NBER WORKING PAPER SERIES

THE WELFARE EFFECTS OF COORDINATED ASSIGNMENT: EVIDENCE FROM THE NYC HS MATCH

Atila Abdulkadirolu
Nikhil Agarwal
Parag A. Pathak

Working Paper 21046
http://www.nber.org/papers/w21046

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2015

We thank Neil Dorosin, Jesse Margolis, and Elizabeth Sciabarra for their expertise and for facilitating access to the data used in this study. A special thanks to Alvin E. Roth for his collaboration, which made this study possible. Jesse Margolis provided extremely helpful detailed comments. We also thank Josh Angrist, Steve Berry, Stefan Bonhomme, Eric Budish, Francesco Decarolis, Matt Gentzkow, Jerry Hausman, Phil Haile, Ali Hortacsu, Derek Neal, Ariel Pakes, Jesse Shapiro, and seminar participants at the University of Chicago, Harvard University, 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. Pathak acknowledges support from the Alfred P. Sloan Foundation (grant BR2012- 068) and National Science Foundation (grant SES-1056325). Abdulkadiroglu and Pathak are on the Scientific Advisory Board of the Institute for Innovation in Public School Choice. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w21046.ack

The Welfare Effects of Coordinated Assignment: Evidence from the NYC HS Match
Atila Abdulkadirolu, Nikhil Agarwal, and Parag A. Pathak
NBER Working Paper No. 21046
March 2015, Revised October 2015
JEL No. C78,D47,D50,D61,I21

ABSTRACT

Centralized and coordinated school assignment systems are a growing part of recent education reforms. This paper describes how uncoordinated offers in New York City's old high school assignment mechanism were responsible for significant mismatch. More than a third of applicants were unassigned after the main round and ultimately administratively assigned to schools much worse than what they ranked. We then evaluate the distributional effects of the new coordinated system compared to the old system and other possible assignment mechanisms, taking advantage of the new mechanism's straightforward incentives to recover the distribution of student preferences. Our estimates suggest that the new mechanism achieves 80% of the possible gains from a no-choice neighborhood extreme to an idealized utilitarian benchmark. The elimination of congestion through offer coordination dominates the allocative effects of further modifications of the matching algorithm.

Atila Abdulkadirolu
Duke University
Department of Economics
Durham, NC 27708
atila.abdulkadiroglu@duke.edu

Nikhil Agarwal
MIT
Department of Economics, E17-224
77 Massachusetts Ave
Cambridge, MA 02139
and NBER
agarwaln@mit.edu

Parag A. Pathak
Department of Economics, E17-240
MIT
77 Massachusetts Avenue
Cambridge, MA 02139
and NBER
ppathak@mit.edu

An Online appendix is available at:
http://www.nber.org/data-appendix/w21046

1 Introduction

In the last fifteen years, the theory of market design 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 placement of 8$^{\text{th}}$ graders into high schools in New York City (NYC). NYC's new system has not only led to widespread scientific and popular acclaim (Nobel 2012, Tullis 2014, Roth 2015), but has also been a template for reforms in other cities.[1] Despite the widespread adoption and apparent consensus on the value of market-design inspired centralized assignment schemes, there is remarkably little evidence on whether, why, and how much a coordinated assignment system affects the allocation of pupils to schools or on 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 and the quantitative aspects of particular design trade-offs also remain open questions.[2]

Characterizing the state of the market prior to the new mechanism is a major challenge for such analysis because decentralized and uncoordinated systems do not usually generate systematic data. This paper surmounts this hurdle by exploiting new data on school placements in the system used in New York before 2003 to study the effects of moving from an uncoordinated assignment system to a coordinated single-offer system on the allocation of students to schools. Using our rich micro-data on applications, assignment, and enrollment, we describe how students are placed into high school in both systems and whether they are placed into one of their preferred choices. Tracking students from application to assignment allows for a comprehensive description of the drawbacks of NYC's previous system across students, but it still leaves open questions. First, we do not know whether the reform was able to realize most of the possible gains associated with a new assignment system and how those gains were distributed across applicants. Second, we know little about the magnitude of further algorithmic improvements considered by the market design literature compared to other aspects of the design. 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, aspiring 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 in turn were allowed to accept only one school and one wait list offer, and the cycle of offers and acceptances repeated two more times. Students not assigned in these rounds were assigned through an administrative process, which manually placed students at nearby schools. Since admissions offers were not coordinated across schools, we refer to this as the **uncoordinated mechanism.** In Fall 2003, the system was replaced by a single-offer assignment system, based on the student-

[1] Cities with new coordinated matching systems include Denver, New Orleans, Newark, and Washington DC. Many former Department of Education officials subsequently took leadership positions in other districts which put new centralized assignment systems in place inspired by the NYC experience including in Louisiana (Gabriela Fighetti, John White) and Newark (Cami Anderson).

[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, which relaxes the stability constraint in New York's system Abdulkadiroğlu, Che, Pathak, Roth, and Tercieux (2015). Abdulkadiroğlu, Che, and Yasuda (2015) argue that ordinal strategy-proof mechanisms may not lead to improvements in cardinal utility.

proposing deferred acceptance algorithm (DA) for the main round. Applicants were allowed to rank up to 12 programs for enrollment in 2004-05, and a supplementary round placed those unassigned in the main round. Since the central office coordinated offers across schools into a single offer, we refer to this new mechanism as the **coordinated** mechanism. The mechanisms could produce different allocations for three main reasons: 1) the new mechanism allows students to rank up to 12 choices, whereas the old mechanism only allowed for five; 2) a limited number of rounds of offers and acceptances in the old mechanism lead to congestion, where students hold on to less preferred choices while waiting to be offered seats at more preferred schools once others decline; and 3) the old 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 is an improvement compared to the old one. In the old mechanism, 18.6% of students matriculate at different schools from their assignment at the end of the match compared to 11.4% 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 obtain no first round offer and 59% of these applicants are administratively assigned. The take-up rates for students assigned administratively 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% of applicants left the district after assignment in the old mechanism, while only 6.4% left under the new mechanism. Students are also traveling 0.69 miles further to their new assigned school. While suggestive of welfare improvements, it is necessary to estimate econometric models of school demand to quantify the distribution of student welfare effects and the relevance of design issues at the heart of literature on school matching market design.

The fact that the new mechanism is based on DA, which is strategy-proof, motivates treating stated preferences as true preferences and sidesteps challenges associated with inferring preferences from highly manipulable systems. As far as we know, our paper is the first to fit econometric models of school demand using data generated by DA, despite the fact that many have argued for strategy-proof mechanisms because they generate credible preference data for guiding policy.[3] Our preference estimates characterize the heterogeneous nature of student preferences, which allows us to quantify what aspects of school choice market design are most important for allocative efficiency. Robustness of our estimates to variations on our assumptions on ranking behavior and evidence of in-sample and out-of-sample fit reassure us about the suitability of using stated preferences 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 extremes: a neighborhood assignment allocation, where each student

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

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 is 80% of the way to this idealized benchmark. Next, we find relatively modest gains from relaxing 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 have improved by 0.58%. An ordinally Pareto efficient matching, which abandons the stability constraints in the current mechanism, is equivalent to an improvement of about 3.3% 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 effects of simply having choice in a coordinated system, as measured by the range between neighborhood assignment and the assignment from the coordinated mechanism.

We then use our demand estimates to evaluate the transition from an uncoordinated to a coordinated mechanism. We examine three different interpretations of the data from the uncoordinated mechanism. Under each assumption, the new mechanism has made it easier for students to obtain a choice they want. Even though school assignments are further away, based on our estimated distribution of preferences, the amount that students prefer these assignments more than compensates for this difference. Our preferred estimate is that admissions coordination represents 58% of the range from no choice to utilitarian optimum. Students across all demographic groups, boroughs, and baseline achievement levels obtain a more preferred assignment on average from the new mechanism. The largest gains are for student groups who were more likely to be unassigned after the old mechanism's main round, suggesting that congestion and ad-hoc placement of unassigned students in the old mechanism are primarily responsible for misallocation. This comparison also shows that the elimination of congestion through offer coordination dominates the allocative effects of further modifications to the matching algorithm within the coordinated system.

This paper brings together two distinct literatures on school choice and matching mechanisms. We share a focus with papers interested in understanding how choice affects the assignment and sorting of students (Epple and Romano 1998, Urquiola 2005) rather than the competitive effects of choice on student achievement (Hoxby 2003, Rothstein 2006). We also concentrate our study on allocative efficiency rather than effects of choice options on subsequent achievement (Abdulkadiroğlu, Angrist, Dynarski, Kane, and Pathak 2011, Deming, Hastings, Kane, and Staiger 2014, Walters 2014, Neilson 2014). The allocative issues we focus on 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 Guell 2014, Hwang 2014, Burgess, Greaves, Vignoles, and Wilson 2015), typically using data from highly manipulable mechanisms like the Boston mechanism under strong assumptions on information and agent sophistication. While some of these papers have compared Boston and DA, ours is the first to examine congestion in an uncoordinated school assignment system. An approach based on estimated preferences is complementary to 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 compared to the strategic draft mechanism used at the Harvard Business School. Finally, our work is related to comparisons of decentralized and centralized medical labor markets by Niederle and Roth (2003), who compare the distance between medical school and residency locations for gastroenterologists before and after a centralized clearinghouse.

2 High School Choice in NYC

2.1 School Options

Aspiring high school students are eligible to apply to any school or program throughout New York City. Programs exist within schools and have curricula ranging from the arts to sciences to vocational training. The 2002-03 High School directory describes program types. Specialized High Schools, such as Stuyvesant and Bronx Science, have only one type of program, which admits students by admissions test performance on the Specialized High Schools Admissions Test (SHSAT).[4] There are three ways in which non-Specialized high schools differ in how they screen students. **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 graded, standardized/diagnostic test scores; attendance and punctuality; interviews; and essays. They 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. Allocation of seats in each half targets a distribution of student ability: 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 DOE closed and 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 the peak year of small high school openings was 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 small effect on overall enrollment patterns during our study period, which focuses on school options available in 2002-03 and 2003-04.

2.2 Uncoordinated Admissions in 2002-03

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

[4] Abdulkadiroğlu, Angrist, and Pathak (2014) describe their admissions process in more detail.

and the LaGuardia High School of Music & Art and Performing Arts, however, have been traditionally administered as a separate process from regular high schools and did not change with the new mechanism.[5] 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 preference in the fall. In the **Main round** in 2002-03, 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 accept, 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 accept 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 (VAS) process. In the Supplementary round, students could rank up to eight choices, and they were assigned based on the negotiation of seat availability between the enrollment office and high school superintendents. After replies to the second letter were received, a third round of letters were mailed. New offers did not necessarily go to wait-listed students in a predetermined order. Unassigned students were either placed at their zoned programs or placed administratively by the central office as close to home as possible. We refer to this final stage as the **Administrative round**.[6]

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 others 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 priority in admissions to students who ranked them first on their application form. The high school directory advises that when ranking schools, students should "determine what your competition is for a seat in this program" (DOE 2002). This recommendation puts strategic pressure on ranking decisions. Students have to both consider the limited number of potential applications and whether the school only considers first-choice applicants.

Third, a number of schools managed to conceal capacity to fill seats later on with better

[5] 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 and Performance Arts. City officials indicated that this law prohibits including these schools within the common application system without an act of the state legislature.

[6] 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 arrived to the process after the high school match. Arvidsson, Fruchter, and Mokhtar (2013) provide further details on the over-the-counter round.

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). Overall, critics alleged that the old mechanism disadvantaged low-achieving students and those without sophisticated parents (Hemphill and Nauer 2009).

2.3 Coordinated Admissions in 2003-04

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 to 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.[7]

In the **Main round**, students can rank up to twelve regular school programs in their applications, which are due in November. The DOE advised parents: "You must now rank your 12 choices according to your true preferences" because this round is built on Gale and Shapley (1962)'s 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 in the preference lists they were ranked. Schools return orderings of applicants to the central enrollment office. Schools that prioritize applicants using geographic or other criteria have those criteria applied by the central office. That office uses a single lottery to break ties among students with the same priority, generating a strict ordering of students at each school.

Assignment is determined by the student-proposing deferred acceptance algorithm, with student preferences over the schools, school capacities, and schools' strict ordering of students as parameters. The algorithm is run with all students in February. In this first round, only students who receive a Specialized High School offer receive a letter indicating their regular school assignment, and they are asked to choose one. After they respond, students who accept an offer are removed, school capacities are adjusted, and the algorithm is re-run with the remaining students. All students receive a letter notifying them of their assignment or whether they are unassigned after the Main round.

Unassigned students from the Main round are provided a list of programs with vacancies and are asked to rank up to twelve of these programs. In 2003-04, the admissions criteria at the remaining school seats were ignored in this **Supplementary round**. Students are ordered by their random number, and the student-proposing deferred acceptance algorithm 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**.

[7] There is very limited overlap between the specialized round and subsequent rounds. In 2003-04, 4,175 out of 4,553 of those offered a specialized high in our sample accepted that offer.

3 Data and Sample Definitions

3.1 Students

The NYC DOE provided us with several data sets for this study, each linked by a unique student identification number: information on student choices and assignments, student demographics, and October student enrollment. For 2002-03, the assignment files record a student's 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 2003-04, the assignment files contain students' choice schools in order of preference, priority information for each school, assignments at the end of each of the rounds, and final assignment as of early August 2004. The student demographic file for both years contains information on home address, gender, race, limited English proficiency, special education status, and performance on 7th grade citywide tests. We use addresses to compute the road distance between each student and school and to place each student in a census block group.[8] We also have access to similar files for 2004-05. Further details are in the Data Appendix.

Our analysis sample makes three 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 with the new mechanism. Third, we consider applicants who are given an assignment at the conclusion of the process (i.e., those who have not left midway). Given these restrictions, we have two main analysis files: the *mechanism comparison* sample and the *demand estimation* sample.

The mechanism comparison sample is used for comparisons of 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 as of the time of the printing of the high school directory. 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 2003-04 sample, a difference mainly due to the students assigned to schools created after the printing of the high school directory or to closed schools (as shown in Appendix Table C2).

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

The demand sample contains participants in the Main round of the new mechanism in the assignment files. The school choices expressed by these students represent the overwhelming

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

majority of students. Among the set of Main round participants, we exclude a small fraction of students who are classified as the top 2 percent because 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 Data Appendix.

3.2 Schools

Data on schools were taken from New York State report card files provided by NYC DOE. Information on programs comes from the official NYC High School Directories made available to students before the application process. Table 2 summarizes school and program characteristics across years. The number of schools increases from 215 to 235, and the number of students enrolled per school decreases by about 40 students. This is driven by the replacement of some large schools with smaller schools that took place concurrently in 2003-04, as described above. Despite this increase, there is little change in the average achievement levels of schools and school demographic composition as measured by report card data. We are not aware of other significant changes in school inputs, recruitment campaigns, and materials, including the format of the high school directory.[9]

Students in New York can choose 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 program, 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 into Arts, Humanities, Math and Science, Vocational, or Other. Not all programs have specialties, though about 70% fall into one of these classes. (Details on our classification scheme are in the Data Appendix). The menu of language program offerings or program specialties changes little across years.

4 Congestion and Changes in Assignments

The similarity of student and school attributes in Tables 1 and 2 suggest that there were not systematic changes in participant attributes and school supply across years. Moreover, there does not appear to be a large-scale change in student locations across years, as shown in Figure 1, which maps both student and school locations. These facts motivate attributing differences in allocations between 2002-03 and 2003-04 primarily to the assignment mechanism rather than changes in student participation or the menu of school options.

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

4.1 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 more students obtain their final assignment in the Administrative round of the uncoordinated mechanism than in the first round. Panel A of the Table shows that 37% of students are assigned administratively compared to 34% in the first round. Since 33,894 students obtained one or more first round offers (shown in Panel B), but only 23,867 students were finalized in the first round, 10,027 students with a first round offer were finalized with offers made in subsequent rounds. The processing of these students took place as schools revised offers based on first round rejections and made offers anew in the second and third rounds. However, the relatively small number of students placed in the second and third round implies that three rounds 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 new mechanism, where 82% of students were placed in the Main round.[10]

These observations about the old mechanism are characteristic of congestion, as described in Roth and Xing (1997)'s study of the labor market for entry-level clinical psychologists. Offers for training positions in that market were made in an uncoordinated fashion during a 7-hour window, and Roth and Xing (1997) argue that uncoordinated processing and a small market-clearing window led to mismatch. In NYC, the low number of rounds and the serial-processing of batches of offers, whereby programs waited for previous offers to be rejected before making new offers, combined to have a similar effect. In addition to insufficient processing of offers, the small number of applications allowed in the old mechanism also led to situations where students fell through the cracks if they applied to oversubscribed schools. Since rank order lists were short, the mechanism considered a smaller number of alternate choices for these students compared to a mechanism which allowed students to rank more choices. 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 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: 1) the Main round, where only the top five choices are considered and there is no restriction in the number of rounds, and 2) the Main round with twelve choices, but only three sets of proposals from the deferred acceptance algorithm.[11] The first is intended to isolate the role of five choices, while the second isolates the role of few offer-processing rounds. Since we do not model behavioral responses by students, we only intend this exercise to shed light on mechanical features generating administrative assignments in the uncoordinated mechanism. With that caveat in mind, we find that the five-choice constraint with an unlimited number of rounds leaves about one quarter of applicants unassigned, while the unconstrained mechanism with three proposal rounds leaves roughly

[10]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%, were placed in the Main round (shown in Appendix Table B1).

[11]Even though there are multiple possible implementations of deferred acceptance, our simulation considers the simultaneous-proposing version, where a round is defined by a set of proposals by students who are not tentatively held or have not exhausted their rank order list.

half of applicants unassigned. Relative to the uncoordinated mechanism, 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 perhaps less so than few offer processing rounds.

4.2 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' residence and their assigned school 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 to Far Rockaway at the easternmost tip of Brooklyn.[12] 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 assignment. 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 a nearby school.

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

Student enrollment patterns documented in Table 3, however, indicate that student assignments in the uncoordinated mechanism, particularly those made in the Administrative round, are undesirable relative to those 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. 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 is higher in the uncoordinated mechanism (8.5% compared to 6.4%), and the fraction of students who enroll at a school other than their assignment is higher (18.6% compared to 11.4%).

In the uncoordinated mechanism, students assigned in earlier rounds appear more satisfied with their assignment than those assigned in later rounds. The fraction of students who exit NYC public schools is 13.3% among administrative placements, compared to 5.2% 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

[12] Our analysis focuses on road distance, which is highly correlated with subway distance. Appendix B presents a detailed comparison of both measures.

