Design Trade-Offs.—A typical school choice problem can be modeled as a matching problem between students and schools. Each student ranks schools on her application. Each student is also granted an assignment priority at every school for which she is eligible. Priorities may differ among schools and may be determined, for instance, by whether the student resides in a particular borough or has attended the school's information session.

An assignment of students to schools matches each student with at most one school for which she is eligible, or leaves her unassigned, and capacity is not

[24] For each applicant, we use 20,000 draws after burning the first 1,000 draws. This step is identical to the first step in the Gibbs sampler described in online Appendix A.

exceeded at any school. An assignment mechanism produces an assignment using students' reported preferences, school priorities, school capacities, and a lottery to break ties amongst equal priority students to produce an assignment. Once the ties are broken, every school ranks students in a linear order, first by priority at the school and then by lottery numbers, which we will also refer to as tiebreaker. A mechanism is strategy-proof if submitting true preferences is a (weakly) dominant strategy for every student.

To facilitate comparison of alternative allocations, we consider two benchmarks. *Neighborhood assignment* is an allocation where students are assigned to the closest possible school subject to capacity constraints. We compute this assignment by running deferred acceptance (DA), described in Section IC, with applicants simply ranking schools in increasing order of distance and schools ranking students in increasing order of distance. This benchmark corresponds to an extreme without school choice, where the market's geographic scope is limited in a similar fashion as the administrative round of the uncoordinated mechanism.

The second benchmark, *utilitarian assignment*, assigns students to obtain the greatest possible average student utility given school capacities. Given school capacities c_j at each program j, it solves the following program:

$$\max_{a} \sum_{ij} u_{ij} a_{ij} \quad \text{subject to} \quad \sum_{j} a_{ij} \leq 1, \ \sum_{j} a_{ij} \leq c_j, \ a_{ij} \in \{0,1\},$$

where a is an $|\mathcal{I}| \times |\mathcal{J}|$ matrix with (i,j) element a_{ij} denoting that student i is assigned to program j. Since no other feasible allocation yields higher average utility, the utilitarian assignment represents an extreme, where a planner implements the best possible allocation with knowledge of the cardinal distribution of student preferences.

While the utilitarian assignment yields the maximal possible student welfare, it is difficult to achieve for two reasons. First, there are over 200 screened programs in New York City, so implementing this allocation would ignore school-side rankings in those programs. For instance, the utilitarian assignment could allow a lower-scoring student or a student in a lower priority group to be assigned to a school over a student with a higher score or higher priority who also wanted that school.[25] More formally, a student-school pair blocks an assignment if the student strictly prefers the school to his assignment, is eligible for it, and the school either has an available seat or is assigned another student with lower priority. An assignment is stable if it is not blocked by any student-school pair, and if every student prefers her assignment to remaining unassigned. The utilitarian assignment may not be stable.

Second, the utilitarian assignment uses cardinal information, which is not elicited by the coordinated mechanism.[26] An alternative notion of efficiency would only consider students' ordinal rankings of schools. An assignment *Pareto dominates* another, if in the latter assignment every student is assigned to a weakly-preferred

[25] Budish et al. (2013) and He et al. (forthcoming) describe alternative utilitarian mechanisms that take into account how programs rank students.

[26] Bogomolnaia and Moulin (2001, p. 297) argue that focusing on ordinal mechanisms can be "justified by the limited rationality of agents participating in the mechanism."

choice but some students are assigned to a strictly preferred choice than in the former assignment. An assignment is Pareto efficient if it is not Pareto dominated by another assignment; in other words, if there is no way to match some students with better schools in their choice lists without hurting other students.

When priorities do not include any ties, the student-proposing DA used in the coordinated mechanism produces the unique student-optimal stable matching (Gale and Shapley 1962). That is, it produces a stable matching that is not Pareto dominated by any stable assignment. Furthermore, DA is strategy-proof (Dubins and Freedman 1981; Roth 1982). When priorities include ties, however, DA must include a tiebreaker. Tiebreaking opens the door to multiple student-optimal stable matchings and DA may fail to find a student-optimal stable matching (Erdil and Ergin 2008; Abdulkadiroğlu, Pathak, and Roth 2009).

Based on this observation, Erdil and Ergin (2008) suggest an algorithm to find a student-optimal stable matching when priorities include ties. The algorithm uses a stable matching (such as the one produced by DA) as its input and looks for *stable improvement cycles* to improve on this matching. A stable improvement cycle is an ordered set of students such that (i) every student is assigned a school but prefers the assigned school of the next student, and the last student prefers the assigned school of the first student (so that we obtain a cycle); (ii) moreover, each student in the set has the highest priority at the assigned school of the next student among all students who prefer that school to their assignment.

By modifying the original assignment by moving each student to the school to which the next student in this cycle is assigned, we can find an assignment that is both stable and Pareto dominates the original assignment. Consequently, they propose the following *stable improvement cycles algorithm*:

Step 1: Start with the assignment produced by the deferred acceptance algorithm.

Step $k > 1$: If a stable improvement cycle exists in the assignment produced by step $k - 1$, obtain a new stable assignment by moving each student to the assigned school of the next student in the cycle.

Since each step improves upon the assignment of each student in the cycle without creating blocking pairs, the resulting new assignment is stable and Pareto dominates the previous one. The algorithm terminates when no stable improvement cycle is found. While the stable improvement cycles assignment is a student-optimal stable matching, no mechanism is student optimal stable and strategy-proof when priorities include ties (Erdil and Ergin 2008).

A student-optimal stable assignment may still fail to be Pareto efficient (Roth 1982), thus leaving room for improvement in student assignments at the expense of creating blocking pairs. In order to measure the efficiency loss from imposing stability, we find a Pareto efficient assignment that Pareto dominates the stable improvement cycles assignment by employing *Gale's top trading cycles algorithm* (Shapley and Scarf 1974) as follows:[27]

[27] This version of the top trading cycles algorithm starts with the DA outcome in contrast to the version defined by Abdulkadiroğlu and Sönmez (2003).

Step 1: Start with a student-optimal stable assignment. Initialize all students to be available.

Step $k > 1$: Given an assignment, each available student points to her most preferred school, each school points to the highest ranked available student that is assigned to the school, and if no available student is assigned the school, it points to the highest ranked available student.

A top trading cycle is an ordered list of student-school pairs such that each student points to her paired school and the school points to the student of the next pair, where "next pair" for the last pair is the first pair in the ordered list.

Since the number of students and schools is finite, a top trading cycle exists, and each student in the cycle is assigned the school he points to and is then removed.

The algorithm terminates when no more top trading cycles are found.

This mechanism is not strategy-proof, and in general, there is no strategy-proof mechanism that Pareto dominates the deferred acceptance or stable improvement cycles mechanism (Abdulkadiroğlu, Pathak, and Roth 2009; Kesten 2010; Kesten and Kurino 2012).

Quantifying Mechanism Design Trade-offs.—The first column of Table 8 shows that the difference in distance-equivalent utility between the neighborhood and utilitarian assignment is 18.96 miles. In other words, the average student is willing to travel 18.96 miles further for the school she is assigned under the utilitarian assignment compared to the neighborhood assignment. To compute the coordinated mechanism, we run DA for 100 different lottery draws.[28] The coordinated mechanism achieves about 80 percent of the 18.96 mile range, since the difference in willingness to travel from the utilitarian assignment for the average student is 3.73 miles shown in column 2.[29] The gains from a choice system are smaller for whites and Asians compared to blacks and Hispanics. Gains are also considerably smaller for residents of Staten Island, which has on average better-performing schools than other boroughs.

As discussed above, DA does not produce a student-optimal stable matching when priorities involve ties, but no strategy-proof mechanism does. Any potential gains from a Pareto-dominant stable assignment can therefore be viewed as the cost of providing straightforward incentives for students. We quantify this cost by computing a student-optimal stable assignment that Pareto dominates the DA assignment. To this end, we run the stable improvements cycle algorithm for each of the 100 lottery draws. Across these draws, 2,348 students on average obtain a better assignment in a student-optimal stable matching without harming other students. The difference in distance-equivalent utility is 0.11 miles on average compared to the assignment produced by the coordinated mechanism. Because the student-optimal matching Pareto dominates the DA assignment for every lottery draw, this

[28] For students unassigned after the main round, we implement NYC's supplementary round by using preferences from the demand model and assigning students under a serial dictatorship according to the lottery number.

[29] The standard error across parameter draws is 0.04. For a fixed parameter draw, the variance of this difference is small because our sample has 69,907 students.

Table 8—Welfare Comparison of Alternative Mechanisms Compared to Utilitarian Assignment

Assignment mechanism:	Neighborhood assignment	School choice		
		Coordinated mechanism	Student-optimal matching	Ordinal Pareto efficient matching
	(1)	(2)	(3)	(4)
All	−18.96	−3.73	−3.62	−3.11
Female	−18.90	−3.71	−3.59	−3.07
Asian	−18.08	−3.53	−3.43	−3.01
Black	−19.43	−3.89	−3.79	−3.25
Hispanic	−19.37	−3.80	−3.67	−3.10
White	−17.07	−3.21	−3.11	−2.82
Bronx	−21.39	−4.63	−4.46	−3.72
Brooklyn	−18.48	−3.21	−3.14	−2.70
Manhattan	−20.07	−5.40	−5.25	−4.43
Queens	−18.02	−3.39	−3.29	−2.96
Staten Island	−13.82	−1.25	−1.10	−1.03
High baseline math	−18.53	−3.29	−3.18	−2.61
Low baseline math	−19.40	−4.28	−4.18	−3.63
Subsidized lunch	−19.16	−3.78	−3.66	−3.12
Bottom neighborhood income quartile	−19.89	−4.25	−4.12	−3.46
Top neighborhood income quartile	−17.44	−3.63	−3.51	−3.15
Special education	−19.41	−4.83	−4.73	−4.11
Limited English proficient	−19.81	−3.74	−3.64	−3.16
SHSAT test-takers	−19.13	−4.17	−4.05	−3.41

Notes: Utility from alternative assignments relative to utilitarian optimal assignment computed ignoring all school-side constraints except capacity. Utility computed using estimates in column 3 of Table 7. Mean utility from the utilitarian optimal assignment normalized to zero. Column 1 is computed by running the student-proposing deferred acceptance algorithm where applicants simply rank schools in order of distance. Column 2 is from 100 lottery draws of student-proposing deferred acceptance with single tie-breaking using the demand estimation sample. If a student is unassigned, we mimic the supplementary round by assigning students according to a serial dictatorship using preferences drawn from the preference distribution estimated in column 3 of Table 7. Student optimal matching in column 3 computed by taking each deferred acceptance assignment and applying the Erdil and Ergin (2008) stable improvement cycles algorithm to find a student-optimal matching. Ordinal Pareto efficient matching in column 4 computed by applying Gale's top trading cycles to the economy where the student-optimal matching determine student endowments, followed by the Abdulkadiroğlu and Sonmez (2003) version of top trading cycles with counters.

difference is non-negative for every realization of utilities and is therefore statistically significant.

A stable assignment constrains student welfare through its treatment of school priorities and preferences. We next compute welfare from a Pareto-efficient assignment that dominates the student-optimal stable assignment via Gale's top trading cycles algorithm. Since this mechanism does not produce a stable outcome, it is possible that schools benefit by offering students seats outside of the assignment process. The difference in aggregate student welfare under this Pareto-efficient assignment and the student-optimal stable matching may therefore be viewed as the cost of providing incentives for schools for participating in the system.[30]

For each of the 100 lottery draws, we calculate a Pareto-efficient matching that dominates each student-optimal stable matching and report the average welfare

[30] Balinski and Sönmez (1999) and Abdulkadiroğlu and Sönmez (2003) provide an alternative equity rationale for stability. No stable mechanism eliminates strategic participation issues for schools (Sönmez 1997; Ekmekci and Yenmez 2016), although this may not be an issue in markets with many participants (Kojima and Pathak 2009).

relative to the utilitarian assignment in column 4 of Table 8. A total of 10,882 students obtain a more preferred assignment at a Pareto-efficient matching. An ordinal Pareto-efficient allocation still produces substantially lower welfare than the utilitarian optimum. The utility difference for the average student is 3.11 miles. Relative to the coordinated mechanism, the cost of limiting the scope for strategizing by schools (i.e., by imposing stability) is 0.62 miles.

