WZB
Wissenschaftszentrum Berlin
für Sozialforschung

Inácio Bó
C.-Philipp Heller

Strategic Schools under the Boston Mechanism Revisited

Discussion Paper

SP II 2016–204

May 2016

WZB Berlin Social Science Center

Research Area
Markets and Choice

Research Unit
Market Behavior

Wissenschaftszentrum Berlin für Sozialforschung gGmbH
Reichpietschufer 50
10785 Berlin
Germany
www.wzb.eu

Inácio Bó, C.-Philipp Heller
Strategic Schools under the Boston Mechanism Revisited

Affiliation of the authors:

Inácio Bó
WZB

C.-Philipp Heller
Humboldt-Universität zu Berlin

Wissenschaftszentrum Berlin für Sozialforschung gGmbH
Reichpietschufer 50
10785 Berlin
Germany
www.wzb.eu

Abstract

Strategic Schools under the Boston Mechanism Revisited

by Inácio Bó and C.-Philipp Heller*

We show that Ergin & Sönmez's (2006) results which show that for schools it is a dominant strategy to truthfully rank the students under the Boston mechanism, and that the Nash equilibrium outcomes in undominated strategies of the induced game are stable, rely crucially on two assumptions. First, (a) that schools need to be restricted to find all students acceptable, and (b) that students cannot observe the priorities set by the schools before submitting their preferences. We show that relaxing either assumption eliminates the strategy dominance, and that Nash equilibrium outcomes in undominated strategies for the simultaneous induced game in case (a) and subgame perfect Nash equilibria in case (b) may contain unstable matchings. We also show that when able to manipulate capacities, schools may only have an incentive to do so if students submit their preferences after observing the reported capacities.

Keywords: Mechanism Design, Two-Sided Matching, Boston Mechanism, School Choice

JEL classification: C78, D63, D78, D82

* E-mail: inacio.bo@wzb.eu, hellerkr@hu-berlin.de..

1 Introduction

Having their children attend a good school is an important concern for many parents. Until recently, in many countries parents did not have much choice when it came to choosing the public school they would attend, as most children were administratively assigned to a school nearby. In their seminal contribution Abdulkadiroğlu and Sönmez (2003) analyzed school assignment procedures in Boston and other cities in the U.S., and also suggested the student-proposing deferred acceptance mechanism (DA, Gale and Shapley, 1962) and the top trading cycles mechanism (TTC, Roth and Postlewaite, 1977) as two alternatives to the methods being used in those cities. One argument in favor of the DA and TTC over the mechanism used in Boston at that time (denoted *Boston mechanism*) is that reporting preferences truthfully under these mechanisms is a dominant strategy. Under the Boston mechanism, on the other hand, students may obtain better assignments by strategically manipulating the preferences over the schools they report.

The Boston mechanism lets each student apply to schools in the order of their reported preferences. Schools in each round consider all applications received in that round, accept the applications from the highest-ranked students according to the schools' preferences or some exogenous priority ordering of the students, until their capacity is filled, and reject the remaining applicants. In subsequent rounds rejected students apply to their next highest-ranked schools. The schools which still have seats available again accept the applications from the highest-ranked students until their capacity is filled.

This paper revisits the question of the possibility of manipulations by schools under the Boston mechanism. Results of Ergin and Sönmez (2006) showed that for schools it is a dominant strategy to truthfully rank the students. This paper shows that for these results to hold, schools need to be restricted to find all students acceptable, and that students cannot observe the priorities set by the schools before submitting their preferences. If schools are allowed to deem students unacceptable, we show that doing so may yield a better allocation from the schools' point of view. In addition, Ergin and Sönmez (2006) showed that, taking schools' priorities as given, the set of Nash equilibrium outcomes under the Boston mechanism equals the set of stable matchings. Since the (student-proposing) DA algorithm always yields the student-optimal stable matching, this implies that the Boston mechanism is Pareto dominated by DA if all students behave strategically. We show that if schools are strategic and can deem students as unacceptable then there may be Nash equilibrium outcomes in undominated strategies that are not stable with respect to the true preferences of students and schools. Moreover, we show that if students are allowed to observe the priorities set by the schools prior to submitting their preferences, it is also the case that schools may have an incentive to manipulate the priorities over students or their capacities, even if they cannot deem any of them as unacceptable, and that the set of subgame perfect Nash equilibria of the induced game may contain unstable matchings. The result we mentioned from Ergin and Sönmez (2006), therefore, are sensitive to assumptions about the game induced by the Boston mechanism.

The Boston mechanism and its properties have been analyzed by many papers in the literature, starting from Roth (1991), which studies a variety of matching procedures for regional medical labor markets in the UK. While some regions used stable mechanisms (Edinburgh and Cardiff) others use priority mechanisms (Newcastle, Birmingham and Edinburgh), of which the Boston mechanism is a special case.

While the use of the Boston mechanism has been criticized beginning with Abdulkadiroğlu and Sönmez (2003) due to the possibility of gains from manipulation by students, a number of authors show that it has some desirable properties. Abdulkadiroğlu, Che and Yasuda (2011) and Miralles (2009) both argue in favor of the Boston mechanism on the grounds of *ex ante* cardinal efficiency. In particular, when schools' priorities are random and uniform and students rank schools similarly but prefer them with different intensities, the equilibrium outcome under the Boston mechanism yields higher expected welfare than the deferred acceptance algorithm. The reason is that students who put a particularly high cardinal value on a school are more likely to rank this school highly and thereby obtain a seat there. Hence, the equilibrium under the Boston mechanism makes use of information on preference intensities, while the deferred acceptance mechanism does not.

Pathak and Sönmez (2008) analyze the Boston mechanism when there are two types of students with different levels of strategic sophistication, while schools' priorities are taken as fixed. They show that sophisticated students are better off under the Boston mechanism than naïve students. This is taken by these authors as justifying the move toward non-manipulable mechanisms. Kojima and Ünver (2014) give two characterizations of the Boston mechanism while allowing for schools to deem some students unacceptable.