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. These students also travel further to their final assignment or enrolled school. In contrast, the majority of students with no offers are assigned through the Administrative round, and this 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 of that round.[13]

4.3 Mismatch in Administrative Round

To further evaluate the assignments of students processed in the Administrative round, we compare the attributes of schools that students wanted (or ranked) to the attributes of schools to which they were assigned. Students processed in earlier rounds are assigned to schools that have attributes more similar to the schools they ranked than students processed in later rounds. Table 4 shows that for those assigned in the Main round, with the exception of distance, ranked schools tend to have similar or slightly better attributes than assignments in both mechanisms. For instance, ranked schools have higher Math and English performance, more students attending four-year colleges, and higher attendance rates. They 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 compared to the Main round. In the uncoordinated mechanism, for instance, the gap in Math performance between ranked and assigned schools is 0.7 percentage points for those assigned in the Main round, while it is 2.5 points in the Supplementary round. The gap between ranked and assigned alternatives for 9th grade size is quite pronounced under both mechanisms. For example, in the uncoordinated mechanism, ranked schools have about 200 fewer 9th graders than the schools where students are assigned. Since students participate in the Supplementary round when they did not obtain a Main round assignment, it is not surprising that the difference between what students wanted and what they received widens.

The most striking pattern in Table 4, however, is for students who are administratively assigned. We've already seen that there are three times more students assigned in this round in the uncoordinated mechanism. Panel C shows that students who are assigned in that round ranked schools on average 5.1 miles away from home and were assigned to schools only 1.6 miles away in the uncoordinated mechanism, a much larger gap than 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 is now 4.4 points, and there is a similar widening in the fraction going on to four-year colleges. The difference in 9th grade size is also considerable: in the Supplementary round, students wanted schools with on average with roughly 700 9th graders, and they were assigned

[13] In the second year of the mechanism, the average distance to the assigned school is 4.07 miles and the average exit rate is 6.4% (shown in Table B1). The take-up rate is higher than the first year and the fraction in the Administrative round decreases to about 5%.

to schools with more than 900 9th graders. In the uncoordinated mechanism's Administrative round, they are assigned to schools with nearly 1,200 9th graders.

The differences between ranked and assigned schools are also large in the coordinated mechanism, though in some cases, they are not as stark. Differences between ranked and assigned schools on Math and English achievement or 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 is that being processed in the Administrative round is undesirable for students in both mechanisms. As a result, it is reasonable to expect that a significant fraction of the changes in student welfare will be driven by the reduction in the number of students who enter this round in the coordinated mechanism.

4.4 Offer Processing by Student Characteristics

Table 5 reports the attributes of students across rounds compared to the overall population of applicants in the uncoordinated mechanism. Students from Manhattan, those with high math baseline scores, and those who have applied to exam schools (indicated as taking the SHSAT) tend to obtain offers earlier in the uncoordinated mechanism. They 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 neighborhoods tend to systematically obtain offers later in the process. Compared to the overall population, these groups are overrepresented in the Administrative round. About a quarter of students in the Administrative round are white, compared to 15% of participants overall.

The coordinated mechanism distributed 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 they are not as overrepresented as in the old mechanism.

The coordinated mechanism assigned 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 2002-03. 1,455 students were assigned to the Louis Brandeis High School, a struggling Manhattan high school whose four-year graduation rates were among the lowest in the city, but only 911 students were assigned there in 2003-04.[14] The upward sloping line means that if a school is more oversubscribed in the old mechanism, the new mechanism tended to assign more students to that school. This phenomenon suggests that the coordinated mechanism was able to use the submitted preferences more effectively to place children into the

[14]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 9th graders. This high school had a longstanding reputation for violence and disorder, and it was eventually closed in 2008.

schools that they wanted. The extent to which this represents an improvement in student welfare depends on the heterogeneity of student preferences, an issue we turn to next.

5 Estimating Student Preferences

5.1 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 NYC High School directory, a booklet that includes 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.[15] The first row of the table shows that 20% of applicants rank 12 school choices; the majority rank nine or fewer choices, and nearly 90% rank at least three 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, most students do not simply rank the school closest to home first. For the typical student, the first choice is 0.44 miles closer than her second choice, and her second choice is 0.25 miles closer than her third choice. Distance increases monotonically until the 9th 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 a higher proportion 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 in high income and non-minority 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 ranking behavior by baseline ability and neighborhood income. High-achieving students tend to rank schools with high Math achievement relative to low achievers, though both groups place less emphasis on achievement further down their preference list. Similarly, students from low-income neighborhoods tend to put less weight on Math achievement than students from high-income neighborhoods, but both groups rank higher achieving schools

[15] Table B3 provides additional information on school assignments. 31.9% of students receive their top choice, 15.0% receive their second choice, and 2.4% receive a choice ranked 10th, 11th, or 12th. 17.5% of students are asked to participate in the Supplementary round because they are unassigned in the Main round.

higher. These differences suggest the importance of allowing for tastes for school achievement to differ by baseline achievement and income groups in the demand model.

The characteristics of schools ranked in the uncoordinated mechanism, also shown in Table 6, are decreasing in a school's achievement and income. For distance, however, preferences are compressed relative to the coordinated mechanism. For instance, the distance to a students' first choice 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 its incentive properties and the fact 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 it led to assignments that are worse on average than the old mechanism. On the other hand, students may prefer schools outside of their neighborhood because they are a better fit. We must therefore weigh the greater travel distance in the new mechanism against changes in other dimensions 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.

5.2 Model and Estimation

The comparison of the attributes of choices ranked higher to those ranked lower or not at all provides rich information to identify how a student values school features compared to distance. To quantify these trade-offs, we work with a random utility model, which allows for factors that are not observed in our dataset to influence ranking decisions. Let i index students and j index programs. Since all of the schools we study are free, we treat distance as our numeraire for our welfare analysis. Specifically, we project student i's indirect utility for program j on student and school characteristics as follows:

$$\begin{aligned} u_{ij} &= \delta_j + \sum_l \alpha^l z_i^l x_j^l + \sum_k \gamma_i^k x_j^k - d_{ij} + \varepsilon_{ij}, \qquad (1) \\ \text{with} \quad \delta_j &= x_j \beta + \xi_j, \end{aligned}$$

where z_i is a vector of student characteristics, x_j is a vector of program j's characteristics, d_{ij} is distance between student i's home address and program j, ξ_j is a program-specific unobserved vertical characteristic, γ_i is a vector of random coefficients that capture idiosyncratic tastes for program characteristics and ε_{ij} captures idiosyncratic tastes for programs.

Because we would like the model to capture heterogenous preferences but still be computationally tractable, we employ an ordered version of the choice model in Rossi, McCulloch, and Allenby (1996). They describe a class of parametric distributions for unobserved characteristics and idiosyncratic preferences that allows for estimation via Gibbs' sampling.[16] Specifically, we

[16]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 chapter 10.1 in van der Vaart (2000)).

assume that

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

and conjugate priors for β, α, Σ_γ, σ_ξ^2, and σ_ε^2. The specific distributions and additional details are described in the Computational Appendix.

Our specification is motivated in part by the fact that students may have idiosyncratic tastes for schools that are not captured by the observable student characteristics in our dataset. This specification allows for arbitrary correlation between the k dimensions of γ_i, and therefore exploits the richness of the rank-ordered data. Berry, Levinsohn, and Pakes (2004), for instance, argue that data on top and second choices can be used to estimate these parameters by exploiting common characteristics between subsequent rankings for a given student. Rank order list data also allow us to relax the common assumption that random coefficients on choice characteristics are independent. For instance, our model allows for correlated unobserved tastes for school size and achievement levels.

The parametric assumptions are made for computational tractability as the distribution of indirect utilities is non-parametrically identified in our setting.[17] Given the independence of idiosyncratic tastes, the two primary restrictions in our specification is that the distance to school is additively separable in the indirect utility function and that taste shocks are independently distributed. All else being equal, if students dislike traveling to school, then the coefficient of -1 is an equivalent scale normalization to the common practice of setting the variance of ε to 1, which allows for a convenient distance metric for utilities.

We do not explicitly include an outside option and instead normalize, without loss of generality, the value of δ for an arbitrarily chosen school to zero. This assumption is motivated by our primary interest in 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 assume that students who have not ranked all 12 choices prefer their outside option to a NYC high school. However, Table B4 shows that roughly three-quarters of Supplementary round students enroll in their offer for that round, and the majority enroll in some other NYC high school.

The demand sample for 2003-04 contains rankings of 69,907 participants across 497 programs in 235 schools, representing a total of 542,666 school choices. Our specifications follow other models of school demand 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, percent subsidized lunch, percent white, and 9th grade size.

We start by assuming that rankings reflect true preferences. This benchmark is natural because of the straightforward incentive properties of the mechanism, and because of the advice that the NYC DOE provides in the 2003-04 High School Directory and through their information

[17] Assumptions needed to identify preferences for choice characteristics in binary and multinomial settings have been examined in Ichimura and Thompson (1998), Lewbel (2000) and Briesch, Chintagunta, and Matzkin (2002), though ordered choice data contains additional information. Agarwal and Somaini (2014) study identification in the school choice context with a potential manipulable mechanism. Non-parametric identification results in these settings carry over to our case.

and outreach campaign. For instance, the DOE guide advises participants to "rank your twelve selections in order of your true preferences" with the knowledge that "schools will no longer know your rankings." Nonetheless, truthful behavior is a strong assumption that we revisit in Section 8.

5.3 Preference Estimates

We report select estimates for six specifications of our demand model in Table 7 (the full set of estimates is in Table A1). We interpret our estimates as describing student utilities from various schooling options and not as the casual effect of changing school attributes on student preferences. The first specification includes school characteristics (high Math achievement, percent subsidized lunch, percent white, and 9th grade size), but does not include interactions of school and student characteristics. We do not include additional achievement characteristics examined in Table 4, such as high English achievement and percent attending a four-year college, because they are closely related to high Math achievement. The second specification includes student-school interactions. The next four specifications each add random coefficients. They also vary which ranks are used in estimation: all choices, only the top three choices, all except the last choice, and choices for students who rank fewer than 12. Each specification with student interactions include dummies for Spanish, Asian and Other Language Programs, interacted with a student's English proficiency status and whether they are Hispanic or Asian. Computational constraints prevent us from estimating all of these models on the full sample, so we only report estimates using the full sample for the specification with student-school interactions and student-specific random coefficients; the rest of these models are estimated on a 10% random sample of the data, which contains about 7,000 students.

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 broadly similar across the four specifications with random coefficients. Third, many of the random coefficients are significantly estimated, suggesting the importance of a flexible specification in accounting for the underlying heterogeneity in student preferences. We report further on measures of model fit in Section 8. Since it fully exploits all of the choice data in the most flexible model, the estimates in the third column will be for our primary calculations, though we will revisit these other specifications in Section 8.

6 Comparing Alternative Allocation Mechanisms

6.1 Measuring Welfare

Our estimates allow us to compute measures of welfare across assignments, assuming that student preferences do not change. We measure the difference between two assignments for a student in terms of the distance-equivalent utility, or the amount a student is willing to travel for the more preferred school assignment. To compare the welfare associated with two assignments, we compute the average welfare in distance-equivalent utility. Consider a group of students in set G

and a matching μ, which specifies the program for each student as $\mu(i)$. Define average welfare for students in G as a function of parameters θ as

$$W_G^{\mu}(\theta) = \frac{1}{|G|}\sum_{i\in G} u_{i\mu(i)}(\theta),$$

where u_{ij} is the utility student i associated with assignment to program j. For two matchings, μ and μ', and students in group $G(\mu)$ and $G(\mu')$, the difference in welfare between the two matchings is given by:

$$W_G^{\mu}(\theta) - W_G^{\mu'}(\theta) = \frac{1}{|G(\mu)|}\sum_{i\in G(\mu)} u_{i\mu(i)}(\theta) - \frac{1}{|G(\mu')|}\sum_{i\in G(\mu')} u_{i\mu'(i)}(\theta).$$

In the coordinated mechanism, we observe the rankings submitted by students. Under the assumption that these reports are truthful, the k^{th} ranked program yields the k^{th} highest utility. For a given student and estimate of θ, the observed ranking contains information about unobserved tastes ε_{ij}. For students in the coordinated mechanism who submit rank order list r_i, we calculate the expected utility for each ranked and unranked choice by simulating the utility from the estimated preference distribution, conditional on the relationships implied by the submitted ranks.

In particular, we compute the expected utility of a program ranked k^{th} by student i, denoted r_{ik}, as

$$\mathbb{E}\left[u_{ij} \mid u_i^{(k+1)} < u_{ij} < u_i^{(k-1)},\, r_{ik} = j\right] \tag{2}$$

and the expected utility for all unranked schools as

$$\mathbb{E}\left[u_{ij} \mid u_{ij} < u_i^{(|r_i|)},\, j \notin \cup_k\{r_{ik}\}\right], \tag{3}$$

where $u_i^{(k)}$ is the k^{th} highest value of $\{u_{ij}\}_{j=1}^{J}$ and $|r_i|$ is the number of ranks submitted by student i.[18] With these definitions, the expected welfare difference for group G from assignments μ and μ' is

$$W_G^{\mu} - W_G^{\mu'} = \frac{1}{|G(\mu)|}\sum_{i\in G(\mu)} \mathbb{E}[u_{i\mu(i)}|r_i] - \frac{1}{|G(\mu')|}\sum_{i\in G(\mu')} \mathbb{E}[u_{i\mu'(i)}|r_i],$$

where $\mathbb{E}[u_{ij}|r_i]$ denotes the conditional expectations in equations (2) and (3) above.

6.2 Evaluating Mechanism Design Features

Even though many of the new mechanism's features were designed to address issues in the old mechanism, the coordinated mechanism still involves several design constraints. Here, we ask: how far is the allocation produced by the coordinated mechanism from the best possible one for students, and what are the quantitative effects of particular design decisions? The alternatives we

[18] For students who are in the mechanism comparison sample, but not the demand sample (so we don't observe their rank order lists), we use the mean utility conditional on observables alone in the welfare calculations. We follow the same approach for students assigned to choices not ranked in the Main round.

consider vary the mechanism, holding fixed the set of schools, students, and residential choices, and therefore are best interpreted as short-run effects of alternative mechanisms.

We begin by using the demand model estimates to assess two benchmark allocations intended to capture the range of what is achievable with a choice system. The first is a **neighborhood assignment**, which models schools as prioritizing students in order of distance and placing students at their closest possible school. We compute this assignment by running the deferred acceptance algorithm with applicants simply ranking schools in order of distance and vice-versa. This allocation further restricts the market's geographic scope as in the Administrative round of the uncoordinated mechanism.

The second benchmark allocation, the **utilitarian assignment**, maximizes the sum of student utility subject to the feasibility constraints of the assignment.[19] Given the estimated distribution of student preferences, no other allocation can yield higher average utility. This allocation therefore represents the other extreme, where a planner implements the best possible allocation by taking full advantage of the cardinal distribution of student preferences.[20]

To compare the coordinated mechanism to these two benchmarks, in Table 8 we report the difference with the utilitarian outcome. The first column shows that the difference in distance-equivalent utility between the neighborhood and utilitarian assignment is 18.96 miles. This difference is a theoretical upper bound on the gains from school choice. Across student groups, the gains from a choice system are smaller for whites and Asians compared to blacks and Hispanics. They are also considerably smaller for residents of Staten Island, which has on average better performing schools than other boroughs. The coordinated mechanism achieves about 80% of the possible gains from a choice system, since the difference in distance-equivalent utility with the utilitarian assignment for the average student is 3.73 miles.[21]

The utilitarian optimal assignment is an idealized benchmark, but 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 at those programs. This allocation may also allow for situations of justified envy at programs that only use coarse priorities, including those based on geography. Second, this assignment uses cardinal information, which is not elicited by the coordinated mechanism.[22] Therefore, we next turn to more efficient outcomes for students that do not completely abandon school priorities and only use ordinal information in student rank order lists.

[19] We solve for this allocation following Shapley and Shubik (1971). Specifically, we solve the following linear program:

$$\max_a \sum_{ij} u_{ij} a_{ij} \text{ s.t. } \sum_j a_{ij} \leq 1, \sum_i a_{ij} \leq c_j, a_{ij} \in \{0,1\},$$

where c_j is program j's capacity and a is $N \times J$ matrix with (i,j) element a_{ij}. The utilitarian allocation implies equal weights on students.

[20] Using illustrative examples, Pycia (2014) argues that the welfare loss of ordinal mechanisms relative to cardinal ones can be arbitrarily large.

[21] We implement the coordinated mechanism by drawing 100 sets of lottery numbers and re-run the student-proposing deferred acceptance algorithm given student's choices. 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.

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

The deferred acceptance algorithm in the coordinated mechanism need not produce a student-optimal stable matching because it must resolve ties between students when they have identical priorities at a school. This tie-breaking can lead DA to produce stable outcomes, which are not student-optimal. Deferred acceptance cannot be Pareto-improved upon without sacrificing strategy-proofness for students (Erdil and Ergin 2008, Abdulkadiroğlu, Pathak, and Roth 2009, Kesten 2010, Kesten and Kurino 2012). We therefore quantify the cost of providing straightforward incentives for students by computing a student-optimal stable assignment, which improves the DA assignment by placing students higher in their choice lists while also respecting stability for strict school priorities. Such an assignment can be computed by the stable improvement cycles algorithm developed by Erdil and Ergin (2008), which iteratively finds Pareto-improving swaps for students while still respecting the stability requirement for underlying weak priority ordering of schools (i.e., prior to tie-breaking). A total of 2,348 students in the demand sample can obtain a better assignment in a student-optimal stable matching. The difference in distance-equivalent utility is 0.11 miles on average compared to the assignment produced by the coordinated mechanism. The cost is that the underlying mechanism is not based on a strategy-proof algorithm.[23]

Another limitation of the coordinated mechanism is that it constrains student welfare due to its treatment of school priorities and preferences. To quantify the importance of this constraint while still retaining the mechanism's ordinal message space, we compute the welfare of students under a Pareto-efficient assignment that dominates DA by transferring students from their assigned schools to their higher-ranked choices via the Gale's Top Trading Cycles algorithm (Shapley and Scarf 1974). 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 incentives for schools to participate in the system.[24]

We calculate a Pareto-efficient matching that dominates each simulated student-optimal stable matching and report average utility 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 represents a substantial difference between the utilitarian optimum. The utility difference for the average student 3.11 miles. Relative to the current mechanism, the cost of limiting the scope for strategizing by schools (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 not having a choice system at all is much larger than possible welfare gains from fine-tuning the coordinated mechanism. That is, the ability to choose schools generates substantial student welfare when preferences are heterogeneous. Within a coordinated matching system, further optimizing the matching algorithm produces relatively little gain in the best case, even if it were possible to implement cardinal allocation

[23] Azevedo and Leshno (2011) provide an example where the equilibrium assignment of the stable improvement cycles mechanism is Pareto inferior to the assignment from deferred acceptance when students are strategic.

