

Leveling the Playing Field: Sincere and Strategic Players in the Boston Mechanism*

Parag A. Pathak[†] and Tayfun Sönmez[‡]

This version: August 2006

Abstract

Motivated by empirical evidence on different levels of sophistication among students in the Boston Public School (BPS) student assignment plan, we analyze the Nash equilibria of the preference revelation game induced by the Boston mechanism when there are two types of players. Sincere players are restricted to report their true preferences, while strategic players play a best response. The set of Nash equilibrium outcomes is characterized in terms of the set of stable matchings of an economy with a modified priority structure, where sincere students lose their priority to sophisticated students. While there are multiple equilibrium outcomes, a sincere student receives the same assignment in all equilibria. Finally, the assignment of any strategic student under the Pareto-dominant Nash equilibrium of the Boston mechanism is weakly preferred to her assignment under the student-optimal stable mechanism, which was recently adopted by BPS for use in 2006.

*Still preliminary. For financial support, Pathak thanks the National Science Foundation, the Spencer Foundation, and the Division of Research at Harvard Business School and Sönmez thanks the National Science Foundation. Fuhito Kojima and Alvin E. Roth provided helpful discussions. We are grateful to seminar participants at the Harvard theory lunch for comments.

[†]Harvard University, e-mail: ppathak@fas.harvard.edu

[‡]Boston College, e-mail: sonmezt@bc.edu

1 Introduction

In May 2005, Thomas Payzant, the Superintendent of Boston Public Schools, recommended to the public that the existing school choice mechanism in Boston (henceforth the **Boston mechanism**) should be replaced with an alternative mechanism that removes the incentives to “game the system” that handicapped the Boston mechanism. Following his recommendation, the Boston School Committee voted to replace the mechanism in July 2005 and adopt a new mechanism for the 2005-06 school year.

Over 75,000 students have been assigned to public schools in Boston under the mechanism before it was abandoned. The Boston mechanism is also widely used to allocate seats at public schools at several school districts throughout the US including Cambridge MA, Charlotte-Mecklenburg, Denver, Miami-Dade, Minneapolis, Rochester NY, Tampa-St. Petersburg, and White Plains NY.

The major difficulty with the Boston mechanism is that students may find it in their interest to submit a rank order list that is different from their true underlying preferences. Loosely speaking, the Boston mechanism attempts to assign as many students as possible to their first choice school, and only after all such assignments have been made does it consider assignments of students to their second choices, etc. The problem with this is that if a student is not admitted to her first choice school, her second choice may already be filled with students who listed it as their first choice. That is, a student may fail to get a place in her second choice school that would have been available had she listed that school as her first choice. If a student is willing to take a risk with her first choice, then she should be careful to rank a second choice that she has a chance of obtaining.

Some parents and families understand these features of the Boston mechanism and have developed rules of thumb for how to submit preferences. For instance, the West Zone Parents Group (WZPG), a well-informed group of approximately 180 members who meet regularly prior to admissions time to discuss Boston school choice at the K2 level, recommends two types of strategies to its members. Their introductory meeting minutes on 10/27/2003 state:

One school strategy is to find a school you like that is undersubscribed and put it as a top choice, OR, find a school that you like that is popular and put it as a first choice and find a school that is less popular for a “safe” second choice.

Using data on stated choices from Boston Public Schools from 2000-2004, Abdulkadiroğlu, Pathak, Roth and Sönmez (2006) describe several empirical patterns which suggest that there are different levels of sophistication among participants in the mechanism. Some fraction of

parents behave as the WZPG suggest, and avoid ranking two overdemanded schools, especially when they do not receive priority at their top choice school. On the other hand, nearly 20% of students list two overdemanded schools, and 27% of these students end up unassigned. This empirical evidence, together with the theoretical evidence in Abdulkadiroğlu and Sönmez (2003) and the experimental evidence in Chen and Sönmez (2006) was instrumental in the decision to replace the Boston mechanism with the **student-optimal stable mechanism** (Gale and Shapley 1962).

One of the remarkable properties of the student-optimal stable mechanism is that it is strategy-proof: truth-telling is a dominant strategy for each student under this mechanism (Dubins and Freedman 1981, Roth 1982). Whether a family has access to advice on how to strategically modify their rank order list, either through the WZPG or through a family resource center, does not matter under the new mechanism. This feature was an important factor in the Superintendent Payzant’s recommendation to change the mechanism. The BPS Strategic Planning team, in their 05/11/2005 dated recommendation to implement a new BPS assignment algorithm, emphasized:¹

A strategy-proof algorithm “levels the playing field” by diminishing the harm done to parents who do not strategize or do not strategize well.

In this paper, our aim is to formalize the intuitive idea that replacing the highly manipulable Boston mechanism with the strategy-proof student-optimal stable mechanism “levels the playing field.” To do so we consider a model with both sincere and strategic families,² analyze the Nash equilibria of the preference revelation game induced by the Boston mechanism (or simply the the Nash equilibria of the **Boston game**), and compare the equilibrium outcomes with the dominant-strategy outcome of the student-optimal stable mechanism. In Proposition 1, we characterize the equilibrium outcomes of the Boston game in terms of the set of stable matchings of a modified problem where sincere students lose their priorities to strategic students. This result implies that there exists a Nash equilibrium outcome where each student weakly prefers her assignment to any other equilibrium assignment. Hence, the Boston game is a coordination game among strategic students.

We next examine properties of equilibrium. Sincere students may gain priority at a school at the expense of another sincere student by ranking the school higher on their preference list

¹See Recommendation to Implement a New BPS Algorithm - May 11, 2005 at <http://boston.k12.ma.us/assignment/>.

²This is consistent with the experimental findings of Chen and Sönmez (2006) who have shown that about 20% of the subjects in the lab utilize the suboptimal strategy of truth-telling under the Boston mechanism.

than other sincere students. This will allow for some sincere students to benefit from the Boston mechanism, at the expense of other sincere students. In Proposition 2, we show that a sincere student receives the same assignment at all equilibria of the Boston game.