In summary, these comparisons show that the difference in student welfare between having a choice system with the coordinated mechanism and neighborhood assignment is much larger than possible welfare gains from fine-tuning the algorithm used in the coordinated mechanism. That is, the ability to choose schools generates substantial student welfare when preferences are heterogeneous. Further optimizing the matching algorithm in NYC is likely to produce relatively little gain in the best case, even if it were possible to implement cardinal allocation schemes. This conclusion does not imply that the matching scheme is unimportant especially in light of the large number administratively assigned in the uncoordinated mechanism. To see where the uncoordinated mechanism lies in the 18.96 mile range, we next turn to analyzing its properties.

VI. Comparing the Coordinated and Uncoordinated Mechanisms

A. *Approach*

We use the estimated preference parameters and observed data to quantify the difference in welfare between assignments produced by the coordinated and uncoordinated mechanisms. Results presented in the previous section used information embedded in a student's submitted ranking when evaluating the expectation of her utility for the observed assignment using equation (3). This exercise relied on the fact that the coordinated mechanism has straightforward incentives. Interpreting rankings submitted in the uncoordinated system, however, is significantly more challenging since the uncoordinated mechanism is not strategy-proof and, consequently, submitted rankings depend on applicant beliefs and behavior.[31] This difficulty is not unique to our setting and is likely a challenge in evaluating other assignment systems where the incentive properties of the mechanism are not well understood and agent behavior is difficult to model. We therefore analyze the uncoordinated mechanism using two extreme models of agents' behavior: *truthful reporting* and *optimal reporting*. Neither extreme is a plausible characterization of behavior under the uncoordinated mechanism, but the former is focal, and the latter represents a best-case analysis for the uncoordinated system and therefore represents a lower bound on the difference between the coordinated and uncoordinated mechanisms.

Truthful reporting parallels the assumption made for the coordinated mechanism. Although we do not presume all participants were truthful, it is a natural definition

[31] Recent work by Kapor, Neilson, and Zimmerman (2017) surveys parents to elicit their beliefs about admissions probabilities in the New Haven choice mechanism and incorporate them as part of an empirical strategy to simulate alternative mechanisms. Budish and Cantillon (2012) utilize survey data from a manipulable mechanism to make statements about changes in mechanism design. Unfortunately, similar survey data does not exist in our setting.

of unsophisticated behavior.[32] Suboptimal reporting is consistent with the fact that about one-third of applicants who submitted preferences end up unassigned, shown in Table 3. Had applicants anticipated they would be unassigned in the main round, they may have ranked schools that would admit them.

To describe our approach, let $\mathcal{I}$ be the set of students in the coordinated mechanism and $\mathcal{I}'$ be the set in the uncoordinated mechanism. Suppose μ is the matching produced by the coordinated mechanism and μ' is the matching produced by the uncoordinated mechanism. Given our estimate of θ from the coordinated mechanism, we estimate the difference in average student welfare assuming truthful reporting in the uncoordinated mechanism as follows:

$$\bar{W}^T(\mu) - \bar{W}^T(\mu') = \frac{1}{|\mathcal{I}|}\sum_{i\in\mathcal{I}} E[u_{i\mu(i)}|\,\mathbf{r}_i] - \frac{1}{|\mathcal{I}'|}\sum_{i\in\mathcal{I}'} E[u_{i\mu'(i)}|\,\mathbf{r}_i], \tag{5}$$

where each expectation is computed as equation (3).[33]

Optimal reporting assumes that applicant rank-order lists are optimal given their (correct) beliefs on admissions probabilities. While it is unlikely that all applicants are sophisticated enough to respond optimally, considering this alternative generates a lower bound on the difference in average student welfare between the coordinated and uncoordinated mechanism. To see why, let O_i be the set of schools that make an offer to student i in the main round of the uncoordinated mechanism. For each applicant i, let $p_i^j(\mathbf{r}_i)$ be the probability that student i is offered a seat at program j given report $\mathbf{r}_i$ in the main round. We assume that students have private information about their preferences, and correct forecasts about their admissions chances at various schools, which depend on the ranking strategies of all other students. We omit the dependence on these ranking strategies for notational simplicity. Probability $p_i^j(\mathbf{r}_i)$ is zero if program j is not ranked in $\mathbf{r}_i$. Since the uncoordinated mechanism is not a single-offer mechanism, O_i may have more than one element. Each student picks her most preferred option in O_i. Therefore, given O_i, her utility from the most preferable assignment in the main round is $\max_{j\in O_i} u_{ij}$. If an applicant does not receive any offers, she expects to participate in the supplementary and administrative rounds. Let $\mathbf{q}_i$ be a probability vector giving the odds student i is assigned to program j, q_i^j, in these rounds. The expected utility from submitting $\mathbf{r}_i$ is, therefore,

$$EU(\mathbf{r}_i;\mathbf{u}_i) = E\left[\max_{j\in O_i} u_{ij}\,\middle|\,\mathbf{r}_i;\mathbf{u}_i\right] + \Pr(O_i = \emptyset|\,\mathbf{r}_i)\sum_j q_i^j u_{ij}. \tag{6}$$

The first term in the right-hand side of expression (6) reflects the fact that applicants obtain utility from their most preferred school when they have more than one offer. The second term represents the expected utility when an applicant is not offered any school in the main round. Let $u_i^{(k)}$ be the utility for the kth best program given $\mathbf{u}_i$,

[32] Experimental, empirical, and theoretical studies of the Boston mechanism assume truthful behavior is focal. See, e.g., Chen and Sönmez (2006); Hastings, Kane, and Staiger (2009); Pathak and Sönmez (2008).

[33] Under the assumption of truthful reporting, it would be possible to estimate demand directly from rank-order lists submitted in the uncoordinated mechanism. We did not do that because some applicants may not have submitted truthful reports, and so we felt more comfortable using the preference estimates obtained from the coordinated mechanism. Instead, we interpret the results in this section as providing a range for the plausible answers.

and let $p_i^{(k)}$ be the probability of an offer at that program. Expression (6) can be written without the O_i notation as

$$(7)\quad EU(\mathbf{r}_i;\mathbf{u}_i) = \sum_{k=1}^{J} u_i^{(k)} \prod_{\ell=1}^{k-1} (1 - p_i^{(\ell)}(\mathbf{r}_i)) p_i^{(k)}(\mathbf{r}_i) + \left[\prod_{k=1}^{J} (1 - p_i^{(k)}(\mathbf{r}_i))\right] \sum_{j=1}^{J} q_i^j u_{ij}.$$

Suppose a student knows the probability of receiving an offer from a program given her rank-order list. The rank-order list $\mathbf{r}_i^*$ is an *optimal report if* it maximizes her expected utility:

$$(8)\qquad \mathbf{r}_i^* \in \arg\max_{\mathbf{r}_i'} EU(\mathbf{r}_i';\mathbf{u}_i).$$

The average student welfare of the uncoordinated assignment when each student submits an optimal report is

$$\bar{W}^* = \frac{1}{|\mathcal{I}'|} \sum_{i\in\mathcal{I}'} E[EU(\mathbf{r}_i^*;\mathbf{u}_i)],$$

where the expectation is simulated by drawing from the estimated distribution of $\mathbf{u}_i$ given $(\mathbf{z}_i, \mathbf{x}_j, \mathbf{d}_i, \theta, \xi)$ and solving for $\mathbf{r}_i^*$.[34] The definition of $\mathbf{r}_i^*$ implies that

$$E[EU(\mathbf{r}_i^*;\mathbf{u}_i)] \geq E[EU(\mathbf{r}_i;\mathbf{u}_i)] \qquad \text{for any } \mathbf{r}_i.$$

The average student welfare when each applicant submits an optimal report, therefore, provides a best-case scenario across possible application strategies in the uncoordinated mechanism.

This exercise assumes that students in 2002–2003 were submitting optimal reports, and these reports generated $\mathbf{p}$ and $\mathbf{q}$, the equilibrium reduced-form admissions probabilities, which we estimate. The approach imposes consistency of the admissions probabilities in the data and ranking strategies. It does not imply that the welfare associated with the old mechanism is the best possible equilibrium given the 2002–2003 mechanism. Students might have, for instance, all done better by coordinating and ranking the school they obtain from the student-optimal stable matching. However, strategies from such an equilibrium would not be optimal given the strategies and admissions rules that generated the data. This approach of computing $\bar{W}^*$ differs from our approach under truthful reporting, which uses data on applicant rank-order lists and assignments from the uncoordinated mechanism directly. Given the large number of choices and no analytic formula for what rankings imply about unobserved tastes in the uncoordinated mechanism, we only use data from the uncoordinated mechanism to measure $(\mathbf{p}, \mathbf{q})$.

While optimal reporting provides a valuable lower bound, we need to tackle two issues in solving the optimization problem given by equation (8). First, we must specify the information students have about $(\mathbf{p}, \mathbf{q})$. Second, given $(\mathbf{p}, \mathbf{q})$ solving

[34] θ is set to the posterior mean. For 423 programs that existed in 2003–2004, ξ is set to the posterior mean. For the remaining programs, ξ is drawn from $\mathcal{N}(0, \sigma_\xi^2)$. For each applicant, we use 1,000 draws of $(\gamma_i, \varepsilon_i)$ from their respective distribution given θ.

equation (8) requires iterating through all possible rank orderings of length up to five to find the maximum for each applicant. When the number of schools is large, this is computationally infeasible.[35]

We estimate $(\mathbf{p}, \mathbf{q})$ using data from the uncoordinated mechanism, using a simple reduced form for school admissions decisions. Given a dataset of applications $\mathbf{r}_i$ and offers, we fit the logit equation

$$p_i^j(\mathbf{r}_i) = G\Big(\tau_j + \mathbf{z}_i\lambda_1 + \sum_l \eta^l z_i^l x_j^l + 1\{r_{i1} = j\}(1, n_j^r, n_j^a)\,\lambda_2\Big), \tag{9}$$

where G is the standard logistic function, τ_j is a program fixed-effect, $\mathbf{z}_i$ is a vector of the same student characteristics as in the demand models with comformable parameter λ_1, $\mathbf{x}_j$ are program-type dummies with parameter η, $1\{r_{i1} = j\}$ is an indicator if applicant i ranked program j first, n_j^r and n_j^a are the number of students that ranked and were admitted to program j, and λ_2 is a three-dimensional column vector. The admission probability depends on the program (via the program fixed effect), student characteristics, and their interaction (via the interaction of program type and student characteristics). Since schools could observe the student's rank-order lists, it also distinguishes between the probability of an offer from a school based on whether it was ranked first, and differentially so based on the number of students that ranked the program and were admitted.[36] We estimate $\mathbf{q}_i$ as the empirical frequency of school assignments for unassigned students in the same geographic area defined by student i's zoned high school. For the vast majority of students, the student's zoned high school predicts a greater than three-quarters chance of assignment into a single program. The zone school is almost a perfect predictor of the school to which a student will be administratively assigned.

