

Authors are encouraged to submit new papers to INFORMS journals by means of a style file template, which includes the journal title. However, use of a template does not certify that the paper has been accepted for publication in the named journal. INFORMS journal templates are for the exclusive purpose of submitting to an INFORMS journal and should not be used to distribute the papers in print or online or to submit the papers to another publication.

Guiding School-Choice Reform Through Novel Applications of Operations Research

Peng Shi

MIT Operations Research Center, Cambridge, Massachusetts 02139, pengshi@mit.edu

In January 2012, Boston initiated a school-assignment reform. After attending several community meetings, I formulated the reform as an optimization problem of finding school-choice menus and priorities that induce the best combination of equity of access, proximity to home, predictability, and community cohesion. Using previous school-choice data, I fit a discrete choice model of how families select schools, and created a simulation engine that estimated a variety of outcome measures for any plan. I also proposed several new plans, and helped Boston Public Schools analyze a short list of plans in detail. In March 2013, Boston adopted one of these plans, which will affect the 9,500 children who apply to Boston elementary schools each year.

Key words: education; public service; school choice; policy analysis.

History: This paper was refereed.

Many public school districts in the United States have implemented school-choice systems to offer families more control in selecting the schools their children will attend, and to help school districts balance supply and demand across schools. Typically, families are offered a menu of potential school options, and apply by submitting a rank-order list of preferences for their children. A centralized algorithm assigns children to one of their top-choice schools if capacities allow; otherwise, it breaks ties using a system of priorities and

lottery numbers. As of 2013, a partial list of public school districts that have implemented such a system includes Boston, Cambridge, Charlotte-Mecklenburg, Chicago, Denver, Minneapolis, Miami-Dade, New York City, New Orleans, and San Francisco.

Most of these districts include schools in which the demand is higher than the capacity, because families tend to select the best-performing schools based on standardized test scores and other quality metrics. Determining which children should be selected to attend which schools is a nontrivial task. Allowing all children to access all schools (i.e., giving each child the possibility of attending any school in the school system) and using a lottery to break ties is arguably the most equitable; however, this may result in unsustainably high transportation costs if a city must pay for school busing, because the school buses would need to pick up children from a large geographic area. Thus, a city has to navigate the delicate balance between offering sufficient access to quality schools and a variety of choices, while keeping choice (in the remainder of this paper, choice refers to school choice) menus manageable. It must also consider families' preferences, because access to a particular school depends on both the number of seats available and the number children competing for these seats.

The Boston public school system (BPS) was facing this problem. Since 1988, its choice system for elementary and middle schools had employed a three-zone plan, which partitioned the city into three geographic zones. A family could seek admittance for its children to any school in its zone and any school within a walk zone (i.e., one mile) around its home by ranking that school. Thus, each family was given a large number of choices on its menu; partly because of this, the assignment system resulted in high transportation costs to the city, low predictability for families, and scattering children from the same neighborhood across the city. Furthermore, the quality of schools between the zones was a

concern; the zone boundaries had been drawn in 1988 and demographics and the quality of some schools had changed since then. In January 2012, Boston launched an initiative to reform the three-zone plan, and a mayor-appointed external advisory committee (EAC) held public meetings to deliberate and to elicit feedback from the community. Important considerations for the new plan included equity of access to quality schools, predictability, variety of choice, proximity to home, transportation savings, transparency, community cohesion, and ability to adapt to future change. Despite the political push, there was no guarantee that the reform would result in a satisfactory solution, because similar reform attempts in 2004 and 2009 had failed as a result of concerns over the equity of access to quality schools in the new proposals.

I became aware of the reform by providence as I read online news articles about it, and was drawn to the issue because it aligned with my research interest of market design—creating allocation systems with the right incentives to induce the desired outcomes. The internal EAC meetings were open to the public. So, I attended many of these meetings, and met concerned parents, BPS staff, city committee members, and activists. Initially, I mostly listened to what other attendees were saying. Through repeated interactions with them, I gained a clearer understanding of the reform, and formulated it as an optimization problem: find the choice menus and priorities that would induce the desired outcomes based on a variety of metrics set by the EAC. In October 2012, at a public meeting to elicit alternative proposals from the community, I gave a short presentation on an optimization-based proposal, which attracted positive interest. Based on feedback from city committee members and concerned parents, I simplified my proposal and presented it at an EAC meeting in November, at which it was well received by both BPS and the EAC. Subsequently, I began working more closely with BPS staff and was given access to

student-level data on how families had previously ranked schools. I also began collaborating with Parag Pathak from the School Effectiveness and Inequality Initiative (SEII) at the Massachusetts Institute of Technology. In December, the EAC officially commissioned us to provide a rigorous analysis of a short list of proposed plans.

Using previous choice data, I fit a discrete choice model for how families would select schools under various choice menus, and built a simulation engine that predicted outcome metrics in any given plan. BPS used this engine to refine its proposals, and I used it to analyze new proposals. The quality of each school was ranked based on a metric specified by BPS and the EAC. In one of these proposals, home-based A, a family's menu of options could include (1) any school within its one-mile walk zone, (2) the two closest schools ranked in the top 25 percent, (3) the four closest schools ranked in the top 50 percent, (4) the six closest schools ranked in the top 75 percent, and (5) the three closest capacity schools, which were designated to help BPS meet excess demand. On January 30, 2013, we published our commissioned report, which provided detailed analyses of the short list of proposals selected by BPS, showing differences by race, grade, student type, and neighborhood. The EAC deliberated on this list of proposals, elicited public feedback, and voted to adopt the home-based A-plan because the choices in the menus were close to students' homes, while also providing equity of access to quality schools based on a variety of thresholds; in addition, it could automatically adapt to future changes in school quality. On March 13, 2013, the Boston School Committee, the regulatory body for the school board, approved the plan for implementation in 2014.

This project illustrates the power of quantitative analysis and simulation in bypassing rhetorical gridlock in public policy and informing the debate. It shows the importance of listening to various stakeholders to formulate the correct quantitative problem, and of

constraining the solution so that it is within the organization’s implementation capabilities. It also suggests the potential of further applications of simulation and optimization techniques in public service allocation.

Background

In 2012, about 9,500 students applied to BPS to attend one of the kindergarten grades, K0, K1, or K2. K2 is the main entry grade to elementary school, and corresponds to the age when mandatory-schooling laws take effect in Massachusetts. BPS had 77 elementary schools, and each school had at least a regular education (Reg. Ed.) program for K2; some also had also K1 or K0 programs. In addition to Reg. Ed. programs, some schools had specialized programs for English language learners (ELL), ELL students of a specific home language, or special education (SPED) students.

Since 1988, families in Boston had been applying to BPS via its three-zone choice plan, which partitioned Boston into three geographically contiguous zones, called the East, the West, and the North zones (see Figure 1). Each child’s family could apply on that child’s behalf to any school within the family’s geographic zone and any school within a one-mile radius of its home (i.e., walk zone). Each school could have multiple programs for the same grade, such as Montessorri, ELL, or SPED. When a child applied to a school, that child could apply to any program for which he (she) was eligible. These programs together constituted the child’s choice menu. For K2 students, the choice menu contained between 24 and 42 Reg. Ed. programs, depending on the student’s geographic zone and walk zone, with additional programs for ELL or SPED students.