Finally in Proposition 3, we compare the Boston game to the dominant strategy outcome of the student-optimal stable mechanism. We show that any strategic student weakly prefers her assignment under the Pareto-dominant Nash equilibrium outcome of the Boston game over the dominant-strategy outcome of the student-optimal stable mechanism. When only some of the students are strategic, the Boston mechanism gives a clear advantage to strategic students provided that they can coordinate their strategies at a favorable equilibrium. This result might explain why, in testimony from the community about the Boston mechanism on 06/08/2005, the leader of the WZPG opposed changing the mechanism:

There are obviously issues with the current system. If you get a low lottery number and don't strategize or don't do it well, then you are penalized. But this can be easily fixed. When you go to register after you show you are a resident, you go to table B and the person looks at your choices and lets you know if you are choosing a risky strategy or how to re-order it.

Don't change the algorithm, but give us more resources so that parents can make an informed choice.

Moreover, this result stands in contrast to Ergin and Sönmez (2006), who show that the set of Nash equilibrium outcomes of the Boston game is equivalent to the set of stable matchings of the underlying problem when all students are strategic.³ Their result implies that a replacement of the Boston mechanism with the student-optimal stable mechanism should not be opposed since it is in all students' best interest.

The layout of the paper is as follows. Section 2 defines the model and Section 3 characterizes the set of equilibrium. Section 4 identifies comparative statics and Section 5 concludes.

2 The Model

In a **school choice problem** (Abdulkadiroğlu and Sönmez 2003) there are a number of students each of whom should be assigned a seat at one of a number of schools. Each student has a strict preference ordering over all schools as well as remaining unassigned and each school has a strict priority ranking of all students. Each school has a maximum capacity.

³Kojima (2006) extends this result to a model with substitutable (Kelso and Crawford 1981) priorities.

Formally, a school choice problem consists of:

1. a set of students $I = \{i_1, \dots, i_n\}$,
2. a set of schools $S = \{s_1, \dots, s_m\}$,
3. a capacity vector $q = (q_{s_1}, \dots, q_{s_m})$,
4. a list of strict student preferences $P_I = (P_{i_1}, \dots, P_{i_n})$, and
5. a list of strict school priorities $\pi = (\pi_{s_1}, \dots, \pi_{s_m})$.

For any student i , P_i is a strict preference relation over $S \cup \{i\}$ where $sP_i i$ means student i strictly prefers a seat at school s to remaining unassigned. For any student i , let R_i denote the “at least as good as” relation induced by P_i . For any school s , the function $\pi_s : \{1, \dots, n\} \rightarrow \{i_1, \dots, i_n\}$ is the priority ordering at school s where $\pi_s(1)$ indicates the student with highest priority, $\pi_s(2)$ indicates the student with second highest priority, and so on. Priority rankings are determined by the school district and schools have no control over them. We fix the set of students, the set of schools and the capacity vector throughout the paper; hence the pair (P, π) denotes a school choice problem (or simply an **economy**).

The school choice problem is closely related to the well-known college admissions problem (Gale and Shapley 1962). The main difference is that in college admissions each school is a (possibly strategic) agent whose welfare matters, whereas in school choice each school is a collection of indivisible goods to be allocated and only the welfare of students is considered.

The outcome of a school choice problem, as in college admissions, is a **matching**. Formally a matching $\mu : I \rightarrow S \cup I$ is a function such that

1. $\mu(i) \notin S \Rightarrow \mu(i) = i$ for any student i , and
2. $|\mu^{-1}(s)| \leq q_s$ for any school s .

We refer $\mu(i)$ as the assignment of student i under matching μ .

A matching μ **Pareto dominates** (or is a **Pareto improvement** over) a matching ν , if $\mu(i)R_i\nu(i)$ for all $i \in I$ and $\mu(i)P_i\nu(i)$ for some $i \in I$. A matching is **Pareto efficient** if it is not Pareto dominated by any other matching.

A mechanism is a systematic procedure that selects a matching for each economy.

The Boston Student Assignment Mechanism

For any economy, the outcome of the Boston mechanism is determined in several rounds with the following procedure:

Round 1: In Round 1, only the first choices of students are considered. For each school, consider the students who have listed it as their first choice and assign seats of the school to these students one at a time following their priority order until there are no seats left or there is no student left who has listed it as their first choice.

In general, at

Round k : Consider the remaining students. In Round k , only the k^{th} choices of these students are considered. For each school with still available seats, consider the students who have listed it as their k^{th} choice and assign the remaining seats to these students one at a time following their priority order until there are no seats left or there is no student left who has listed it as his k^{th} choice.

The procedure terminates when each student is assigned a seat at a school.

As any mechanism, the Boston mechanism induces a preference revelation game among students. We refer this game as the **Boston game**.

Sincere and Strategic Students

Motivated by the empirical observations in Abdulkadiroğlu, Roth, Pathak and Sönmez (2006), we assume that there are two types of students: sincere and strategic. Let N, M denote sets of sincere and strategic students, respectively. We have $N \cup M = I$ and $N \cap M = \emptyset$. Sincere students are unaware about the strategic aspects of the student assignment process and they simply reveal their preferences truthfully. The strategy space of each sincere student is a singleton under the Boston game and it consists of truth-telling. Each strategic student, on the other hand, recognizes the strategic aspects of the student assignment process, her strategy space is all strict preferences over the set of schools plus remaining unassigned. Each strategic student selects a best response to the other students. We focus on the Nash equilibria of the Boston game where only strategic students are active players.

Stability

The following concept that plays a central role in the analysis of two-sided matching markets will be useful to characterize the Nash equilibria of the Boston game.

A matching μ is **stable** (Gale and Shapley 1962) if

1. it is **individually rational** in the sense that there is no student i who prefers remaining unassigned to her assignments $\mu(i)$, and
2. there is no student-school pair (i, s) such that,
 - (a) student i prefers s to her assignment $\mu(i)$, and
 - (b) either school s has a vacant seat under μ or there is a lower priority student j who nonetheless received a seat at school s under μ .

Gale and Shapley (1962) show that the set of stable matchings is non-empty and there exists a stable matching, the **student-optimal stable matching**, that each student weakly prefers to any other stable matching. We refer the mechanism that selects this stable matching for each problem as the **student-optimal stable mechanism**. Dubins and Freedman (1981) and Roth (1982) show that for each student, truth telling is a dominant strategy.