Mennle and Seuken (2014) distinguish between the 'naïve' Boston mechanism and the 'adaptive' Boston mechanism. The latter differs from the former in that students in later rounds never apply to schools that have already been filled up. The mechanism designer thereby improves the behavior on behalf of the students. Since the authors consider a house allocation problem, their adaptive Boston mechanism additionally involves randomly drawing priorities over the students. As such, there is no scope for strategic behavior on the part of the schools in their paper. Dur (2015) considers the modified Boston mechanism which also involves students never applying to schools that were filled in a previous round. Unlike Mennle and Seuken (2014) he allows for schools' priorities to be exogenously given, rather than randomly generated. The author also does not consider a strategic role for the schools. Dur (2015) shows that the modified Boston mechanism is less manipulable than the Boston mechanism in the sense of Pathak and Sönmez (2013) and that the set of Nash equilibria induced by the (complete information) preference revelation game equals the set of stable matchings under the true preferences.

Pais and Pintér (2008) experimentally compare the performance of the Boston, DA, and TTC mechanisms in terms of efficiency and manipulability. As expected from the theory, the frequency of manipulation under the Boston mechanism is greater than under the other two strategy-proof mechanisms, especially when participants (taking the role of the students) are given more information.

Schools' ability to independently determine priorities over students is present in many school choice procedures currently being used. Procedures for matching students to elementary schools in Ireland (Chen, 2016) and for secondary education in Amsterdam (De Haan, Gautier, Oosterbeek and Van der Klaauw, 2015) and Berlin (Basteck, Huesmann and Nax, 2016), for example, explicitly allow for schools to determine, sometimes subject to approval by the district school board, the criteria to be used to select students when demand exceeds the number of seats. Other than Ergin and Sönmez (2006), however, to our knowledge only two other papers consider the question of manipulation by schools of the Boston mechanism. Kojima (2008) generalized the earlier results of Ergin and Sönmez (2006) to a model in which restrictions on the possible preferences of the schools are relaxed. Treating schools' preferences as given, he shows that if schools' preferences satisfy a substitutability condition then the Nash equilibrium outcomes under the Boston mechanism are stable. He also shows that stable matchings can be supported by Nash equilibria under more general priority structures. He further provides an example in which a school may profitably manipulate the Boston mechanism when its preferences satisfy substitutability, but not responsiveness. Since we show that schools with responsive preferences may profitably manipulate the Boston mechanism by declaring some student unacceptable, this result is implied by our paper. More specifically, our result shows that it is not necessary to extend schools' preferences beyond responsiveness to obtain those incentives. Ehlers (2008) considers manipulations of priority mechanisms (manipulations by schools under the Boston mechanism being a special case of them) under incomplete information. The author shows that when agents have symmetric (incomplete) information, any non-truncation strategy is stochastically dominated by a truncation[1] of the true preferences of that agent. That does not imply our results, however, since it relies on his specific assumptions over beliefs. This strategic behavior under symmetric information is explored experimentally by Featherstone and Mayefsky (2011). Finally, the game induced by schools' ability to manipulate their capacities is analyzed in a series of papers for the case of stable mechanisms (Ehlers, 2010; Konishi and Ünver, 2006; Romero-Medina and Triossi, 2013). None of these results, however, imply our results on capacity manipulation.

We introduce our model and the variants of the Boston mechanism that we consider in section 2. Results concerning manipulability by schools and the stability of the Nash equilibria under the simultaneous Boston mechanism are obtained in section 3. The sequential Boston mechanism is analyzed in section 4. We conclude with a discussion in section 5.

2 Model

A **two-sided matching market** consists of:

1. A finite set of **students** $I = \{i_1, \ldots, i_n\}$,
2. A finite set of **schools** $S = \{s_1, \ldots, s_m\}$,
3. A **capacity vector** $q = (q_{s_1}, \ldots, q_{s_m})$,
4. A list of strict **student preferences** $P_I = (P_{i_1}, \ldots, P_{i_n})$ and
5. A list of strict **school preferences** $P_S = (P_{s_1}, \ldots, P_{s_m})$.

[1] A truncation strategy leaves the true ranking over students unchanged, but might drop some acceptable students.

Preference relations P_i for students are over the set of schools and the option of remaining unassigned, that is, $S \cup \{\varnothing\}$. Preference relations of the schools P_s are over sets of students. For each school $s \in S$ and any positive integer q_s the preference relation P_s is **responsive with capacity** q_s if for any set of students $J \subset S$ and students $i, i' \notin J$, $\{i\} P_s \{i'\}$ implies $J \cup \{i\} P_s J \cup \{i'\}$, and any set of students exceeding the capacity q_s is unacceptable. Let $\mathcal{P}_I$ be the set of all possible strict preferences over schools and $\mathcal{P}_S$ be the set of all possible responsive preferences over sets of students. Moreover, for a student i, let $\mathcal{P}_{-i}$ be the set of all possible strict preference profiles over schools for students $I \backslash \{i\}$, and $\mathcal{P}_{-s}$ and q_{-s} defined accordingly for school s. We will from here on abuse notation by denoting singleton sets by i or s and letting $\mathcal{P}_S$ and P_S denote the ranking over students associated with the underlying preferences over sets of students. Furthermore, we say that student i is **unacceptable** for school s if $\varnothing P_s i$. Unacceptable schools are defined analogously for students.

A **matching** μ is a function from $I \cup S$ to subsets of $I \cup S$ such that:

- $\mu(i) \in S \cup \{\varnothing\}$ and $|\mu(i)| = 1$ for every student i[2],
- $\mu(s) \subseteq I$ for every school s,
- $\mu(i) = s$ if and only if $i \in \mu(s)$.

A matching μ is **feasible** if for all $s \in S$, $|\mu(s)| \leqslant q_s$. The set of feasible matchings is denoted by $\mathcal{M}$. A matching is **individually rational** if for every student i, $\mu(i) \succ_i \varnothing$ and for every school s and every student $i' \in \mu(s)$, $i' \succ_s \varnothing$. A matching μ is **blocked** by a student i and school s if $sP_i\mu(i)$ and there is a set $I' \subseteq \mu(s) \cup \{i\}$ such that $i \in I'$ and $I' P_s \mu(s)$. A matching μ is **stable** if it is individually rational and is not blocked. A **school choice mechanism** Ψ is a mapping from the set of students' preferences, schools' capacities, and ranking over students to the set of matchings, i.e., $\Psi : \mathcal{P}_{\mathcal{I}} \times \mathcal{P}_S \times \mathbb{N}^m \to \mathcal{M}$. A mechanism is stable if it yields a stable matching for every profile of agents' preferences and schools' capacities.