In January of each year, a family planning to enroll a child for September of that year had to submit a rank-ordered list of programs within its choice menu, in the order of

Figure 1 This map shows the three zones in Boston’s three-zone plan.



its preferences, and could rank and submit as many programs as it would like. A computer algorithm based on a system of priorities and lottery numbers assigned children to programs, taking as input the families’ choices and the programs’ available capacities.

Since 1999, the priorities for school selection had been the following:

1. Continuing students (i.e., students who had been enrolled at the school for the previous grade and were advancing a grade) had the highest priority at the schools at which they were continuing.
2. Among students who were tied based on continuing status, those with currently enrolled siblings at the school were given priority.
3. Among students who were tied based on continuing and sibling status, those who lived in the school’s walk zone (an approximate one-mile radius around the school) received priority; however, this priority was applicable to only 50 percent of the seats in each program. The remaining 50 percent of the seats were open; no consideration was given to whether students came from inside or outside the walk zone.

4. A random lottery number for each student determined the remaining ties. Each student received only one lottery number, which was applicable to the child’s application to all schools.

Given students’ choices and schools’ priorities, BPS still needed a way to compute the final assignment in a fair and efficient way; this method of determining the assignment is called the assignment algorithm. Since 2006, Boston had been using the student-proposing deferred acceptance (DA) algorithm, as described below.

1. Find an unassigned student with a nonempty choice list, and let the student apply to his (her) first-choice program.

2. Let the program tentatively accept the student; if this acceptance causes the program to violate its capacity limit, then let the program find the tentatively accepted student with the least priority and reject that student. Remove the program from that student’s choice list and unassign that student.

3. Go back to Step 1 until each unassigned student has an empty choice list.

The output of this algorithm does not depend on the order in which students apply in Step 1 (Roth and Sotomayor 1990). Gale and Shapley (1962) originally proposed this algorithm, and Abdulkadiroğlu and Sönmez (2003) adapted it to school choice. Based on a theorem by Dubins and Freedman (1981) and Roth (1982), Abdulkadiroğlu and Sönmez (2003) showed that as long as the priorities are predetermined and do not depend on students’ choice submissions, the above algorithm cannot be influenced by any strategy (i.e., regardless of what other students submit, no student can benefit by manipulating his (her) own choice submissions). This property influenced Boston to adopt the algorithm in 2006 (Abdulkadiroğlu et al. 2006). Alvin Roth, one of the economists responsible for introducing this algorithm to school choice in Boston and other cities in the United States,

as well using it for medical-residence matching, and Lloyd Shapley, who first proposed this algorithm, won the Nobel Prize in Economics in 2012.

After BPS had run the DA algorithm, some students could still be unassigned. For students applying to K0 or K1, BPS would place them on a wait list and accept them later if spaces opened up. Many of these students might not be accommodated that year because BPS did not guarantee placement in K0 or K1; however, it was required by law to accommodate every K2 student. Therefore, after the algorithm terminated, BPS would administratively assign unassigned K2 students to the remaining available seats based on distance; if necessary, it would open new classrooms before the school year began to make space for all K2 students. If all rooms in a school building were occupied, BPS might add outdoor modular classrooms.

Each year, about 80 percent of the K0, K1, and K2 applicants participated in the January-February round of the assignment algorithm, which BPS called Round 1. BPS also executed three smaller assignment rounds between February and June. At the completion of these rounds, any unassigned students were assigned manually based on idiosyncratic considerations.

Boston provided transportation to any elementary school student who lived further than a one-mile straight-line distance from his (her) school. During the 2011-2012 school year, it bused about 18,500 elementary school students, including students up to grade 5; the average straight-line distance between home and school for the bused students was 2.0 miles.

Assignment Plan: Problems

The major criticisms of the three-zone plan were inequitable access to quality schools, high transportation costs, low predictability of assignment, and community dispersion.

Inequitable access to quality schools: The original zone boundaries were drawn in 1988 to achieve racial and socioeconomic balance (Alves and Willie 1990); however, by 2012, demographics had shifted; the East zone had a higher concentration of poverty than the other zones. The percentages of elementary school students eligible for a free lunch because they lived in a household with a gross income of not greater than 130 percent of Federal poverty guidelines were 58, 53, and 52 percent for the East, North, and West zones, respectively. BPS statistics published in 2012 show that schools in which students exhibited lower performance and slower advancement in the Massachusetts comprehensive assessment system (MCAS) tests were also more concentrated in the East zone. Some argued that the inequity was exacerbated by the walk-zone priority, because students living outside the walk zones of higher-performing schools had lower priority for admittance to these schools. The debate about the correct proportion of seats at each school to which the walk-zone priority should be applied was ongoing because the proportion of seats fluctuated; for example, it changed from 50 percent in 1988 to 75 percent in 1990, then to 100 percent in 1996, and back to 50 percent in 1999. This illustrated the difficulty of achieving public consensus over questions of access and equity.

High transportation costs: In 2011-2012, Boston spent more than \$80 million on school busing; this represented almost 10 percent of its total school board budget (Russell and Ebbert 2011). One reason for this high cost was that the city was required by law to pay for transportation of elementary school children not only to public schools, but also to charter, private, and parochial schools, and for expensive door-to-door transportation for certain SPED students. Nevertheless, a significant share of this cost resulted from having to bus Reg. Ed. students to BPS schools, which represented about 35 percent of the routes (Boston Public Schools 2013). The geographically large zones required BPS to send buses to cover an area potentially as large as one-third of the city.

Low predictability of assignment: In 2011, the Boston Globe published a series of articles about the experiences of 13 families who had participated in the choice process (Russell et al. 2011). One salient theme in a majority of these articles is that families were frustrated with their inability to predict whether their children would be admitted to one of their top choices. If their children were not admitted to one of their top-choice schools, some families would consider moving to suburban districts so that their children could attend schools in these districts; however, moving is a major decision for a family to make. Therefore, these families wanted to know the school-selection outcome for their children as early as possible. For some families, however, the uncertainty was prolonged because they hoped their children would be admitted to a school (i.e., kindergarten) in a later round. Some families were discontent; they had exerted much effort into researching many options; however, their children were not selected to attend any of their top-choice schools. This prompted some parents and activists to label the system as an illusion of choice, because being able to rank a school did not imply being selected to attend that school. Unpredictability was a central reason that the Seattle school board changed from a choice lottery system to a neighborhood school system in 2009, with school assignments determined primarily by home location (Lilly 2009, Woodward 2011).