[24] Balinski and Sönmez (1999) and Abdulkadiroğlu and Sönmez (2003) provide an alternative equity rationale for stability. Note that no stable mechanism eliminates strategic maneuvers by schools (Sönmez 1997), although this may not be an issue in markets with many participants (Kojima and Pathak 2009).

schemes. This conclusion does not imply that the matching scheme is not important because we have seen the large number of those processed administratively in the uncoordinated mechanism. To see where in the spectrum the uncoordinated mechanism lies, we next turn to analyzing its properties.

7 Comparing the Coordinated and Uncoordinated Mechanisms

Unlike the variations on the coordinated mechanism we've just examined by assuming that applicants report their true preferences, there is no simple way to model behavior in the uncoordinated mechanism. To evaluate it, we adopt an alternative approach that takes advantage of the fact that we observe the rankings and assignments of students in that mechanism. This approach requires two important assumptions.

First, since we do not observe the same student in both the uncoordinated and coordinated mechanism, we have to assume that we can use preference estimates from the coordinated mechanism to make statements about the previous year. One potential concern is that schools have changed significantly because of the mechanism, so a student's valuation of the school likely has changed alongside the mechanism. To probe this issue, in Figure A1, we plot the average characteristics of the schools, as measured by the attributes of peers, in both mechanisms.[25] The Figure shows that for measures of poverty, racial make-up, and baseline math scores, there is relatively little change in school attributes despite the change in the mechanism. Consistent with the increased travel distance in Table 3, the last panel of the Figure shows that schools differ when measuring the travel distance of enrolled students.

The second assumption involves our interpretation of the rankings submitted in the uncoordinated mechanism. For the counterfactuals in Table 8, to compute the utility associated with an assignment for the new mechanism, we condition on the rank order list submitted by the student, adjusting for what that rank order list implies about unobserved tastes under the assumption that the student reported her true preferences. The old mechanism's uncoordinated nature makes the relationship between preferences and the final assignments less straightforward. Computing the expected utility, conditional on the assignments produced by the uncoordinated mechanism, requires strong assumptions about that mechanism, its properties, and agent behavior and expectations. 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.[26] Instead, our approach exploits our data on observed rankings from the uncoordinated mechanism's first round.

The approach we start with, **unordered applications**, assumes that choices submitted in the uncoordinated mechanism have the property that any ranked choice is preferred to an unranked choice, but it does not assume that higher-ranked choices are preferred to lower- ranked ones.

[25] We find qualitatively similar results for 25th, 50th and 75th percentiles of the distribution of these quantities.

[26] 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. Agarwal and Somaini (2014) present an approach to uncovering preferences from a class of single-offer manipulable mechanisms that involve lotteries; because the uncoordinated mechanism is a multiple-offer system, it falls outside of this class, and their methods are not applicable in our setting.

Specifically, we assume that the expected utility of program j is $\mathbb{E}\left[u_{ij}|u_{ij} > u_{ij'}, j' \notin \cup_k\{r_{ik}\}\right]$ if it is ranked by student i, and $\mathbb{E}\left[u_{ij}|u_{ij} < u_{ij'}, j' \in \cup_k\{r_{ik}\}\right]$ otherwise.[27] This assumption has the benefit of not assuming rank order lists submitted in the uncoordinated mechanism are truthful, while still using some of the information contained in the preference submission. A weakness of this assumption is that it excludes the possibility that students have omitted choices from their rank order list that they prefer, but do not expect to be admitted. However, in Table 6, we saw that less than half of participants rank all five choices in the old mechanism, suggesting that if these participants wanted to rank more schools, the mechanism does not constrain them. Nonetheless, we also consider two other assumptions later in this section.

Under the unordered applications, selection assumption from the uncoordinated mechanism, and assuming true preferences from the coordinated mechanism, we find that most students are better off under the coordinated mechanism. Figure 4 plots the overall distribution of student welfare from the two mechanisms. The average student improvement in welfare is equivalent to 10.96 miles. The figure shows a bimodal pattern of utility in the uncoordinated mechanism due to the students who are assigned in the Administrative round. Most of the mass in the first mode shifts rightward in the coordinated mechanism, a phenomenon driven by the sharp reduction in the number of students assigned administratively.

For each student group shown in Table 9, there is a positive gain from the new mechanism. Across racial groups, whites and Asians gain more than blacks and Hispanics. There are more pronounced differences across boroughs. Students from Staten Island and Queens gain the most, while the smallest gain of any student group in the table comes from Manhattan residents. Low baseline math students gain more than high baseline math students or SHSAT test takers.

It's worthwhile to compare the difference in utility to the difference in travel distance to decompose the welfare effects of the improved school match compared to distance. On average, the improved match ignoring distance is equal to 11.65 miles. This implies that the improved school match is worth thirteen times the costs associated with increased travel distance. The lowest gains are for Manhattan residents, a group which experiences no increase in travel distance. However, the welfare gains are not solely driven by changes in distance. For instance, across boroughs, Staten Island pupils only travel 0.34 miles further, but they experience the largest improvement of any borough at 22.56 miles. This suggests that 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.[28] The magnitude of our estimates of improved school matches suggest that even if there are not enough good schools to assign everyone their top choice, preference heterogeneity generates an important role for the coordinated mechanism.

The welfare gains in the coordinated mechanism are larger for many disadvantaged groups, a pattern consistent with Hemphill and Nauer (2009)'s claim that the uncoordinated mechanism advantaged high-achieving students and those with sophisticated parents. For instance, welfare

[27] We simulate these expectations using a Gibbs sampler.

[28] In the uncoordinated mechanism, there are 1,054 students who ranked Staten Island Technical, a highly sought-after screened school. Only 16% are assigned there, and about 75% do not obtain a Main round offer and are subsequently administratively assigned.

gains are larger for low baseline math students than for high baseline math students. They 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 large role in the fact that whites and rich neighborhoods experience larger welfare gains than blacks and Hispanics and poorer neighborhoods.

Differences across student groups closely track the rounds in which students were processed in the uncoordinated mechanism. For instance, substantially more high baseline students were assigned in the Main round of the old mechanism compared to low baseline students. Student groups with higher fractions assigned in the Administrative round, shown in column 7, tend to experience the largest gains. Figure 5 reports on the relationship between the likelihood to be assigned administratively and student welfare. To generate this figure, we fit a probit model to describe whether a student is administratively assigned on all student characteristics in the demand model for the uncoordinated mechanism sample. The specification includes census tract dummies to account for neighborhoods that may or may not have high schools, which are a guaranteed fallback option for some students. We then use this estimated relationship to compute each student's propensity to be assigned administratively in both the uncoordinated and coordinated mechanism. For each decile of this propensity, we compare the utility from assignment across both mechanisms. However, there is a clear monotonic pattern between administrative assignment propensity and the welfare improvement, whether comparing utility differences either including distance or net of distance.

A possible threat to our welfare calculation is that an initially bad allocation was subsequently undone post-assignment through the aftermarket, the period between when offers where made and the school year starts.[29] 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. The coordination of admissions occurred with greater central control of enrollment, which may have made it possible that the more rigid aftermarket in the coordinated mechanism is actually worse for students. To examine this possibility, we also compute the utility associated with the schools students enroll at in October of the following school year. Compared to assignments, the gains from the coordinated mechanism measured by enrollment are somewhat smaller, but are still large. For instance, the distance-equivalent utility for the average student is 9.67 miles, which is 88% of the gain from the assignment. The change in travel distance to enrolled school is also lower than the change in travel distance to assignment. Though a smaller gain from enrollment suggests that 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.

Finally, it's worth noting how the mechanism design matters when we compare the uncoordinated and coordinated mechanisms. That difference represents nearly 60% of the total range between the neighborhood and utilitarian assignments, which is much larger than the range associated with tweaks to the matching algorithm. This finding informs a broader debate in the

[29]Relatedly, since the exit rate in the coordinated mechanism is lower than the uncoordinated mechanism, more students preferred accepting their coordinated offer to enrolling in a high school outside of the system. This fact suggests that our welfare estimate understates the overall effect for all public school 8th graders.

market design literature about the importance of sophisticated market clearing mechanisms. In the context of auctions, Klemperer (2002) 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.

To see how robust our conclusions are, in Table 10, we repeat the analysis in Table 9 for three alternative assumptions. In each case, we break the asymmetric treatment of selection rules in Table 9, and use the same assumption about what rank order lists imply about unobserved tastes in our welfare calculation for both mechanisms. In the first variation, we do not adjust the utility based on rank order lists submitted in either the uncoordinated or coordinated mechanism. This **unselected applications** (on unobservables) selection rule simply assumes that the idiosyncratic component of tastes has no influence on utility. Specifically, we compute the expected utilities given only observable school and student characteristics, i.e. for both ranked and unranked programs, $\mathbb{E}\left[u_{ij}\right] = x_j\beta + \sum_l \alpha^l z_i^l x_j^l - d_{ij}$, with the parameters α and β set to their posterior means. While this assumption may be unappealing given the importance of unobserved determinants of applicant preferences, examining welfare differences under it indicates how much our estimates are driven by observable dimensions of school and student characteristics. Column 1 of Table 10 reports that the difference in welfare is more than one half of that implied by unordered applications selection rule. The average student is going to a school that is about eight times more valuable compared to the increased travel distance. The qualitative patterns are similar to those under unordered applications selection rules: all student subgroups benefit and the gains are larger for groups that were more likely to be administratively assigned. For instance, gains are larger for Asians and whites; they are largest for Staten Island residents and lowest for Manhattan. One difference with the unordered applications selection rule is for comparisons across baseline achievement groups: high baseline math students benefit more than low baseline math students, though low baseline students still gain slightly more than SHSAT test-takers.

Columns 3 and 4 present estimates using the unordered application selection rule for both the coordinated and uncoordinated mechanism. The average welfare difference of 9.02 miles is smaller than that reported in Table 9, but the patterns across student groups is similar. Finally, in columns 5 and 6, we report estimates from the **truthful applications** selection rule, which assumes that the rank order list submitted in the either the uncoordinated or coordinated mechanism reflect true applicant preferences, with expected utilities computed as in equations (2) and (3). Under this assumption, the estimated welfare gains are slightly smaller than those reported in Table 9, at 10.62 miles on average. The qualitative patterns mirror those from the unordered application rule, though they tend to be slightly larger.

It is reassuring that the estimates we obtain under different selection rules using information from submitted rank order lists are similar between Table 9 and 10. The smaller estimate under the unselected applications selection rule in Table 10 implies that unobserved tastes are important for our quantitative conclusions. Except for comparisons across baseline achievement groups, the qualitative comparisons between student groups are also similar under each of the assumptions

8 Model Fit and Alternate Behavioral Assumptions

8.1 Model fit

Since our goal is to make statements about welfare, it is important to examine how well our econometric model matches the data. We first investigate within-sample fit to see what our estimates imply for the aggregate patterns by rank in Table 6. Figure A2 reports on measures of fit using the specification with student-level random coefficients and student-school interactions from 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 12th choice is 16.7 to 10.4. Our estimates imply that for first choices, school's fraction high math achievement is 18.4, while it drops to 11.8 for the 12th choice. The 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've estimated with random coefficients. It's 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 mixture 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 school's size may be indicative of taste for large schools. In Table A2, 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, comparing the 2003-04 preference estimates to examine the correlation pattern of choice made in 2004-05, we also see that the correlation pattern in our main specification is closer to the observed pattern than from a demand model without student interactions.

8.2 Behavioral Assumptions on Ranking

8.2.1 Stability of Preferences

Motivated by the incentive properties of the deferred acceptance algorithm, the preference estimates that we reported come from models assuming that students are truthful in the coordinated mechanism. We revisit some potential objections to this assumption in this subsection.

In an influential experiment, Hastings and Weinstein (2008) show parent preferences are malleable when provided with direct information on school test-scores in Charlotte-Mecklensberg's choice system. Such a finding might suggest that students could be overwhelmed by the prospect of evaluating over 500 school programs and that preferences are an unreliable guide for welfare analysis. However, if preferences were generated without much of a systematic component, then we'd expect that most of our point estimates to be imprecisely estimated or have unintuitive patterns, contrary to what we've seen. Moreover, in Figure A4, we plot school market shares between the first year of the coordinated mechanism and its second year. This plot illustrates the extent of variation in market shares, which may occur with new information such as expanded high school fairs across time. The market shares of most programs are very similar across both years of the new system. This fact suggests that aggregate preferences have a substantial stable component across time.

Another way to see that choices are consequential for participant welfare is to examine enrollment decisions by choice received. Table B4 reports assignment and enrollment decisions for students who are assigned in the Main round. The table shows that 92.7% of students enroll in their assigned choice, and this number varies from 88.4% to 94.5%, 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 fact suggests that either families are indifferent between later choices and simply enroll where they obtain an offer or that families have deliberately investigated 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 the top three choices of applicants in column 4 of Table 7.

8.2.2 Assumptions on 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. Table B4 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 a school in which they have a higher chance of admissions last. For instance, Calsamiglia, Haeringer, and Kljin (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 these schools in case they are rejected by their higher choices. If students consider applying and commuting to schools further away from their neighborhood for the school's achievement level, 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 twelve 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 in column 5 of Table 7.

Another issue with the assumption of 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 her choice set down to at most 12 schools. If a student is only interested in 11 or fewer schools, this constraint in principle should not influence her 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 in the application form. However, 20.3% of students in our demand sample rank 12 schools. Some of these students may drop highly sought-after schools from 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 twelve choices in column 6 of Table 7.

In Table A3, we report on our evaluation of mechanism design choices under these three specifications: 1) using only the top 3 choices, 2) excluding applicants who have ranked all 12 programs, and 3) dropping the last choice of applicants. Because of computational constraints, we estimate the models on a 10% random sample, but we use the full sample for the counterfactuals. For all three demand models, the coordinated mechanism in column 2 is roughly 81% 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 A3 reports on how the comparison between mechanisms varies with our demand specification using all of the ranking information of participants. 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 heterogeneity (column 1 in Table A3) and without random coefficients (column 2 in Table A3). The neighborhood assignment is more appealing according to those two demand models, since they are only 15.5 and 16.2 miles away from the utilitarian assignment, compared to 21.5 miles from the main specification in the 10% sample. Moreover, the difference between neighborhood assignment and the coordinated mechanism is smaller in specifications that do not allow for student interactions or random coefficients than in the main specification that includes both student interactions and random coefficients.

9 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 enrollment. We find that the new coordinated mechanism is an improvement relative to the old uncoordinated mechanism on a 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 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 what 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 to their assignments. However, the benefit of being assigned through the new mechanism is at least eight times the cost of additional travel, and is often larger depending on the assumption about the information revealed about unobserved tastes from rank order lists submitted under the uncoordinated mechanism. The gains are positive on average for students from all boroughs, demographic groups, and baseline achievement categories. Welfare improvements are also seen 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 the preferences submitted in both the uncoordinated and coordinated mechanism.

These gains are measured by a rich specification of student demand that implies significant estimated heterogeneity in willingness to travel for school. Preference heterogeneity is important for measuring the allocative effects of choice when there is a shortage of good schools. Our estimates reveal that the benefits from having coordinated choice are much larger than than those associated with modifications to the assignment algorithm within the coordinated mechanism. This does not imply that the design of the mechanism 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 calls to unify enrollment across school sectors (Vaznis 2013, Fox 2015). The relative value of policies such as common timelines, common applications, single vs. multiple offers, sophisticated matching algorithms, and good information and decision aides is an interesting avenue for future research.

Finally, it is worth emphasizing that our analysis has focused on the allocative aspects of school choice and different school assignment procedures. An important question is whether allocative changes contribute to changes in the productive dimensions of assignment. Recent work on school value added models (Deming 2014, Angrist, Hull, Pathak, and Walters 2015) may provide a route to measuring how achievement would change under the coordinated mechanism. Extrapolation based on these methods may require restrictive assumptions about heterogeneous effects. We'd anticipate achievement effects to be driven by the reduction in the number of students assigned to poorly performing large zoned schools in the coordinated mechanism. A more comprehensive examination of these issues is left for future work.

Table 1. Characteristics of Student Sample

	Mechanism Comparison		Demand Estimation
	Uncoordinated Mechanism	Coordinated Mechanism	Coordinated Mechanism
	(1)	(2)	(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: Means unless otherwise noted. Uncoordinated mechanism refers to 2002-03 mechanism and coordinated mechanism refers to the 2003-04 mechanism based on deferred acceptance. Neighborhood income is the median census block group family income from the 2000 census - table reports the mean neighborhood income across students. SHSAT stands for Specialized High School Achievement Test.

Table 2. Descriptive Statistics for Schools and Programs

	Uncoordinated Mechanism (1)	Coordinated Mechanism (2)
	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
Fraction Inexperienced Teachers	54.7	55.6
Attendance Rate (out of 100)	85.5	85.7
Percent Subsidized Lunch	62.5	62.6
Size of 9th 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
	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 2002-03 mechanism and coordinated mechanism refers to the 2003-04 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 student that scored more than 85% on the Math A and English Regents tests in New York State Report Cards, respectively. Inexperienced teachers are those that have taught for less than two years.

Table 3. Offer Processing across Mechanisms

	Number of Students	Distance to School (in miles)		Exit from NYC Public Schools	Enrolled in School Other than Assigned
		Assignment	Enrollment		
	(1)	(2)	(3)	(4)	(5)
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%
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%
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 2002-03 mechanism and coordinated mechanism refers to the 2003-04 mechanism based on deferred acceptance. Student distance calculated as road distance using ArcGIS. Assignment is the school assigned at the conclusion of the high school assignment process. Enrollment is the school a student enrolls in October following application. Assigned student exits 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 NYC Public, 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 first made.

Table 4. Ranked vs. Assigned Schools by Student Assignment Round

	Uncoordinated Mechanism		Coordinated Mechanism	
	Ranked Schools	Assigned	Ranked Schools	Assigned
	(1)	(2)	(3)	(4)
		A. Main Round		
Distance (in 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
Fraction Inexperienced Teachers	45.3	45.6	46.6	43.8
Attendance Rate (out of 100)	85.1	84.6	85.7	83.8
Percent Subsidized Lunch	60.0	60.5	57.6	56.7
Size of 9th grade	694.3	698.8	675.0	819.2
Percent White	15.1	14.7	16.7	17.8
		B. Supplementary Round		
Distance (in 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
Fraction Inexperienced Teachers	46.0	41.5	45.3	36.9
Attendance Rate (out of 100)	85.1	82.2	87.4	83.2
Percent Subsidized Lunch	62.0	61.8	53.5	51.0
Size of 9th grade	685.3	908.0	638.5	1129.7
Percent White	13.8	13.3	17.4	15.3
		C. Administrative Round		
Distance (in 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
Fraction Inexperienced Teachers	41.9	39.4	47.8	42.1
Attendance Rate (out of 100)	85.8	80.8	86.7	82.9
Percent Subsidized Lunch	53.8	50.4	57.2	53.1
Size of 9th grade	760.6	1181.9	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 with observed assignments. Uncoordinated mechanism refers to the 2002-03 mechanism and coordinated mechanism refers to the 2003-04 mechanism based on deferred acceptance. Main round in the uncoordinated mechanism corresponds to the first round. Rankings used are those submitted in the main round of the process. Student distance calculated as road distance using ArcGIS. See Table 2 notes for details on school characteristics.