Given our estimate of $(\mathbf{p}, \mathbf{q})$, to solve the optimization problem, we approximate student welfare from optimal reports by limiting the set of reports a student considers to include only a subset of programs based on her preferences. In particular, we assume that agents consider the set of programs which give the highest utility for a student given our estimate θ. Given a utility vector $\mathbf{u}_i$ for student i, let $J_{i\ell}$ be the set of the ℓ-highest utility schools for applicant i. Define the optimal report, given that only the top ℓ choices are considered, as

$$\mathbf{r}_i^{*,\ell} \in \underset{\mathbf{r}_i', \text{ s.t. } r_{ik}' \in J_{i\ell}}{\arg\max}\, EU(\mathbf{r}_i', \mathbf{u}_i) \tag{10}$$

and the corresponding student welfare as

$$\bar{W}^{*,\ell} = \frac{1}{|\mathcal{I}'|} \sum_{i \in \mathcal{I}'} E\big[EU(\mathbf{r}_i^{*,\ell}, \mathbf{u}_i)\big]. \tag{11}$$

[35] With 612 programs, this requires a search over $612 \times 612 \times 611 \times 610 \times 609$ (since it is possible to submit an incomplete list) or more than 85 trillion possible rank orderings.

[36] Specifications that allowed the probability to differ if a program is ranked second show little evidence of a systematic second-choice effect.

As ℓ increases, the $J_{i\ell}$ approaches the set of all schools. Therefore, the expected utility of the optimal report from this restricted set, $\mathbf{r}_i^{*,\ell}$, approaches the expected utility from the optimal report when considering all programs, $\mathbf{r}_i^*$. As a result, our approximation of $\bar{W}^*$ improves as ℓ increases and equals the best-case average student welfare when ℓ equals the size of the choice set. In the analysis that follows, we consider three different values of $\ell = 10$, 15, and 20. We have not been able to compute equation (10) for larger values of ℓ, but seeing how the estimate changes for the three values of ℓ we can compute informs us about what to expect from larger values of ℓ.

B. *Estimates*

Under the truthful reporting assumption for the uncoordinated mechanism, the improvement in average student welfare between the uncoordinated and coordinated mechanism equals 10.62 miles with a standard error of 0.64 miles. Figure 4 shows the distribution of welfare from the two mechanisms. The bimodal distribution of utility in the uncoordinated mechanism is caused by students who are assigned in the administrative round. In the coordinated mechanism, most of the mass in the first mode shifts rightward, a phenomenon driven by the sharp reduction in administrative placements. The shift in the distribution is broad-based: each student group shown in Table 9 experiences a positive gain from the coordinated mechanism.

We've seen in Table 3 that the new mechanism assigned students 0.69 miles further from home. The magnitude of the difference in average student welfare far exceeds this increased travel distance. The lowest gains in Table 9 are for Manhattan residents, who experience no increase in travel distance. However, the welfare gains are not solely driven by changes in distance. For instance, Staten Island pupils only travel 0.34 miles further, but they experience the largest improvement of any borough at 22.62 miles. This suggests mismatch was particularly severe in Staten Island and is consistent with the substantially larger fraction of Staten Island residents who are administratively assigned in the uncoordinated mechanism.[37]

The welfare gains in the coordinated mechanism are larger for many disadvantaged groups, as seen in Table 9, a pattern consistent with Hemphill and Nauer's (2009) claim that the uncoordinated mechanism advantaged high-achieving students and those with sophisticated parents. For instance, welfare gains are larger for low baseline math students than for high baseline math students. Gains are also higher for limited English proficient students than for SHSAT test-takers. However, the difference for Staten Island, which has a larger white population and wealthier neighborhoods, plays a significant role in the fact that whites and those from rich neighborhoods experience larger welfare gains than blacks, Hispanics, and those from poorer neighborhoods.

[37] The 22.62 mile difference in Staten Island is larger than the 13.82 mile difference between the neighborhood and utilitarian assignment in Table 8, a result that suggests the uncoordinated assignment is actually worse than neighborhood assignment in Staten Island. In the uncoordinated mechanism, there are 1,054 students who ranked Staten Island Technical, a highly sought-after screened school. Only 16 percent are assigned there, and about 75 percent do not obtain a main round offer and are subsequently administratively assigned. In the coordinated mechanism, there is also a large reduction in the number assigned to four main Staten Island schools: New Dorp, Tottenville, Port Richmond, and Curtis. Staten Island students submitting rankings were unlikely to highly rank these schools.

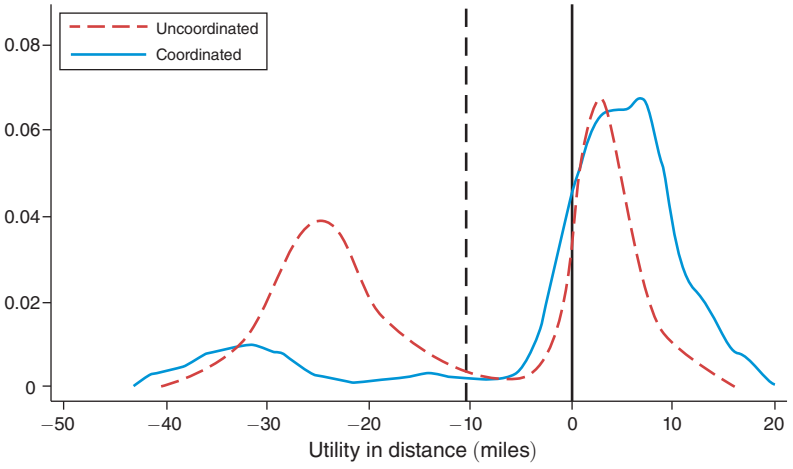

FIGURE 4. STUDENT WELFARE FROM UNCOORDINATED AND COORDINATED MECHANISM

Notes: Distribution of utility (measured in distance units) from assignment based estimates in column 3 of Table A1 with mean utility in 2003–2004 normalized to zero. Top and bottom 1 percent are not shown in figure. Line fit from Gaussian kernel with bandwidth chosen to minimize mean integrated squared error.

The estimates discussed above may be optimistic because they are not based on the best-case scenario for the uncoordinated mechanism. Table 10 reports the extent to which truthful reporting may be biased in favor of the uncoordinated mechanism. Specifically, the table reports

$$\Delta^{\ell} = \frac{1}{|\mathcal{I}'|}\sum_{i\in\mathcal{I}'} E\big[EU(\mathbf{r}_i^{*,\ell}, \mathbf{u}_i)\big] - \frac{1}{|\mathcal{I}'|}\sum_{i\in\mathcal{I}'} E\big[EU(\mathbf{r}_i^{T}, \mathbf{u}_i)\big],$$

for $\ell = 10, 15$, and 20 where $\mathbf{r}_i^T$ is the truthful report, i.e., r_{ik}^T is the program with the kth highest utility in $\mathbf{u}_i$, and a student only ranks a school if it is preferred to remaining unassigned in the main round. The difference between our approximation of optimal reporting and truthful reporting in the uncoordinated mechanism indicates that the range under which our behavioral assumptions about rankings submitted in the uncoordinated mechanism may alter the conclusions about the two mechanisms.[38]

There is a large difference in behavior between estimates that assume truthful reporting and our approximation of optimal reporting. Using our approximation computed from the top ten choices, only about one-tenth of applicants submit the same rank ordering as they would if they submitted preferences truthfully. As we improve our approximation by considering larger choice sets, the fraction of applicants who have the same optimal report increases. Roughly half of applicants

[38] If applicants submitted truthful reports in the uncoordinated mechanism and $(\mathbf{p}, \mathbf{q})$ represents the corresponding admissions probabilities, then

$$\bar{W}^T(\mu') = \frac{1}{|\mathcal{I}'|}\sum_{i\in\mathcal{I}'} E\left[EU(\mathbf{r}_i^T, \mathbf{u}_i)\right],$$

where μ' is assignment in the uncoordinated mechanism.

Table 9—Welfare Comparison between Coordinated and Uncoordinated Mechanism

	Change in utility net distance (miles)		Change in distance (miles)	
	Assignment (1)	Enrollment (2)	Assignment (3)	Enrollment (4)
All students	10.62	9.25	0.69	0.38
Female	10.01	8.59	0.68	0.37
Asian	11.91	10.11	0.63	0.33
Black	8.97	7.83	0.74	0.45
Hispanic	10.21	9.11	0.66	0.39
White	15.46	13.03	0.56	0.22
Bronx	9.46	8.54	0.93	0.64
Brooklyn	10.57	9.49	0.52	0.33
Manhattan	5.10	4.33	0.00	−0.09
Queens	11.30	8.78	1.13	0.57
Staten Island	22.62	22.42	0.34	0.00
High baseline math	8.85	6.85	0.53	0.20
Low baseline math	10.61	9.83	0.57	0.33
Subsidized lunch	10.13	8.93	0.65	0.38
Bottom neighborhood income quartile	8.81	8.18	0.57	0.42
Top neighborhood income quartile	12.02	9.71	0.71	0.25
Special education	10.30	9.10	0.76	0.43
Limited English proficient	11.62	10.58	0.60	0.38
SHSAT test-takers	6.91	5.51	0.55	0.25

Notes: Utilities are in distance units (miles) averaged across students in the mechanism comparison sample in Table 1 using preference estimates in column 3 of Table 7. Utility estimates assume truthful reports. Assignment is the school assigned at the conclusion of the high school assignment process. Enrollment is the school student in which the student enrolls in October following application. If a student enrolls in the assigned school, we use the assigned program to compute the utility of enrollment. If a student enrolls at another school, we use the program-size weighted average of utilities from all programs at that school. 2002–2003 offer process reports the fraction of students with row characteristic first offered school finally assigned in the main round (rounds 1–3), the supplementary round, or the administrative round. Student distance calculated as road distance using ArcGIS. High baseline math students score above the seventy-fifth percentile for seventh grade relative to citywide distribution; low baseline math students score below the twenty-fifth percentile. Subsidized lunch, not available pre-assignment, comes from enrolled students as of the 2004–2005 school year. Neighborhood income is median census block group family income from the 2000 census.

submit the same rank list when optimizing over their top 10 choices compared to their top 15 choices. That is, $\mathbf{r}_i^{*,10} = \mathbf{r}_i^{*,15}$ for about half of the applicants. When we compare $\mathbf{r}_i^{*,15}$ and $\mathbf{r}_i^{*,20}$, nearly three quarters have the same choices. Consistent with this fact, the difference between truthful and optimal reporting from a menu of top 20 choices is $\Delta^{20} = 2.08$ miles, which is only slightly larger than the difference with optimal reporting and a menu of top 15 choices $\Delta^{15} = 1.94$. These findings suggest that we would not find a large difference if we were able to approximate optimal reports from the full choice menu.