A mechanism Ψ is **manipulable** by schools if there is a school s and school rankings P_s, P_s', capacities q_s, q_s' such that $q_s' \leqslant q_s$, $P_I \in \mathcal{P}_I$ and $P_{-s} \in \mathcal{P}_{-s}$ such that:

$$\Psi\left(P_I, \left(P_s', P_{-s}\right), \left(q_s', q_{-s}\right)\right) P_s \Psi\left(P_I, P_S, q\right)$$

Notice that by responsiveness no manipulation yielding a matching that is not feasible would make a mechanism manipulable. Manipulability by students is defined analogously except that students do not report capacities. In the case of mechanisms that are not manipulable, each agent has a weakly dominant strategy of submitting their true preferences. We distinguish between three types of manipulations by schools. First, a mechanism Ψ is **manipulable by declaring students unacceptable** if it is manipulable by a pair (P_s', q_s') such that $q_s' = q_s$ and there exists some $i \in I$ such that $iP_s\varnothing$ and $\varnothing P_s' i$. Second, a mechanism Ψ is **manipulable by a ranking change** if it is manipulable by a pair (P_s', q_s') such that $q_s' = q_s$ and for all $i \in I$, $iP_s\varnothing$ implies $iP_s'\varnothing$. Third, a mechanism Ψ is **manipulable via capacities** if it is manipulable by a pair (P_s', q_s') such that $P_s' = P_s$ and $q_s' < q_s$. If a

[2] We abuse notation and consider $\mu(i)$ as an element of S, instead of a set with an element of S.

mechanism is not manipulable by schools we say that for that mechanism truth-telling is a dominant strategy for the schools.

For manipulable mechanisms one cannot rely on the truthful revelation of preferences by all agents. Instead, we will examine the induced complete information preference revelation game and consider its Nash equilibria. For any agent $x \in I \cup S$ and her preference P_x, we can derive the corresponding weak preference relation R_x, where $cR_xc' \iff cP_xc'$ or $c = c'$.

Definition 1. A strategy profile $\left(\tilde{P}_I, \tilde{P}_S, \tilde{q}\right)$ is a **Nash equilibrium** of the game induced by the mechanism Ψ under preferences (P_I, P_S, q) if the following conditions hold:

(i) for all $s \in S$, $\hat{q}_s \in \mathbb{N}$ and $\hat{P}_s \in \mathcal{P}_s$, it holds that:

$$\Psi\left(\tilde{P}_I, \tilde{P}_S, \tilde{q}\right) R_s \Psi\left(\tilde{P}_I, \left(\hat{P}_s, \tilde{P}_{-s}\right), (\hat{q}_s, \tilde{q}_{-s})\right)$$

(ii) for all $i \in I$ and $\hat{P}_i \in \mathcal{P}_i$, it holds that:

$$\Psi\left(\tilde{P}_I, \tilde{P}_S, \tilde{q}\right) R_i \Psi\left(\left(\hat{P}_i, \tilde{P}_{-i}\right), \tilde{P}_S, \tilde{q}\right)$$

We say that $\Psi\left(\tilde{P}_I, \tilde{P}_S, \tilde{q}\right)$ is a **Nash equilibrium outcome** for preferences (P_I, P_S) and capacity q if $\left(\tilde{P}_I, \tilde{P}_S, \tilde{q}\right)$ is a Nash equilibrium of the game induced by the mechanism Ψ under (P_I, P_S, q).

In some of our results, we consider allowing students to observe reports by the schools before submitting their own preferences. In that case the strategy of a student is no longer simply the choice of a preference ordering but it is a preference ordering for each possible pair ranking/capacities reported by the schools. We denote by $f_i : \mathcal{P}_S \times \mathbb{N}^m_+ \to \mathcal{P}_i$ a strategy for a student in a preference revelation game in which schools move first. We let $\mathcal{F}_{\mathcal{I}}$ be the set of strategy profiles for all students. We say that a mechanism Ψ is a **sequential (school choice) mechanism** if students can observe the schools' reported preferences before themselves submitting preferences.[3] To analyze schools' incentives in a sequential mechanism we restrict the strategies that students play to those that are optimal against the other students' strategies given the submitted preferences and capacities of the schools. In other words, given the schools' reports the students are assumed to play Nash equilibrium strategies.

Definition 2. A strategy profile of students $f_I \in \mathcal{F}_I$ is **sequentially rational** (with respect to a sequential mechanism Ψ) if for all $(P_S, q) \in \mathcal{P}_S \times \mathbb{N}^m_+$ and for all $i \in I$ and $P_i \in \mathcal{P}_i$ it holds that:

[3] There is an alternative concept of sequential mechanisms in the school choice literature due to Dur and Kesten (2014). In their analysis schools are not strategic agents. There are two sets of schools that are assigned sequentially. In the first round students are assigned to one of the schools in the first set, based solely on their preferences over those schools. In the second round, students who were left unassigned in the first round are assigned to the second set of schools. In their case "sequential" thus refers to sequentially making an allocation decision. This allows us to use different allocation rules, such as the Boston mechanism, top trading cycles or deferred acceptance for different rounds. In contrast, "sequential" in our paper refers to schools submitting their preferences before the students with a fixed allocation rule.

$$\Psi\left(f_I\left(\tilde{P}_S,\tilde{q}\right),\tilde{P}_S,\tilde{q}\right)R_i\Psi\left(P_i,f_{-i}\left(\tilde{P}_S,\tilde{q}\right),\tilde{P}_S,\tilde{q}\right)$$

Definition 3. A sequential mechanism Ψ is **sequentially manipulable by schools** if there is a school s and school rankings P_s, P_s', capacities q_s, q_s' such that $q_s' \leqslant q_s$ and $P_{-s} \in \mathcal{P}_{-s}$ such that for all sequentially rational $f_I \in \mathcal{F}_I$:

$$\Psi\left(f_I\left(\left(P_s',P_{-s}\right),\left(q_s',q_{-s}\right)\right),\left(P_s',P_{-s}\right),\left(q_s',q_{-s}\right)\right)P_s\Psi\left(f_I\left(P_S,q\right),P_S,q\right)$$

The above definition of manipulability of a sequential mechanism restricts the set of strategies that students can play to those that are optimal for all possible reports made by the school and strategies chosen by the other schools.[4] More importantly, if a mechanism is sequentially manipulable by schools, it is not a dominant strategy for schools to submit their true rankings and/or capacities. While this definition is not standard, it captures the notion that schools can manipulate the sequential Boston mechanism even if students react optimally to their manipulations.[5] The definitions of **sequentially manipulable by declaring students unacceptable, sequentially manipulable by a ranking change** and **sequentially manipulable via capacities** are straightforwardly derived from the ones for the simultaneous game.