An Illustrative Example

Since a student “loses” her priority to students who rank a school higher in their preferences, the outcome of the Boston mechanism is not necessarily stable. However, Ergin and Sönmez (2006) have shown that any Nash equilibrium outcome of the Boston game is stable when all students are strategic. Based on this result they have argued that a change from the Boston mechanism to the student-optimal stable mechanism shall be embraced by all students for it will result in a Pareto improvement. This is not what happened in summer 2005 when Boston Public Schools gave up the Boston mechanism and adopted the student-optimal stable mechanism. The leader of an organized parents group in Boston, the West Zone Parents Group (WZPG), publicly opposed the change in the mechanism. A simple example can illustrate how Ergin and Sönmez (2006) result is affected when only part of the students are strategic and provide an explanation to the resistance of the WZPG to the change of the mechanism.

Example 1. There are three schools a, b, c each with one seat and three students i_1, i_2, i_3 . The priority list $\pi = (\pi_a, \pi_b, \pi_c)$ and student utilities representing their preferences $P = (P_{i_1}, P_{i_2}, P_{i_3})$ are as follows:

| | | | | |
|-----------|----------|----------|----------|---------------------------|
| | <i>a</i> | <i>b</i> | <i>c</i> | |
| u_{i_1} | 1 | 2 | 0 | $\pi_a : i_2 - i_1 - i_3$ |
| u_{i_2} | 0 | 2 | 1 | $\pi_b : i_3 - i_2 - i_1$ |
| u_{i_3} | 2 | 1 | 0 | $\pi_c : i_2 - i_3 - i_1$ |

Students i_1 and i_2 are strategic whereas student i_3 is sincere. Hence the strategy space of each of students i_1, i_2 is $\{abc, acb, bac, bca, cab, cba\}$ whereas the strategy space of student i_3 is the singleton $\{abc\}$. Hence we have the following $6 \times 6 \times 1$ Boston game for this simple example:

| | <i>abc</i> | <i>acb</i> | <i>bac</i> | <i>bca</i> | <i>cab</i> | <i>cba</i> |
|------------|------------|------------|----------------|----------------|------------|------------|
| <i>abc</i> | (0,0,1) | (0,0,1) | (1,2,0) | (1,2,0) | (1,1,1) | (1,1,1) |
| <i>acb</i> | (0,0,1) | (0,0,1) | (1,2,0) | (1,2,0) | (1,1,1) | (1,1,1) |
| <i>bac</i> | (2,0,0) | (2,0,0) | (0,2,2) | (0,2,2) | (2,1,2) | (2,1,2) |
| <i>bca</i> | (2,0,0) | (2,0,0) | (0,2,2) | (0,2,2) | (2,1,2) | (2,1,2) |
| <i>cab</i> | (0,0,1) | (0,0,1) | (0,2,2) | (0,2,2) | (0,2,2) | (0,2,2) |
| <i>cba</i> | (0,0,1) | (0,0,1) | (0,2,2) | (0,2,2) | (0,2,2) | (0,2,2) |

where the row player is student i_1 and the column player is student i_2 .

There are four Nash equilibrium profiles of the Boston game (indicated in boldface) each with a Nash equilibrium payoff of (1,2,0) and a Nash equilibrium outcome of

$$\mu = \begin{pmatrix} i_1 & i_2 & i_3 \\ a & b & c \end{pmatrix}.$$

We have the following useful observations about the equilibria:

1. Truth-telling is not a Nash equilibrium of the Boston game.
2. Unlike in Ergin and Sönmez (2006), the Nash equilibrium outcome μ is not a stable matching of the economy (P, π) . The sincere student i_3 not only prefers school b to her assignment $\mu(i_3) = c$ but also she has the highest priority there. Nevertheless, by being truthful and ranking b second, she has lost her priority to student i_2 at equilibria.
3. The unique stable matching of the economy (P, π) is

$$\nu = \begin{pmatrix} i_1 & i_2 & i_3 \\ a & c & b \end{pmatrix}.$$

Matchings μ and ν are not Pareto ranked. While the strategic student i_1 is indifferent between the two matchings, the strategic student i_2 is better off under matching μ and

the sincere student i_3 is better off under matching ν . That is, the strategic student i_2 is better off under the Nash equilibria of the Boston game at the expense of the sincere student i_3 .

We next characterize the Nash equilibrium outcomes of the Boston game which will be useful to generalize the above observations.

3 Characterization of Nash Equilibrium Outcomes

Given an economy (P, π) , we will construct an augmented economy that will be instrumental in describing the set of Nash equilibrium outcomes of the Boston game.

Given an economy (P, π) and a school s , partition the set of students I into m sets as follows:

I_1 : Strategic students and sincere students who rank s as their first choices under P ,

I_2 : sincere students who rank s as their second choices under P ,

I_3 : sincere students who rank s as their third choices under P ,

\vdots \vdots

I_m : sincere students who rank s as their last choices under P .

Given an economy (P, π) and a school s , construct an **augmented priority ordering** $\tilde{\pi}_s$ as follows:

- each student in I_1 has higher priority than each student in I_2 , each student in I_2 has higher priority than each student in I_3 , \dots , each student in I_{m-1} has higher priority than each student in I_m , and
- for any $k \leq m$, priority among students in I_k is based on π_s .

Define $\tilde{\pi} = (\tilde{\pi}_s)_{s \in S}$. We refer to the economy $(P, \tilde{\pi})$ as the **augmented economy**.

Example 1 continued. Let us construct the augmented economy for Example 1. Since only student i_3 is sincere, $\tilde{\pi}$ is constructed from π by pushing student i_3 to the end of the priority

ordering at each school except her top choice a (where she has the lowest priority to begin with):

$$\begin{aligned}\pi_a : i_2 - i_1 - i_3 &\Rightarrow \tilde{\pi}_a : i_2 - i_1 - i_3 \\ \pi_b : i_3 - i_2 - i_1 &\Rightarrow \tilde{\pi}_b : i_2 - i_1 - i_3 \\ \pi_c : i_1 - i_3 - i_2 &\Rightarrow \tilde{\pi}_c : i_1 - i_2 - i_3\end{aligned}$$

The key observation here is that the unique Nash equilibrium outcome μ of the Boston game is the unique stable matching for the augmented economy $(P, \tilde{\pi})$.

While the uniqueness is specific to the above example, the equivalence is general. We are ready to present our first result.

Proposition 1: The set of Nash equilibrium outcomes of the Boston game under (P, π) is equivalent to the set of stable matchings under $(P, \tilde{\pi})$.