The difference between assuming truthful reporting and our approximation of optimal reporting changes the 10.62 mile estimate in Table 9 by about 2.08 miles overall, or about 20 percent. The conclusion that the coordinated mechanism generated significant welfare improvement over the uncoordinated mechanism appears robust to a best-case analysis for the uncoordinated mechanism. The gains are largest for groups that experienced high rates of administrative assignment in the uncoordinated mechanism. Nearly all of the patterns across demographic groups remain the same when considering optimal reports among the top 20: Manhattan residents

Table 10—Welfare Comparison for Alternative Selection Rules

	Difference in expected utility from truthful reporting		
	Top 10	Top 15	Top 20
	(1)	(2)	(3)
Optimal reporting computed from choice menu consisting of			
All students	1.57	1.94	2.08
Female	1.51	1.85	1.98
Asian	2.18	2.68	2.86
Black	1.39	1.73	1.87
Hispanic	1.39	1.71	1.84
White	2.07	2.52	2.66
Bronx	1.43	1.76	1.89
Brooklyn	1.66	2.07	2.21
Manhattan	1.24	1.52	1.63
Queens	1.77	2.19	2.36
Staten Island	1.56	1.82	1.90
High baseline math	1.60	1.91	2.01
Low baseline math	1.53	1.93	2.10
Subsidized lunch	1.49	1.83	1.97
Bottom neighborhood income quartile	1.29	1.59	1.71
Top neighborhood income quartile	1.81	2.22	2.36
Special education	1.56	1.96	2.12
Limited English proficient	1.62	2.01	2.16
SHSAT test-takers	1.60	1.94	2.05

Notes: Utilities are in distance units (miles) averaged across students in the mechanism comparison sample in Table 1 using preference estimates in column 3 of Table 7. Table reports difference in expected utility between coordinated and uncoordinated mechanism and an estimate of optimal reports from 2002–2003 where students best respond to the empirical distribution given the top 10, top 15, and top 20 choices based on the preference distribution estimated from 2003–2004. 11.6 percent of applicants have the same application from truthful reporting as optimal reporting optimizing over a choice menu with the applicant's top 10 choices. 55.8 percent of applicants have the same application from optimal reporting optimizing over a choice menu with the applicant's top 10 choices and the applicant's top 15 choices. 75.5 percent of applicants have the same application from optimal reporting optimizing over a choice menu with the applicant's top 15 choices and the applicant's top 20 choices.

gained the least, while Staten Island residents gained the most; low baseline students gained more than SHSAT test takers. Since optimal reporting provides the best case for the uncoordinated mechanism and there is evidence that our approximation may be close to optimal reporting, the average improvement in student welfare is likely more than 8.6 miles.

The large gains associated with assignment offers need not correspond to differences in matriculation patterns observed earlier, especially if the uncoordinated mechanism's aftermarket is more flexible.[39] Coordinating admissions occurred with greater central control over enrollment, which may mean the more rigid aftermarket in the coordinated mechanism is actually worse for students. Table 3 shows that students enrolled in schools further away on average than where they were assigned in the uncoordinated mechanism, but the opposite pattern is true in the coordinated mechanism. Column 2 of Table 9 also reports the utility associated with the schools at which students enroll in October of the following school year. Compared to assignments, the gains from the coordinated mechanism measured by enrollment are

[39] Similar arguments are often made in the context of auction design in the presence of resale markets. See, e.g., Milgrom (2004).

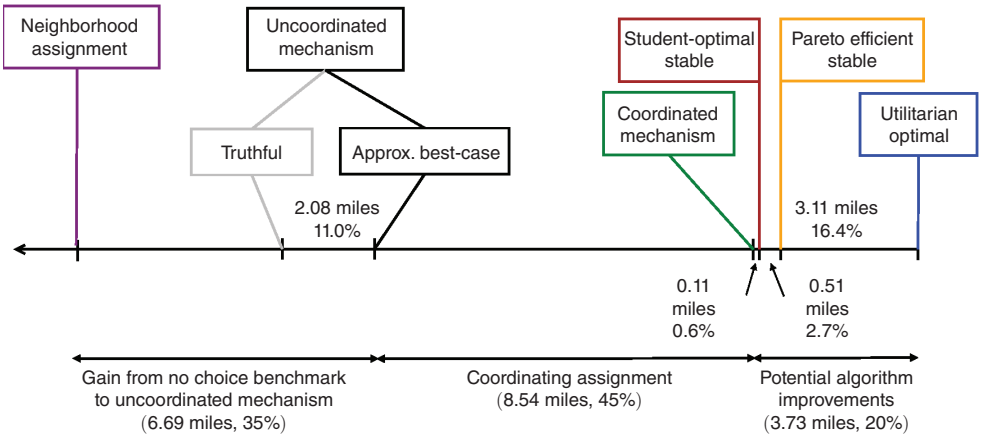

FIGURE 5. COORDINATING ASSIGNMENTS VERSUS ALGORITHM IMPROVEMENTS

somewhat smaller, but they are still large. For instance, the distance-equivalent utility for the average student is 9.25 miles (with standard error 0.56 miles). The change in distance to enrolled school is also lower than the change in distance to assignment. Though a smaller gain from enrollment suggests some of the old mechanism's mismatch was undone in its aftermarket, these facts weigh against the argument that post-market reallocation has undone a large fraction of misallocation. Relatedly, since the exit rate in the coordinated mechanism is lower than in the uncoordinated mechanism, more students preferred accepting their coordinated offer over enrolling in a high school outside of the system. This finding suggests that our welfare estimate may understate the overall effect for all public school eighth graders.

Figure 5 summarizes the comparisons across the alternative mechanisms. The scale corresponds to 18.96 miles from neighborhood to utilitarian assignment. Under our approximation to the best-case for the uncoordinated mechanism, the difference between it and the coordinated mechanism represents 45 percent of the total range. This is more than double the possible range associated with further tweaks to the matching algorithm, which is at most 20 percent. This finding informs a broader debate in the market design literature about the importance of sophisticated market clearing mechanisms. In the context of auctions, Klemperer (2002, p. 170) argued that "most of the extensive auction literature is of second-order importance for *practical* auction design," and that "good auction design is mostly good elementary economics." Consistent with this point of view, for school matching market design, coordinating admissions produces much larger gains than algorithm refinements within the coordinated system.

VII. Comparison for the Administratively Assigned

A key difference between mechanisms is the number of students administratively assigned. Table 4 shows that being administratively assigned is undesirable: students are assigned to schools that differ substantially from the schools they ranked. These facts suggest that students who are administratively assigned loom large in comparisons between mechanisms. In this section, we investigate what our demand estimates imply for this group, and we also examine achievement outcomes.

To begin, we estimate the likelihood that a student is administratively assigned in the uncoordinated mechanism based on demographic characteristics and geographic location. Denote a_i as an indicator if a student is administratively assigned. We fit a probit model:

$$a_i = \Phi(\mathbf{z}_i \rho + \pi_t), \tag{12}$$

where $\mathbf{z}_i$ is a vector of student characteristics (same as the demand model) with coefficient vector ρ, π_t are census tract effects for tracts indexed by t, and Φ is the standard normal CDF. The specification includes census tract effects to account for neighborhoods that may or may not have zoned high schools, which are a guaranteed fallback option for some students.

Equation (13) allows us to construct an index, based on each student's observed characteristics, that quantifies a student's risk of being administratively assigned. With this index, we compare students from the uncoordinated mechanism with students in the coordinated mechanism that are similar on this dimension.

A. *Welfare*

Figure 6 shows that students who were likely to be administratively assigned in the old mechanism realized higher student welfare under the new mechanism. To produce this figure, we compute the expected utility for student i under mechanism m, where $m \in \{\text{coordinated}, \text{uncoordinated}\}$ from equation (3). Then we relate the fitted values $\hat{a}_i$ to this expected utility using a flexible functional form:

$$y_{im} = g_m(\hat{a}_i) + \epsilon_{im}, \tag{13}$$

where g_m is estimated via local linear regressions with normal kernel, with twice Silverman's rule-of-thumb bandwidth.

For a grid of points $a \in \{a_1, \ldots, a_K\}$, Figure 6 reports the difference in g_m for the coordinated and uncoordinated mechanism. This provides a nonparametric estimate of the difference in utility for students with administratively assignment propensity a.[40] Consistent with the descriptive facts in Tables 3 and 4, there is a clear monotonic pattern between administrative assignment propensity and welfare improvement, whether comparing utility differences including either distance or net of distance.

B. *Test Scores and Graduation*

We next examine whether the allocative differences on which we've focused influence educational outcomes. This exercise not only complements our focus on allocative effects, but it is also of independent interest. As far as we know, there is no previous evidence on how changes in school assignment mechanisms translate into differences in downstream educational outcomes. Comparisons between mechanisms are challenging because they are aggregate market-wide shocks, making

[40] These results are similar if the expected utility under the uncoordinated assignment mechanism is computed under the assumption of optimal reporting rather than under truthful reporting.

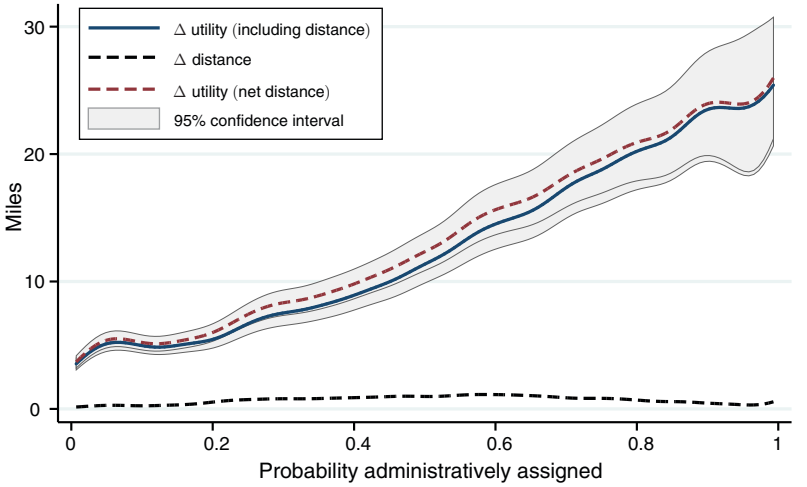

FIGURE 6. CHANGE IN STUDENT WELFARE BY PROPENSITY TO BE ADMINISTRATIVELY ASSIGNED IN THE UNCOORDINATED MECHANISM

Notes: Probability of administrative assignment estimated from probit of administrative assignment indicator on student census tract dummies and all student characteristics in the demand model except for distance in the uncoordinated mechanism. If student lives in a tract where either all or no students are administratively assigned, all students from those tracts are coded as administratively assigned. For a grid of points of administrative assignment propensity, we plot the difference between local linear regression fits of utility including distance and net of distance from coordinated and uncoordinated mechanism computed as in Table 9. Standard errors constructed using 100 draws of parameter values taken from the posterior distribution and re-estimated utility distributions. For each draw of the parameter, we updated ξ and obtained utility draws consistent with observed rank-ordered data using a Gibbs' sampler. Programs that are not ranked by anyone are assigned a draw of ξ from the unconditional distribution.

it difficult to disentangle changes in the mechanism from other contemporaneous changes. Our approach considers groups of students who are more likely to benefit from the new mechanism based their likelihood of being assigned administratively in the old mechanism. Because our findings suggest that effects are largest for those most likely to have been administratively assigned, we anticipate that the downstream consequences are largest for that group.

Figure 7 reports estimates of math and English Regents and graduation based on the probability a student is administratively assigned, constructed in the same way as Figure 6. The top panel shows the difference in Regents math and English achievement is largest for students who were most likely to be administratively assigned in the uncoordinated mechanism, and the difference in achievement mirrors the difference in utility shown in Figure 6. The bottom panel shows that these differences translate into differences in graduation rates, with a nearly 10 percent graduation increase for students who were most likely to be formerly administratively assigned.

VIII. Model Fit and Alternate Behavioral Assumptions

A. *Model Fit*

Since our goal is to make statements about welfare, it is important to examine how well our demand estimates match the data. We first investigate within-sample fit to see what our estimates imply for the aggregate patterns by rank in Table 6.