Definition 4. A profile $\left(f_I,\tilde{P}_S,\tilde{q}\right)$ is a **subgame perfect Nash equilibrium** of a mechanism Ψ under preferences (P_I,P_S) if the following conditions hold:

(i) for all $s\in S$, $\hat{P}_s\in\mathcal{P}_s$ and $\hat{q}_s\in\mathbb{N}$ we have:

$$\Psi\left(f_I\left(\tilde{P}_S,\tilde{q}\right),\tilde{P}_S,\tilde{q}\right)R_s\Psi\left(f_I\left(\left(\hat{P}_s,\tilde{P}_{-s}\right),\left(\hat{q}_s,\tilde{q}_{-s}\right)\right),\left(\hat{P}_s,\tilde{P}_{-s}\right),\left(\hat{q}_s,\tilde{q}_{-s}\right)\right)$$

(ii) for all $i\in I$ and $\hat{f}_i\in\mathcal{F}_i$ we have

$$\Psi\left(f_I\left(\tilde{P}_S,\tilde{q}\right),\tilde{P}_S,\tilde{q}\right)R_i\Psi\left(\left(\hat{f}_i\left(\tilde{P}_S,\tilde{q}\right),f_{-i}\left(\tilde{P}_S,\tilde{q}\right)\right),\tilde{P}_S,\tilde{q}\right)$$

We focus here on one particular mechanism, which was used in the Boston Public School Match before it was changed to the student-proposing deferred acceptance mechanism of Gale and Shapley (1962) in 2005 (Abdulkadiroğlu, Pathak, Roth and Sönmez, 2005a). Many other cities have used similar mechanisms to allocate school seats to students.

The Boston mechanism: Each student and school reports a preference order over each other and the option of being unmatched. Schools additionally report their capacities. Students and schools are matched in rounds.

[4]If one allowed students to play arbitrary strategies in a sequential mechanism then the requirement of non-manipulability would be too strong. For example, one could specify strategies that call for all students to rank a particular school first if that school is not truthful but to rank it last if the school reports truthfully. Under most matching algorithms this would yield an incentive not to tell the truth.

[5]The idea behind this definition is not that it is a property that is necessarily of independent interest. Rather it is a definition that allows us to precisely discuss how the Boston mechanism may give schools an incentive to misrepresent their preferences if students observe the schools' reports before submitting theirs.

- **Round 1:** Each student applies to the highest-ranked school according to the reported preference order. Each school accepts the students who applied to it in order of their priority according to the preference order reported by the school until there are either no more students who applied to the school or the school reaches its capacity. Students whose applications are unsuccessful are rejected.
- **Round $k \geqslant 2$:** Each student who was rejected in round $k-1$ applies to her k highest-ranked school if it is acceptable to the student. Otherwise the student is assigned to the outside option. Each school with remaining spots accepts students who applied to it in round k in order of their priority according to the preference order reported by the school until there are either no more students who applied or the school has reached its capacity. Students whose applications are unsuccessful are rejected.

The procedure ends when all students are either assigned a seat at a school or the outside option.

The Boston mechanism as described above will be referred to as the **two-sided simultaneous Boston mechanism**. The reason is that both sides of the market, students and schools, are taken as strategic agents who report preferences simultaneously. A commonly analyzed variant is the **one-sided Boston mechanism** in which only the students are seen as strategic agents and the schools' preferences and capacities are seen as administrative priorities which are simply observed or directly chosen by the market designer. We further consider the **two-sided sequential Boston mechanism,** which only differs from the two-sided simultaneous Boston mechanism in that students observe the rankings and capacities reported by the schools before reporting their own preferences. It is often more realistic to suppose that schools move first in school choice mechanisms since schools' priorities are often set in advance and communicated to prospective applicants. For example, a school could indicate that it accepts students according to their residence location or a specified weighted average of the students' grades in an exam. Since we are mainly concerned with the strategic behavior of schools under the Boston mechanism, we focus on the two-sided variants of the Boston mechanism, although we will make use of some existing results of the one-sided Boston mechanism.

In a setting in which schools are restricted to reporting rankings where every student is acceptable,[6] Ergin and Sönmez (2006) prove the following results:

Theorem. (Theorem 2 in Ergin and Sönmez, 2006) In the two-sided (*simultaneous*) version of the Boston mechanism, it is a dominant strategy for any school s to rank students based on its true preferences P_s. Moreover, any other dominant strategy of school s is outcome equivalent to truthfully ranking students based on P_s.

Theorem. (Theorem 1 in Ergin and Sönmez, 2006) Let P_I be the list of true student preferences, and consider the preference revelation game induced by the (*one-sided*) Boston mechanism. The set of Nash equilibrium outcomes of this game is equal to the set of stable matchings under the true preferences P_I.

[6]This rules out manipulating by declaring students unacceptable.

In the next section we show that the conclusions of the two theorems no longer hold if schools are allowed to declare some students unacceptable.

3 The Simultaneous Boston Mechanism

First, we show that schools may improve the set of students who are matched to them by declaring some students unacceptable:

Proposition 1. *The simultaneous two-sided Boston mechanism is manipulable by declaring students unacceptable.*

Proof. Consider the following two-sided matching market:

$$\begin{array}{rl} I = \{i_1, i_2, i_3\} & \qquad S = \{s_1, s_2\},\ q_1 = q_2 = 1 \\ P_{i_1} : s_2\ s_1 & \qquad P_{s_1} : i_1\ i_2\ i_3 \\ P_{i_2} : s_1\ s_2 & \qquad P_{s_2} : i_3\ i_1\ i_2 \\ P_{i_3} : s_2\ s_1 & \end{array}$$

The outcome of the Boston mechanism when students and schools submit their rankings truthfully is μ:

$$\mu = \begin{pmatrix} s_1 & s_2 & \varnothing \\ i_2 & i_3 & i_1 \end{pmatrix}$$

If school s_1 submits, instead, the ranking $P'_{s_1} : i_1\ i_3$ and the same capacity, the outcome of the Boston mechanism is μ' as follows:

$$\mu' = \begin{pmatrix} s_1 & s_2 & \varnothing \\ i_1 & i_3 & i_2 \end{pmatrix}$$

Since s_1 prefers student i_1 to i_2, school s_1 is better off after that manipulation. □

The rationale behind the manipulation of school s_1 is as follows. In the first round of the Boston mechanism only student i_2 applies to it. However, in the next round there will be an application by student i_1 which school s_1 prefers to i_2. By declaring i_2 unacceptable school s_1 can prevent student i_2 from taking its one seat and can then accept student i_1 in the second round.