Proof:

\Leftarrow :

Fix an economy (P, π) and let μ be stable under $(P, \tilde{\pi})$. Let preference profile Q be such that $Q_i = P_i$ for all $i \in N$ and $\mu(i)$ is the first choice under Q_i for all $i \in M$. Matching μ is stable under $(Q, \tilde{\pi})$ as well. Let ν be the outcome of the Boston mechanism under (Q, π) . We first show, by induction, that $\nu = \mu$.

Consider any student j who does not receive her first choice s_j^1 under Q at matching μ . By construction of Q , student j is sincere. Since μ is stable under $(Q, \tilde{\pi})$ and since student j does not lose priority to any student at school s_j^1 when priorities change from π to $\tilde{\pi}$, she has lower priority under $\pi_{s_j^1}$ than any student who has received a seat at s_j^1 under μ . Each of these students rank s_j^1 as their first choices under Q and school s_j^1 does not have empty seats under μ for otherwise (j, s) would block μ under $(Q, \tilde{\pi})$. Therefore $\nu(j) \neq s_j^1$. So a student can receive her first choice under Q at matching ν only if she receives her first choice under Q at matching μ . But then, since the Boston mechanism is Pareto efficient, matching ν is Pareto efficient under (Q, π) which in turn implies that $\nu(i) = \mu(i)$ for any student i who receives her first choice under Q at matching μ .

Next given $k > 1$, suppose

1. any student who does not receive one of her top k choices under Q at matching μ does not receive one of her top k choices under Q at matching ν either, and
2. for any student i who receives one of her top k choices under Q at matching μ , $\nu(i) = \mu(i)$.

We will show that the same holds for $(k + 1)$ and this will establish that $\nu = \mu$. Consider any student j who does not receive one of her top $k + 1$ choices under Q at matching μ . By construction of Q , student j is sincere and by assumption she does not receive one of her top k choices under Q at matching ν . Consider $(k + 1)^{\text{th}}$ choice s_j^{k+1} of student j under Q_j . Since μ is stable under $(Q, \tilde{\pi})$, there is no empty seat at school s_j^{k+1} for otherwise pair (j, s_j^{k+1}) would block matching μ under $(Q, \tilde{\pi})$. Moreover since μ is stable under $(Q, \tilde{\pi})$, for any student i with $\mu(i) = s_j^{k+1}$ one of the following three cases should hold:

1. $i \in M$ and by construction s_j^{k+1} is her first choice under Q_i ,
2. $i \in N$ and s_j^{k+1} is one of her top k choices under Q_i ,
3. $i \in N$, she has ranked s_j^{k+1} as her $(k + 1)^{\text{th}}$ choice under Q_i , and she has higher priority than j under $\pi_{s_j^{k+1}}$.

If either of the first two cases holds, then $\nu(i) = s_j^{k+1}$ by inductive assumption. If Case 3 holds, then student i has not received one of her top k choices under Q_i at matching ν by the inductive assumption and furthermore she has ranked school s_j^{k+1} as her $(k + 1)^{\text{th}}$ choice under Q_i . Since she has higher priority than j under $\pi_{s_j^{k+1}}$, $\nu(j) = s_j^{k+1}$ implies $\nu(i) = s_j^{k+1}$. Therefore considering all three cases, $\nu(j) = s_j^{k+1}$ implies $\nu(i) = s_j^{k+1}$ for any student i with $\mu(i) = s_j^{k+1}$ and since school s_j^{k+1} does not have empty seats under μ , $\nu(j) \neq s_j^{k+1}$. So a student can receive one of her top $k + 1$ choices under Q at matching ν only if she receives one of her top $k + 1$ choices under Q at matching μ . Moreover matching ν is Pareto efficient under (Q, π) and therefore $\nu(i) = \mu(i)$ for any student i who receives her $(k + 1)^{\text{th}}$ choice under Q_i at μ completing the induction and establishing $\nu = \mu$.

Next we show that Q is a Nash equilibrium profile and hence ν is a Nash equilibrium outcome. Consider any strategic student $i \in M$ and suppose $sP_i\nu(i) = \mu(i)$ for some school $s \in S$. Since $\nu = \mu$ is stable under $(Q, \tilde{\pi})$ and since student i gains priority under $\tilde{\pi}_s$ over only students who rank s second or worse under Q , not only any student $j \in I$ with $\nu(j) = s$ ranks school s as her first choice under Q_i but she also has higher priority under π_s . Therefore regardless of what preferences student i submits, each student $j \in I$ with $\nu(j) = s$ will receive a seat at school s . Moreover by stability of $\nu = \mu$ under $(Q, \tilde{\pi})$ there are no empty seats at school s and hence student i cannot receive a seat at s regardless of her submitted preferences. Therefore matching ν is a Nash equilibrium outcome.

\Rightarrow :

Suppose matching μ is not stable under $(P, \tilde{\pi})$. Let Q be any preference profile where $Q_i = P_i$ for any sincere student i and where μ is the outcome of the Boston mechanism under (Q, π) . We will show that Q is not a Nash equilibrium strategy profile of the Boston game under (P, π) .

First suppose μ is not individually rational under $(P, \tilde{\pi})$. Then there is a student $i \in I$ with $iP_i\mu(i)$. Since the Boston mechanism is individually rational, student i should be a sophisticated student who has ranked the unacceptable school $\mu(i)$ as acceptable. Let P_i^0 be a preference relation where there is no acceptable school. Upon submitting P_i^0 , student i will profit by getting unassigned. Hence Q cannot be an equilibrium profile in this case.

Next suppose there is a pair (i, s) that blocks μ under $(P, \tilde{\pi})$. Since μ is the outcome of the Boston mechanism under (Q, π) , student i cannot be a sincere student. Let P_i^s be a preference relation where school s is the first choice. We have two cases to consider:

Case 1: School s has an empty seat at μ .

Recall that by assumption μ is the outcome of the Boston mechanism under (Q, π) . Since s has an empty seat at μ , there are fewer students who rank s as their first choice under Q than the capacity of school s . Therefore upon submitting the preference relation P_i^s , student i will profit by getting assigned a seat at school s . Hence Q cannot be an equilibrium profile.

Case 2: School s does not have an empty seat at μ .