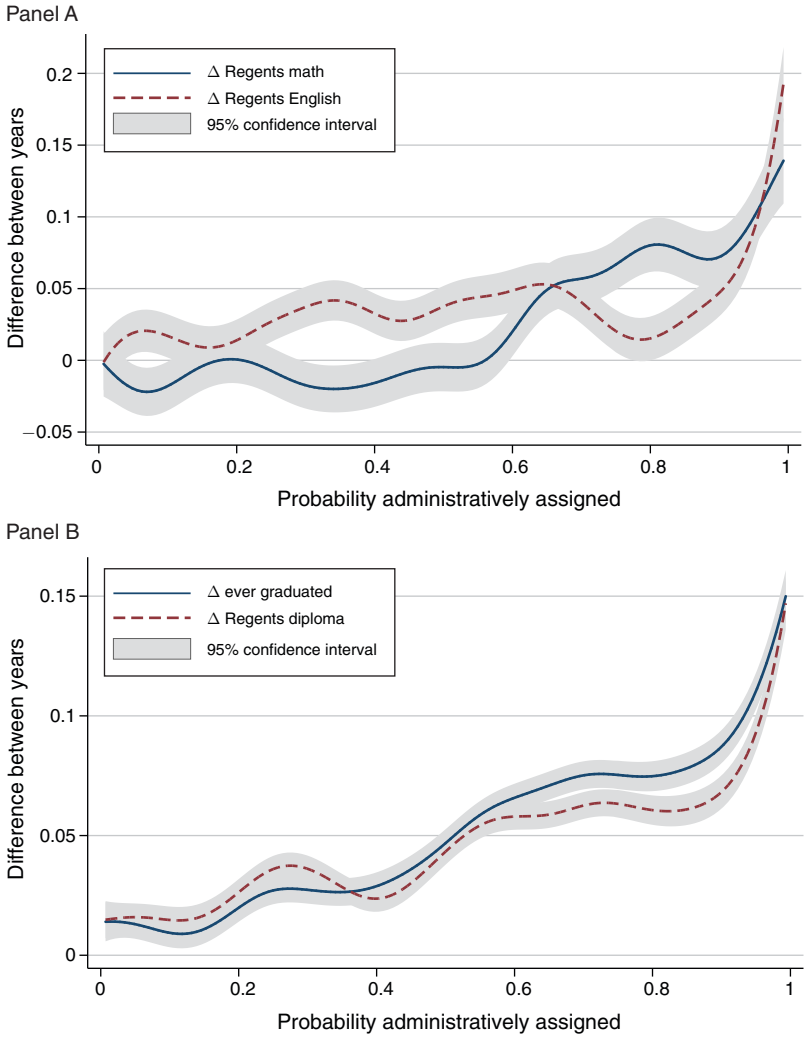

FIGURE 7. CHANGE IN STUDENT ACHIEVEMENT AND GRADUATION BY PROPENSITY TO BE ADMINISTRATIVELY ASSIGNED IN THE UNCOORDINATED MECHANISM

Notes: Probability of administrative assignment estimated from probit of administrative assignment indicator on student census tract dummies and all student characteristics in the demand model except for distance in uncoordinated mechanism. If student lives in a tract where either all or no students are administratively assigned, all students from those tracts are coded as administratively assigned. For a grid of points of administrative assignment propensity, we plot the difference between local linear regression fits of math, English, and graduation outcomes from coordinated and uncoordinated mechanisms. Online Appendix contains details on Regents and graduation outcomes.

Figure A2 reports on measures of fit using the main specification in column 3 of Table 7. We plot the observed versus predicted pattern of three school characteristics—high math achievement, percent subsidized lunch, and percent white—as we go down a student's choice list. The panels include three pairs of lines for the entire sample and for the low and high baseline math applicants. For these three characteristics, our estimates capture the broad pattern of the choices, matching both the level and slope of these characteristics. For instance, the average high math

achievement is 10.0, and the range from the top choice to the twelfth choice is 16.7 to 10.4. Our estimates imply that for first choices, a school's fraction high math achievement is 18.4, which drops to 11.8 for the twelfth choice. Our model also captures high baseline math applicants' greater sensitivity of high baseline math applicants to school math performance is also captured by our model. Furthermore, first choices for percent white are 19.1, while we predict them to be 19.8 on average, and the average percent white across New York's schools is only 10.8. Relative to the average attributes of schools, the model fit is much closer to the actual ranked distribution.

In the last panel of Figure A2, we report the fit for distance. Here, we find that while the increase in distance observed for lower-ranked choices mirrors that predicted by our model, there is a greater divergence in the level of distance. This pattern appears in all of the models we have estimated with random coefficients. It is worth noting that the difference in levels between our model and the data is small compared to the difference between the average distance to a high school in New York (12.7 miles from home) and the closest school (less than a mile from home).

Berry, Levinsohn, and Pakes (2004) emphasize the importance of random coefficients models in the context of rank data for automobiles. In particular, they emphasize that when examining the within-consumer relationship between the attributes of alternatives ranked first and second, models without random coefficients do a poor job. This concern may be particularly important in our context. For instance, a high correlation between the first and second ranked schools' size may indicate taste for large schools. In Table A1, we report on the correlation between the first and second choice, the first and third choice, and the second and third choice. Consistent with earlier work, we see that the observed correlation between choices is much closer in our preferred specification than in the simpler model within sample. When we examine a more demanding out-of-sample test, which compares the 2003–2004 preference estimates to examine the correlation pattern of choice made in 2004–2005, we also see that the correlation pattern in our main specification is closer to the observed pattern than that from a demand model without student interactions.

B. *Behavioral Assumptions on Ranking*

Motivated by the detailed information in school brochures and the incentive properties of the deferred acceptance algorithm, the preference estimates that we reported come from models that assume students are well-informed and truthful in the coordinated mechanism. We examine some potential objections to this assumption in this subsection.

Stability of Preferences.—In an influential field experiment, Hastings and Weinstein (2008) sent Charlotte-area parents clear information about schools. The percentage of applicants who requested to change schools increased by about 6.6 percent (with a standard error of 3.6 percent). The authors interpret this finding as showing that how school information is framed can influence school choices. In NYC, the information available to participants about school characteristics was similar in the uncoordinated mechanism and the first few years of the coordinated

mechanism. Consistent with this fact, Figure A4 shows that school market shares are fairly stable in the first and second year of the new mechanism.

Enrollment patterns also support the idea that choices are deliberate and contain welfare-relevant information. Online Appendix Table B4 reports assignment and enrollment decisions for students who are assigned in the main round. The table shows that 92.7 percent of students enroll in their assigned choice, and this number varies from 88.4 percent to 94.5 percent, depending on which choice a student receives. Interestingly, take-up is higher for students who receive lower-ranked choices, while the fraction of students who exit is highest among students who obtain one of their top three choices. This suggests that either families are unconcerned with differences among later choices and simply enroll where they obtain an offer or that families deliberately investigate later choices and are therefore willing to enroll in lower-ranked schools. If families are more uncertain about lower-ranked choices, then using all submitted ranks may provide a misleading account of student preferences. To examine how sensitive our conclusions are to this assumption, we fit a demand model that considers only students' top three choices in column 4 of Table 7.

Assumptions about Ranking Behavior.—A second concern with treating submitted rankings as truthful is that parents rank schools using heuristics carried over from the previous system. Despite the theoretical motivation and the DOE's advice, parents might still deviate from truth-telling because of misinformation. Online Appendix Table B3 shows that students are more likely to be assigned their last choice than their penultimate choice. This pattern may be caused by strategic behavior if students apply to schools that they like, and, as a safety option, rank last a school in which they have a higher admissions chance. For instance, Calsamiglia, Haeringer, and Klijn (2010) present laboratory evidence that a constraint on rank-order lists encourages students to rank safer options. However, it may also be fully consistent with truth-telling. For example, students usually obtain borough priority or zone priority for schools in their neighborhoods. Ranking these schools improves their likelihood of being assigned to them in case they are rejected by their higher choices. If students consider applying to higher-achieving schools further away from their neighborhoods, they may as well stop ranking schools below their neighborhood schools once such considerations no longer justify the cost of their commute. Alternatively, search costs may induce parents to stop their search for schools before they identify 12 schools for their children and rank their neighborhood school as last choice. This preference pattern would produce the observed assignment pattern. To examine how sensitive our conclusions are to this assumption, we fit demand models that drop the last choice of each student.

Another issue with assuming truthful preferences is that students can rank at most 12 programs on school applications. When a student is interested in more than 12 schools, she has to carefully reduce the choice set down to 12 schools. If a student is only interested in 11 or fewer schools, this constraint in principle should not influence ranking behavior (Abdulkadiroğlu, Pathak, and Roth 2009; Haeringer and Klijn 2009). It is a weakly dominant strategy to add an acceptable school to a rank-order list as long as there is room for additional schools on the application

form. However, 20.3 percent of students in our demand sample rank 12 schools. Some of these students may drop highly sought-after schools from the top of their choice lists because of this constraint. To examine how sensitive our conclusions are to this assumption, we fit a demand model that drops students who have ranked all 12 choices.

In Table A2, we report on our evaluation of mechanism design choices under these alternative specifications: (i) using choices among eligible programs; (ii) using only the top 3 choices; (iii) excluding applicants who have ranked all 12 programs; and (iv) dropping applicants' last choice. (Table A3 reports the corresponding preference estimates.) For all demand models, the coordinated mechanism in column 2 is more than three quarters of the way from the neighborhood assignment to the utilitarian assignment. It therefore appears that our conclusions on the value of choice relative to changes within the coordinated mechanism are robust to these alternative ways of using the submitted rank-order lists in the coordinated mechanism.

Panel B of Table A2 reports on how the comparison between mechanisms varies with our demand specification using participants' full ranking information. Given that Staten Island has the highest fraction of students who are assigned administratively and experiences the largest welfare gain, we re-estimate preferences excluding any applicant from Staten Island. Our conclusions are unaltered by this modification. The table does show, however, that preference heterogeneity generates a larger role for school choice compared to neighborhood assignment. This phenomenon can be seen by comparing the estimates from our main specification to those from specifications without student interactions and without random coefficients. The neighborhood assignment is more appealing according to those two demand models, since they are only 15.3 and 16.0 miles away from the utilitarian assignment, compared to 19.0 miles from the main specification.

IX. Conclusion

The reform of NYC's high school assignment system provides a unique opportunity to study the effects of centralizing and coordinating school admissions with detailed data on preferences, assignments, and enrollments. We find that the new coordinated mechanism is an improvement relative to the old uncoordinated mechanism in a of variety dimensions. More than a third of students were assigned through an ad-hoc administrative process in the uncoordinated mechanism after multiple offers with few choices and few rounds of market-clearing left a large number of students without offers after the main round. Students placed in the administrative round were assigned to schools with considerably worse characteristics than the schools they ranked. The new mechanism relieved this congestion and assigned more students to schools where they applied.

The coordinated mechanism assigns students 0.69 miles further from home compared to the uncoordinated mechanism. However, the benefit of being assigned through the coordinated mechanism is significantly larger than the cost of additional travel. The gains are positive on average for students from all boroughs, demographic groups, and baseline achievement categories. Welfare improvements are present whether utility is measured based on assignments made at the end of the high school match or subsequent school enrollment. The largest gains are for

students who were more likely to be processed in the Administrative round of the uncoordinated mechanism. These conclusions are robust to alternative behavioral assumptions on preferences submitted in both the uncoordinated and coordinated mechanism. Cross-sectional differences in Regents test performance and graduation also coincide with the fact that students who were most likely to be assigned administratively in the old mechanism experienced the largest gains.

These gains are measured by a rich specification of student demand that implies significant estimated heterogeneity in willingness to travel for school. Understanding preference heterogeneity is important for measuring the allocative effects of choice. Our estimates reveal that the benefits of coordinated choice with deferred acceptance are much larger than than those associated with modifications to the assignment algorithm within the coordinated mechanism. This does not imply that the mechanism design is not important, however, because the gap in average student welfare between the uncoordinated and coordinated mechanism is large.