Manipulating by declaring some students unacceptable appears to have been featured in some real-life markets. Consider, for example, the school assignment procedure used in New York City before the change to the Student-Proposing Deferred Acceptance Algorithm of Gale and Shapley (1962) in 2003. Before the change, there was no central authority coordinating the assignment. Strictly speaking, there was no use of the Boston mechanism in New York. However, the procedure used is roughly comparable to the Boston mechanism. Students could send a letter to up to five schools in the first round. Schools, upon receiving an application could decide whether to accept the student, put her on a waiting list or reject the student (Abdulkadiroğlu, Pathak and Roth, 2005b). After the change to a new mechanism it was reported that: "Before you might have a situation where a school was going to take 100 new children for ninth grade, they might have declared only 40 seats, and then placed the other 60 outside of the process" (*New York Times*, November 19, 2004). While it may appear as though schools may have manipulated capacities, this is not the case.[7] By declaring only 40 seats and then filling another 60 seats at a later stage, the schools effectively declared some students as unacceptable. Furthermore, the type of manipulation that has occurred suggests that the incentives to manipulate that we identified in Proposition 1 are of empirical relevance and importance. This suggests that the incentive to strategically withhold seats in order to accept more preferred applicants at a later stage, which we have identified under the Boston mechanism, is practically relevant. Another method by which schools could declare students as unacceptable is by setting grade thresholds below which students are not accepted or by setting other admissions requirements. Both of these manipulations may be argued to be due to an objective necessity of students satisfying such requirements.

Theorem 1. *If schools are allowed to deem students as unacceptable, the set of Nash Equilibrium outcomes in undominated strategies of the game induced by the two-sided simultaneous Boston mechanism may contain unstable matchings. Moreover, the resulting equilibrium outcome may Pareto dominate all stable matchings for the schools.*

Proof. Consider the following two-sided matching market:

$$\begin{aligned} I = \{i_1, i_2, i_3, i_4\} \quad & S = \{s_1, s_2, s_3\}, \; q_1 = 2, \; q_2 = q_3 = 1 \\ P_{i_1} : s_2 \; s_1 \; s_3 \quad & P_{s_1} : i_1 \; i_2 \; i_3 \; i_4 \\ P_{i_2} : s_1 \; s_2 \quad & P_{s_2} : i_2 \; i_1 \; i_3 \\ P_{i_3} : s_2 \; s_3 \; s_1 \quad & P_{s_3} : i_4 \; i_1 \; i_3 \\ P_{i_4} : s_1 \; s_3 \quad & \end{aligned}$$

The following strategy profile $\tilde{P}$ is a Nash equilibrium:

$$\begin{aligned} \tilde{P}_{i_1} : s_2 \; s_1 \; s_3 \quad & \tilde{P}_{s_1} : i_1 \; i_3 \\ \tilde{P}_{i_2} : s_2 \; s_1 \quad & \tilde{P}_{s_2} : i_2 \; i_1 \; i_3 \\ \tilde{P}_{i_3} : s_2 \; s_3 \; s_1 \quad & \tilde{P}_{s_3} : i_4 \; i_1 \; i_3 \\ \tilde{P}_{i_4} : s_3 \; s_1 \quad & \end{aligned}$$

The outcome in that equilibrium in the Boston mechanism is the matching μ as follows:

$$\mu = \begin{pmatrix} s_1 & s_2 & s_3 \\ i_1, i_3 & i_2 & i_4 \end{pmatrix}$$

[7] When manipulating capacities, a mechanism cannot assign more students to a school than its stated capacities. In New York schools clearly received more students than was possible in their initially declared capacity quota.

To see that this is an equilibrium, consider the deviations that schools and students could have:

- Student i_1 cannot be accepted by s_2 at the first step, since i_2 is also applying there at that step and has higher priority. Since there are no other seats at s_2, i_1 cannot profitably deviate.
- Student i_2 is not acceptable to school s_1, and therefore cannot profitably deviate.
- Student i_3 cannot be accepted at school s_3, since i_4 is applying there at the first step and has higher priority. She also cannot be accepted at s_2, since she has a lower priority than i_1 and i_2, who apply there at the first step. Therefore, i_3 also cannot profitably deviate.
- Student i_4 is not acceptable at school s_1, and therefore cannot profitably deviate.
- School s_1 would only be able to be better off if it got students i_1 and i_2. Since i_2 ranks school s_2 first and is accepted in the first step, the report s_1 makes does not affect where i_2 is allocated.
- Schools s_2 and s_3 already have their most preferred students, and so have no incentive to deviate.

The matching μ is not stable, however, since school s_1 and student i_2 form a blocking pair. To show that μ Pareto dominates for the schools all stable matchings, it suffices to show that μ Pareto dominates for the schools the school-optimal stable matching μ^S as follows:

$$\mu^S = \begin{pmatrix} s_1 & s_2 & s_3 \\ i_2, i_3 & i_1 & i_4 \end{pmatrix}$$

Given the schools' preferences, $\mu(s_1) P_{s_1} \mu^S(s_1)$, $\mu(s_2) P_{s_2} \mu^S(s_2)$ and $\mu(s_3) R_{s_3} \mu^S(s_3)$. Therefore, μ Pareto dominates μ^S for the schools.

It remains to show that the strategy profile $\tilde{P}$ is undominated for the schools and the students. First, a strategy that differs from an initial strategy only by declaring some unacceptable student/school to be acceptable cannot dominate the initial strategy. Second, for a school no strategy is dominated by another strategy that declares fewer students than the school's capacity as acceptable.