Table 5. Offer Processing by Student Type

	Uncoordinated Mechanism			Coordinated Mechanism		
	Main Round	Supplementary	Administrative	Main	Supplementary	Administrative
	(1)	(2)	(3)	(4)	(5)	(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: Uncoordinated mechanism refers to 2002-03 mechanism and coordinated mechanism refers to the 2003-04 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 family income from the 2000 census.

Table 6. School Characteristics by Rank of Student Choice

Choice	Mechanism	1st	2nd	3rd	4th	5th	6th	9th	12th
					A. All Students				
Students Ranking Choice	Coordinated	69,907	93.4%	88.7%	82.8%	76.2%	69.1%	49.7%	20.3%
	Uncoordinated	59,277	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	5.65	5.12
	Uncoordinated	4.80	4.91	4.94	4.88	4.79			
Median	Coordinated	3.51	3.95	4.20	4.37	4.57	4.63	4.78	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	11.5	10.4
	Uncoordinated	14.1	13.3	12.8	12.1	11.7			
Fraction Subsidized Lunch	Coordinated	51.4	53.4	54.5	56.2	57.4	58.7	61.3	63.1
	Uncoordinated	56.6	58.0	59.1	60.7	62.0			
Size of 9th Grade	Coordinated	713.4	708.1	689.3	668.0	655.3	635.9	608.8	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	10.4	9.0
	Uncoordinated	14.6	13.4	12.5	11.4	10.8			
High Math Achievement					*B. Student Subgroups*				
Students with Low Baseline Math	Coordinated	10.9	10.9	10.5	10.1	10.0	9.7	9.4	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	15.2	12.8
	Uncoordinated	21.5	19.0	17.8	16.9	16.1			
Neighborhood Income									
Students from Bottom Neighorhood	Coordinated	11.4	10.9	10.5	10.4	10.1	9.9	9.6	8.7
Income Quartile	Uncoordinated	9.5	9.6	9.5	9.1	8.7			
Students from Top Neighorhood	Coordinated	23.3	20.7	19.6	18.7	17.7	16.8	15.0	12.7
Income Quartile	Uncoordinated	21.4	18.5	17.6	16.5	16.1			

Notes: Uncoordinated mechanism refers to 2002-03 mechanism and coordinated mechanism refers to the 2003-04 mechanism based on deferred acceptance. Student distance calculated as road distance using ArcGIS. High Math achievement is the fraction of students scoring over 85% on the Math A regents in New York State Report Card. High baseline math students score above the 75th percentile for 7th grade middle school math, low baseline math students score below the 25th percentile. Neighborhood income is median family income from the 2000 census.

Table 7. Selected Preference Estimates for Different Demand Specifications

Specifications:		No Student Interactions (1)	Without Random Coefficients (2)	School Characteristics x Student Characteristics / Models with Random Coefficients: All Choices (3)	Top Three Choices (4)	All except last choice (5)	Students that rank less than 12 (6)
High Math Achievement							
Main effect		0.061***	0.048***	-0.029	-0.029	-0.013	-0.014
Baseline Math			0.028***	0.039***	0.062***	0.042***	0.036***
Percent Subsidized Lunch							
Main effect		-0.004	-0.017**	-0.069***	-0.014	-0.011	-0.047***
Size of 9th Grade (in 100s)							
Main effect		-0.029	0.044	-0.113**	0.175	0.300***	0.001
Percent White							
Main effect		0.071***	0.115***	0.062***	0.177***	0.139***	0.119***
Asian			-0.049***	-0.075***	-0.135***	-0.083***	-0.100***
Black			-0.090***	-0.124***	-0.233***	-0.155***	-0.169***
Hispanic			-0.041***	-0.084***	-0.133***	-0.097***	-0.114***
Standard Deviation of ε		7.291***	7.473***	7.858***	9.753***	8.603***	8.414***
Standard Deviation of ξ		3.207***	2.783***	3.676***	4.889***	3.729***	3.679***
Random Coefficients (Covariances)							
Size of 9th Grade (in 100s)	Size of 9th Grade (in 100s)			1.584***	14.552***	11.379***	14.210***
Size of 9th Grade (in 100s)	Percent White			-0.006***	-0.009	-0.007	-0.006
Size of 9th Grade (in 100s)	Percent Subsidized Lunch			-0.002***	-0.019***	-0.008**	-0.011**
Size of 9th Grade (in 100s)	High Math Achievement			-0.011***	-0.021*	-0.012*	-0.009
Percent White	Percent White			0.008***	0.026***	0.015***	0.017***
Percent White	Percent Subsidized Lunch			-0.001***	0.001**	0.000	0.000
Percent White	High Math Achievement			0.005***	0.006***	0.004***	0.004***
Percent Subsidized Lunch	Percent Subsidized Lunch			0.002***	0.015***	0.007***	0.008***
Percent Subsidized Lunch	High Math Achievement			-0.000**	0.002**	0.000	0.000
High Math Achievement	High Math Achievement			0.016***	0.044***	0.024***	0.025***
Number of Students		69,907	69,907	69,907	12,007	65,310	55,695
Number of Ranks		542,666	542,666	542,666	23,545	472,759	372,122

Notes: Select estimates of demand system with submitted ranks over 497 program choices in 235 schools. Student distance calculated as road distance using ArcGIS. Dummies for missing school attributes are estimated with separate coefficients. Estimates use all submitted ranks except in columns 4-6. Column 1 contains no interactions between student and school characteristics. Column 2 contains interactions of baseline achievement, gender, race, special education, limited English proficiency, subsidized lunch, and median 2000 census block group family income with school characteristics. Columns 3-6 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 student that scored more than 85% on the Math A in New York State Report Cards. Models estimate the utility differences amongst inside options only, with an arbitrarily chosen school's mean utility normalized to zero (without loss of generality). * significant at 10%; ** significant at 5%; *** significant at 1%

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

		School Choice		
Assignment Mechanism:	Neighborhood Assignment	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 using actual preferences 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-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 Abdulkadiroglu-Sonmez (2003) version of top trading cycles with counters.

Table 9. Welfare Comparison between Coordinated and Uncoordinated Mechanism

	Change in Utility (in miles)		Change in Distance (in miles)		2002-03 Offer Processing		
	Assignment	Enrollment	Assignment	Enrollment	Main Rounds	Supplementary	Administrative
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
All Students	10.96	9.67	0.69	0.38	48.5%	14.5%	37.1%
Female	10.41	9.08	0.68	0.37	51.0%	14.3%	34.7%
Asian	12.40	10.68	0.63	0.33	46.1%	5.4%	48.5%
Black	9.27	8.20	0.74	0.45	53.2%	18.4%	28.4%
Hispanic	10.56	9.53	0.66	0.39	51.2%	17.3%	31.5%
White	15.76	13.47	0.56	0.22	31.5%	3.8%	64.7%
Bronx	9.82	8.96	0.93	0.64	53.3%	20.2%	26.5%
Brooklyn	10.86	9.85	0.52	0.33	49.8%	16.2%	33.9%
Manhattan	11.62	9.20	0.00	-0.09	43.1%	8.3%	48.6%
Queens	5.80	5.14	1.13	0.57	66.8%	19.2%	14.0%
Staten Island	22.56	22.38	0.34	0.00	11.9%	0.0%	88.1%
High Baseline Math	9.52	7.66	0.53	0.20	57.3%	7.4%	35.3%
Low Baseline Math	10.80	10.07	0.57	0.33	46.8%	19.8%	33.4%
Subsidized Lunch	10.49	9.35	0.65	0.38	51.8%	15.9%	32.3%
Bottom Neighborhood Income Quartile	9.17	8.61	0.57	0.42	55.4%	23.3%	21.3%
Top Neighborhood Income Quartile	12.34	10.15	0.71	0.25	41.3%	8.1%	50.6%
Special Education	10.43	9.32	0.76	0.43	38.9%	18.8%	42.3%
Limited English Proficient	11.88	10.89	0.60	0.38	46.9%	16.3%	36.8%
SHSAT Test-Takers	7.55	6.24	0.55	0.25	61.9%	10.3%	27.8%

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 estimated from unordered application selection from the uncoordinated mechanism and the truthful application selection rule from the coordinated mechanism. Assignment is the school assigned at the conclusion of the high school assignment process. Enrollment is the school 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. 2002-03 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 75th percentile for 7th grade relative to citywide distribution, while low baseline math students score below the 25th percentile. Subsidized lunch not available pre-assignment and comes from enrolled students as of 2004-05 school year. Neighborhood income is median census block group family income from the 2000 census.

Table 10. Welfare Comparison for Alternative Selection Rules

Selection assumption:	Unselected Applications		Unordered Applications		Truthful Applications	
	Assignment	Enrollment	Assignment	Enrollment	Assignment	Enrollment
	(1)	(2)	(3)	(4)	(5)	(6)
All Students	4.99	4.67	9.02	7.66	10.62	9.25
Female	4.89	4.52	8.37	6.99	10.01	8.59
Asian	6.52	5.83	10.73	8.99	11.91	10.11
Black	3.90	3.60	7.17	6.03	8.97	7.83
Hispanic	4.61	4.29	8.47	7.37	10.21	9.11
White	8.66	8.03	14.52	12.09	15.46	13.03
Bronx	4.28	4.00	7.46	6.46	9.46	8.54
Brooklyn	4.66	4.35	8.84	7.77	10.57	9.49
Manhattan	2.46	2.10	3.14	2.30	5.10	4.33
Queens	5.10	4.43	10.27	7.84	11.30	8.78
Staten Island	14.23	14.77	21.68	21.47	22.62	22.42
High Baseline Math	5.60	4.90	7.30	5.41	8.85	6.85
Low Baseline Math	4.45	4.38	8.95	8.14	10.61	9.83
Subsidized Lunch	4.76	4.44	8.40	7.21	10.13	8.93
Bottom Neighborhood Income Quartile	3.88	3.73	6.83	6.15	8.81	8.18
Top Neighborhood Income Quartile	6.34	5.70	10.86	8.58	12.02	9.71
Special Education	4.02	4.10	9.08	7.66	10.30	9.10
Limited English Proficient	4.99	4.72	9.84	8.80	11.62	10.58
SHSAT Test-Takers	4.38	3.81	5.45	4.11	6.91	5.51

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. Assignment is the school assigned at the conclusion of the high school assignment process. Enrollment is the school student enrolls in October following application. Results from the unselected applications selection rule in columns 1 and 2 do not include the idiosyncratic taste shock in utility calculations for both mechanisms. Results from unordered application selection rule in columns 3 and 4 compute utility for an assignment conditional on the schools listed on a rank order list being preferred to those not listed for both mechanisms. Results from the truthful application selection rule in columns 5 and 6 compute utility for an assignment conditional on a student's rank order list assuming its truthful for both mechanisms. High baseline math students score above the 75th percentile for 7th grade relative to citywide distribution, while low baseline math students score below the 25th percentile. Neighborhood income is median census block group family income from the 2000 census.

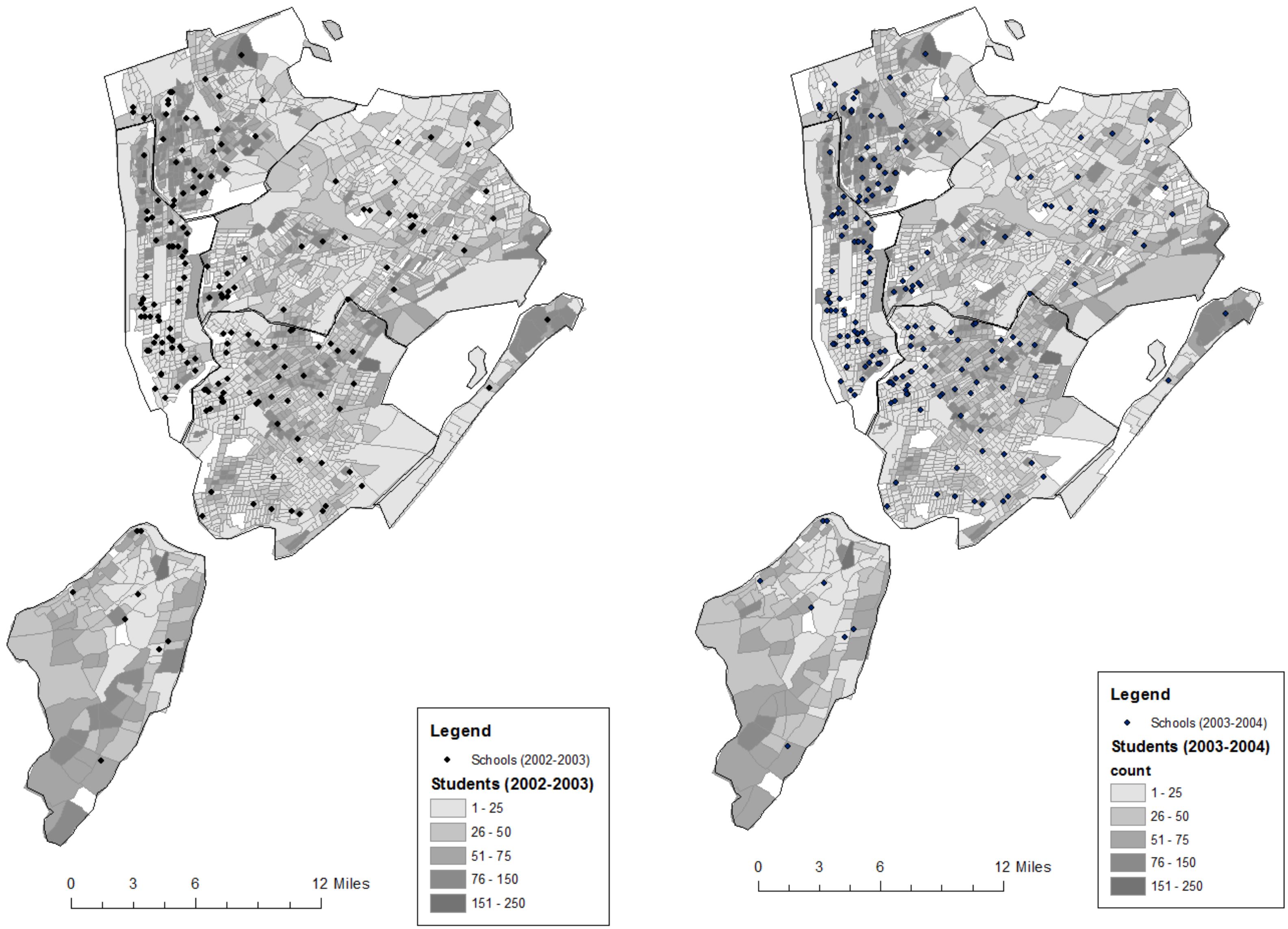

Figure 1. School Locations and Students by New York City Census Tract in 2002-03 and 2003-04

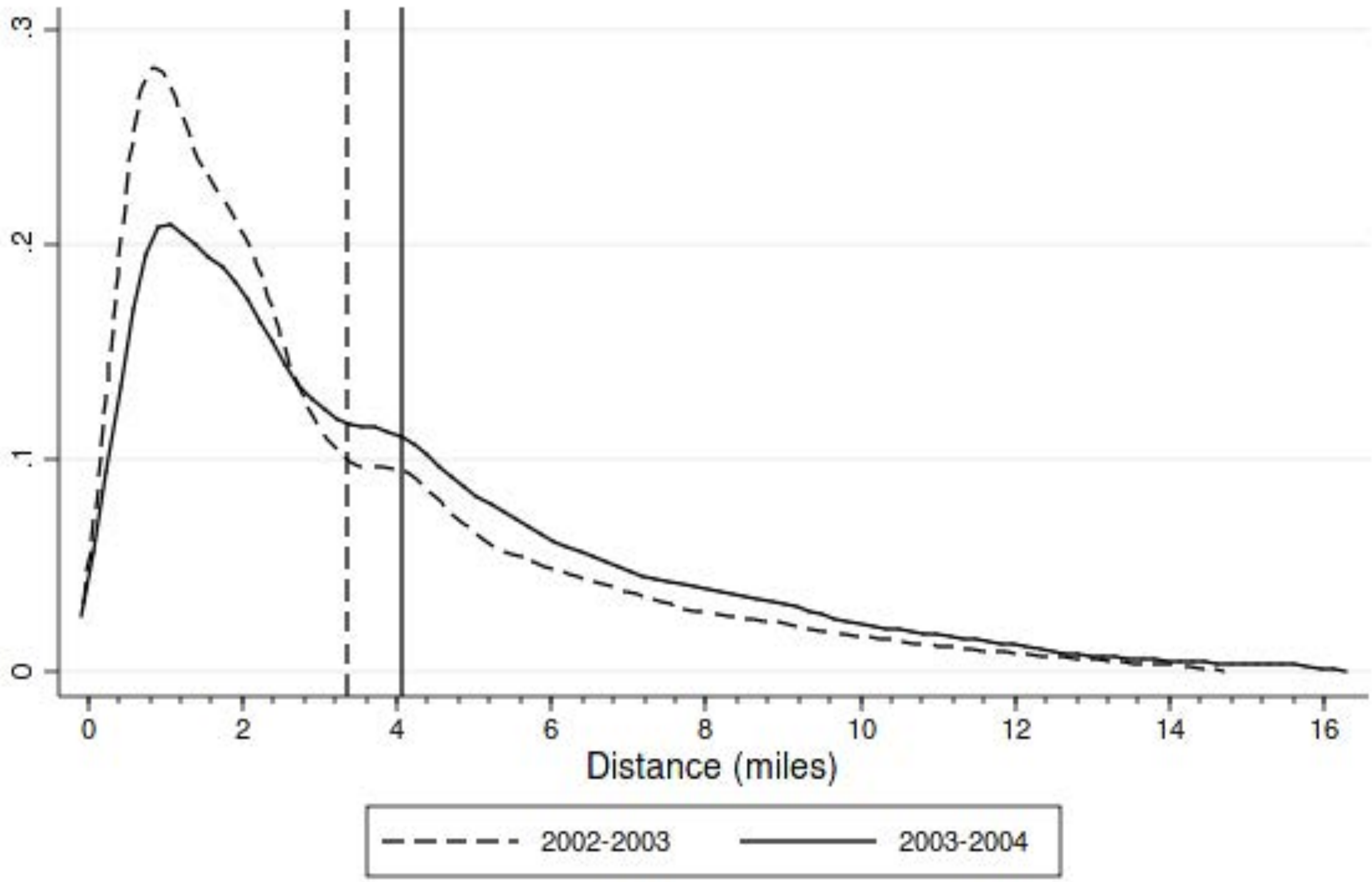

Figure 2. Distribution of Distance to Assigned School in Uncoordinated (2002-03) and Coordinated (2003-04) Mechanism

Mean (median) travel distance is 3.36 (2.25) miles in 2002-03 and 4.05 (3.04) miles in 2003-04. Top and bottom 1% are not shown in figure. Line fit from Gaussian kernel with bandwidth chosen to minimize mean integrated squared error using STATA's kdensity command.