The increase in student welfare due to the new mechanism illustrates that there are considerable frictions to exercising choice in poorly designed assignment systems. The 2003 change in NYC took place in an environment where participants already had some familiarity with choice since both the uncoordinated and coordinated system had a common application. In other cities, the school choice market is even less well organized, without readily available information on admissions processes and application timelines. For instance, admissions in Boston's growing charter sector are uncoordinated, and the schools have only recently adopted a standardized application timeline. Recently, there have been efforts to unify enrollment across school sectors (Vaznis 2013; Fox 2015). The relative value of policies such as common timelines, common applications, single versus multiple offers, sophisticated matching algorithms, and good information and decision aides is an interesting avenue for future research.

Appendix

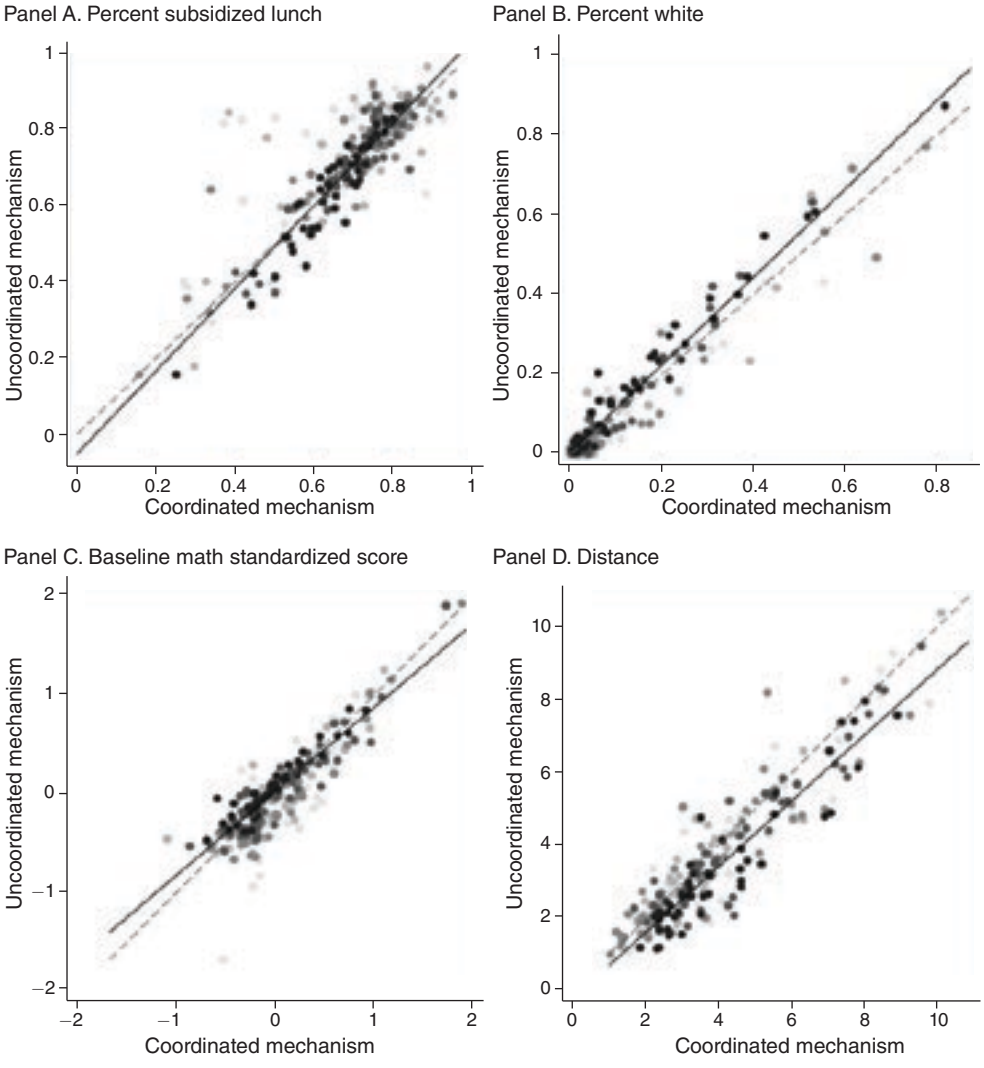

Figure A1. Comparison of Characteristics of Enrolled Students at Each School between Uncoordinated and Coordinated Mechanism

Notes: This figure reports school characteristics measured by the attributes of students enrolled at each school across mechanisms. The dotted line is the 45 degree line; the solid line is the least squares line fit.

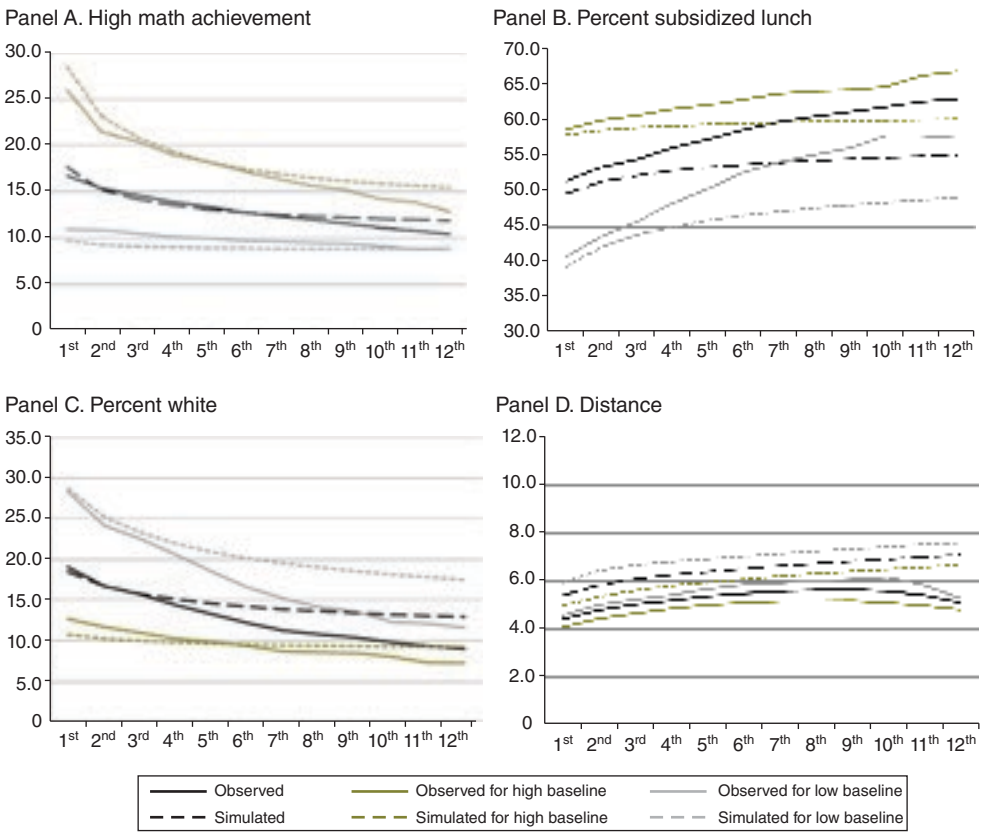

FIGURE A2. MODEL FIT

Notes: This figure reports the observed and estimated school characteristics for different student ranked choices. The estimates are from the main specification in column 3 of Table 7.

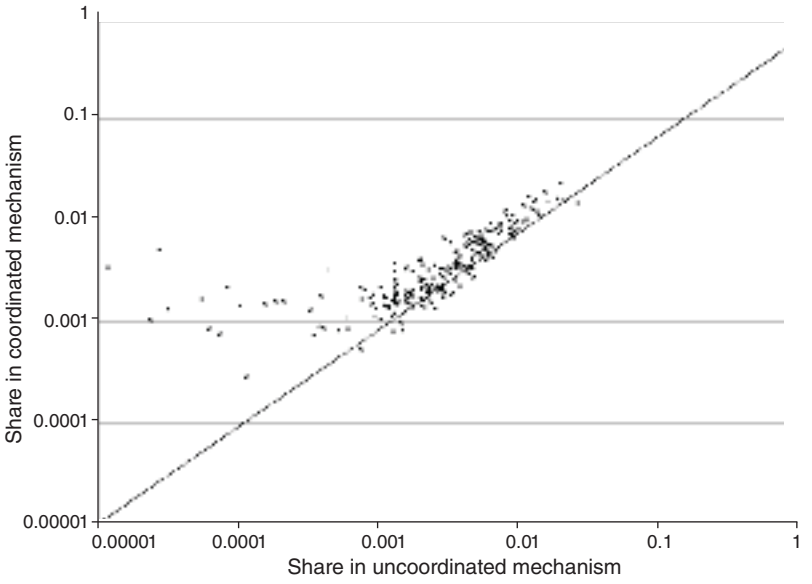

FIGURE A3. COMPARISON OF SCHOOL MARKET SHARES BETWEEN 2002–2003 UNCOORDINATED MECHANISM AND 2003–2004 COORDINATED MECHANISM

Notes: This figure plots school market shares defined as the count of applicants ranking a program at a given school divided by the total number of choices expressed for schools to which students can apply to in 2002–2003 and 2003–2004. Market shares are normalized within this set to sum to one.

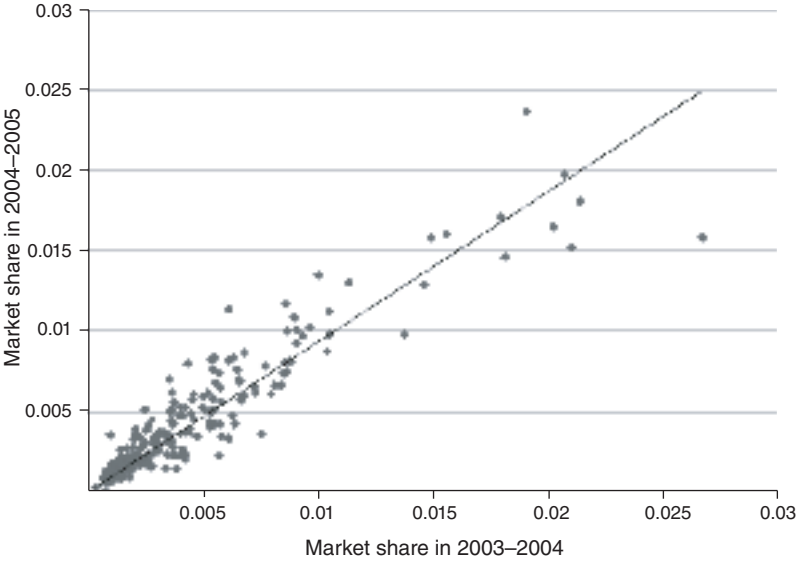

FIGURE A4. COMPARISON OF SCHOOL MARKET SHARES BETWEEN 2003–2004 COORDINATED MECHANISM AND 2004–2005 COORDINATED MECHANISM

Notes: This figure plots school market shares defined as the count of applicants ranking a program at a given school divided by the total number of choices expressed for schools to which students can apply in 2003–2004 and 2004–2005. Market shares are normalized within this set to sum to one. The slope of the line fit is 0.93 and the R^2 is 0.86.