Third, truth-telling is not a dominated strategy for the students. To see this, consider for some student i a strategy profile for the schools so that only student i acceptable. Then any strategy of student i that does not truthfully reveal i's most preferred school does worse than truth-telling. Next consider for student i a strategy profile for the schools such that i's most preferred school finds i unacceptable, while all other schools find only i acceptable. Then no strategy that doesn't rank student i's second-most preferred school second or better can dominate truth-telling. This argument can be continued down to i's least preferred acceptable school.

Fourth, truth-telling is not a dominated strategy for the schools with capacity 1. To see this, consider a strategy profile for the students such that all students rank school s first. If school s does not truthfully reveal its most preferred student it does not get it, while it would under truth-telling. Hence truth-telling does better in that case than not revealing the most preferred student. Next, consider a strategy profile where all students, except the most preferred student of school s rank school s first. Then any strategy that ranks the second-best student for school s worse than under truth-telling does worse than truth-telling. This argument can be continued down to school s's least preferred acceptable student.

From the above it follows that the strategies under $\tilde{P}$ for schools s_2 and s_3 as well as for students i_1 and i_3 are undominated. For student i_2, the only strategy that could dominate $\tilde{P}_{i_2}$ is truth-telling as there are only two ways to rank the two schools acceptable to student i_2. Given $\tilde{P}_{-i_2}$ truth-telling would result in i_2 being unassigned, implying that $\tilde{P}_{i_2}$ is undominated. Similarly, for student i_4 the only strategy that could dominate $\tilde{P}_{i_4}$ is truth-telling. To see that it does not, consider a profile where student i_1 ranks s_3 first, s_1 declares i_4 to be unacceptable and i_4 and s_3 are truthful. Then i_4 is left unassigned, but would be allocated to s_3 if instead she played $\tilde{P}_{i_4}$.

Finally it remains to show that $\tilde{P}_{s_1}$ is not dominated. For school s_1, given its capacity of two seats, any preference profile that considers only i_1 and i_3 acceptable is outcome equivalent to $\tilde{P}_{s_1}$ and therefore does not dominate it. Given $\tilde{P}_{-s_1}$, any other strategy for s_1 that also declares two students acceptable makes s_1 strictly worse off and therefore cannot dominate $\tilde{P}_{s_1}$. Hence we only need to consider strategies that make more students acceptable. For any such strategy, consider a profile of preferences such that all students except i_1 apply to s_1 in the first round, while i_1 ranks another school, that considers i_1 acceptable, first and school s_1 second. Then the best possible outcome for s_1 is to obtain students i_2 and i_3. However, playing $\tilde{P}_{s_1}$ would yield i_1 and i_3 for school s_1, which it strictly prefers. Hence no strategy that finds more students acceptable than i_1 and i_3 dominates $\tilde{P}_{s_1}$. Therefore the equilibrium strategy profile $\tilde{P}$ consists only of undominated strategies. □

It's still the case that for any stable matching there exist preference reports such that the Nash equilibrium outcome of the two-sided simultaneous Boston mechanism equals the stable matching. This can be supported by simply letting each student state that only the school to which it is allocated under a stable matching is acceptable. Schools likewise rank only those students which are matched to them under the stable matching. If all agents play such strategies then no agent can gain by deviating.

While schools may have an incentive to manipulate their priorities by declaring some students unacceptable, there are no incentives to understate capacities.

Proposition 2. *The two-sided simultaneous Boston mechanism is not manipulable via capacities.*

Proof. Fix students' and schools' reported preference profile. Consider a school s with true capacity q_s that reports $\hat{q}_s < q_s$, while keeping all other schools' reports constant. In the first round of the Boston mechanism students apply to their favorite school. If school s receives less than $\hat{q}_s$ applications from students then the reported capacities does not affect which students get accepted by any school. The mechanism then moves to the next round. Let a_k be the number of students school s has accepted just before round k. As long as a_k plus the number of new acceptable applicants in round k is below $\hat{q}_s$, the reported capacities do not affect which student gets accepted by any school and the mechanism moves to the next round. However, if a_k plus the number of new acceptable applicants in round k exceeds $\hat{q}_s$ then s will simply reject up to $q_s - \hat{q}_s$ students who would be matched to it if it reported capacities truthfully. In any subsequent round school s cannot accept any more students if it reported $\hat{q}_s$. Therefore, by reporting $\hat{q}_s$, school s is being assigned a subset of the students who would be matched to it if it had reported its capacity truthfully. Since all students matched to s under the true capacity are acceptable to s, responsiveness of preferences imply that school s cannot prefer its assignment when reporting a smaller capacity. □

The intuition behind this result is similar to that behind Theorem 1 in Ergin and Sönmez (2006). By reporting a capacity below the true capacity, a school will still receive the same sequence of applications but it can accept only a subset of the applicants. Therefore, there is no opportunity to gain from pretending to have a smaller capacity.

4 The Sequential Boston Mechanism

In many school choice applications students who apply to schools are aware of how schools' priorities over students are formed. For example, in Boston it was well-known that students with a sibling attending a school would be given higher priority. The analysis so far has assumed that both students and schools report their preferences simultaneously. We now consider the case in which schools first submit their ranking over students, and second, students submit their ranking over schools after having observed how the schools rank students. This allows students to report different preferences depending on the observed priorities reported by the schools. The following proposition shows that this allows schools to manipulate the sequential Boston mechanism without having to declare a student unacceptable. This implies that the non-manipulability of the Boston mechanism requires both that schools cannot declare students unacceptable and that preferences and capacities are submitted simultaneously.

Proposition 3. *The two-sided sequential Boston mechanism is sequentially manipulable by ranking changes.*

Proof. We assume that students' strategy profiles are sequentially rational. Therefore, each student's strategy is chosen to be optimal against other students' strategies for all possible profile of schools' reported preferences. By Theorem 1 in Ergin and Sönmez (2006), this implies that every equilibrium outcome is stable with respect to schools' *reported* rankings and students' *true* preferences. Moreover, notice that sequential rationality fully (though not necessarily uniquely) specifies students' strategies at each subgame.