By assumption μ is the outcome of the Boston mechanism under (Q, π) and there is a student j with $\mu(j) = s$ although i has higher priority than j under $\tilde{\pi}_s$. If school s is not j 's first choice under Q_j then there are fewer students who rank s as their first choice under Q than the capacity of school s , and upon submitting the preference relation P_i^s , student i will profit by getting assigned a seat at school s contradicting Q being an equilibrium profile. If on the other hand school s is j 's first choice under Q_j , then either j is sophisticated or j is sincere and s is her first choice under P_j . In either case i having higher priority than j under $\tilde{\pi}_s$ implies i having higher priority than j under π_s . Moreover since $\mu(j) = s$, the capacity of school s is strictly larger than the number of students who both rank it as their first choice under Q and also has higher priority than j under π_s . Therefore the capacity of school s is strictly larger than the number of students who both rank it as their first choice under Q and also has higher priority than i under π_s . Hence upon submitting the preference relation P_i^s , student i will profit by getting assigned a seat at school s contradicting Q being an equilibrium profile.

Since there is no Nash equilibrium profile Q for which μ is the outcome of the Boston mechanism under (Q, π) , μ cannot be a Nash equilibrium outcome of the Boston game under (P, π) . \square

Therefore at Nash equilibria strategic students gain priority at the expense of sincere students. Another implication of Proposition 1 is that the set of equilibrium outcomes inherits some of the properties of the set of stable matchings. In particular there is a Nash equilibrium outcome of the Boston game that is weakly preferred to any other Nash equilibrium outcome by any student. We refer this outcome as the **Pareto-dominant Nash equilibrium outcome**.

Equilibrium Assignments of Sincere Students

The student-optimal stable mechanism replaced the Boston mechanism in Boston in 2005. In the following section we will compare the equilibrium outcomes of the Boston game with the dominant-strategy equilibrium outcome of the student-optimal stable mechanism. One of the difficulties in such comparative static analysis is that the Boston game has multiple equilibria in general. Nevertheless, as we present next, multiplicity is not an issue for sincere students.

Proposition 2: Let μ, ν be both Nash equilibrium outcomes of the preference revelation game induced by the Boston mechanism. For any sincere student $i \in N$, $\mu(i) = \nu(i)$.

Proof: Fix an economy (P, π) . Let μ, ν be both Nash equilibrium outcomes of the preference revelation game induced by the Boston mechanism. By Proposition 1, μ, ν are stable matchings under $(P, \tilde{\pi})$. Let $\bar{\mu} = \mu \vee \nu$ and $\underline{\mu} = \mu \wedge \nu$ be the join and meet of the stable matching lattice. That is, $\bar{\mu}, \underline{\mu}$ are such that, for all $i \in I$,

$$\bar{\mu}(i) = \begin{cases} \mu(i) & \text{if } \mu(i)R_i\nu(i) \\ \nu(i) & \text{if } \nu(i)R_i\mu(i) \end{cases} \quad \underline{\mu}(i) = \begin{cases} \nu(i) & \text{if } \mu(i)R_i\nu(i) \\ \mu(i) & \text{if } \nu(i)R_i\mu(i) \end{cases}$$

Since the set of stable matchings is lattice (attributed to John Conway by Knuth 1976), $\bar{\mu}$ and $\underline{\mu}$ are both stable matchings under $(P, \tilde{\pi})$.

Let $T = \{i \in I : \bar{\mu}(i) \neq \underline{\mu}(i)\}$. That is, T is the set of students who receive a different assignment under $\bar{\mu}$ and $\underline{\mu}$. If $T \subseteq M$, then we are done. So suppose there exists $i \in T \cap N$. We will show that this leads to a contradiction. Let $s = \bar{\mu}(i)$, $s^* = \underline{\mu}(i)$, and $j \in \bar{\mu}^{-1}(s^*) \cap T$. Such a student $j \in I$ exists because by the rural hospitals theorem of Roth (1985) the same set of students and the same set of seats are assigned under any pair of stable matchings. Note that $j \in \bar{\mu}^{-1}(\underline{\mu}(i))$.

Claim: $j \in N$.

Proof of the Claim: By construction of $\bar{\mu}$ and $\underline{\mu}$, $sP_i s^*$ and therefore school s^* is not i 's first choice. Moreover by Roth and Sotomayor (1989) each student in $\underline{\mu}(s^*) \setminus \bar{\mu}(s^*)$ has higher priority

under $\tilde{\pi}_{s^*}$ than each student in $\bar{\mu}(s^*) \setminus \underline{\mu}(s^*)$, and hence i has higher priority than j under $\tilde{\pi}_{s^*}$. But since i is sincere by assumption and since s^* is not her first choice, student j has to be sincere as well for otherwise she would have higher priority under $\tilde{\pi}_{s^*}$. \diamond

Next construct the following directed graph: Each student $i \in T \cap N$ is a node and there is a directed link from $i \in T \cap N$ to $j \in T \cap N$ if $j \in \bar{\mu}^{-1}(\underline{\mu}(i))$. By the above Claim there is at least one directed link emanating from each node. Therefore, since there are finite number of nodes, there is at least one cycle in this graph. Pick any such cycle. Let $T_1 \subseteq T \cap N$ be the set of students in the cycle, and let $|T_1| = k$. Relabel students in T_1 and their assignments under $\bar{\mu}, \underline{\mu}$ so that the restriction of matchings $\bar{\mu}$ and $\underline{\mu}$ to students in T_1 is as follows:

$$\bar{\mu}_{T_1} = \begin{pmatrix} i^1 & i^2 & \dots & i^k \\ s^1 & s^2 & \dots & s^k \end{pmatrix} \quad \underline{\mu}_{T_1} = \begin{pmatrix} i^1 & i^2 & \dots & i^{k-1} & i^k \\ s^2 & s^3 & \dots & s^k & s^1 \end{pmatrix}$$

Note that a school may appear more than once in a cycle so that schools s^t, s^u does not need to be distinct for $t \neq u$ (although they would have if the cycle we pick is minimal). This has no relevance for the contradiction we present next.