Community dispersion: Ebbert and Russell (2011) document 19 school-aged children on one street in Boston who traveled to 15 different schools. They argue that this resulted in a loss of a sense of community within a neighborhood, because “families are less likely to know one another when their children don’t attend schools together.” Having schools with strong ties to the community was also important to Boston, because it allowed the city to provide health, unemployment, and parental education services through the school. Boston’s mayor, Thomas Menino, elaborated on community dispersion under the three-zone plan in his 2012 State of the City Address:

”But something stands in the way of taking our system to the next level: a student assignment process that ships our kids to schools across our city. Pick any street. A dozen children probably attend a dozen different schools. Parents might not know each other; children might not play together. They can’t carpool, or study for the same tests. We won’t have the schools our kids deserve until we build school communities that serve them well” (Menino 2012).

School Assignment Reform

Since 2003, Mayor Menino had made several attempts to reform the three-zone plan. In December 2003, he appointed a student-assignment task force, which developed seven assignment plans, each with smaller zones. The proposed plans had 4, 6, or 12 zones for elementary schools, each of which covered a smaller geographic area. As a result, fewer buses would be needed to cover each school, thus reducing transportation costs and community dispersion; however, previous city committees had not approved any of the zone reforms because of concerns about equity of access to quality schools (Landsmark et al. 2004).

In April 2009, the mayor launched another initiative to review student assignments. BPS presented a five-zone plan; however, concerns over equity of access to quality schools (some zones had fewer high-quality schools) and unpopular proposed changes to some schools caused heated public opposition, which halted this attempt at reform.

In 2012, Mayor Menino again tried to reform the assignment plan. He appointed the EAC discussed above to develop a new plan. The EAC held community meetings to elicit families’ input, and requested BPS to provide data on supply and demand, school quality, student demographics, and choice patterns. In September 2012, BPS proposed new zone-based plans with 6, 9, 11, and 23 zones, and a plan that assigned each child to the closest school (Vaznis 2012). Shortly after the publication of these plans, an outside analysis

highlighted serious inequities in them in terms of access to quality schools, especially for children in the poorest neighborhoods (Burge 2012, Levinson et al. 2012). Doubts arose about whether Boston would be able to reform the three-zone system this time, or would have to abort this reform attempt, as it had abandoned its 2004 and 2009 tries.

Problem Formulation

From February to November 2012, I attended many EAC meetings and engaged parents, community organizations, activists, BPS staff, EAC members, and other academics. These interactions helped me to iteratively refine my understanding of the key issues, and eventually formulate a precise quantitative problem.

As I understood it, the primary flaw of the three-zone plan was that the choice menus provided families with too many options. For example, offering distant schools as choices meant that some children might have to be bused very long distances, contributing to the high transportation costs. Furthermore, having menus with a high number of options contributed to unpredictability because a child might be assigned to any of more than 25 schools; this also imposed a burden on families to research the many potential options. Moreover, the large number of options increased the likelihood that students from the same neighborhood would be scattered across the city.

Reducing the number of school options, however, might result in children living in inner-city neighborhoods having less access to schools that were perceived as higher quality, because the inner city was associated with lower-performing schools. In addition, because such neighborhoods also had a greater proportion of children of color, any policy change that might reduce access to quality schools for such students would increase racial inequity, which would be politically unacceptable.

Thus, the problem was how to allocate choice among various neighborhoods so that the total number of choices is not so high that it increases transportation costs and community dispersion, but gives students living near lower-performing schools enough choices to provide them with sufficient access to quality schools. We can view this as an optimization problem; the decision variables are the choice menus and priorities of students from various neighborhoods; the objective is a combination of equitable access to quality schools, predictability, proximity to home, community cohesion, racial and socioeconomic diversity in schools, and variety of choice. The challenge was how to quantify these concepts and how to relate the decision variables of choice menus and priorities to these outcomes.

The plan needed to be simple, within the capabilities of BPS to implement, and understandable to families. I learned to appreciate these constraints only after some trial and error. At a community meeting in October 2012, I proposed a solution that used a linear program to optimize for proximity to home and small choice menus, while guaranteeing equitable access to quality in a quantitatively precise way. Despite its theoretical appeal, this plan was perceived as too complex and too much of a black box. BPS viewed it as too technically complex to implement. This is understandable; its objective of educating children does not normally require a technically sophisticated operations research team. Some city staff members told me that for any plan to be implementable, I should be able to explain it to a fifth grader. Simpler solutions also tend to be more robust. A technically sophisticated solution may be optimal for a given model; however, in a real-world scenario, which often involves multiple and hard-to-quantify goals, human behavior and perception, and uncertain long-term ramifications, any model is at best a crude approximation. It may be more important for a solution to be simple and intuitive, thus allowing policy makers to take ownership of it and adapt it as necessary when new issues emerge, than for the solution to be optimal for a quantitative model.

I decided on the following approach to solve the problem:

1. Use previous choice data to fit and validate a demand model for how families would choose schools under new choice menus. (Because the student-proposing DA could not be influenced by any strategy, I assumed that the previous choice data reflected the true preferences of families.)
2. Precisely define outcome metrics and build a simulation engine that can forecast any system of choice menus and priorities.
3. Work with BPS to propose a set of plans that reduces the choice menus, protecting equity of access, and that BPS can implement. Use the simulation engine to explore variations and narrow the number of variations to a short list.
4. Present a detailed analysis of the plans to BPS and the EAC, showing differences by neighborhood, grade, student type, race, and socioeconomic status.

The final decision would still result in debates among various stakeholders to determine suitable trade-offs among various goals. As a result, the decision makers would more accurately and precisely understand the trade-offs, and would be able to discuss them in a more informed way.

Data and Analysis

The project used two data sources. The first was publicly available data posted on the BPS *Improving School Choice* website (<http://bostonschoolchoice.org/>), which BPS created as part of its public engagement for the 2012-2013 reform. The second was student-level data from BPS (BPS made the data anonymous, such that no student was identifiable). I received access to these data following an October 2012 meeting at which I presented to the EAC a rough analysis using aggregate data, and petitioned for access to student-level data to permit more refined modeling.

The publicly available data contained a table of school characteristics for each school, including its location, enrollment numbers as of December 2011, proportion of students of each race, English proficiency status and SPED status of each student, aggregate MCAS results for the previous two years, and the school's BPS rank, which was the key quality metric used during the reform. BPS computed this rank by ranking its schools using a weighted average of student performance levels and performance growth (i.e., a metric that measures the improvement in a student's test score relative to average improvement) in the previous two years' MCAS tests; it gave performance a weight of 67 percent and growth a weight of 33 percent.

BPS stored location data using an internal geocoding system that partitioned the city into 868 small regions, which it called geocodes. The average area of a geocode was about 0.1 square miles. The geocode of a student's home determined the student's zone and the schools in the student's walk zone.

The student-level data BPS provided were the Round 1 choice data for grades K0, K1, and K2 in 2010, 2011, and 2012. In each year and for each applicant, the data contained the student's application grade (i.e., K0, K1, or K2), geocode, race, lunch status (a proxy for socioeconomic status), and first-10 school choices. For each choice, the data included the school code, the program type (e.g., Reg. Ed., ELL, SPED), and the student's priority status relative to that school (e.g., continuing, sibling status, walk zone). For children whose families submitted rankings, the median number of choices was three, five, and four, for grades K0, K1, and K2, respectively. For those who ranked more than 10 school options, the data showed only the first 10; however, because 94 percent of the applications ranked fewer than 10 options, this was not a major issue.