Table A1—Model Fit of Correlation between Choices across Demand Specifications

School characteristic	Correlation between Choice (1)	Choice (2)	Coordinated mechanism (2003–2004) Observed (3)	Specification: Main specification (4)	Specification: No student interactions (5)	Coordinated mechanism (2004–2005) Observed (6)	Specification: Main specification (7)	Specification: No student interactions (8)
Distance	1	2	0.47	0.16	0.09	0.50	0.17	0.07
	1	3	0.39	0.15	0.09	0.41	0.16	0.07
	2	3	0.44	0.17	0.12	0.47	0.16	0.07
High math performance	1	2	0.35	0.39	0.03	0.39	0.42	0.03
	1	3	0.32	0.36	0.03	0.35	0.38	0.03
	2	3	0.34	0.33	0.03	0.39	0.36	0.03
Percent free lunch	1	2	0.61	0.49	0.30	0.66	0.49	0.34
	1	3	0.54	0.46	0.28	0.60	0.47	0.33
	2	3	0.55	0.43	0.26	0.62	0.45	0.31
Percent white	1	2	0.55	0.55	0.29	0.60	0.56	0.37
	1	3	0.47	0.52	0.26	0.54	0.54	0.36
	2	3	0.48	0.48	0.24	0.56	0.51	0.35
Size of ninth grade	1	2	0.29	0.60	0.07	0.34	0.59	0.09
	1	3	0.21	0.58	0.06	0.24	0.57	0.09
	2	3	0.27	0.56	0.06	0.33	0.56	0.08

Notes: Table reports the observed correlation between the school characteristic of the choice in column 1 with the choice in column 2 for the main specification (shown in column 3 of Table 7) and the specification with no student interactions (shown in column 1 of Table 7).

Table A2—Welfare Comparisons for Alternative Demand Specifications

	Neighborhood assignment (1)	School choice: Coordinated mechanism (2)	Student optimal matching (3)	Ordinal Pareto efficient matching (4)
Panel A. Alternative behavioral assumptions				
Choice among eligible programs	−25.22	−5.48	−5.33	−4.68
Top three choices	−19.47	−3.82	−3.70	−3.19
Excluding full lists	−18.81	−3.71	−3.59	−3.09
Excluding last choice	−18.70	−3.69	−3.58	−3.07
Panel B. Alternative samples and demand model interactions				
Main specification	−18.96	−3.73	−3.62	−3.11
Excluding Staten Island	−19.68	−3.93	−3.82	−3.28
No student interactions	−15.32	−3.29	−3.19	−2.76
No random coefficients	−15.95	−3.22	−3.12	−2.67
Number of students reassignments relative to column 2			2,344	10,881

Notes: Utility from alternative assignments relative to utilitarian assignment computed using actual preferences and ignoring all school-side constraints except capacity. See notes to Table 8 for details on mechanism calculations. All mechanism counterfactuals used these estimates for all applicants in the mechanism comparison sample. Choice among eligible programs restricts each applicant's choice set to include eligible programs. If an applicant ranked an ineligible program, that program is included in the choice set. Top three choices refers to estimates that only use the top three choices of applicants. Excluding full lists refers to estimates that only use rankings of students who rank fewer than 12 choices. Excluding last choice refers to estimates that use all rankings except the last one. No student interactions and no random coefficient refers to the specification in column 1 and 2 of Table 7, respectively.

TABLE A3—POSTERIOR MEANS OF PREFERENCE ESTIMATES FOR ALTERNATIVE DEMAND SPECIFICATIONS

			Models with random coefficients					
	No student interactions	Without random coefficients	All choices	Top three choices	All except last choice	Students that ranked less than 12 schools	Drop Staten Island students	Choice among eligible programs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High math achievement								
Main effect	0.016	0.027	−0.029	−0.010	−0.026	−0.024	−0.032	−0.058
	(0.016)	(0.014)	(0.018)	(0.018)	(0.017)	(0.017)	(0.045)	(0.039)
Baseline math		0.031	0.039	0.051	0.039	0.038	0.040	0.050
		(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Baseline English		0.025	0.039	0.047	0.039	0.041	0.040	0.049
		(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
Subsidized lunch		−0.008	−0.016	−0.015	−0.016	−0.016	−0.018	−0.018
		(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Neighborhood income (in 1,000s)		0.004	0.012	0.012	0.011	0.011	0.013	0.012
		(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)
Limited English proficiency		−0.002	0.000	−0.009	0.000	0.000	−0.001	−0.003
		(0.002)	(0.003)	(0.005)	(0.003)	(0.004)	(0.003)	(0.004)
Special education		−0.007	−0.006	0.003	−0.006	0.002	−0.005	−0.008
		(0.002)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.005)
Percent subsidized lunch								
Main effect	−0.085	−0.057	−0.069	−0.044	−0.064	−0.067	−0.103	−0.113
	(0.007)	(0.004)	(0.009)	(0.011)	(0.006)	(0.007)	(0.062)	(0.058)
Asian		−0.009	−0.012	−0.022	−0.012	−0.012	−0.016	−0.021
		(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Black		0.005	0.009	−0.004	0.008	0.007	0.005	−0.012
		(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Hispanic		0.031	0.043	0.049	0.043	0.043	0.041	0.043
		(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Subsidized lunch		0.007	0.011	0.013	0.011	0.011	0.011	0.012
		(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Neighborhood income (in 1,000s)		−0.004	−0.008	−0.011	−0.008	−0.009	−0.008	−0.010
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Size of ninth grade (in 100s)								
Main effect	−0.164	−0.092	−0.113	−0.151	−0.077	−0.039	−0.188	−0.153
	(0.036)	(0.032)	(0.048)	(0.054)	(0.043)	(0.043)	(0.192)	(0.178)
Baseline math		−0.017	−0.026	−0.036	−0.026	−0.022	−0.027	−0.031
		(0.002)	(0.008)	(0.010)	(0.008)	(0.009)	(0.008)	(0.008)
Baseline English		−0.038	−0.066	−0.108	−0.066	−0.075	−0.067	−0.087
		(0.002)	(0.008)	(0.011)	(0.008)	(0.010)	(0.008)	(0.009)
Subsidized lunch		0.016	0.038	0.039	0.036	0.050	0.032	0.062
		(0.003)	(0.012)	(0.015)	(0.011)	(0.014)	(0.013)	(0.013)
Neighborhood income (in 1,000s)		−0.011	−0.012	−0.005	−0.013	−0.019	−0.010	−0.017
		(0.001)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Special education		−0.031	−0.048	−0.087	−0.048	−0.044	−0.049	−0.058
		(0.005)	(0.022)	(0.028)	(0.021)	(0.027)	(0.022)	(0.025)
Percent white								
Main effect	−0.002	0.070	0.062	0.104	0.069	0.076	0.062	0.093
	(0.014)	(0.012)	(0.016)	(0.016)	(0.014)	(0.014)	(0.059)	(0.062)
Asian		−0.054	−0.075	−0.094	−0.075	−0.079	−0.082	−0.100
		(0.002)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)
Black		−0.084	−0.124	−0.157	−0.124	−0.132	−0.139	−0.189
		(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
Hispanic		−0.047	−0.084	−0.095	−0.084	−0.089	−0.099	−0.119
		(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)

(*Continued*)

TABLE A3—POSTERIOR MEANS OF PREFERENCE ESTIMATES FOR ALTERNATIVE DEMAND SPECIFICATIONS (*Continued*)

			Models with random coefficients					
	No student interactions	Without random coefficients	All choices	Top three choices	All except last choice	Students that ranked less than 12 schools	Drop Staten Island students	Choice among eligible programs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Spanish language program								
Limited English proficient		14.042	15.437	16.004	15.415	16.213	15.769	20.022
		(0.282)	(0.307)	(0.543)	(0.312)	(0.378)	(0.328)	(0.412)
Limited English proficient x		−9.452	−10.502	−10.184	−10.509	−10.889	−10.765	−14.152
		(0.407)	(0.445)	(0.672)	(0.449)	(0.503)	(0.472)	(0.577)
Asian language program								
Limited English proficient		10.623	11.814	12.466	11.811	11.994	12.066	15.221
		(0.108)	(0.120)	(0.219)	(0.119)	(0.152)	(0.125)	(0.158)
Limited English proficient x		−6.373	−7.091	−8.299	−7.086	−6.870	−7.259	−8.960
		(0.342)	(0.378)	(0.657)	(0.385)	(0.417)	(0.390)	(0.490)
Other language program		6.538	7.448	8.538	7.450	7.656	7.597	9.873
Limited English proficient		(0.183)	(0.204)	(0.320)	(0.206)	(0.249)	(0.210)	(0.269)
Standard deviation of ε	7.226	7.385	7.858	8.032	7.857	7.869	8.025	10.059
	(0.010)	(0.011)	(0.013)	(0.022)	(0.013)	(0.015)	(0.015)	(0.022)
Standard deviation of ξ	3.519	2.954	3.676	3.936	3.609	3.554	4.155	5.151
	(0.121)	(0.100)	(0.129)	(0.131)	(0.121)	(0.119)	(0.817)	(0.650)
Random coefficients (covariances)								
Size of ninth grade (in 100s)			1.584	2.351	1.584	1.967	1.651	1.837
			(0.009)	(0.016)	(0.009)	(0.013)	(0.010)	(0.012)
Percent white			−0.006	−0.010	−0.006	−0.005	−0.007	−0.009
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Percent subsidized lunch			−0.002	−0.004	−0.002	−0.003	−0.003	−0.002
			(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
High math achievement			−0.011	−0.015	−0.011	−0.009	−0.012	−0.015
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Percent white			0.008	0.010	0.008	0.008	0.009	0.013
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Percent subsidized lunch			−0.001	0.000	−0.001	−0.001	−0.001	−0.002
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
High math achievement			0.005	0.006	0.005	0.004	0.005	0.007
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Percent subsidized lunch			0.002	0.005	0.002	0.003	0.003	0.003
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
High math achievement			0.000	0.000	0.000	0.000	0.000	−0.001
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
High math achievement			0.016	0.021	0.016	0.016	0.017	0.022
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Program type dummies			X	X	X	X	X	X
Program specialty dummies			X	X	X	X	X	X
Number of students	69,907	69,907	69,907	69,907	69,907	55,695	65,518	69,907
Number of ranks	542,666	542,666	542,666	197,245	55,323	372,122	530,861	542,666

Notes: Estimates of demand system with submitted ranks over 497 program choices in 235 schools. All columns use all students. See Table 7 and Table A2 for additional details.

REFERENCES

Abdulkadiroğlu, Atila, Nikhil Agarwal, and Parag A. Pathak. 2017. "The Welfare Effects of Coordinated Assignment: Evidence from the New York City High School Match: Dataset." *American Economic Review.* https://doi.org/10.1257/aer.20151425.

Abdulkadiroğlu, Atila, Joshua D. Angrist, Susan M. Dynarski, Thomas J. Kane, and Parag A. Pathak. 2011. "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots." *Quarterly Journal of Economics* 126 (2): 699–748.

Abdulkadiroğlu, Atila, Joshua Angrist, and Parag Pathak. 2014. "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools." *Econometrica* 82 (1): 137–96.

Abdulkadiroğlu, Atila, Yeon-Koo Che, Parag Pathak, Alvin Roth, and Olivier Tercieux. 2017. "Minimizing Justified Envy in School Choice: The Design of New Orleans' OneApp." National Bureau of Economic Research Working Paper 23265.

Abdulkadiroğlu, Atila, Yeon-Koo Che, and Yosuke Yasuda. 2015. "Expanding 'Choice' in School Choice." *American Economic Journal: Microeconomics* 7 (1): 1–42.

Abdulkadiroğlu, Atila, Weiwei Hu, and Parag A. Pathak. 2013. "Small High Schools and Student Achievement: Lottery-Based Evidence from New York City." National Bureau of Economic Research Working Paper 19576.

Abdulkadiroğlu, Atila, Parag A. Pathak, and Alvin E. Roth. 2005. "The New York City High School Match." *American Economic Review* 95 (2): 364–67.

Abdulkadiroğlu, Atila, Parag A. Pathak, and Alvin E. Roth. 2009. "Strategy-Proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match." *American Economic Review* 99 (5): 1954–78.

Abdulkadiroğlu, Atila, Parag Pathak, Alvin E. Roth, and Tayfun Sönmez. 2006. "Changing the Boston School Choice Mechanism." National Bureau of Economic Research Working Paper 11965.