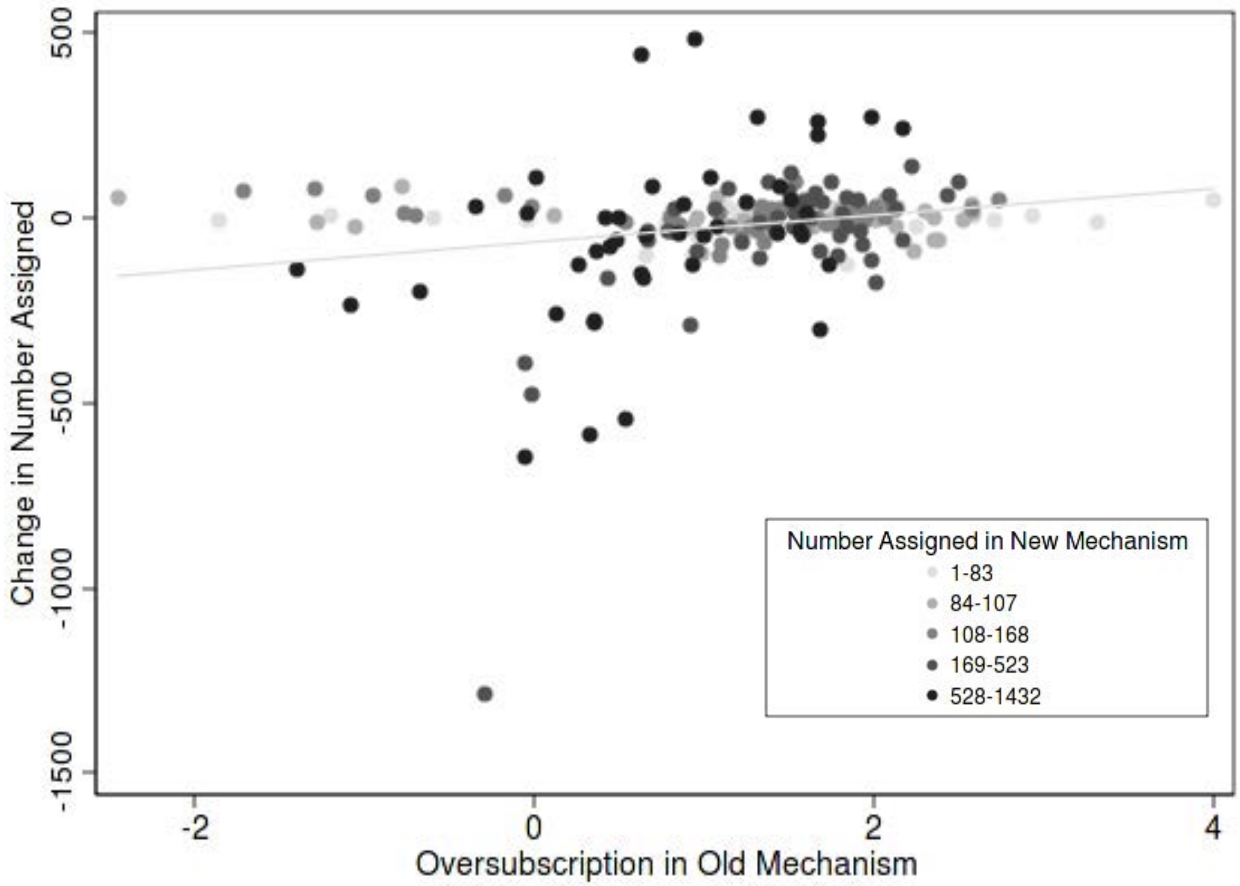

Figure 3. Change in Number Assigned by Oversubscription in Uncoordinated Mechanism

The figure the change in the number 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.

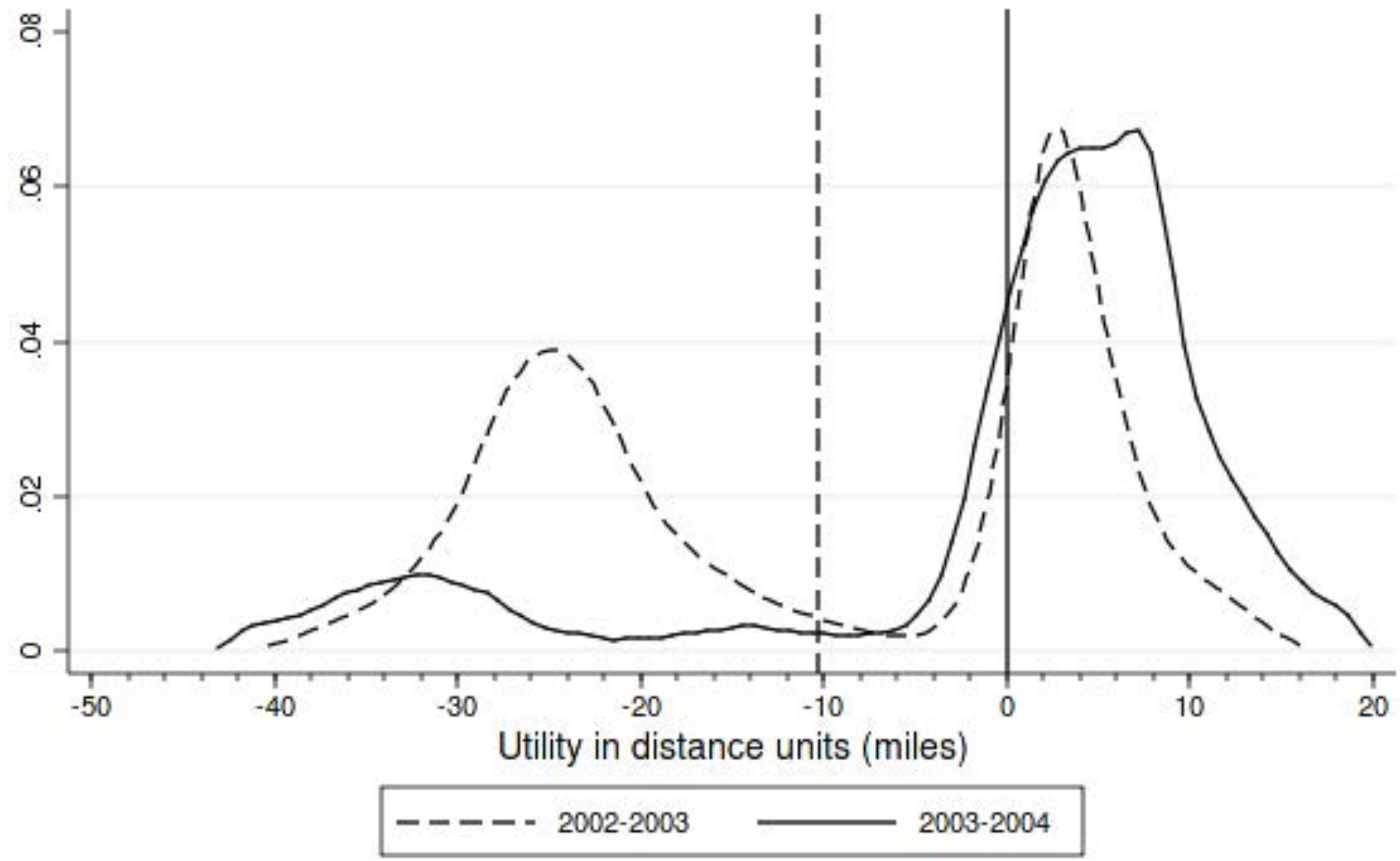

Figure 4. Student Welfare from Uncoordinated and Coordinated Mechanism

Distribution of utility (measured in distance units) from assignment based estimates in column 3 of Table A1 with mean utility in 2003-04 normalized to 0. Top and bottom 1% are not shown in figure. Line fit from Gaussian kernel with bandwidth chosen to minimize mean integrated squared error using STATA's kdensity command.

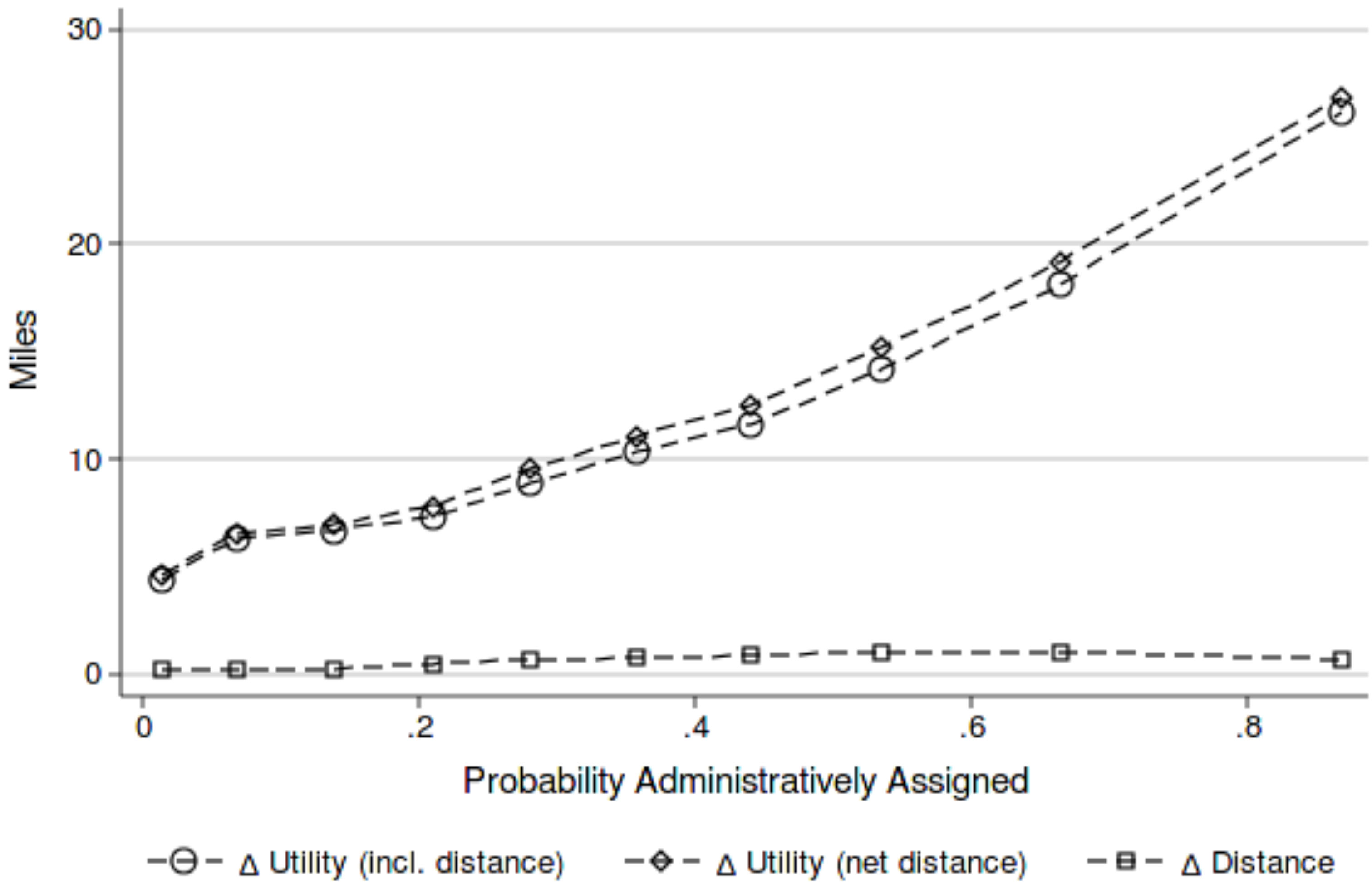

Figure 5. Change in Student Welfare by Propensity to be Administratively Assigned in the Uncoordinated Mechanism

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. If student lives in tract where either all students are administratively or no students are administratively assigned, all students from those tracts are coded as administratively assigned. Each student is assigned to one of ten deciles of probability of administrative assignment based on these estimates. Differences across deciles in distance-equivalent utility including distance, distance-equivalent utility net of distance, and distance are plotted, where preference estimates come from column 3 of Table 7, under the selection assumption in Table 9.

References

Abdulkadiroğlu, A., J. Angrist, and P. Pathak (2014): "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools," Econometrica, 82(1), 137–196.

Abdulkadiroğlu, A., J. D. Angrist, S. M. Dynarski, T. J. Kane, and P. 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, A., Y.-K. Che, P. Pathak, A. Roth, and O. Tercieux (2015): "Designing the New Orleans School Match: TTC vs. DA," Working Paper, MIT.

Abdulkadiroğlu, A., Y.-K. Che, and Y. Yasuda (2015): "Expanding "Choice" in School Choice," American Economic Journal: Microeconomics, 7(1), 1–42.

Abdulkadiroğlu, A., P. A. Pathak, and A. E. Roth (2005): "The New York City High School Match," American Economic Review, Papers and Proceedings, 95(2), 364–367.

——— (2009): "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the New York City High School Match," American Economic Review, 99(5), 1954–1978.

Abdulkadiroğlu, A., P. A. Pathak, A. E. Roth, and T. Sönmez (2006): "Changing the Boston Public School Mechanism: Strategy-proofness as Equal Access," NBER Working paper, 11965.

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

Abulkadiroğlu, A., W. Hu, and P. A. Pathak (2013): "Small High Schools and Student Achievement: Lottery-Based Evidence from New York City," NBER Working Paper, 19576.

Agarwal, N., and P. Somaini (2014): "Demand Analysis using Strategic Reports: An Application to a School Choice Mechanism," NBER Working Paper, 20775.

Ajayi, K. (2013): "School Choice and Educational Mobility: Lessons from Secondary School Applications in Ghana," Unpublished working paper, Boston University.

Angrist, J. D., P. Hull, P. Pathak, and C. Walters (2015): "Leveraging Lotteries for Value-Added: Testing and Estimation," Working paper, MIT.

Arvidsson, T. S., N. Fruchter, and C. Mokhtar (2013): "Over the Counter, Under the Radar: Inequitably Distributing New York City's Late-Enrolling High School Students," Annenberg Institute for School Reform at Brown University, Available at: http://annenberginstitute.org/publication/over-counter-under-radar-inequitably-distributing-new-york-city%E2%80%99s-late-enrolling-high-sc, Last accessed: October 2013.

Azevedo, E. M., and J. D. Leshno (2011): "Can we make school choice more efficient? An example," Unpublished mimeo, Harvard University.

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

Berry, S., J. Levinsohn, and A. 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, A., and H. Moulin (2001): "A New Solution to the Random Assignment Problem," Journal of Economic Theory, 100, 295–328.

Briesch, R., P. Chintagunta, and R. Matzkin (2002): "Semiparametric Estimation of Brand Choice Behavior," Journal of the American Statistical Association, 97, 973–982.

Budish, E., and E. Cantillon (2012): "The Multi-Unit Assignment Problem: Theory and Evidence from Course Allocation at Harvard," American Economic Review, 102(5), 2237–2271.

Burgess, S., E. Greaves, A. Vignoles, and D. Wilson (2015): "What Parents Want: School Preferences and School Choice," Economic Journal, 125(587), 1262–1289.

Calsamiglia, C., C. Fu, and M. Guell (2014): "Structural Estimation of a Model of School Choices: the Boston Mechanism vs. Its Alternatives," Unpublished paper, UAB.

Calsamiglia, C., G. Haeringer, and F. Kljin (2010): "Constrained School Choice: An Experimental Study," American Economic Review, 100(4), 1860–1874.

Deming, D. (2014): "Using School Choice Lotteries to Test Measures of School Effectiveness," American Economic Review, Papers and Proceedings, 104(5), 406–411.

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

DOE (2002): "Directory of the New York City Public High Schools: 2002-2003," New York City Department of Education.

Epple, D., and R. Romano (1998): "Competition Between Private and Public Schools, Vouchers, and Peer Group Effects," American Economic Review, 88(1), 33–62.

Erdil, A., and H. Ergin (2008): "What's the Matter with Tie-Breaking? Improving Efficiency in School Choice," American Economic Review, 98(3), 669–689.

Fox, J. C. (2015): "Plan would simplify Boston public school enrollments," Boston Globe, Metro Section, September 17.

Gelman, A., and D. Rubin (1992): "Inference from Iterative Simulation Using Multiple Sequences," Statistical Science, 7(4), 457–472.

Gendar, A. (2000): "Simpler HS Entry Eyed Ed Board Aims to Trim Selection Process, Improve Schools," New York Daily News, October 21.

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

Hastings, J., T. J. Kane, and D. O. Staiger (2009): "Heterogenous Preferences and the Efficacy of Public School Choice," Working paper, Yale University.

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

He, Y. (2012): "Gaming the Boston Mechanism in Beijing," Unpublished mimeo, Toulouse.

Hemphill, C. (2007): New York City's Best Public High Schools: A Parent's Guide, Third Edition. Teachers College Press.

Hemphill, C., and K. Nauer (2009): "The New Marketplace: How Small-School Reforms and School Choice have Reshaped New York City's High Schools," Center for New York City Affairs, Milano the New School for Management and Urban Policy.

Herszenhorn, D. M. (2004): "Council Members See Flaws in School-Admissions Plan," New York Times, Metropolitan Desk, November 19.

Hoxby, C. (2003): "School Choice and School Productivity (Or, Could School Choice be a Rising Tide that Lifts All Boats)," in The Economics of School Choice, ed. by C. Hoxby. Chicago: University of Chicago Press.

Hwang, S. I. M. (2014): "A Robust Redesign of High School Match," University of Chicago, Working Paper.

Ichimura, H., and T. S. Thompson (1998): "Maximum Likelihood Estimation of a Binary Choice Model with Random Coefficients of Unknown Distribution," Journal of Econometrics, 86(2), 269–295.

Jacob, B., and L. Lefgren (2007): "What Do Parents Value in Education? An Empirical Investigation of Parents' Revealed Preferences for Teachers," Quarterly Journal of Economics, 122(4), 1603–1637.

Kerr, P. (2003): "Parents Perplexed by High School Admissions Rule: Reply," New York Times, November 3.

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

Kesten, O., and M. Kurino (2012): "On the (im)possibility of improving upon the student-proposing deferred acceptance mechanism," WZB Discussion Paper SP II 2012-202.