Let $r_{i,s}$ be the ranking of school s in P_i (so $r_{i,s} = \ell$ means that s is i 's ℓ th choice). By Roth and Sotomayor (1989) i^k has higher priority at school s^1 than i^1 under $\tilde{\pi}_{s^1}$, and since i^1, i^k are both sincere,

$$r_{i^k, s^1} \leq r_{i^1, s^1}$$

Similarly

$$\begin{aligned} r_{i^1, s^2} &\leq r_{i^2, s^2} \\ &\vdots \\ r_{i^{k-1}, s^k} &\leq r_{i^k, s^k} \end{aligned}$$

Moreover since $\bar{\mu}(i)P_i\underline{\mu}(i)$ for each $i \in T$,

$$\begin{aligned} s^1 P_{i^1} s^2 &\Rightarrow r_{i^1, s^1} < r_{i^1, s^2} \\ s^2 P_{i^2} s^3 &\Rightarrow r_{i^2, s^2} < r_{i^2, s^3} \\ &\vdots \\ s^{k-1} P_{i^{k-1}} s^k &\Rightarrow r_{i^{k-1}, s^{k-1}} < r_{i^{k-1}, s^k} \\ s^k P_{i^k} s^1 &\Rightarrow r_{i^k, s^k} < r_{i^k, s^1} \end{aligned}$$

Combining the inequalities, we obtain

$$r_{i^k, s_1} \leq r_{i^1, s^1} < r_{i^1, s^2} \leq r_{i^2, s^2} < r_{i^2, s^3} \leq \dots \leq r_{i^{k-1}, s^{k-1}} < r_{i^{k-1}, s^k} \leq r_{i^k, s^k} < r_{i^k, s^1}$$

establishing the desired contradiction. Hence there exists no $i \in N$ with $\bar{\mu}(i) \neq \underline{\mu}(i)$. But that means there exists no $i \in N$ with $\mu(i) \neq \nu(i)$ either, completing the proof. \square

4 Comparative Statics: Comparing Mechanisms

The outcome of the student-optimal stable mechanism can be obtained with the following **student-proposing deferred acceptance algorithm** (Gale and Shapley 1962):

Step 1: Each student proposes to her first choice. Each school tentatively assigns its seats to its proposers one at a time following their priority order. Any remaining proposers are rejected.

In general, at:

Step k: Each student who was rejected in the previous step proposes to her next choice. Each school considers the students it has been holding together with its new proposers and tentatively assigns its seats to these students one at a time following their priority order. Any remaining proposers are rejected.

The algorithm terminates when no student proposal is rejected and each student is assigned her final tentative assignment. Any student who is not holding a tentative assignment remains unassigned.

Comparing Mechanisms for Sincere Students

Sincere students lose priority to strategic students under the Boston mechanism. They may also be affected by other sincere students, so that some sincere students may benefit at the expense of other sincere students under the Boston mechanism. More precisely, a sincere student may prefer the Boston mechanism to the student-optimal stable mechanism since:

- she gains priority at her first choice school over sincere students who rank it second or lower, and in general

- she gains priority at her k^{th} choice school over sincere students who rank it $(k + 1)^{\text{th}}$ or lower, etc.

Example 2. There are three schools a, b, c each with one seat and three sincere students i_1, i_2, i_3 . Preferences and priorities are as follows:

$$\begin{array}{ll} P_{i_1} : & a \ b \ c & \pi_a : & i_1 - i_2 - i_3 \\ P_{i_2} : & a \ b \ c & \pi_b : & i_1 - i_2 - i_3 \\ P_{i_3} : & b \ a \ c & \pi_c : & i_1 - i_2 - i_3 \end{array}$$

Outcomes of the Boston mechanism and the student-optimal stable mechanism are

$$\begin{pmatrix} i_1 & i_2 & i_3 \\ a & c & b \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} i_1 & i_2 & i_3 \\ a & b & c \end{pmatrix}$$

respectively. Under the Boston mechanism the sincere student i_3 gains priority at her top choice school b over the sincere student i_2 . Hence student i_3 prefers her assignment under the Boston mechanism whereas student i_2 prefers her assignment under the student-optimal stable mechanism.

Comparing Mechanisms for Strategic Students

Unlike a sincere student, a strategic student may be assigned seats at different schools at different equilibrium outcomes of the Boston game. Hence we will concentrate on the Pareto-dominant Nash equilibrium outcome of the Boston game.

Proposition 3: The school a strategic student receives in the Pareto-dominant equilibrium of the Boston mechanism is weakly better than her dominant-strategy outcome under the student-optimal stable mechanism.

Proof: Let μ^I and ν^I be the student-optimal stable matching for economies (P, π) , $(P, \tilde{\pi})$ respectively. Define matching ν_0 as follows:

$$\begin{array}{ll} \nu_0(i) & = \mu^I(i) & \text{for all } i \in M, \\ \nu_0(i) & = i & \text{for all } i \in N. \end{array}$$

If ν_0 is stable under $(P, \tilde{\pi})$, then each sophisticated student $i \in M$ weakly prefers $\nu^I(i)$ to $\nu_0(i) = \mu^I(i)$ and we are done. So w.l.o.g. assume ν_0 is not stable under $(P, \tilde{\pi})$. We will

construct a sequence of matchings $\nu_0, \nu_1, \dots, \nu_k$ where ν_k is stable under $(P, \tilde{\pi})$, and

$$\nu_\ell(i) R_i \nu_{\ell-1}(i) \quad \text{for all } i \in M \text{ and } \ell \geq 1.$$

Consider matching ν_0 . Since ν_0 is not stable but individually rational under $(P, \tilde{\pi})$, there is a blocking pair. Pick any school s^1 in a blocking pair and let i^1 be the highest priority student (sincere or sophisticated) under $\tilde{\pi}_{s^1}$ who strictly prefers s^1 to her assignment under ν_0 .

Claim 1: School s^1 has an empty seat under ν_0 .

Proof of Claim 1: We have three cases to consider.

Case 1: $i^1 \in M$.

By construction $\nu_0^{-1}(s^1) = (\mu^I)^{-1}(s^1) \cap M$ and (i^1, s^1) does not block μ^I under (P, π) . When priorities change from π to $\tilde{\pi}$, no sophisticated student loses priority to any other student and therefore (i^1, s^1) can block ν_0 under $(P, \tilde{\pi})$ only if school s^1 has an empty seat under ν_0 .

Case 2: $i^1 \in N$ and school s^1 is not student i^1 's first choice.

Student i^1 has lower priority under $\tilde{\pi}_{s^1}$ than any sophisticated student. Since $\nu_0^{-1}(s^1) \subseteq M$, pair (i^1, s^1) can block ν_0 under $(P, \tilde{\pi})$ only if school s^1 has an empty seat under ν_0 .