Consider the following two-sided matching market:

$$I = \{i_1, i_2, i_3, i_4, i_5\} \qquad S = \{s_1, s_2, s_3\}, \; q_1 = 3, \; q_2 = q_3 = 1$$
$$P_{i_1} : s_2 \; s_1 \; s_3 \qquad P_{s_1} : i_1 \; i_2 \; i_3 \; i_4 \; i_5$$
$$P_{i_2} - P_{i_5} : s_1 \; s_2 \; s_3 \qquad P_{s_2} : i_3 \; i_2 \; i_1 \; i_4 \; i_5$$
$$P_{s_3} : i_1 \; i_2 \; i_3 \; i_4 \; i_5$$

Consider the matching μ:

$$\mu = \begin{pmatrix} s_1 & s_2 & s_3 \\ i_2, i_3, i_4 & i_1 & i_5 \end{pmatrix}$$

We first show that μ is the unique stable matching with respect to schools' and students' *true* preferences. Since students i_1, i_2, i_3 and i_4 are all allocated to their first choice they cannot be part of a blocking pair. Student i_5 is the least preferred student of all schools and therefore not part of a blocking pair. To see that there is no other stable matching, note that since total capacity equals the number of students and since all students and schools are acceptable, every student has to be matched to a school at any stable matching. Student i_5 has to be matched to school s_3 at every stable matching. If i_5 were matched to another school $s \neq s_3$ then that school s and the student matched to school s_3 would form a blocking pair since student i_5 is the least preferred student of any school and school s_3 is the least preferred school of any student. It remains to be considered whether a student other than i_1 could be matched to school s_2 at some other stable matching. If i_2 were matched to s_2 and i_1, i_3, i_4 were matched to s_1 then (i_2, s_1) would form a blocking pair since $s_1 P_{i_2} s_2$ and $i_2 P_{s_1} i_4$. If i_3 were matched to s_2 then (i_3, s_2) would form a blocking pair since $s_1 P_{i_3} s_2$ and $i_3 P_{s_1} i_4$. Finally, if i_4 were matched to s_2 then (i_1, s_2) would form a blocking pair since $s_2 P_{i_1} s_1$ and $i_1 P_{s_2} i_4$. Hence, there is no other stable matching. Thus, the matching μ is the unique equilibrium outcome when schools report their rankings and capacities truthfully.

Consider the following deviation by school s_1: it reports the same capacity and P'_{s_1} : $i_1 \; i_2 \; i_4 \; i_5 \; i_3$ in the first stage of the sequential Boston mechanism while the other schools report their preferences truthfully. Then the unique stable allocation with respect to schools' *reported* rankings and students' *true* preferences (and therefore the unique equilibrium outcome of that subgame) is μ':

$$\mu' = \begin{pmatrix} s_1 & s_2 & s_3 \\ i_1, i_2, i_4 & i_3 & i_5 \end{pmatrix}$$

To see this, note that schools s_1 and s_2 get their most preferred students according to their reported preferences and thus cannot be part of a blocking pair. School s_3 is the least preferred school by any student and therefore also cannot be part of a blocking pair. To see that there is no other stable matching, note that again no student can be unassigned at a stable matching. Also, i_5 must be matched to s_3. If i_5 were assigned to a some other school $s \neq s_3$ then s and the student assigned to s_3 would block the matching. It remains to be considered whether a student other than i_3 could be matched to school s_2 at some other stable matching. If $i \in \{i_1, i_2, i_4\}$ were matched to s_2 (and thus i_3 is matched to s_1) then (i_5, s_1) would form a blocking pair as $i_5 P'_{s_1} i_3$ and $s_1 P_{i_5} s_3$.

School s_1, therefore, receives students $\mu'(s_1) = \{i_1, i_2, i_4\}$ with the deviation, which is a set of students strictly preferred to $\mu(s_1) = \{i_2, i_3, i_4\}$ for any responsive preferences consistent with school s_1's true ranking over students. □

Ergin and Sönmez (2006) have shown that if schools rank all students truthfully then the set of equilibrium outcomes under the Boston mechanism is equal to the set of stable matchings when schools have no unacceptable students. This leaves open the possibility of unstable equilibrium outcomes which are supported by schools not reporting their preferences truthfully. In such equilibria the schools would be indifferent between their equilibrium strategies and truthfully reporting their preferences. Furthermore, such equilibria require schools to play weakly dominated strategies. The result in Proposition 3 shows that when submitting their rankings before the students, manipulations may be profitable for schools. The reason is that by changing their reports, schools can also affect the students' reported preferences. In the example used to prove Proposition 3 school s_1, by shifting student i_3's ranking to the bottom, induces students to instead first apply to school s_2. This then leads student i_1 to apply first to school s_1.

If we assume that students play optimally in the second stage of the sequential Boston mechanism, the outcome will be a stable allocation with respect to the reported preferences of the schools. This, however, does not necessarily determine the outcome of the second stage since in many examples there is more than one stable matching. If we consider the (arbitrary) equilibrium selection rule that always picks the equilibrium yielding the student optimal stable matching then from the results of Roth (1985) it follows that schools will generally have an incentive to misstate their preferences. In addition, it follows from his paper that by manipulating their preferences, schools may obtain a matching that is preferred to the school-optimal stable matching. This is the logic that underlies our Theorem 2.

Theorem 2. *Holding schools' capacities fixed, the set of subgame perfect Nash equilibrium outcomes of the sequential Boston mechanism may contain matchings that are not stable with respect to agents' true preferences. Moreover, the resulting equilibrium may be weakly preferred by all schools to all stable matchings.*

Proof. Consider the example used to prove Proposition 3 and the matching μ' in that proposition's proof.

$$\mu' = \begin{pmatrix} s_1 & s_2 & s_3 \\ i_1, i_2, i_4 & i_3 & i_5 \end{pmatrix}$$

that results when s_1 reports P'_{s_1} : $i_1\ i_2\ i_4\ i_5\ i_3$ and the other schools report their preferences truthfully. We will show that the schools' reports $(P'_{s_1}, P_{s_2}, P_{s_3})$ constitute a Nash equilibrium. School s_2 gets its most preferred student i_3 and therefore cannot gain by any deviation. School s_3 is the least liked by all the students. Suppose that for some other reported preference profile school s_3 obtains a student $i \in \{i_1, i_2, i_3, i_4\}$. If $i = i_3$, then i and s_2 constitute a blocking pair. If $i \in \{i_1, i_2, i_4\}$ then i and school s_1 constitute a blocking pair because i is among the three most preferred students of s_1 according to P'_{s_1}. Therefore, it cannot be a continuation equilibrium for s_3 to obtain a student other than i_5 irrespective of the report submitted by