Demand Modeling

The first component of the simulation model is a demand model—a model of how families would choose schools when faced with new choices. Demand modeling uses previous choice data to correlate families’ observable characteristics with the choice rankings they submitted; this predicts how families of various neighborhoods, races, and socioeconomic backgrounds would choose schools, given a new set of options. The predictions are statistical in nature; that is, they do not predict the choices of specific individuals, but only the demand patterns of subpopulations in a neighborhood. Below, the data available for building the demand model are listed.

- Distance: the walking distance from the student’s home to the school, estimated using the Google Maps application programming interface and the centroid of the student’s and the school’s geocodes as proxies for exact locations.
- Free-lunch, reduced-lunch, full-fare status: variables indicating the student’s lunch status. Students received free lunches if their family income was below 130 percent of the Federal poverty guidelines. Students received reduced-price lunches if their family income was between 130 and 180 percent of the Federal poverty guidelines.
- Student race: variables indicating race (e.g., black, white, Asian, Hispanic, other).
- Sibling, walk, continuing status: variables indicating whether the student had a sibling at the school, was in the school’s walk zone, or would be a continuing student at the school.
- East Boston: variable indicating whether the student and the school are both in East Boston, which is geographically isolated from the rest of Boston by bridges and tunnels that are inconvenient to cross.
- Racial proportion: percentage of students at the school who are black, white, Asian, or other. Hispanic is not included to avoid multicollinearity.

Parag Pathak and I worked on the demand modeling. We chose to include several combinations of the above variables in the model. The trade-off in choosing the right variables was that if we used more variables, the model might capture more nuanced effects; however, it might not generate as precise an estimate, because it must estimate more parameters using the same number of data points. We tried several combinations and chose one that provided a good trade-off. We validated the model by fitting a version of the model using data from the previous year to predict choice patterns in the next year. Called out-of-sample validation, this ensures that we are not over-fitting. The appendix shows details of the analysis.

Simulation

I used the demand model to build a simulation engine that took as input any assignment plan (i.e., any system of choice menus and priorities); it then used the actual student and capacity data from 2012 to forecast the 2012 outcome had the plan been implemented. In the simulation engine, for each student, I simulated his (her) utility in each program using the demand model, then computed the student's choice menu, and sorted the programs in the choice menu using the utilities, truncating to the top-10 schools according to the simulated utilities. These represented the student's choice submissions. I then independently drew a lottery number for each student and simulated the DA algorithm based on the priorities associated with the plan. In each simulation round, I estimated the following metrics for each student:

- Access to quality schools: the student's chance of receiving a lottery number that would enable that student to gain admittance to a school with an above-average BPS rank, and that the student would rank as a top-10 choice. This implicitly assumed that only the student's first 10 choices were acceptable to that student.

- Distance: the student’s walking distance to the assigned school.
- School-choice rank: the rank of the school choice obtained. For example, being admitted to one’s first-choice school corresponded to a choice rank of 1, and being admitted to one’s second-choice school corresponded to a rank of 2. This was undefined if the student was unassigned.
- Number of neighbors coassigned: the number of other students who lived within .5 miles of the student and who were assigned to the same grade and school. Distances between students’ homes were approximated by the Google Maps walk distance between the centroids of the geocodes of the homes.
- Access to top-three dream schools (defined as top-three choices without restricting the student’s menu): the student’s chance of receiving a lottery number that would enable that student to attend one of his (her) dream schools—the schools that student would have chosen if he (she) could have ranked any school in the BPS. Note that if the dream choices were not in the student’s menu, access would be zero.
- Percentage of peers of a specific race or lunch status: the percentage of other students of a specific race or lunch status who were assigned to the same program.

I averaged the above estimates over many independent rounds of simulation to compute an expected value for each student. In each simulation, the utilities and the lottery numbers were redrawn. For metrics that were only defined for students who were assigned, I computed the expected value, conditional on the student being assigned.

Using these student-level expected values, I computed the following aggregate statistics for the entire population.

- Minimum access to quality schools: over all students, the minimum of their expected access to quality schools. This corresponded to the welfare of the student with the worst access to quality schools.

- Median expected number of neighboring students coassigned.
- Minimum access to capacity: over all students, the minimum of their chances of being assigned to any school.
- Median access to a dream school.
- Average walk distance from student’s home to school.
- Median expected choice rank.
- Standard deviation across students of expected percentage of peers of each race or lunch status.

These metrics measured (1) equity of access to quality schools, (2) community cohesion, (3) supply and demand shortage, (4) availability of desirable-school choice, (5) proximity to home, (6) predictability of assignment, and (7) racial or socioeconomic segregation. For the first four metrics, a larger value is better; for the last three metrics, a smaller value is more desirable.

As a proxy for transportation costs, for each school, I computed the area a school bus would need to cover to pick up students; this was the difference between the area for which the school may appear in a student’s menu and the area of the school’s walk zone. I averaged this across all schools and called it the average bus coverage area.

Designing Simple Plans

For the reform, BPS decided to vary only the proposals for Reg. Ed. programs. For ELL and SPED programs, BPS committed to use a separate six-zone system, which it called ELL clusters and which it would apply on top of any assignment plan.

For the allocation of Reg. Ed. programs, BPS used the simulation engine to tweak its zone-based plan. After three iterations, it presented a 10-zone plan to the EAC. This plan was designed to balance the proportion of schools in each zone that BPS ranked in the top

50 percent. The EAC had previously decided to use these schools (i.e., highest 50 percent) as a proxy for quality schools. Based on EAC feedback, BPS also proposed a modified 11-zone plan. In both plans, a student’s menu consisted of the schools in his (her) zone and possibly additional schools within the one-mile walk zone.

I proposed two plans that did not use zones. In the first plan, which BPS called the home-based-A plan, each family could rank any school within each of the following sets:

- Any school within the one-mile walk zone
- The two closest schools ranked within the highest 25 percent (by BPS rank or another quality metric)
- The four closest schools ranked within the highest 50 percent
- The six closest schools ranked within the highest 75 percent
- The three closest capacity schools—schools to be designated by BPS in which excess capacity is available or in which capacity can be expanded cost effectively

The second plan, which BPS called the home-based-B plan, is similar to the home-based-A plan; however, the parameters two, four, and six (i.e., two, four, and six closest schools) were replaced by three, six, and nine.

I designed these plans based on the following criteria.

- Families should be able to rank all sufficiently close schools.
- Students living near lower-performing schools should be compensated with additional, higher-performing alternatives. Each student should have a sufficient number of options of various quality thresholds.
- Choice menus should include schools with sufficient capacity.
- Choice menus should be determined based on simple rules, which should allow the menus to adjust automatically when a school opens or closes, or its quality changes; otherwise, BPS would require another costly reform every few years.