Case 3: $i^1 \in N$ and school s^1 is student i^1 's first choice.

If $\mu^I(i^1) = s^1$, then by construction the seat i^1 occupies at s^1 under μ^I is empty under ν_0 . If $\mu^I(i^1) \neq s^1$, then all seats at s^1 are occupied under μ^I by higher priority students under π . Since (i^1, s^1) blocks ν^0 under $(P, \tilde{\pi})$, at least one of these students must be a sincere student who ranks s^1 second or worse in her preferences. That is because student i^1 gains priority over only these student under $\tilde{\pi}$. Therefore at least one seat at s^1 must be empty under ν_0 .

This completes the proof of Claim 1. ◇

Construct matching ν_1 by satisfying pair (i^1, s^1) at matching ν_0 :

$$\begin{aligned} \nu_1(i) &= \nu_0(i) & \text{for all } i \in I \setminus \{i^1\}, \\ \nu_1(i^1) &= s^1 \end{aligned}$$

By construction $\nu_1(i) R_i \nu_0(i)$ for all $i \in I$. If ν_1 is stable under $(P, \tilde{\pi})$, then for all $i \in M$,

$$\nu^I(i) R_i \nu_1(i) R_i \underbrace{\nu_0(i)}_{=\mu^I(i)}$$

and we are done. If not, we proceed with the construction of matching ν_2 .

In general for any $\ell > 0$, if ν_ℓ is not stable under $(P, \tilde{\pi})$ construct $\nu_{\ell+1}$ as follows: Pick any school $s^{\ell+1}$ in a blocking pair for ν_ℓ and let $i^{\ell+1}$ be the highest priority student (sincere or sophisticated) under $\tilde{\pi}_{s^{\ell+1}}$ who strictly prefers $s^{\ell+1}$ to her assignment under ν_ℓ . As we prove next, $s^{\ell+1}$ has an empty seat under ν_ℓ . Construct matching $\nu_{\ell+1}$ by satisfying pair $(i^{\ell+1}, s^{\ell+1})$ at matching ν_ℓ :

$$\begin{aligned}\nu_{\ell+1}(i) &= \nu_\ell(i) && \text{for all } i \in I \setminus \{i^{\ell+1}\}, \\ \nu_{\ell+1}(i^{\ell+1}) &= s^{\ell+1}\end{aligned}$$

Claim 2: School $s^{\ell+1}$ has an empty seat under ν_ℓ .

Proof of Claim 2: We have three cases to consider.

Case 1: $i^{\ell+1} \in M$.

By construction $\nu_\ell^{-1}(s^{\ell+1}) \subseteq [(\mu^I)^{-1}(s^{\ell+1}) \cap M] \cup \{i^1, \dots, i^\ell\}$. Since $(i^{\ell+1}, s^{\ell+1})$ does not block μ^I under (P, π) and since student $i^{\ell+1}$ does not gain priority over any sophisticated student when priorities change from π to $\tilde{\pi}$, any student in $(\mu^I)^{-1}(s^{\ell+1}) \cap M$ has higher priority than student $i^{\ell+1}$ under $\tilde{\pi}_{s^{\ell+1}}$. Moreover for any $i^m \in \{i^1, \dots, i^\ell\}$ with $i^m \in \nu_\ell^{-1}(s^{\ell+1})$, we must have $s^{\ell+1} = s^m$ and thus student i^m has higher priority than student $i^{\ell+1}$ at school $s^{\ell+1} = s^m$ under $\tilde{\pi}$ by the choice of blocking pairs. Therefore student $i^{\ell+1}$ has lower priority under $\tilde{\pi}_{s^{\ell+1}}$ than any student in $\nu_\ell^{-1}(s^{\ell+1})$ and hence pair $(i^{\ell+1}, s^{\ell+1})$ can block ν_ℓ under $(P, \tilde{\pi})$ only if school $s^{\ell+1}$ has an empty seat at ν_ℓ .

Case 2: $i^{\ell+1} \in N$ and school $s^{\ell+1}$ is not student $i^{\ell+1}$'s first choice.

By construction $\nu_\ell^{-1}(s^{\ell+1}) \subseteq M \cup \{i^1, \dots, i^\ell\}$. Student $i^{\ell+1}$ has lower priority under $\tilde{\pi}_{s^{\ell+1}}$ than any sophisticated student. Moreover for any $i^m \in \{i^1, \dots, i^\ell\}$ with $i^m \in \nu_\ell^{-1}(s^{\ell+1})$, we must have $s^{\ell+1} = s^m$ and thus student i^m has higher priority than student $i^{\ell+1}$ at school $s^{\ell+1} = s^m$ under $\tilde{\pi}$ by the choice of blocking pairs. Therefore student $i^{\ell+1}$ has lower priority under $\tilde{\pi}_{s^{\ell+1}}$ than any student in $\nu_\ell^{-1}(s^{\ell+1})$ and hence pair $(i^{\ell+1}, s^{\ell+1})$ can block ν_ℓ under $(P, \tilde{\pi})$ only if school $s^{\ell+1}$ has an empty seat at ν_ℓ .

Case 3: $i^{\ell+1} \in N$ and school $s^{\ell+1}$ is student $i^{\ell+1}$'s first choice.

Recall that $\nu_\ell^{-1}(s^{\ell+1}) \subseteq [(\mu^I)^{-1}(s^{\ell+1}) \cap M] \cup \{i^1, \dots, i^\ell\}$. First suppose $\mu^I(i^{\ell+1}) \neq s^{\ell+1}$. For any for any $i^m \in \{i^1, \dots, i^\ell\}$ with $i^m \in \nu_\ell^{-1}(s^{\ell+1})$, we must have $s^{\ell+1} = s^m$ and thus student i^m has higher priority than student $i^{\ell+1}$ at school $s^{\ell+1} = s^m$ under $\tilde{\pi}$ by the choice of blocking pairs. Moreover $s^{\ell+1}$ is $i^{\ell+1}$'s first choice and yet μ^I is stable under (P, π) . Therefore all students in $(\mu^I)^{-1}(s^{\ell+1})$ has higher priority under $\pi_{s^{\ell+1}}$ than $i^{\ell+1}$ does. But student $i^{\ell+1}$ does not gain priority at school $s^{\ell+1}$ over any sophisticated student when priorities change from

π to $\tilde{\pi}$. Therefore any student in $(\mu^I)^{-1}(s^{\ell+1}) \cap M$ has higher priority under $\tilde{\pi}_{s^{\ell+1}}$ than student $i^{\ell+1}$ does, which in turn implies any student in $\nu_\ell^{-1}(s^{\ell+1})$ has higher priority under $\tilde{\pi}_{s^{\ell+1}}$ than student $i^{\ell+1}$ does. Hence pair $(i^{\ell+1}, s^{\ell+1})$ can block ν_ℓ under $(P, \tilde{\pi})$ only if school $s^{\ell+1}$ has an empty seat at ν_ℓ .