Klemperer, P. (2002): "What Really Matters in Auction Design," Journal of Economic Perspectives, 16(1), 169–189.

Kojima, F., and P. A. Pathak (2009): "Incentives and Stability in Large Two-Sided Matching Markets," American Economic Review, 99(3), 608–627.

Lewbel, A. (2000): "Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables," Journal of Econometrics, 97(1), 145–177.

Neilson, C. (2014): "Targeted Vouchers, Competition among Schools and the Academic Achievement of Poor Students," Unpublished working paper, Yale University.

Niederle, M., and A. E. Roth (2003): "Unraveling reduced mobility in a labor market: Gastroenterology with and without a centralized match," Journal of Political Economy, 111(6), 1342–1352.

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

Pathak, P. A., and P. Shi (2014): "Demand Modeling, Forecasting, and Counterfactuals, Part I," NBER Working Paper, 19859.

Pycia, M. (2014): "The Cost of Ordinality," Unpublished working paper, UCLA.

Rossi, P. E., G. M. Allenby, and R. McCulloch (2005): Bayesian Statistics and Marketing. Wiley.

Rossi, P. E., R. E. McCulloch, and G. M. Allenby (1996): "The Value of Purchase History Data in Target Marketing," Marketing Science, 15(4), 321–340.

Roth, A. E. (2015): Who Gets What — And Why: The New Economics of Matchmaking and Market Design. Eamon Dolan.

Roth, A. E., and X. 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, J. (2006): "Good Principals or Good Peers: Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition Among Jurisdictions," American Economic Review, 96(4), 1333–1350.

Shapley, L., and H. Scarf (1974): "On Cores and Indivisibility," Journal of Mathematical Economics, 1, 23–28.

Shapley, L., and M. Shubik (1971): "The Assignment Game I: The core," International Journal of Game Theory, 1(1), 111–130.

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

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

Train, K. (2009): Discrete Choice Methods with Simulation. Cambridge University Press.

Tullis, T. (2014): "How Game Theory Helped Improve New York City's High School Application Process," NY Times, December 5.

Urquiola, M. (2005): "Does school choice lead to sorting? Evidence from Tiebout variation," American Economic Review, 95(4), 1310–1326.

van der Vaart, A. W. (2000): Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics.

Vaznis, J. (2013): "Barros Proposes Single Application for City, Charter Schools," *Boston Globe*, Education Desk, June 5.

Walters, C. (2014): "The Demand for Effective Charter Schools," NBER Working Paper 20640.

Table A1. Posterior Means of Preference Estimates for Different Demand Specifications

	No Student Interactions	School Characteristics x Student Characteristics				
		Without Random Coefficients	Models with Random Coefficients			
Specifications:			All Choices	Top Three Choices	All except last choice	Students that rank less than 12
	(1)	(2)	(3)	(4)	(5)	(6)
High Math Achievement						
Main effect	0.061***	0.048***	-0.029	-0.029	-0.013	-0.014
Baseline Math		0.028***	0.039***	0.062***	0.042***	0.036***
Baseline English		0.031***	0.039***	0.071***	0.053***	0.059***
Subsidized Lunch		-0.011***	-0.016***	-0.022**	-0.021***	-0.026***
Neighborhood Income (in 1000s)		0.004***	0.012***	0.017***	0.014***	0.012***
Limited English Proficient		0.014**	0.000	0.032*	0.026**	-0.021
Special Education		0.009	-0.006	0.039*	0.020	0.010
Percent Subsidized Lunch						
Main effect	-0.004	-0.017**	-0.069***	-0.014	-0.011	-0.047***
Asian		-0.009	-0.012***	-0.024*	-0.012	0.008
Black		0.007	0.009***	0.005	0.016***	0.017**
Hispanic		0.035***	0.043***	0.084***	0.059***	0.064***
Subsidized Lunch		0.006***	0.011***	0.024***	0.015***	0.015***
Neighborhood Income (in 1000s)		-0.005***	-0.008***	-0.019***	-0.013***	-0.011***
Size of 9th Grade (in 100s)						
Main effect	-0.029	0.044	-0.113**	0.175	0.300***	0.001
Baseline Math		-0.012**	-0.026***	-0.007	-0.024	0.042
Baseline English		-0.050***	-0.066***	-0.226***	-0.104*	-0.083
Subsidized Lunch		0.014	0.038***	0.015	0.059	0.102
Neighborhood Income (in 1000s)		-0.011***	-0.012***	-0.010	-0.031*	-0.009
Special Education		0.019	0.052**	0.198	0.091	0.155
Percent White						
Main effect	0.071***	0.115***	0.062***	0.177***	0.139***	0.119***
Asian		-0.049***	-0.075***	-0.135***	-0.083***	-0.100***
Black		-0.090***	-0.124***	-0.233***	-0.155***	-0.169***
Hispanic		-0.041***	-0.084***	-0.133***	-0.097***	-0.114***
Spanish Language Program						
Limited English Proficient		14.281***	15.437***	18.961***	16.887***	16.386***
Limited English Proficient x Hispanic		-9.517***	-10.502***	-18.420***	-11.541***	-12.465***

Asian Language Program							
Limited English Proficient			11.180***	11.814***	16.651***	13.508***	14.139***
Limited English Proficient x Asian			-8.424***	-7.091***	-17.598***	-9.926***	-8.279***
Other Language Program							
Limited English Proficient			6.423***	7.448***	10.023***	7.930***	8.992***
Standard Deviation of ε		7.291***	7.473***	7.858***	9.753***	8.603***	8.414***
Standard Deviation of ξ		3.207***	2.783***	3.676***	4.889***	3.729***	3.679***
Random Coefficients (Covariances)							
Size of 9th Grade (in 100s)	Size of 9th Grade (in 100s)			1.584***	14.552***	11.379***	14.210***
Size of 9th Grade (in 100s)	Percent White			-0.006***	-0.009	-0.007	-0.006
Size of 9th Grade (in 100s)	Percent Subsidized Lunch			-0.002***	-0.019***	-0.008**	-0.011**
Size of 9th Grade (in 100s)	High Math Achievement			-0.011***	-0.021*	-0.012*	-0.009
Percent White	Percent White			0.008***	0.026***	0.015***	0.017***
Percent White	Percent Subsidized Lunch			-0.001***	0.001**	0.000	0.000
Percent White	High Math Achievement			0.005***	0.006***	0.004***	0.004***
Percent Subsidized Lunch	Percent Subsidized Lunch			0.002***	0.015***	0.007***	0.008***
Percent Subsidized Lunch	High Math Achievement			-0.000**	0.002**	0.000	0.000
High Math Achievement	High Math Achievement			0.016***	0.044***	0.024***	0.025***
Number of Students		69,907	69,907	69,907	12,007	65,310	55,695
Number of Ranks		542,666	542,666	542,666	23,545	472,759	372,122

Notes: Estimates of demand system with submitted ranks over 497 program choices in 235 schools. Student distance calculated as road distance using ArcGIS. Dummies for missing school attributes are estimated with separate coefficients. Estimates use all submitted ranks except in columns 4-6. Column 1 contains no interactions between student and school characteristics. Column 2 contains interactions of baseline achievement, gender, race, special education, limited English proficiency, subsidized lunch, and median 2000 census block group family income with school characteristics. Columns 3-6 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 student that scored more than 85% on the Math A in New York State Report Cards. Models estimate the utility differences amongst inside options only, with an arbitrarily chosen school's mean utility normalized to zero (without loss of generality). * significant at 10%; ** significant at 5%; *** significant at 1%

Table A2. Model Fit of Correlation between Choices across Demand Specifications

	Correlation between		Coordinated Mechanism (2003-04)			Coordinated Mechanism (2004-05)		
				Specification			Specification	
				Main	No Student		Main	No Student
School Characteristic	Choice	Choice	Observed	Specification	Interactions	Observed	Specification	Interactions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(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 9th 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 A1) and the specification with no student interactions (shown in column 1 of Table A1).

Table A3. Welfare Comparisons for Alternative Demand Specifications

	Neighborhood Assignment	School Choice		
		Coordinated Mechanism	Student Optimal Matching	Ordinal Pareto Efficient Matching
	(1)	(2)	(3)	(4)
		A. Alternative Behavioral Assumptions		
10% Sample, Top 3 choices	-24.91	-4.65	-4.51	-3.88
10% Sample, Excl. Full Lists	-21.14	-4.13	-4.00	-3.46
10% Sample, Excl. Last Choice	-21.46	-4.13	-4.01	-3.45
		B. Alternative Samples and Demand Model Interactions		
Full Sample, Main specification	-18.96	-3.73	-3.62	-3.11
10% Sample, Main specification	-21.45	-4.09	-3.96	-3.41
10% Sample, No student interactions	-15.48	-3.29	-3.19	-2.75
10% Sample, No random coefficients	-16.17	-3.23	-3.13	-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 ignoring all school-side constraints except capacity. See notes to Table 8 for details on mechanism calculations. 10% sample represents random 10% of sample of applicants to estimate preferences. All mechanism counterfactuals used these estimates for all applicants in the mechanism comparison sample. Top 3 choices refers to estimates that only use the top 3 choices of applicants. Excl. Full lists refers to estimates that only use rankings of students who rank fewer than 12 choices. Excl. 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.

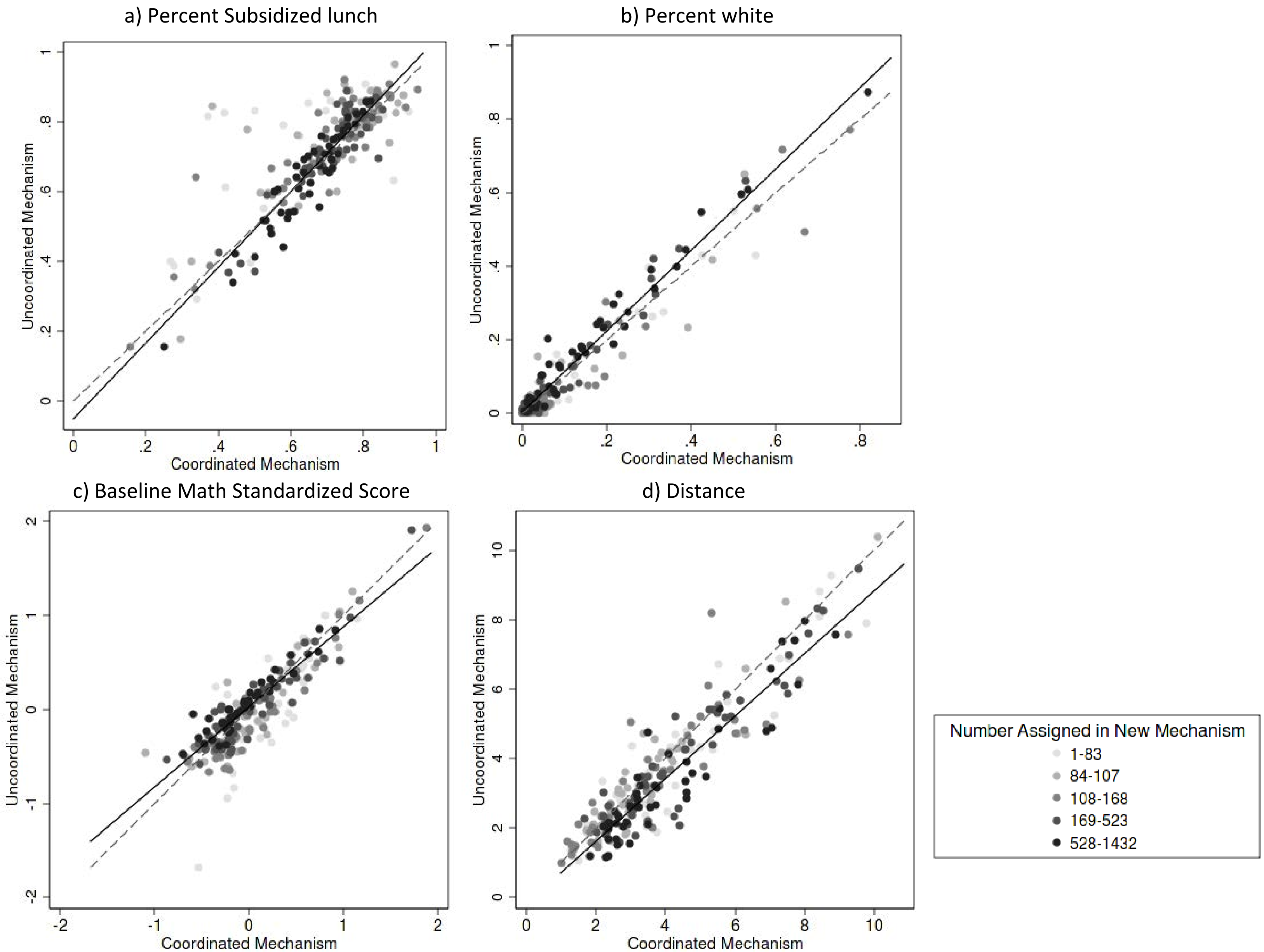

Figure A1. Comparison of Characteristics of Enrolled Students at Each School between Uncoordinated and Coordinated Mechanism

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, while the solid line is the least squares line fit.

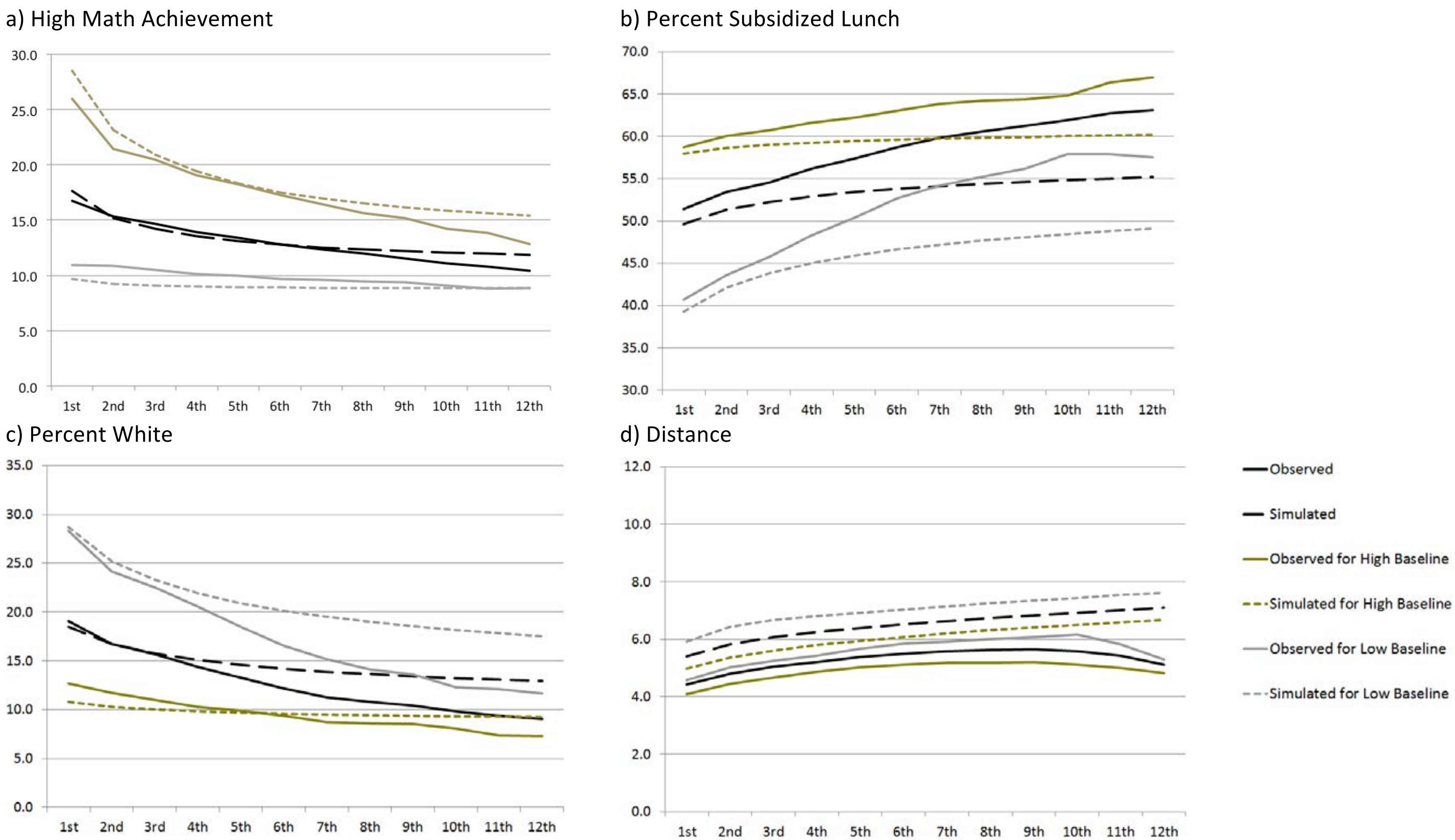

Figure A2. Model Fit

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.