The parameters for the home-based-A plan were chosen so that if quality were evenly distributed geographically and if capacity were widely available, the choice menus would essentially become the closest eight schools, in expectation that the choice menu would include two schools ranked in the highest 25 percent, four schools in the highest 50 percent, six schools in the highest 75 percent, and at least three schools within the set of the closest eight schools, which have excess capacity. However, quality schools were not yet distributed evenly; therefore, the menus were designed so that students who lived near the top schools would have much of their sets intersect, because a school in the top 25 percent school would simultaneously be in the top 50 percent and in the top 75 percent, and a student's options would essentially include the closest six schools. However, a student whose closest schools did not have a good BPS ranking might have to travel farther than the closest six schools to find a school ranked in the highest 25 percent or 50 percent of schools; therefore, that student would be given additional, although more distant, options as compensation. Hence, to maintain equitable access to quality schools in the long term, the plans provided some guarantee of the quality of the schools that students could rank, regardless of the current geographic distribution of the schools. Provided that the chosen quality metric averaged a few years' of data and did not fluctuate erratically, the adjustment in choice menu from year to year would also be smooth, because only a small subset of any family's menu would change. This was in contrast to zone-based plans, in which any small change in zone boundaries would drastically affect those living on the boundary. Furthermore, the home-based menus varied smoothly across geographies; however, in a zone-based plan, menus might change abruptly across zone boundaries, such that neighboring students might have very different choice menus. The capacity schools helped BPS meet supply and demand by pooling excess demand and directing them toward local centers of supply (i.e. larger schools with greater capacities or greater potential to expand).

This framework can be generalized to include the closest set of other types of schools, including regional magnets, early-learning centers, or advanced-work classes, if BPS were to decide to allow each family to rank such options.

Note that this plan guaranteed only the ability to rank certain schools, but did not guarantee placement. A student's chances of gaining admittance to a school would depend on his (her) priority and the level of competition for this school. I had proposed an earlier plan that guaranteed equity of probabilities of placement; however, that plan required solving a linear program and significantly changing the assignment process, and did not appeal to the EAC or BPS because of its complexity and black-box nature.

Results

Using the simulation engine, I analyzed the short-listed plans selected by the BPS: the status quo (3-zone), 10-zone, modified 11-zone, home-based-A, and home-based-B plans. I used 25 rounds of independent simulations for each. For the same round of simulation, I chose to use the same idiosyncratic taste shocks (i.e., random numbers drawn in the utility model to capture unexplained preferences) and lottery numbers across the plans, to ensure that any differences found would result primarily from the different menus, and not from different random numbers. I presented the results to BPS, who simplified them and presented them to the EAC. The most important student pool for analysis was K2 noncontinuing students with no siblings enrolled in a BPS school, because K2 was the main entry grade to elementary school, and most seats were assigned to this category. Moreover, the assignment of continuing students and siblings was essentially predetermined by the school at which the student was continuing or the school that the student's siblings attended, both of which were artifacts of the previous assignment system. Applying these filters, the relevant pool contained 1,659 students, 34 percent of whom were ELL students.

Table 1 shows the aggregate simulation results for these students in each of the short-listed plans. BPS created the 10-zone and modified 11-zone plans, and I proposed the home-based plans. Although no plan completely dominated the others, the home-based-A plan performed reasonably well in equity of access to quality and proximity to home, which were the most important criteria. The simulations considered all students; however, the statistics reported were for this pool.

The metrics that the EAC considered to be most important were minimum access to quality schools (i.e., to schools ranked in the 50th percentile), and average distance (i.e., walking distance from home to school). As Table 1 shows, in both of these metrics, all the new plans were better than the status quo. Furthermore, the new plans reduced the bus coverage area by a factor of three to four, and increased community cohesion by 20 to 34 percent, as measured by the number of neighboring children (i.e., children living within .5 mile) in the same grade and coassigned to the same school. The new plans, however, because of the smaller menus, were projected to increase supply challenges (because restricting choice menus reduced BPS' flexibility in assigning children), to reduce the variety of choice (because median access to dream-school choices was lower), and to decrease racial and socioeconomic diversity within classrooms. Nevertheless, BPS and the EAC considered the overall trade-offs to be positive: the supply challenges could be overcome by adding seats in constrained regions, and the reductions in choice and in diversity were reasonably mild. In terms of predictability, as measured by median rank obtained, the new zone-based plans and the home-based-A plan performed better than the status quo plan, and the home-based-B plan was worse; however, the differences were fairly minor.

The primary differences in the new plans were as follows: the home-based-A and home-based-B plans significantly outperformed the 10-zone and 11-zone plans based on minimum

Table 1 The table shows simulated performance of the assignment plans that the EAC considered for its final vote in February 2013.

Plan	Status quo	10-zone	Modified 11-zone	Home-based-A	Home-based-B
Minimum access to quality					
Top 33% of schools	14.6%	0.7%	0.9%	12.7%	13.2%
Top 50% of schools	19.5%	22.7%	22.7%	22.5%	27.8%
Top 67% of schools	31.4%	22.7%	22.7%	32.9%	32.9%
Median # of coassigned neighbor students	3.12	3.84	4.20	3.92	3.80
Minimum access to capacity	47.2%	31.8%	31.9%	36.0%	35.6%
Median access to dream school	42.4%	32.0%	31.4%	31.0%	32.1%
Average distance (miles)	2.03	1.24	1.19	1.25	1.30
Median rank obtained	2.83	2.65	2.62	2.78	2.91
Standard deviation across students					
% Peers: free lunch	9.0	11.5	11.5	11.2	11.1
% Peers: black	14.1	15.4	15.5	15.5	15.4
% Peers: white	9.4	12.2	12.2	11.2	11.0
% Peers: Hispanic	17.7	19.2	19.3	18.5	18.2
% Peers: Asian	9.0	10.0	10.0	10.3	10.4
Bus coverage area per school (sq. miles)	24.71	6.96	6.41	6.58	8.63

access to quality (i.e., top one-third and top two-thirds) schools. Although the zone-based plans were optimized for the one threshold (i.e., highest 50 percent), the home-based plans considered a range of thresholds. This illustrated the brittleness of the zone-based plans in maintaining equity: they only worked for the one metric for which they were optimized. If quality changed significantly in the future, these plans could not guarantee that equity would be maintained without modifying the zone boundaries, which might require another reform.

The initial report contained simulation results broken down by grade, race, neighborhood, lunch status, and English-learner status. It also included sensitivity analysis on the percentage of seats at each school for which to apply walk-zone priorities (Pathak and Shi 2013). After the simulation results were made public, the EAC met several times to deliberate the various trade-offs, and held several community-engagement meetings to elicit feedback from the public. Parent groups, community organizations, activist groups, and local and state politicians participated in these discussions and voiced comments.

On February 25, 2013, the EAC met to take its the final vote on the five plans shown in Table 1. The EAC overwhelmingly favored the home-based-A plan: of 23 members present for the vote, 20 voted for this plan, 1 voted for the home-based-B plan, and 2 did not take a position by voting present. According to the EAC’s recommendation letter to the superintendent of the BPS, it choose this plan because it “ensures every family has high-quality schools on its list of options,” and that “it adapts to changes in quality over time” (EAC 2013, p. 4). In a follow-up statement by the BPS superintendent, additional reasons behind the EAC decision included increased predictability, increased equity of access, significantly lower travel distances, more opportunities for families to understand their choices (because of the reduced menus), greater community cohesion, elimination of the zone system, and the association of choices with school quality (Johnson 2013).