Abdulkadiroğlu, Atila, and Tayfun Sönmez. 2003. "School Choice: A Mechanism Design Approach." *American Economic Review* 93 (3): 729–47.

Agarwal, Nikhil, and Paulo Somaini. 2014. "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism." National Bureau of Economic Research Working Paper 20775.

Ajayi, Kehinde F. 2013. "School Choice and Educational Mobility: Lessons from Secondary School Applications in Ghana." http://people.bu.edu/kajayi/Ajayi_EducationalMobility.pdf (accessed October 2017).

Arvidsson, Toi Sin, Norm Fruchter, and Christina Mokhtar. 2013. *Over the Counter, Under the Radar: Inequitably Distributing New York City's Late-Enrolling High School Students*. Providence, RI: Annenberg Institute for School Reform at Brown University. http://annenberginstitute.org/publications/over-counter-under-radar-inequitably-distributing-new-york-city's-late-enrolling-high-s (accessed October 2013).

Balinski, Michel, and Tayfun Sönmez. 1999. "A Tale of Two Mechanisms: Student Placement." *Journal of Economic Theory* 84 (1): 73–94.

Berry, Steven, James Levinsohn, and Ariel Pakes. 2004. "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market." *Journal of Political Economy* 112 (1): 68–105.

Bogomolnaia, Anna, and Herve Moulin. 2001. "A New Solution to the Random Assignment Problem." *Journal of Economic Theory* 100 (2): 295–328.

Briesch, Richard A., Pradeep K. Chintagunta, and Rosa L. Matzkin. 2002. "Semiparametric Estimation of Brand Choice Behavior." *Journal of the American Statistical Association* 97 (460): 973–82.

Budish, Eric, and Estelle Cantillon. 2012. "The Multi-Unit Assignment Problem: Theory and Evidence from Course Allocation at Harvard." *American Economic Review* 102 (5): 2237–71.

Budish, Eric, Yeon-Koo Che, Fuhito Kojima, and Paul Milgrom. 2013. "Designing Random Allocation Mechanisms: Theory and Applications." *American Economic Review* 103 (2): 585–623.

Burgess, Simon, Ellen Greaves, Anna Vignoles, and Deborah Wilson. 2015. "What Parents Want: School Preferences and School Choice." *Economic Journal* 125 (587): 1262–89.

Calsamiglia, Caterina, Chao Fu, and Maia Güell. 2014. "Structural Estimation of a Model of School Choices: the Boston Mechanism vs. Its Alternatives." http://www.maiaguell.com/research/Barcelona.pdf (accessed October 2017).

Calsamiglia, Caterina, Guillaume Haeringer, and Flip Klijn. 2010. "Constrained School Choice: An Experimental Study." *American Economic Review* 100 (4): 1860–74.

Chen, Yan, and Tayfun Sönmez. 2006. "School Choice: An Experimental Study." *Journal of Economic Theory* 127 (1): 202–31.

Deming, David J., Justine S. Hastings, Thomas J. Kane, and Douglas O. Staiger. 2014. "School Choice, School Quality, and Postsecondary Attainment." *American Economic Review* 104 (3): 991–1013.

Department of Education (DOE). 2002. *Directory of the New York City Public High Schools: 2002–2003*. New York: New York City Department of Education.

Department of Education (DOE). 2003. *Directory of the New York City Public High Schools: 2003–2004*. New York: New York City Department of Education.

Dubins, L. E., and D. A. Freedman. 1981. "Machiavelli and the Gale-Shapley Algorithm." *American Mathematical Monthly* 88 (7): 485–94.

Economic Sciences Prize Committee of the Royal Swedish Academy of Sciences. 2012. "Scientific Background on the Svergies Riksbank Prize in Economic Sciences in Memory of Alfred Nobel: Stable Allocations and the Practice of Market Design."

Ekmekci, Mehmet, and M. Bumin Yenmez. 2016. "Integrating Schools for Centralized Admissions." http://dx.doi.org/10.2139/ssrn.2474700 (accessed October 19, 2017).

Epple, Dennis, and Richard E. Romano. 1998. "Competition between Private and Public Schools, Vouchers, and Peer-Group Effects." *American Economic Review* 88 (1): 33–62.

Erdil, Aytek, and Haluk Ergin. 2008. "What's the Matter with Tie-Breaking? Improving Efficiency in School Choice." *American Economic Review* 98 (3): 669–89.

Fack, Gabrielle, Julien Grenet, and Yinghua He. 2015. "Beyond Truth-Telling: Preference Estimation with Centralized School Choice." CEPR Discussion Paper 10907.

Fox, J. C. 2015. "Plan Would Simplify Boston Public School Enrollments." *Boston Globe*, September 17.

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

Gendar, Alison. 2000. "Simpler HS Entry Eyed Ed Board Aims to Trim Selection Process, Improve Schools." *New York Daily News*, October 21. http://www.nydailynews.com/archives/news/simpler-hs-entry-eyed-ed-board-aims-trim-selection-process-improve-schools-article-1.883046.

Haeringer, Guillaume, and Flip Klijn. 2009. "Constrained School Choice." *Journal of Economic Theory* 144 (5): 1921–47.

Hastings, Justine S., Thomas J. Kane, and Douglas O. Staiger. 2009. "Heterogeneous Preferences and the Efficacy of Public School Choice." http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=F2D51D8E574C2BBCB17C9CC845EB4E72?doi=10.1.1.700.3796&rep=rep1&type=pdf (accessed October 19, 2017).

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

He, Yinghua. 2012. "Gaming the Boston Mechanism in Beijing." TSE Working Paper 12-345.

He, Yinghua, Antonio Miralles, Marek Pycia, and Jianye Yan. Forthcoming. "A Pseudo-Market Approach to Allocation with Priorities." *American Economic Journal: Microeconomics*.

Hemphill, Clara. 2007. *New York City's Best Public High Schools: A Parent's Guide*, 3rd ed. New York: Teachers College Press.

Hemphill, Clara, and Kim Nauer. 2009. *The New Marketplace: How Small-School Reforms and School Choice Have Reshaped New York City's High Schools*. New York: Center for New York City Affairs, Milano the New School for Management and Urban Policy.

Herszenhorn, David M. 2004. "Council Members See Flaws in School-Admissions Plan." *New York Times*, November 19. http://www.nytimes.com/2004/11/19/education/council-members-see-flaws-in-schooladmissions-plan.html.

Hoxby, Caroline M. 2003. "School Choice and School Productivity: Could School Choice Be a Tide that Lifts All Boats." In *The Economics of School Choice*, edited by Caroline M. Hoxby, 287–341. Chicago: University of Chicago Press.

Hwang, Sam Il Myoung. 2014. "A Robust Redesign of High School Match." http://econ.sites.olt.ubc.ca/files/2015/01/pdf_HwangJMP-Nov2014.pdf (accessed October 19, 2017).

Ichimura, Hidehiko, and T. Scott Thompson. 1998. "Maximum Likelihood Estimation of a Binary Choice Model with Random Coefficients of Unknown Distribution." *Journal of Econometrics* 86 (2): 269–95.

Jacob, Brian A., and Lars Lefgren. 2007. "What Do Parents Value in Education? An Empirical Investigation of Parents' Revealed Preferences for Teachers." *Quarterly Journal of Economics* 122 (4): 1603–37.

Kapor, Adam, Christopher Neilson, and Seth Zimmerman. 2017. "Heterogeneous Beliefs and School Choice." https://www.dropbox.com/s/u6ztqyafl64g858/KNZ_NHPS_Current.pdf?dl=0 (accessed October 2017).

Kerr, P. 2003. "High School Admissions." *New York Times*, Editorial Desk, p. 18. November 3.

Kesten, Onur. 2010. "School Choice with Consent." *Quarterly Journal of Economics* 125 (3): 1297–1348.

Kesten, Onur, and Morimitsu Kurino. 2012. "On the (im)possibility of Improving Upon the Student-Proposing Deferred Acceptance Mechanism." WZB Discussion Paper SP II 2012-202.

Klemperer, Paul. 2002. "What Really Matters in Auction Design." *Journal of Economic Perspectives* 16 (1): 169–89.

Kojima, Fuhito, and Parag A. Pathak. 2009. "Incentives and Stability in Large Two-Sided Matching Markets." *American Economic Review* 99 (3): 608–27.

Lewbel, Arthur. 2000. "Semiparametric Qualitative Response Model Estimation with Unknown Heteroscedasticity or Instrumental Variables." *Journal of Econometrics* 97 (1): 145–77.

Milgrom, Paul. 2004. *Putting Auction Theory to Work*. Cambridge, UK: Cambridge University Press.

Narita, Yusuke. 2016. "Match or Mismatch: Learning and Inertia in School Choice." http://economics.yale.edu/sites/default/files/narita_update_mismatch.pd_.pdf (accessed October 2017).

Neilson, Christopher. 2013. "Targeted Vouchers, Competition among Schools, and the Academic Achievement of Poor Students." https://economics.sas.upenn.edu/system/files/event_papers/Neilson_2013_JMP_current.pdf (accessed October 2017).

Niederle, Muriel, and Alvin E. Roth. 2003. "Unraveling Reduces Mobility in a Labor Market: Gastroenterology with and without a Centralized Match." *Journal of Political Economy* 111 (6): 1342–52.

Pathak, Parag A., and Peng Shi. 2014. "Demand Modeling, Forecasting, and Counterfactuals, Part I." National Bureau of Economic Research Working Paper 19859.

Pathak, Parag A., and Tayfun Sönmez. 2008. "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism." *American Economic Review* 98 (4): 1636–52.

Rossi, Peter E., Robert E. McCulloch, and Greg M. Allenby. 1996. "The Value of Purchase History Data in Target Marketing." *Marketing Science* 15 (4): 321–40.

Roth, Alvin E. 1982. "The Economics of Matching: Stability and Incentives." *Mathematics of Operations Research* 7 (4): 617–28.

Roth, Alvin E. 2015. *Who Gets What and Why.* Boston: Houghton Mifflin Harcourt.

Roth, Alvin E., and Xiaolin Xing. 1997. "Turnaround Time and Bottlenecks in Market Clearing: Decentralized Matching in the Market for Clinical Psychologists." *Journal of Political Economy* 105 (2): 284–329.

Rothstein, Jesse M. 2006. "Good Principals or Good Peers? Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition among Jurisdictions." *American Economic Review* 96 (4): 1333–50.

Shapley, Lloyd, and Herbert Scarf. 1974. "On Cores and Indivisibility." *Journal of Mathematical Economics* 1 (1): 23–37.

Sönmez, Tayfun. 1997. "Manipulation via Capacities in Two-Sided Matching Markets." *Journal of Economic Theory* 77 (1): 197–204.

Sönmez, Tayfun. 2013. "Bidding for Army Career Specialties: Improving the ROTC Branching Mechanism." *Journal of Political Economy* 121 (1): 186–219.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation*. Cambridge, UK: Cambridge University Press.

Tullis, Tracy. 2014. "How Game Theory Helped Improve New York City's High School Application Process." *New York Times*, December 5. https://www.nytimes.com/2014/12/07/nyregion/how-game-theory-helped-improve-new-york-city-high-school-application-process.html.

Urquiola, Miguel. 2005. "Does School Choice Lead to Sorting? Evidence from Tiebout Variation." *American Economic Review* 95 (4): 1310–26.

van der Vaart, A. W. 2000. *Asymptotic Statistics*. Cambridge, UK: Cambridge University Press.

Vaznis, James. 2013. "Barros Proposes Single Application for City, Charter Schools." *Boston Globe*, June 5.

Walters, Christopher R. 2014. "The Demand for Effective Charter Schools." National Bureau of Economic Research Working Paper 20640.