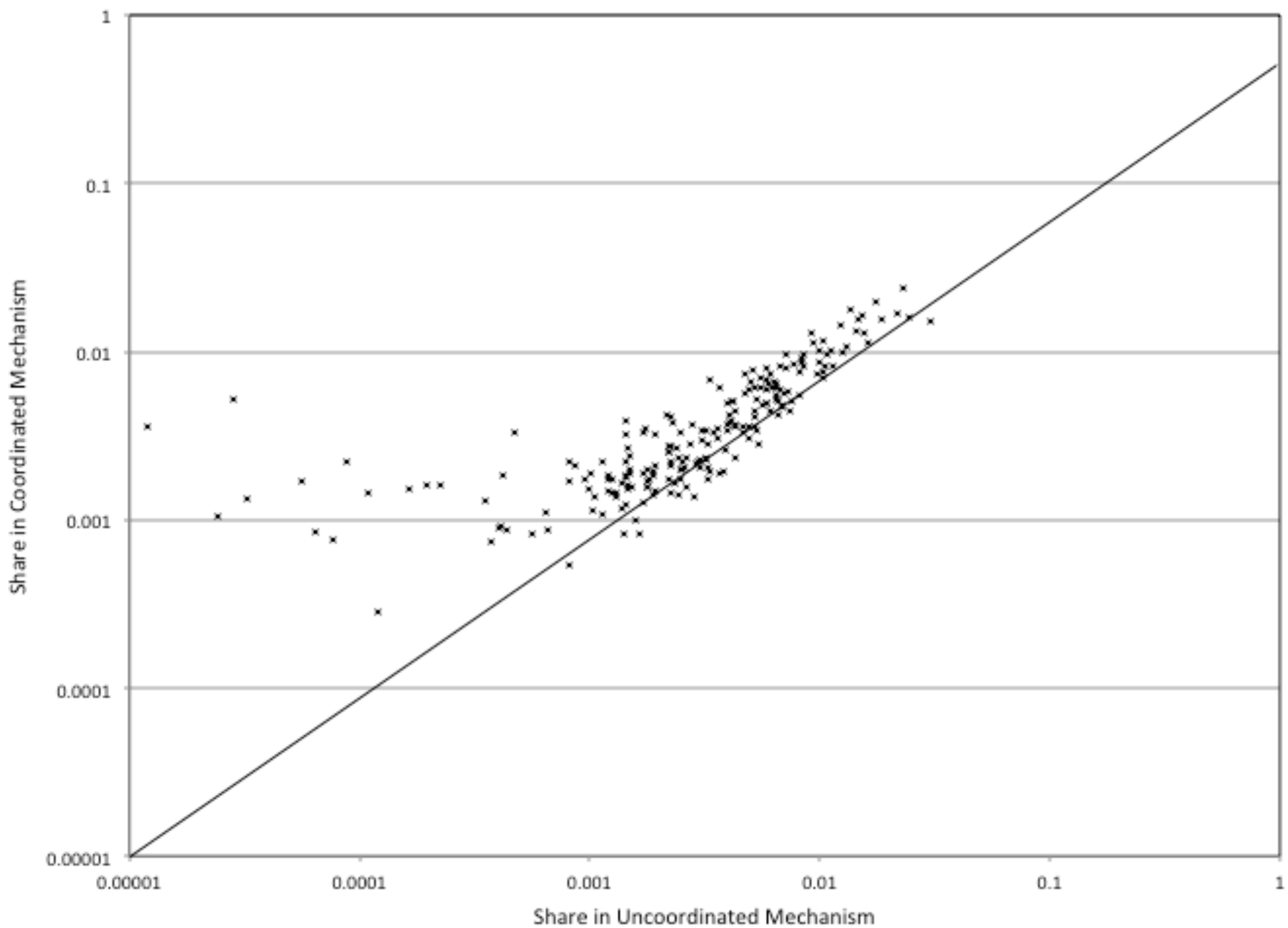

Figure A3. Comparison of School Market Shares between 2002-03 Uncoordinated Mechanism and 2003-04 Coordinated Mechanism

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 that students can apply to in 2002-03 and 2003-04. Market shares are normalized within this set to sum to 1.

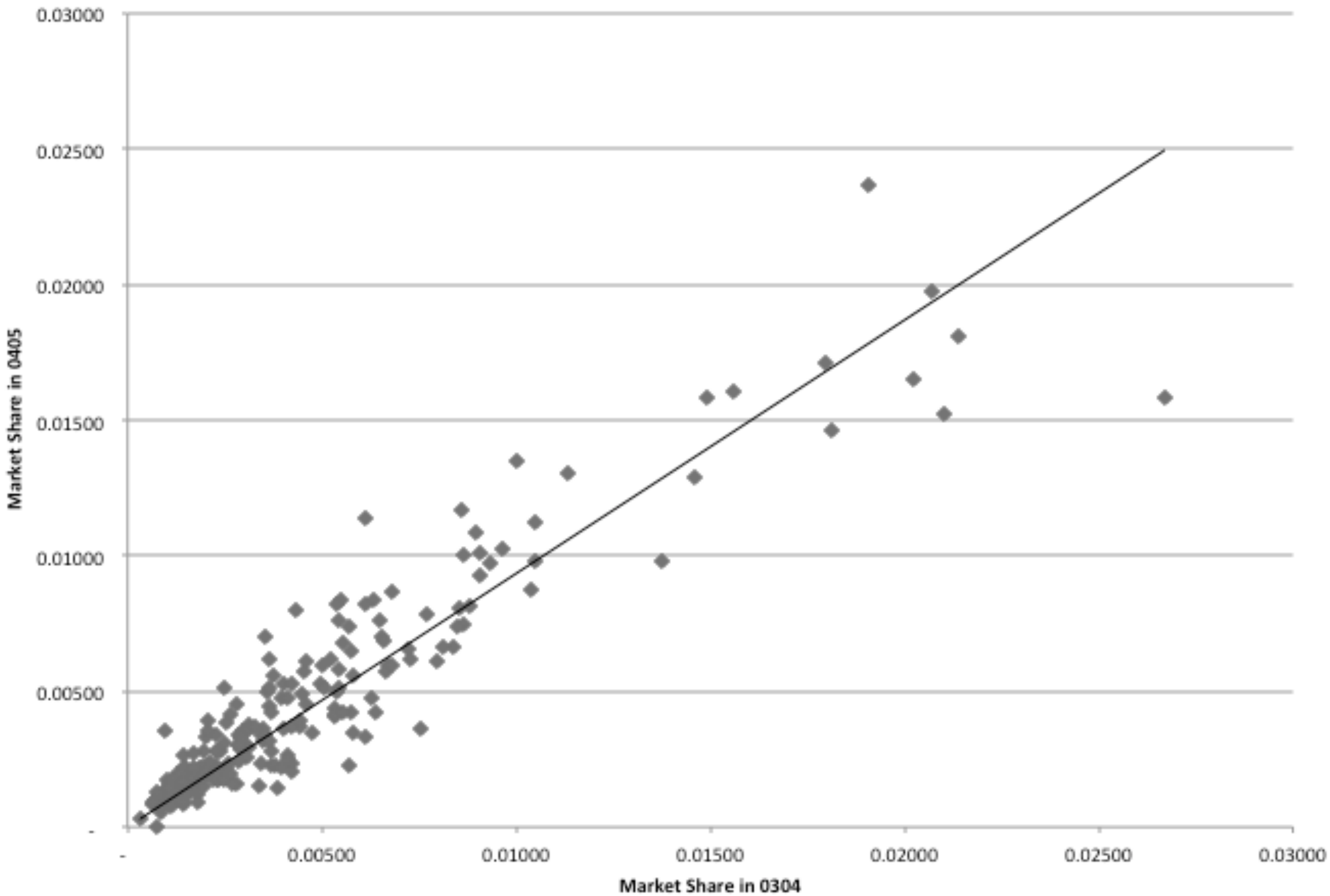

Figure A4. Comparison of School Market Shares between 2003-04 Coordinated Mechanism and 2004-05 Coordinated Mechanism

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 that students can apply to in 2003-04 and 2004-05. Market shares are normalized within this set to sum to 1. The slope of the line fit is 0.93 and the R2 is 0.86.

Online Appendix for
"The Welfare Effects of Coordinated Assignment:
Evidence from the New York City High School Match"

Atila Abdulkadiroğlu Nikhil Agarwal Parag A. Pathak

October 19, 2015

A Computational Appendix (Not for Publication)

The demand model is an ordered version of the model in Rossi, McCulloch, and Allenby (1996). We assume that the utility for student i for program j can be written as:

$$\begin{aligned} u_{ij} &= \delta_j + \sum_l \alpha^l z_i^l x_j^l + \sum_k \gamma_i^k x_j^k - d_{ij} + \varepsilon_{ij}, \\ \text{with} \quad \delta_j &= x_j \beta + \xi_j. \end{aligned}$$

We parametrize the random coefficients as follows:

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

The priors for β, α, Σ_γ, σ_ξ^2, and σ_ε^2 are as follows:

$$\beta \sim \mathcal{N}(0, \bar{\Sigma}_\beta), \qquad \alpha \sim \mathcal{N}(0, \bar{\Sigma}_\alpha)$$

$$\Sigma_\gamma \sim \text{IW}(\bar{\Sigma}_\gamma, \nu_\gamma), \quad \sigma_\xi^2 \sim \text{IW}(\bar{\sigma}_\xi^2, \nu_\xi), \quad \text{and} \quad \sigma_\varepsilon^2 \sim \text{IW}(\bar{\sigma}_\varepsilon^2, \nu_\varepsilon),$$

where IW is the inverse Wishart distribution. Following Chapter 5 of Rossi, Allenby, and McCulloch (2005), we set diffuse priors as follows: the prior variances of β and α are 100 times the identity matrix, and

$$(\bar{\Sigma}_\gamma, \nu_\gamma) = ((3 + \dim(\gamma_i)) I_{\dim(\gamma_i)}, 3 + \dim(\gamma_i)),$$

$$(\bar{\sigma}_\xi^2, \nu_\xi) = (1, 2) \quad \text{and} \quad (\bar{\sigma}_\varepsilon^2, \nu_\varepsilon) = (3 + J, 3 + J),$$

where I_k is the identity matrix of dimension k.

The Gibbs sampler iterates through the following steps (where we omit conditioning on the observed data and the priors for notational simplicity). First, we iterate through the observed rank ordered lists to update the values of u_{ij}. We then draw utilities for the unranked options by observing that their indirect utility must be at most the indirect utility of the lowest ranked option. This step can be written as

$$u_{ij} | u_{i-j}, r_i, \beta, \xi, \gamma_i, \alpha,$$

where each simulation is from a (two-sided) truncated normal.

Given the utilities, the posteriors of ξ, β and α are multivariate normal distributions that can be computed as follows:

$$\begin{aligned} \xi &\mid u, \gamma, \beta, \alpha, \sigma_\xi^2, \\ \beta &\mid u, \gamma, \xi, \alpha, \bar{\Sigma}_\beta, \\ \alpha &\mid u, \gamma, \beta, \alpha, \bar{\Sigma}_\alpha, \end{aligned}$$

where u and γ stack the utilities and random coefficients for all students. We then update the student-specific random coefficients:

$$\gamma_i | u_i, \beta, \xi, \alpha, \Sigma_\gamma.$$

The priors and distribution of ε_{ij} imply that a posterior is a multivariate normal distribution for each student. Finally, we sample from the posteriors $\sigma_\varepsilon^2|\varepsilon$, $\sigma_\xi^2|\xi$ and $\Sigma_\gamma|\gamma$, which are given by inverse Wishart distributions.

For the estimates for the Full sample, main specification, we iterate through the Markov Chain 1.25 million times, and discard the first 0.75 million draws as "burn in" to ensure mixing. We diagnosed mixing by examining trace plots and computing the Potential Scale Reduction Factor (PSRF) following Gelman and Rubin (1992). Because of computational constraints in drawing from separate chains, we split the draws after the burn-in period into three equally sized continguous pieces and computed the PSRF using the first and third pieces. The PSRFs for almost all parameters were within 1.1 and were within 1.3 for all parameters. Trace plots for the few parameters with PSRFs higher than 1.1 did not indicate any obvious convergence issues.

Estimates of the 10% samples were computed by iterating through the Markov Chain 1 million times and discarding the first 0.75 million draws. We obtained estimates from three distinct chains initiated from dispersed starting values. We compared variances within each chain and the variance between chains, by computing both within and across chain values of the PSRF. For nearly all parameters, the PSRF is close to one, suggesting that we've reached the target distribution.

Our estimates report the posterior mean and standard deviations. We examined the histograms of the marginal distributions of the posteriors to assess the skew. These histograms indicate that the means, modes and medians of the parameters in the main specification are similar.

B Appendix: Subway Distances (Not for Publication)

In New York, high school students who live within 0.5 miles of a school are not eligible for transportation. If a student lives between 0.5 and 1.5 miles, the Metropolitan Transit Authority provides them with a half-fare student Metrocard that works only for bus transportation. If they reside 1.5 miles or greater, they obtain full-fare transportation with a student Metrocard that works for subways and buses and is issued by the school transportation office.

Since subway is a common mode of transportation in New York City, this appendix assesses how the driving distance measure we utilize in the paper differs from commuting distance using NYC's subway system. Subway distance is defined as the sum of distance on foot to the student's nearest subway station, travel distance on the subway network to a school's nearest subway station, and the distance on foot to the school from that station. To compute these distances, we used ESRI's ArcGIS software and information on the NYC subway system using GIS files downloaded from Metropolitan Transit Authority's website. Details on these sources are in the Data appendix.

The overall correlation between driving distance and total commuting distance for all student-program pairs is 0.96. A regression of commuting distance on driving distance yields a coefficient of 0.77. Table S1 provides a summary of the correlations by the student and school borough. The correlations are higher than 0.84, except for schools in Staten Island, where the subway system is not quite as extensive as in other boroughs. In fact, it may be that driving distances are a more accurate measure of travel costs in Staten Island than subway distance.

Panels B and C show that most students are assigned to schools in their borough in both the uncoordinated and coordinated mechanism. In both mechanisms, a very small number of students who do not live in Staten Island are assigned to schools there, and conversely, only a small number of students living in Staten Island are assigned to schools in a different borough.

Table B1. Offer Processing in Second Year of Coordinated Mechanism (2004-05)

	Number of Students (1)	Distance to School (in miles): Assignment (2)	Distance to School (in miles): Enrollment (3)	Exit from NYC Public Schools (4)	In NYC Public, but at School Other than Assigned (5)
	Coordinated Mechanism - 2004 - 2005				
Overall	69,013	4.07	3.96	6.6%	6.9%
Main Round	60,251	4.11	3.99	6.5%	6.4%
Supplementary Round	5,475	4.16	4.03	8.5%	13.6%
Administrative Round	3,287	3.25	3.26	4.9%	5.4%

Notes: Columns 2-5 report means. Coordinated mechanism for 2004-05 based on deferred acceptance. Student distance calculated as road distance using ArcGIS. Assignment is the school assigned at the conclusion of the high school assignment process. Enrollment is the school a student enrolls in October following application. Assigned student exits 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 NYC Public, 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 first made.

Table B2. Subway and Driving Distance and Cross-Borough Travel

	School Borough					
	Bronx	Brooklyn	Manhattan	Queens	Staten Island	Total
Student Borough	(1)	(2)	(3)	(4)	(5)	(6)
	A. Correlation between Subway and Driving Distance					
Bronx	0.90	0.93	0.97	0.91	0.76	…
Brooklyn	0.90	0.91	0.95	0.91	0.92	…
Manhattan	0.96	0.95	0.98	0.95	0.76	…
Queens	0.91	0.91	0.95	0.87	0.85	…
Staten Island	0.84	0.92	0.85	0.89	0.73	…
	B. Cross-Borough Travel in Uncoordinated Mechanism					
Bronx	15,187	41	1,382	66	1	16,677
Brooklyn	13	20,877	1,073	502	12	22,477
Manhattan	89	42	8,604	24	1	8,760
Queens	15	493	586	16,498	0	17,592
Staten Island	2	13	59	4	4,774	4,852
	C. Cross-Borough Travel in Coordinated Mechanism					
Bronx	13,335	85	2,049	84	8	15,561
Brooklyn	39	20,035	1,858	846	40	22,818
Manhattan	238	108	7,492	52	7	7,897
Queens	26	584	1,028	14,972	9	16,619
Staten Island	3	37	69	4	3,913	4,026

Notes: Panel A reports on the correlation between student-school distance as computed by road distance and by subway distance. Subway distance is the sum of distance on foot to the student's nearest subway station, travel distance on the subway network to a school's nearest subway station, and the distance on foot to the school from that location. Both measures of distance computed using ArcGIS. Panels B and C report on the number of students in each borough who are assigned school in each borough.

Table B3. Main Round Assignments in Coordinated Mechanism, by Length of Rank Order List

		Length of Rank Order List											
Choice Assigned	All	1	2	3	4	5	6	7	8	9	10	11	12
Total	69,907	4,597	3,282	4,128	4,622	4,952	4,776	4,406	4,390	4,558	6,135	9,849	14,212
1	31.9%	88.6%	40.7%	35.2%	31.9%	27.9%	28.6%	27.1%	25.7%	25.6%	25.4%	26.2%	25.2%
2	15.0%		39.8%	17.7%	15.1%	14.8%	14.6%	13.7%	13.9%	13.9%	15.2%	14.7%	14.6%
3	10.2%			24.3%	11.6%	11.6%	10.6%	10.0%	10.8%	9.9%	10.4%	10.4%	10.5%
4	7.3%				18.0%	9.3%	8.1%	7.9%	8.0%	7.6%	7.6%	7.8%	8.2%
5	5.4%					12.8%	7.0%	7.0%	6.3%	6.1%	6.6%	6.2%	6.7%
6	3.9%						10.2%	5.7%	4.9%	5.0%	4.9%	4.8%	5.3%
7	2.9%							8.1%	4.3%	4.4%	4.0%	4.1%	4.3%
8	2.0%								5.8%	3.4%	3.3%	2.9%	3.5%
9	1.5%									4.0%	2.8%	2.7%	2.8%
10	1.1%										3.2%	2.3%	2.6%
11	0.8%											2.6%	2.2%
12	0.5%												2.5%
Unassigned	17.5%	11.4%	19.5%	22.8%	23.3%	23.6%	20.9%	20.6%	20.3%	20.1%	16.7%	15.3%	11.6%

Notes: This table reports choices assigned after the main round in coordinated mechanism in 2003-04.

Table B4. Assignment and Enrollment Decisions of Students in Coordinated Mechanism by Rank Order List Length

		Length of Rank Order List											
	All	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th
							A. Students Offered Assignment in Main Round						
Number of Students	57,658	4,072	2,641	3,187	3,545	3,782	3,776	3,497	3,499	3,642	5,113	8,340	12,564
Average Rank of Assignment	3.00	1.00	1.49	1.86	2.21	2.53	2.76	3.04	3.20	3.35	3.49	3.60	3.93
Accept Main Round Assignment	92.7%	91.2%	88.5%	88.4%	90.2%	91.2%	92.3%	91.9%	93.0%	93.6%	94.5%	94.6%	94.3%
Enroll in Private School	2.5%	6.9%	7.4%	6.1%	4.5%	2.9%	2.4%	2.1%	1.9%	1.2%	1.0%	0.7%	1.0%
Remain in Current School	1.2%	1.2%	2.0%	2.3%	1.9%	2.1%	1.4%	1.8%	1.4%	1.2%	0.9%	0.7%	0.6%
Attend Specialized or Alternative School	0.1%	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.2%	0.1%	0.1%	0.2%	0.1%	0.0%
Participate in Supplementary Round	0.3%	0.1%	0.2%	0.2%	0.4%	0.5%	0.6%	0.3%	0.6%	0.3%	0.3%	0.2%	0.3%
							B. Students Unassigned after Main Round						
Number of Students	12,249	525	641	941	1,077	1,170	1,000	909	891	916	1,022	1,509	1,648
Participate in Supplementary Round	52.6%	26.1%	44.8%	54.0%	54.1%	56.2%	55.6%	55.7%	52.7%	46.5%	43.5%	49.6%	68.2%
Enroll at Supplementary Round Assignment	72.9%	73.0%	85.0%	76.0%	75.5%	77.8%	73.0%	75.9%	74.5%	68.8%	71.7%	69.5%	66.3%
Enroll in Private School	2.8%	6.7%	6.1%	4.7%	3.5%	3.8%	2.2%	1.7%	1.6%	1.9%	2.2%	1.4%	1.9%
Remain in Current School	3.2%	6.7%	6.2%	5.6%	5.3%	4.4%	3.2%	3.3%	2.5%	2.0%	1.5%	0.9%	1.8%
Attend Specialized or Alternative School	0.3%	0.8%	0.5%	0.6%	0.2%	0.1%	0.3%	0.3%	0.3%	0.1%	0.2%	0.5%	0.1%

Notes: Assignment and enrollment decisions of students in the demand estimation sample under the coordinated mechanism. Panel A restricts to students that received an assignment in an NYC Public School in the Main Round. Panel B restricts to students that did not receive an assignment in the Main Round.