On March 13, 2013, the Boston school committee approved the home-based-A plan for implementation for the 2014 assignment cycle (Seelye 2013, Vaznis 2013).

Discussion

Following the EAC’s approval of this plan, BPS made several modifications. First, it removed the walk-zone priority, because the schools on the menus were already close to the students’ homes. Second, it changed the capacity schools to be essentially the lowest 25 percent of the schools based on BPS rank, because BPS predicted these schools would have excess capacity. Third, it expanded the menus to include the closest early-learning center (i.e., full-day kindergartens) and the closest schools with advanced-work classes. In addition, it made available several more schools citywide or for specific neighborhoods. This illustrated the value of having a sufficiently simple and flexible plan that the host organization could own and adapt to its needs.

This project highlighted the power of quantitative analysis and simulation to move debates forward by making the objects of discussion precise and by quantifying trade-offs. For example, prior to this analysis, participants in the reform vigorously debated the pros and cons of having a greater percentage of seats with walk-zone priority to help students attend schools closer to home, or a smaller percentage to grant greater equity of access. Both were important objectives; therefore, such debates often evolved to gridlock situations. Perhaps surprisingly, the simulation results showed that the walk-zone priority, as it was implemented in the three-zone plan, did not make a significant difference for the majority of schools, because a greater proportion of families applied from within the walk zone and filled up the 50 percent of seats set aside for walk-zone applicants. After seeing these results, the participants in these discussions were able to bypass this false dichotomy and perceive the actual trade-offs.

One important question this project unfortunately could not address was the long-term impacts of the reform. The demand model was based on choice data collected before the reform; this data might not accurately reflect how people would choose schools following the reform. For example, the reform itself or the new plan’s presentation might significantly alter the behavior of families. Parag and I analyzed the validity of using data collected prior to the reform to build a demand model in a follow-up project (Pathak and Shi 2014), in which we make predictions using these data and the models fitted, and evaluate the prediction accuracy using data collected after the reform. Other issues not yet addressed include complex interactions between the assignment plan and qualities of schools in the future, as well as how the assignment plan may affect student learning and performance, which one might argue are the most important results. A conclusive study of the long-term impacts of the reform is possible only after several years of postreform data become available.

During this project, I constrained the solution to a specific level of simplicity; in the future, increased complexity might be possible if societal tolerance for complexity increases and computer technologies can shield end users from some complexities. For example, BPS chose to implement the home-based-A plan by directing families to a Web applet; the user entered an address and the applet showed a map that includes the choice menu. The families did not need to directly construct their menus using the home-based-A plan rules. Hence, provided that families trust this black box, we may hope to implement a plan that is more sophisticated and better optimized using the same method of presentation. Two follow-up papers explore these possibilities. Ashlagi and Shi (2013) explore an optimization-based lottery-correlation procedure to improve community cohesion, and Ashlagi and Shi (2014) solve a convex optimization problem to obtain the optimal choice menus and priorities to

maximize a weighted combination of efficiency and equity, while staying within an expected busing budget.

The modeling and simulation techniques in this paper and the optimization procedures in these follow-up papers can be applied to school choice in other cities and to other areas of public service allocation; examples include allocating subsidized housing, high-demand college courses, and office space.

Appendix. Demand Modeling Technical Details

We modeled the family choice process using a multinomial logit discrete choice model with school and program-type fixed effects. Let j be a program, with $j = (s, p)$, where s denotes the school and p the program type. We assumed that a student’s choice submission was driven by unobserved utilities of the form,

$$u_{ij} = q_j + \vec{\beta} \cdot \vec{z}_{ij} + \epsilon_{ij},$$

where u_{ij} represents student i ’s utility to program j , q_j represents the quality or desirability of program j , \vec{z}_{ij} is a vector of observed data for the student-program pair, $\vec{\beta}$ represents how these features impact the decision, and ϵ_{ij} represents idiosyncratic preference for the program and is assumed to be independently distributed as a standard Gumbel distribution; the reason for this distribution was to allow tractable estimation. We also assumed the quality term to be additively decomposable as a term for the school and a term for the program type, $q_j = a_s + b_p$. We called \vec{a} the school fixed effects and \vec{b} the program-type fixed effects. We estimated the parameters \vec{a} , \vec{b} , $\vec{\beta}$ using maximum likelihood. Beggs et al. (1981) and Hausman and Ruud (1987) introduced multinomial logit models for ranked-order data, and the estimation procedure is now standard. For more details on the estimation, see the appendix in Pathak and Shi (2013).

Choosing which features to include in the model represents a trade-off between model flexibility and estimation precision. A model with more parameters may be able to capture more nuanced effects; however, estimating more parameters using the same amount of data entails that each parameter may not be estimated with the same precision. We experimented with several models and used out-of-sample testing to determine the final model. We tried four sets of parameters, and we called the sets Simple, Simple2, Medium, and Medium2. The lists below show the features we included in each of the four sets. In the *Demand Modeling* section, we include an explanation of these variables.

- Simple: distance, sibling, walk, continuing.
- Simple2: all features in Simple plus the square root of distance.
- Medium: all features in Simple2, plus pairwise product of the student’s race and the school’s racial demographics, plus the pairwise product of the student’s lunch status and the school’s percentage of students who receive free lunches.
- Medium2: all features in Medium2, plus the East Boston indicator (i.e., whether both the student and the school are in East Boston), plus the product of distance and the student’s lunch status, plus the product of the square root of distance and the student’s lunch status.

Note that the models are nested. Moving from Simple to Simple2, we allow for nonlinear preferences for distance. The square root is to capture the intuition that although a school near a student’s home may be more desirable than a school one mile away, a school that is four miles away may not seem much better than a school that is five miles away, because the child will be on a bus for a long time in either case. Moving from Simple2 to Medium, we allow students of different races and socioeconomic status to have different preferences for aspects of the school correlated with the school’s racial or socioeconomic composition. Such aspects may include school climate, safety, and culture. Moving from Medium to Medium2, we allow differential preferences for distances for students of different socioeconomic status, as well as potential preferences for East Boston students to not have to cross the frequently congested bridges or tunnels to attend school.