Next suppose $\mu^I(i^{\ell+1}) = s^{\ell+1}$. Let $i^m \in \{i^1, \dots, i^\ell\}$ be such that $i^m \in \nu_\ell^{-1}(s^{\ell+1})$. We have $s^{\ell+1} = s^m$ and since student i^m has higher priority than student $i^{\ell+1}$ at school $s^{\ell+1}$ under $\tilde{\pi}$, the same should be true under π as well. That is because, student $i^{\ell+1}$ does not lose priority to any student at school $s^{\ell+1}$ when priorities change from π to $\tilde{\pi}$. Since pair $(i^m, s^m) = (i^m, s^{\ell+1})$ blocks matching ν_{m-1} under $(P, \tilde{\pi})$ by definition, and since assignments only improve as we proceed by the sequence $\nu_0, \nu_1, \dots, \nu_k$,

$$s^{\ell+1} P_{i^m} \nu_{m-1}(i^m) R_{i^m} \nu_0(i^m).$$

So on one hand $\mu^I(i^{\ell+1}) = s^{\ell+1}$ where $i^{\ell+1}$ has lower priority at $s^{\ell+1}$ than i^m under π , and on the other hand i^m strictly prefers $s^{\ell+1}$ to its assignment under ν_0 . That means not only student i^m is sincere, but also $\mu^I(i^m) = s^{\ell+1}$ for otherwise μ^I cannot be stable under (P, π) . So for each $i^m \in \nu_\ell^{-1}(s^{\ell+1})$, there is one empty seat at $s^{\ell+1}$ under ν_0 . In addition there is at least one more empty seat, namely the seat sincere student $i^{\ell+1}$ occupies at school $s^{\ell+1}$ at matching μ^I . Therefore under matching ν_ℓ there must still be at least one empty seat at school $s^{\ell+1}$.

This covers all three cases and completes the proof of Claim 2. \diamond

We are now ready to complete the proof. Since each student weakly prefers matching ν_ℓ to matching $\nu_{\ell-1}$ for any $\ell \geq 0$, eventually the sequence terminates which means no pair blocks the final matching ν_k in the sequence and thus ν_k is stable under $(P, \tilde{\pi})$. Therefore for any strategic student $i \in M$,

$$\nu^I(i) R_i \nu_k(i) R_i \underbrace{\nu_0(i)}_{=\mu^I(i)}$$

where the first relation holds by definition of the student-optimal stable matching. This completes the proof. \square

While the Boston mechanism is easy to describe, it induces a complicated coordination game among strategic students. Therefore, it is important to be cautious in interpreting Proposition 3. When strategic students and their families (such as the members of the WZPG) are able to reach the Pareto-dominant Nash equilibrium assignment, they may prefer keeping the Boston mechanism for they will reap the benefit of their strategic advantage.

5 Conclusion

Boston Public Schools stated that their main rationale for changing their student assignment system is that it levels the playing field. They identified a fairness rationale for a strategyproof system. In this paper, we examined this intuitive notion and showed the assignment of a strategic students in the Pareto-dominant Nash equilibrium of the Boston mechanism is weakly better than their assignment under the student-optimal stable mechanism, providing formal support for BPS's claim.

Despite its theoretical weaknesses, performance in laboratory experiments, and empirical evidence of confused play, the Boston mechanism is the most widely used school choice mechanism in the United States. This paper proposes another theoretical rationale for abandoning the mechanism based on fairness or equal access, which was central in Boston's decision.

The reason why the mechanism continues to be used is a puzzle, which might be related to our analysis here. Chubb and Moe (1999) argue that important stakeholders often control the mechanisms of reform in education policy. In the context of student assignment mechanisms, the important stakeholders may be strategic parents who have invested energy in learning about the mechanism, and the choice of the Boston mechanism may reflect their preferences.

References

- [1] A. Abdulkadiroğlu, P. A. Pathak, A. E. Roth, and T. Sönmez (2006), "Changing the Boston School Choice Mechanism: Strategyproofness as Equal Access," Boston College, Columbia, Harvard working paper.
- [2] A. Abdulkadiroğlu and Tayfun Sönmez (2003), "School Choice: A Mechanism Design Approach," *American Economic Review*, 93: 729-747.
- [3] J. Chubb and T. Moe (1990), *Politics, Markets and America's Schools*. Washington D.C.: Brookings Institution Press.
- [4] Y. Chen and T. Sönmez (2006), "School Choice: An Experimental Study," *Journal of Economic Theory*, 127: 202-231.
- [5] L. E. Dubins and D. A. Freedman (1981), "Machiavelli and the Gale-Shapley Algorithm," *American Mathematical Monthly*, 88: 485-494.

- [6] H. Ergin and T. Sönmez (2006), “Games of School Choice Under the Boston Mechanism,” *Journal of Public Economics*, 90: 215-237.
- [7] D. Gale and L. Shapley (1962), “College Admissions and the Stability of Marriage,” *American Mathematical Monthly*, 69: 9-15.
- [8] A. Kelso and V. Crawford (1981), “Job Matching, Coalition Formation, and Gross Substitutes,” *Econometrica*, 50: 1483-1504.
- [9] D. Knuth (1976), *Marriage Stables*, Les Press de l’Universit e de Montreal, Montreal.
- [10] F. Kojima (2006), “Games of School Choice under the Boston Mechanism with General Priority Structures,” mimeo, Harvard University.
- [11] A. E. Roth (1982), “The Economics of Matching: Stability and Incentives,” *Mathematics of Operations Research*, 7: 617-628.
- [12] A. E Roth and M. Sotomayor(1989), “The College Admissions Problem Revisited,” *Econometrica*, 57, 1989, 559-570.