C Appendix: Data (Not for Publication)

The data for this study comes from the NYC Department of Education (DOE), the 2000 US Census, ArcGIS Business Analyst toolbox and GFTS NYC subway data from the NYC Metropolitan Transit Authority. These sources provide us with data on students, schools, the rank order lists submitted by the students, the assignment of students to schools or the distance between the students and schools on the road network or the subway system. Students and programs are uniquely identified by a number that can be used to populate fields and merge across DOE datasets. We geocode student and school addresses to merge with geo-spatial data.

We also use three samples of students in our analysis: one sample to estimate demand and the other two to infer the welfare effects of the mechanism change. The welfare samples consists of public middle school students who matriculate into NYC Public High Schools in the academic years 2003-04 and 2004-05. The demand sample consists of public middle school students who participated in the Main round of the mechanism in 2003-04. The demand sample and the welfare sample from 2003-04 are not nested because students participating in the mechanism may choose to enroll in a school outside the NYC Public School system, whereas others may be assigned to a public school outside the main assignment process.

C.1 Students

Assignment and Rank Data

Data on the assignment system come from the DOE's enrollment office. The files indicate the final assignment of all students in both years in our analysis. We use these assignments as the basis of our baseline welfare calculations. In addition, the assignment system also provides separate files that detail the rank orders, applications, or processes through which a student is assigned to a given school.

We use the records from the Main round in the new mechanism to obtain the rank order lists submitted by students and the assignment proposed by the mechanism. A total of 87,355 students participated in the main round.

For the old mechanism, the assignment system provides student choice and decision files for the Main round. The former contains the ranked applications submitted by the students and the latter provides the decisions of the schools to accept/reject/waitlist the student and the students response to these offers, if any. A total of 84,272 students participated in the Main round.

The old assignment system also contains several files documenting the supplementary variable assignment process (VAS) round.

Assignment Rounds and Offers in the Old Mechanism

The files in the old mechanism do not directly contain information on how students were assigned to their programs. However, we are able to determine whether a student applied to a particular program/school in the Main process or the supplementary VAS process. We first append fields indicating whether a student applied to her assigned program in the main process. We also append a field indicating whether a student applied to her assigned school in the supplementary

VAS process. It turns out that no final assignment appears in both the main and the VAS files. We therefore categorize the former assignments as main-round assignments and the latter as VAS assignments. We assume that all other assignments are through the Administrative round. Based on conversations with DOE officials, students were typically assigned to the school closest to home that had open seats. Our understanding is that most of the students who participated in the VAS process did not have a default local school. An analysis of the geographic distribution of our definition of students assigned administratively is consistent with this fact: many parts of NYC have no students assigned administratively.

Finally, we also append the number of offers made to a particular student using a file with the initial response of schools to the student application.

Assignment Rounds in the New Mechanism

We use the NYC assignment files described above to determine the process through which a student was assigned a given school.

The assignment files in the new mechanism contain, for every student program pair that is ranked in either the Main round or the Supplementary round, two fields indicating whether the student is eligible for the school and if the student was assigned to that school. A final assignment is treated as a Main round assignment if it appears as an eligible assignment in the Main round. Assignments that are not made in the main round are treated as a supplementary round assignment if they appear in the Supplementary round files. All other assignments are treated as administrative assignments.

Student Characteristics

The records from the NYC Department of Education contain the street address, previous and current grade, gender, ethnicity, and whether the student was enrolled in a public middle school. Each student is identified by a unique number that allows us to merge these data with additional data from the NYC DOE on a student's scores in middle school standardized tests, Limited English Proficiency status, and Special Education status. A separate file indicates subsidized lunch status as of the 2004-05 enrollment. If a student is not in that file, we code the student as not receiving a subsidized lunch.

There are several standardized tests taken by middle school students in NYC. To avoid the concern that two different tests may not be comparable indicators of student achievement, we identify the modal standardized tests in math and reading taken by students in our sample. These are the May tests with codes CTB and TEM respectively. Of the students who did not take either of these tests in May, at most 10% (<2% of the full sample) took a different standardized test in the same subject while in middle school. We verified that the distribution and support of the test scores are similar across the two years in our sample. Some students took the test multiple times. The highest score obtained by a student was used in these instances.

In 2002-03, the math and reading scores are missing for 13.56% and 17.55% students from our final sample respectively. For the 2003-04 welfare sample, scores are missing for 8.29% and 13.57% students respectively for math and reading. In the demand sample the corresponding fractions are 7.13% and 12.56%.

Geographic Data

We use the 2000 US Census to obtain block group family income. The addresses of students and their distance to school were calculated using ArcGIS. Corrections to the addresses, when necessary, were made using Google Map Tools followed by manual checks and corrections.

The final set of addresses were geocoded using ArcGIS geocoder with the address-set in the Business Analyst toolbox (ver. 10.0). We first used an exact match to determine if a student's address can be geocoded precisely to a rooftop. If the results were unreliable, we coded the student to the centroid of the zip-code. The vast majority of students were placed at the rooftop level. The OD Cost matrix tool in the Network Analyst toolbox was used to compute the distance by road for each student-school pair. The road network is also obtained from Business Analyst.

Our computation of subway distances assumes that a student first walks to the closest subway stop, then uses the subway system to travel to the subway stop closest to the school, and finally walks from the subway to the school. The location of the subway stops is taken from the GTFS and geodata data on the NYC Metropolitan Transit Authority website. The network analyst toolbox is used to compute the walking distance and the GTFS data is used to compute the distance on the subway system between every pair of subway stops.

Merging Student Records

Assignment and other DOE files are matched using the unique student identifier linking these data. Each eighth-grade non-private middle school student in the Department of Education records could be merged uniquely with a student in the NYC assignment records. Less than 0.45% of students with known assignments in the records of the NYC assignment system could not be merged with a student in the DOE records. These students were not included in the analysis.

C.2 Applicant Sample Construction

Our goal is to consider first-time applicants to the NYC public (unspecialized) high school system who live in New York City and attend a public middle school 8th grade. Below, we described the procedure used to construct the samples. The selection procedure is also summarized in Table B1.

Welfare Sample

The welfare samples are constructed from the NYC Department of Education's records of all students enrolling in ninth grade at a high school in academic years 2003-04 and 2004-05.

Because our choice set in the demand analysis will be restricted to unspecialized, non-charter high schools in the public school system, we do not include students who matriculated to such schools in the welfare sample.

Of the 92,623 eighth grade students matriculating into ninth grade at a NYC public school in 2002-03, 11,790 (12.73%) students went to a private middle school and were dropped. Another 8,051 (8.69%) of students were not included in the analysis because their assignment was unknown, or because they matriculated at either a specialized high school or a charter school. Finally, we exclude students in schools that were closed (no assignments in the new system).

In 2003-04, about 1.3% students had also participated in the old mechanism, presumably because these students repeated eighth grade. These students were considered a part of the 2002-03 sample and only their 2002-03 assignment into high school is considered in our analysis. We also drop private middle school students and those not assigned to public school. These fractions were similar to the 2002-03 numbers, at 12.21% and 8.13% respectively. We also drop students who were assigned to new schools.

These selections into the sample leave us with 70,358 students in 2002-03 and 66,921 students in 2003-04. Students who may have been assigned to a high school program through a process other than the Main round are included in these samples.

Demand Sample

This sample is sourced from the NYC Assignment system's records on the participants in the Main round of the mechanism. As discussed in the text, we use data only from the Main round of the mechanism because this round has the most desirable incentive properties.

We do not want to exclude students on the basis of final assignment to avoid selecting on the choice to leave the public school system. In order to most closely match the construction of the welfare sample, we select the demand sample only on characteristics that can be considered as exogenous at the time of participation.

Since we focus on first-time applicants in eighth grade, we exclude 747 students who were part of the 2002-03 files, and 5,311 students who were ninth graders. Presumably, these students were held back in eighth or ninth grade. This leaves us with a sample of 81,297 eighth grade students.

Of the eighth-grade participants, 9,301 or 11.44% of students were from private middle schools and were dropped. We also excluded students designated as belonging to the top 2% of their middle school class beause they are prioritized at education option schools, creating incentives to misreport their preferences. These are 2.5% of the non-private eighth grade population.

A total of 216 students did not rank any public schools in our sample. After excluding these students, a total of 69,907 students remain in the sample we will use for the demand analysis.

C.3 Programs/Schools

NYC Department of Education School Report Cards

The school characteristics were taken from the report card files provided by the NYC Department of Education. These data provide information on a school's enrollment statistics, racial composition of student body, attendance rates, suspensions, teacher numbers and experience, and Regents Math and English performance of the graduating class. A unique identifier for each school allows these data to be merged with data from our other sources.

There were significant differences in the file formats and field names across the years. To keep the school characteristics constant across years, we use the data from the 2003-04 report cards as the primary source. Except for data on the math and reading achievement, variable descriptions were comparable across years. For these comparable variables, we used the 2002-03 data only when the 2003-04 data were not available. The coverage of the characteristics for the sample of schools is enumerated in Table B2.

Assignment System and DOE files

Assignment data contain a list of all school programs in the public school system along with an identifier for the associated high school. The Department of Education provided a separate file with data on the school addresses and identifiers that allow a merge with the assignment system database. A second identifier can be used to merge these data with other fields in the department of education records described above.

Across the two years, the high school identifiers in the files were inconsistent for a small number of schools in our sample. These were matched by name and address of the school. One school moved from Brooklyn to Manhattan and was investigated to ensure that the records were appropriately matched.

Program Characteristics

Program characteristics are taken from the DOE's High School Directory, which is made available to students before the application process. Reliable data on program types was not available in 2002-03. For that year, the program types were imputed from the 2003-04 program types if the program was present in both years. Otherwise, the program was categorized as a general program.

There were a very large number of program types. These were aggregated into fewer broad categories. The items in the list below are the aggregated categories that include all of the subcategories as described by the data.

1. Arts: Dance, Instrument Performance, Musical Theater, Performing and Visual, Performing Arts, Theater, Theater Tech, Visual Arts, Vocal Performance.
2. Humanities/Interdisciplinary: Education, Humanities/Interdisciplinary.
3. Business/Accounting: Accounting, Business, Business Law, Computer Business, Finance, International Business, Marketing, Travel Business.
4. Math/Science: Engineering, Engineering – Aerospace, Engineering – Electrical, Environmental, Math and Science, Science and Math.
5. Career: Architecture, Computer Tech, Computerized Mech, Cosmetology, Journalism, Veterinary, Vision Care Technology.
6. Vocational: Auto, Aviation, Clerical, Construction, Electrical Construction, Health, Heating, Hospitality, Plumbing, Transportation.

7. Government/law: Law, Law Enforcement, Law and Social Justice, Public Service.

8. Other: Communication, Expeditionary, Preservation, Sports.

9. Zoned

10. General: General, Unknown.

Finally, some programs adopt a language of instruction other than English. We categorized the languages into Spanish, English, Asian Languages, and Other.

C.4 School Sample Construction

We consider the assignment of eighth grade students in NYC public middle schools into public high schools that are not charters, specialized or parochial. Our analysis uses two school sample, one for each year in our analysis.

To construct these samples, we started with the set of schools and programs in the assignment records. For the analysis of rank data, we added the set of school programs that were ranked by any student in our demand sample. This initial set consists of 743 (301) programs (schools) in 2002-03 and 677 (293) programs (schools) in 2003-04.

In 2003-04, this list contained 62 parochial school programs. We verified that each of the 130 students matriculating to these school programs were private middle-schoolers. These schools were dropped from the analysis because private middle-schoolers are not in the population of interest. Subsequently, we dropped all charter and specialized high school programs and other school programs that do not have assignments and were not ranked by any student in our sample.

A total of 9 continuing student programs accepted students only from their associated middle school. Since these programs cannot be chosen by students who were not in that school in eighth grade, we combine these programs with a generic program (e.g., unscreened, English, general/humanities/math). Rank order lists for students who ranked both the continuing students' only program and the associated program were modified as described below.

Finally, we dropped new and closed schools from the analysis. Closed schools were ones that admitted students in 2002-03, but not in 2003-04. The set of new schools was collected from a separate DOE directory of new schools. These schools were not well advertised and very few students ranked them, making calculations with those schools unreliable.

The number of schools and programs at each stage of our selection procedure is also summarized in Table B2.

C.5 Program Capacities

Program capacities are not provided separately in the data files. We have estimated program capacities from the actual match files and students' final assignments. The capacity of each program is initially set to zero. If a student in our demand sample is assigned a program at the end of the assignment process, the capacity of the program is increased by one. Otherwise, if the student is assigned a program in the Main round, the capacity of the program is increased by one. Finally, if a student is not assigned in the Main round and is assigned a program in the supplementary round, the capacity of the program is increased by one.

Education Option programs are divided into six buckets: High Select, High Random, Middle Select, Middle Random, Low Select and Low Random. The bucket capacities are calculated as above by taking into account the category of the assigned student. For example, if a student of High category is assigned an Education Option program, then the capacity of a High bucket is increased by one. If the current capacity of the High Select bucket is less than or equal to that of High Random, then the capacity of the High Select bucket is increased, otherwise the capacity of the High Random bucket is increased.

C.6 Program Priorities

The type of a program determines how students are priority-ordered for the program. The data contains a list of all programs with program-specific information, including type, building number, street address, etc. When students have the same priority, the tie is broken randomly. The random numbers are generated by computer during our simulations.

The assignment data contains the following fields that determine a student's priority order at programs. Priority group is a number assigned by the NYC Department of Education depending on students' home addresses and location of programs, etc. High school rank is a number assigned by each program. This may reflect an student's ranking among all applicants to an Education Option program, or whether a student attended the information session of an limited unscreened program, etc. These fields are provided for every student at every program that the student ranked. Students applying to Educational Option programs are placed into one of three categories based on their score on the 7th grade reading test: top 16 percent (high), middle 68 percent (middle), and bottom 16 percent (low). Student categories are included in the assignment data.

Unscreened programs order students based on their random numbers only. Limited unscreened and formerly zoned programs order students first by priority group, and then by random number within the priority group. Screened programs order students by priority group, then by high school rank. Each Education Option program orders all applicants for each of six buckets, High Select, High Random, Middle Select, Middle Random, Low Select and Low Random. A high bucket orders high category students first, then middle category students, then low category students. A middle bucket orders middle category students first, then high category students, then low category students. A low bucket orders low category students first, then high category students, then middle category students. A select bucket orders students within each category by priority order, then by high school rank. A random bucket orders students within each category by priority order.

C.7 Miscellaneous Issues

Modifications to the rank order list

1. Some students ranked a program that were either charter schools or specialized high schools in the Main round. These programs are not in the sample of schools we consider and were

likely ranked by the students in error. In such cases, programs were removed from the rank order lists and rank orders lists were made contiguous where all programs ranked below a program not in the sample were moved up in the rank order lists. These programs were observed a total of 795 times in the data. Thirty students ranked only charter or specialized programs.

2. The rank order lists of students who ranked continuing student program were modified as follows: First, the lists of all students who ranked only the continuing student program were modified so that the student ranked the associated generic program instead. When students ranked both the generic program and the associated continuing student program, the list was modified so that only the associated program was ranked, and at the highest of the two ranked positions. All programs ranked at positions below the lower ranked of the two programs were moved up by one. A total of 46 students ranked both the continuing program and the generic program we mapped the continuing program to. In 17 cases, these ranks were not consecutive.

Table C1. Student Sample Selection

	Mechanism Comparison			Demand Analysis	
	Uncoordinated	Coordinated		Coordinated	
	2002-2003	2003-2004	2004-2005	2003-2004	2004-2005
	(1)	(2)	(3)	(3)	(4)
Number in the NYC DOE student file	100,669	97,569	96,327		
Number of students in the rank data				87,355	91,290
Excluding students in both 2002-03 and 2003-04 files from 2003-04		96,275		86,608	
Excluding ninth grade students	92,623	89,062	90,250	81,297	86,514
Excluding private middle school students	80,833	78,183	80,093	71,996	78,439
Excluding students with addresses outside the five boroughs	80,725	78,089	79,977	71,916	78,327
Total number of students with known assignments to sample schools	75,515	73,989	75,049		
Excluding students attending specialized high schools	72,725	70,992	71,861		
Excluding students attending charter schools	72,681	70,886	71,749		
Excluding students in closed and new new schools	70,358	66,921	69,013		
Excluding top 2% students				70,123	76,753
Excluding students that did not rank any sample schools				69,907	76,569

Notes: Uncoordinated mechanism refers to 2002-03 mechanism and coordinated mechanism refers to the mechanism based on deferred acceptance. A student has invalid census information if address is missing, cannot be geocoded or places the student outside of New York City. A distance observation is invalid if it is missing or is greater than 65 miles.

Table C2. Construction of School Sample

	Uncoordinated		Coordinated			
	2002-2003		2003-2004		2004-2005	
	Programs	Schools	Programs	Schools	Programs	Schools
	(1)	(2)	(3)	(4)	(5)	(6)
Programs where NYC public school students assigned	743	301	669	293	658	322
Adding additional programs ranked by students			677	294	764	338
Excluding parochial schools	681	239	677	294	752	331
Excluding specialized schools	669	232	665	287	750	329
Excluding charter schools	667	230	663	285	702	315
Excluding programs with no assignments or ranking	637	225	648	284	691	313
Combining continuing education programs	637	225	639	284	691	313
Excluding closed schools	612	215	639	284	691	313
Excluding schools opened after HS directory printed*	612	215	558	235	661	283
Programs/schools ranked by students in sample			497	234	660	283

Notes: 13 continuining student programs were merged with a generic program at host school. Parochial schools in 2002-03 only have private middle school students assigned to them and are not ranked by students in the demand sample. *A total 20 schools and 23 programs opened before HS directory printed are included in 2003-04.

Table C3. Coverage of School Characteristics

	Uncoordinated 2002-03 (1)	Coordinated 2003-04 (2)	Both Years (3)
Total number of schools in the sample	215	234	215
9th grade enrollment	196	199	189
Race	196	199	189
Attendance Rate	196	199	189
Percent Free Lunch	196	198	189
Percent of teachers less than 2 years experience	219	223	212
High Math Achievement	198	200	191
High English Achievement	180	177	173
Percent Attending College	171	167	165

Notes: Table reports the number of schools with the characteristics from New York State Report cards.