Because the school characteristics were for December 2011, we used only the choice data in the years 2011 and 2012. We designated 2011 as the in-sample data set and 2012 as the out-of-sample data set. We considered only grades K1 and K2, because most school programs started in K1 and only a few seats were allocated in K0. To select the model, we fit the models using 2011 data, and considered the Bayesian information criteria (BIC), which takes into account the goodness of fit and penalizes having too many parameters. In our case, this criteria persuaded us to select the model with more parameters if the incremental gain in log likelihood divided by the number of additional variables was more than $\frac{\ln \# \text{ of observations}}{2} = \frac{\ln 20533}{2} = 4.96$. Moving from Simple to Simple2 and from Simple2 to Medium, the gains per variable were 144 and 24, respectively, which significantly exceeded the threshold of 4.96; however, moving from Medium to Medium2, the gain was 7.1, which was greater than the threshold, but not by much. Furthermore, many of the additional variables in Medium2 turned out to be statistically insignificant. We ultimately chose to be conservative in the number of parameters to include; therefore, we chose the Medium model. We also fit this model using the out-of-sample

Table 2 The table shows estimated coefficients of demand models used to predict how families living in different areas would choose between different schools. Each model uses a different set of variables. The numbers without parenthesis represent the impact of each variable and the numbers in parenthesis are the standard errors (or uncertainty). The absolute value of the impacts are hard to interpret in isolation; however, greater magnitude implies greater relative importance. We found the Medium model to be superior and used it for out-of-sample

testing.

<u>Variable</u>	<u>Model</u>				Out-of-Sample <u>Medium</u>
	<u>Simple</u>	<u>Simple2</u>	<u>Medium</u>	<u>Medium2</u>	
Distance	-0.481 (0.008)	-0.139 (0.021)	-0.152 (0.021)	-0.169 (0.037)	-0.115 (0.017)
Distance x free-lunch				0.055 (0.047)	
Distance x reduced-lunch				0.159 (0.107)	
Distance x full-fare				-0.031 (0.072)	
Sqrt(distance)		-1.198 (0.069)	-1.115 (0.069)	-1.140 (0.111)	-1.082 (0.057)
Sqrt(distance) x free-lunch				0.066 (0.145)	
Sqrt(distance) x reduced-lunch				-0.270 (0.33)	
Sqrt(distance) x full-fare				-0.099 (0.217)	
Sibling	2.984 (0.05)	2.958 (0.05)	2.888 (0.05)	2.891 (0.05)	2.860 (0.046)
Walk	0.350 (0.023)	0.170 (0.026)	0.154 (0.026)	0.161 (0.026)	0.262 (0.022)
Continuing	5.945 (0.085)	5.938 (0.085)	5.884 (0.085)	5.896 (0.085)	5.261 (0.068)
East-Boston				0.293 (0.142)	
School % race x student race	N	N	Y	Y	Y
School % free lunch x student lunch status	N	N	Y	Y	Y
<u>Summary Statistics</u>					
Log likelihood	-49,486	-49,342	-48,766	-48,716	-68,798
Number of parameters	88	89	112	119	112
Number of students	5,758	5,758	5,758	5,758	6,696
Number of choices	20,533	20,533	20,533	20,533	27,899

data, to test for stability of coefficients. Table 2 shows the coefficient estimates of all four models using in-sample data, as well as the chosen model using out-of-sample data. Each model uses a different set of variables. The numbers without parentheses represent the impact of each variable and the numbers within parentheses are the standard errors (or uncertainty). The absolute value of the impacts are hard to interpret in isolation; however, greater magnitude implies greater relative importance. We found the Medium model to be superior and used it for out-of-sample testing.

To validate the Medium model, we compared the forecasted relative market shares of programs as predicted by the demand model fitted using the 2011 data with the actual market shares in 2012. We did this for the highest one, three, five, and seven choices; having several cutoffs gave us a picture of the goodness of our prediction for the entire choice list. The relative top k market share of a program is defined as the proportion of the students' top k choices for this program. Figure 2 shows the results. In the plots, each point corresponds to a program in K1 or K2. Its x -value corresponds to the predicted market share, and its y -value to the actual market share. If predictions were perfect, then all points would lie on the 45 degree line. As the plot shows, although the predictions were not perfect, the majority of points lie close to the 45 degree line, indicating that the predictions were reasonable. Moreover, the predictions were best for the first choices, and less accurate for other choices. An explanation of this is that a multinomial logit model implies a certain type of substitution patterns as we go down the choice list; however, the actual data might not satisfy these patterns. In the literature, this substitution pattern is called independence of irrelevant alternatives (IIA); see McFadden et al. (1977).

Our demand model does not contain outside options, because we had little data on families' preferences for nonBPS alternatives, such as private, charter, parochial, or other schools. We decided to simply assume that students all ranked up to 10 options, which was the maximum number contained in our input data. We did sensitivity analyses on the number of ranked options, and found that 10 was a reasonable choice. The details are in the appendix of Pathak and Shi (2013).

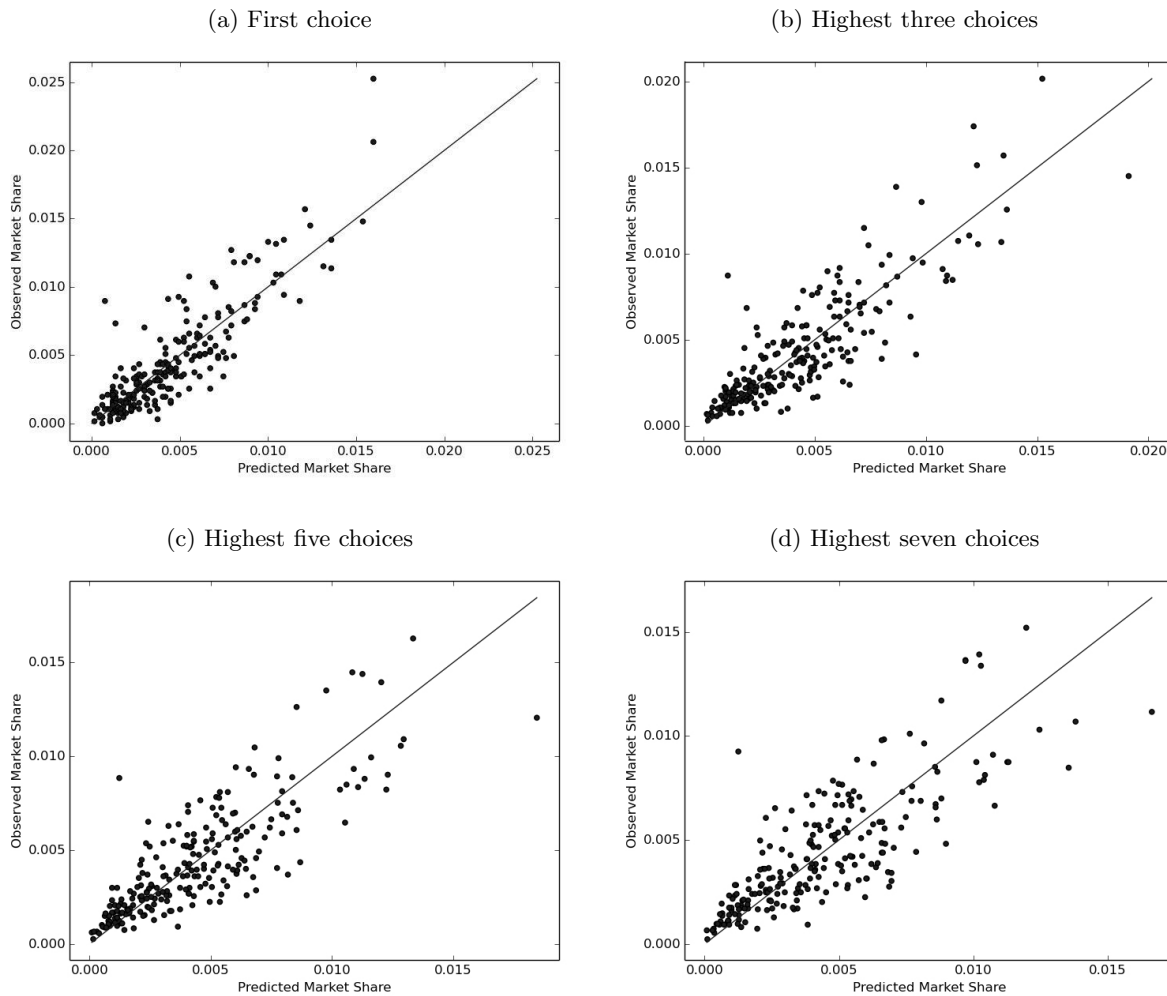
Acknowledgments

I acknowledge the kind and helpful support of many parents and community members, of BPS staff members Carleton Jones, Timothy Nicolette, Peter Sloan, Kamal Chavda, and Jack Yessayan, of BPS superintendent Carol Johnson, of academic colleagues Parag Pathak and Tayfun Sönmez, and of my advisor Itai Ashlagi during this project.

References

- Abdulkadiroğlu A, Pathak PA, Roth AE, Sönmez T (2006) Changing the Boston school choice mechanism. Boston College Working Papers in Economics 639. Boston College Department of Economics, Boston.
- Abdulkadiroğlu A, Sönmez T (2003) School choice: A mechanism design approach. *Amer. Econom. Rev.* 93(3):729–747.

Figure 2 These plots compare the relative market shares of programs as predicted by the demand model fitted using 2011 data with the actual market shares in 2012. Although the predictions were not perfect, they were reasonably good, because most of the dots are near the 45 degree line.



Alves MJ, Willie CV (1990) Choice, decentralization, and desegregation: The Boston 'controlled choice plan'.

Choice and Control in American Education, Volume 2: The Practice of Choice, Decentralization, and School Restructuring (The Falmer Press, New York), 17–75.

Ashlagi I, Shi P (2013) Improving community cohesion in school choice via correlated-lottery implementation.

Accessed August 4, 2014, <http://www.mit.edu/~pengshi/papers/community-cohesion.pdf>.

Ashlagi I, Shi P (2014) Optimal allocation without money: An engineering approach. Accessed April 30,

2014, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2428449.

- Beggs S, Cardell S, Hausman J (1981) Assessing the potential demand for electric cars. *J. Econometrics* 17(1):1–19.
- Boston Public Schools (2013) Preliminary transportation analysis. Accessed February 4, 2014, <http://bostonschoolchoice.files.wordpress.com/2013/02/2013-02-20-transportation-cost-analysis-final.pdf>.
- Burge K (2012) Study finds inequities in schools' zone plans. Accessed October 30, 2012, <http://www.bostonglobe.com/metro/2012/09/30/boston-school-redistricting-proposals-will-create-more-inequity-according-harvard-analysis/eLEAD0QyNBQUTr18yQHV2L/story.html>.
- Dubins LE, Freedman DA (1981) Machiavelli and the Gale-Shapley algorithm. *Amer. Math. Monthly* 88(7):485–494.
- EAC (2013) Recommendations by the external advisory committee on school choice. Accessed March 31, 2013, <http://bostonschoolquality.files.wordpress.com/2013/04/eac-recommendations-final-2-27-2013.pdf>.
- Ebbert S, Russell J (2011) A daily diaspora, a scattered street. Accessed June 4, 2012, http://www.boston.com/news/education/k_12/articles/2011/06/12/on_one_city_street_school_choice_creates_a_gap/.
- Gale D, Shapley LS (1962) College admissions and the stability of marriage. *Amer. Math. Monthly* 69(1):9–14.
- Hausman JA, Ruud PA (1987) Specifying and testing econometric models for rank-ordered data. *J. Econometrics* 34(1):83–104.
- Johnson C (2013) Superintendent's statement on external advisory committee on school choice recommendation. Accessed March 31, 2013, <http://bostonschoolchoice.files.wordpress.com/2013/03/2-27-13-superintendent-eac-statement.pdf>.
- Landsmark T, Dajer H, Gonsalves L, Acevedo S, Bakken M, Chu H, Groover G, et al. (2004) Report and recommendations of the boston public schools. Technical report, 2004 Student Assignment Task Force, Boston.

- Levinson M, Noonan J, Fay J, Mantil A, Johnston W, Buttimer C, Mehta J (2012) First-round analysis of BPS proposed 6-zone, 9-zone, 11-zone, and 23-zone school assignment plans. Technical report, Harvard Graduate School of Education, Boston.
- Lilly D (2009) It's back to neighborhood schools for Seattle. Accessed April 30, 2012, <http://crosscut.com/2009/10/07/seattle-schools/19286/Its-back-neighborhood-schools-for-Seattle/>.
- McFadden D, Talvitie A, Cosslett S, Hasan I, Johnson M, Reid F, Train K (1977) Demand model estimation and validation. Institute of Transportation Studies, University of California, Berkeley.
- Menino T (2012) The honorable mayor Thomas M. Menino state of the city address. Accessed January 31, 2012, <https://data.cityofboston.gov/dataset/State-Of-The-City-Address-2012-Mayor-Thomas-Menino/7w3j-67fs>.
- Pathak PA, Shi P (2013) Simulating alternative school choice options in Boston. Massachusetts Institute of Technology, School Effectiveness and Inequality Initiative, Cambridge, MA.
- Pathak PA, Shi P (2014) Demand modeling, forecasting, and counterfactuals, part i. Technical report, National Bureau of Economic Research, Cambridge, MA.
- Roth AE (1982) The economics of matching: Stability and incentives. *Math. Oper. Res.* 7(4):617–628.
- Roth AE, Sotomayor MAO (1990) *Two-sided Matching: A Study in Game-theoretic Modeling and Analysis* (Cambridge University Press, Cambridge, MA).
- Russell J, Ebbert S (2011) The high price of school assignment. Accessed February 4, 2012, http://www.boston.com/news/education/k_12/articles/2011/06/12/the_high_price_of_school_assignment/.
- Russell J, Ebbert S, Vaznis J (2011) Getting in: Inside Boston's school assignment maze. Accessed February 7, 2012, http://www.boston.com/news/education/specials/school_chance/articles/.
- Seelye KQ (2013) Boston Schools Revise '70s Busing System. *New York Times* (March 15) A17.
- Vaznis J (2012) 5 new plans to assign boston students. *Boston Globe* (September 24) A1.
- Vaznis J (2013) Boston adopts new school assignment plan. *Boston Globe* (March 14) A1.
- Woodward M (2011) Southeast Seattle schools: World war II to present. Accessed August 30, 2012, <http://www.seattle.gov/neighborhoods/preservation/southeastseattle>.