s_3. Thus, s_3 has no incentive to deviate. Lastly, consider the incentives of school s_1 to deviate. To gain from a deviation s_1 needs to be matched to students $\{i_1, i_2, i_3\}$. Suppose that there is some report $\tilde{P}_{s_1}$ that yields this outcome for s_1 while the other schools report truthfully. Since the outcome must be stable with respect to students' true preferences and $(\tilde{P}_{s_1}, P_{s_2}, P_{s_3})$, in a continuation equilibrium yielding this outcome for s_1 we would need i_4 to be matched to s_2 and i_5 to s_3. Otherwise (i_4, s_2) would be a blocking pair. However, in this case (i_1, s_2) would then be a blocking pair. The outcome is thus unstable, which consists of a contradiction with school s_1 profitably deviating from P'_{s_1}. It follows that s_1 also cannot gain by a deviation. Therefore μ' is a subgame perfect Nash equilibrium outcome of the sequential Boston mechanism.

It remains to be noted that μ' is not stable with respect to schools' and students' true preferences. The pair (i_3, s_1) blocks the matching, since $s_1 P_{i_3} \mu'(i_3) = s_2$, $i_4 \in \mu'(s_1)$ and $i_3 P_{s_1} i_4$. Finally, since schools s_1 and s_2 strictly prefer μ' over μ (which is the unique stable matching with respect to the true preferences) and school s_3 is indifferent between them, μ' is a Pareto improvement for the schools over all stable matchings. □

In our proof we relied on the fact that the set of Nash equilibrium outcomes in undominated strategies of the one-sided Boston mechanism is the set of stable matchings with respect to students' true preferences. This is noteworthy because even the student-proposing deferred acceptance mechanism (Gale and Shapley, 1962) has equilibria in undominated strategies that are not stable (Roth and Sotomayor, 1990) when schools report preferences truthfully. In other words, while the Boston mechanism fully implements the set of stable matchings in undominated strategies, that is not the case for the student-proposing deferred acceptance mechanism and so the incentives that schools have under the latter doesn't necessarily translates into incentives in the former. The equivalence of the set of outcomes of the one-sided Boston mechanism to the set of stable matchings simplified our equilibrium analysis since it allowed us to consider the outcomes of continuation equilibria directly without having to explicitly specify and consider the strategies played by the students which give rise to these outcomes.

One feature of the preference relation P'_{s_1} that we use in the proofs of Proposition 3 and Theorem 2 is that it does not declare any student unacceptable. For the simultaneous Boston mechanism Ergin and Sönmez (2006) show that such manipulations cannot yield a better outcome for the schools. What our result highlights is that the timing of preference submission is a critical assumption in that result. The manipulability of the sequential Boston mechanism by declaring a student unacceptable and the instability of its equilibrium outcome can be shown by following the same steps of the proofs of Proposition 3 and Theorem 2 by school s_1 declaring student i_3 to be unacceptable.

Remark 1. The two-sided sequential Boston mechanism is sequentially manipulable by declaring students unacceptable. Moreover, the set of subgame perfect Nash equilibrium (SPNE) outcomes may contain matchings that are not stable with respect to agents' true preferences.

If we consider the case in which students submit their preferences before the schools, Theorem 2 in Ergin and Sönmez, 2006 implies that schools will not have any incentive to misrepresent their rankings. Therefore, students will play the preference revelation game induced by the one-sided version of the Boston mechanism and the corresponding results in Ergin and Sönmez, 2006, hold.

Remark 2. If students submit their preferences first and schools submit theirs afterward under the Boston mechanism, the set of subgame perfect Nash equilibria outcomes equals the set of stable matchings under the true preferences, when all students are required to be acceptable by all schools.

For the simultaneous Boston mechanism we have shown that schools do not have an incentive to manipulate their capacities. We now show that this result also depends on the timing of the game. Specifically, when schools report their capacities before students report their preferences (and capacities are observable by the students) then schools may have an incentive to manipulate their capacities.

Proposition 4. *The two-sided sequential Boston mechanism is sequentially manipulable via capacities.*

Proof. Consider the following example. Preferences for both schools are responsive.[8]

$$I = \{i_1, i_2, i_3, i_4\} \qquad S = \{s_1, s_2\}, \ q_1 = 3, \ q_2 = 2$$
$$P_{i_1} : s_2 \ s_1 \qquad P_{s_1} : \{i_1, i_2, i_3\} \ \{i_1, i_2\} \ \{i_1, i_3\} \ \{i_1\} \ \{i_2, i_3\} \ \{i_1\} \ \{i_2\} \ \{i_3\} \ \{i_4\}$$
$$P_{i_2} : s_1 \ s_2 \qquad P_{s_2} : i_4 \ i_3 \ i_2 \ i_1$$
$$P_{i_3} : s_1 \ s_2$$
$$P_{i_4} : s_2 \ s_1$$

For school s_1 we have only shown preferences that are relevant to our analysis (that is, we do not show how the preferences of school s_1 over sets including student i_4 are). We assume that schools' priorities are fixed. Given the school's preferences and capacities, by sequential rationality and Theorem 1 of Ergin and Sönmez (2006), the outcome of the subgame played by the students after schools report their capacities is stable with respect to their *true* preferences and the schools' *reported* capacities. Hence, it is sufficient to consider the resulting set of stable matchings for each possible combination of capacities reported by the schools. When the reported capacities are $(3, 2)$ it is easy to verify that the unique stable allocation is given by:

$$\mu = \begin{pmatrix} s_1 & s_2 \\ i_2, i_3 & i_1, i_4 \end{pmatrix}$$

When the reported capacities are $(1, 2)$ the unique stable allocation is given by:

$$\hat{\mu} = \begin{pmatrix} s_1 & s_2 \\ i_1 & i_3, i_4 \end{pmatrix}$$

Since $\{i_1\} P_{s_1} \{i_2, i_3\}$, school s_1 gains from understating its capacity by 2. □

[8]For this example it is not necessary to specify the preferences of school s_2 beyond its ranking over singleton sets of students.

Unlike in the simultaneous Boston mechanism, schools may have incentives to misstate their capacity in the sequential Boston mechanism. The intuition is similar to preference manipulations of the sequential Boston mechanism: by misstating its capacity, a school may change the set of stable matchings.

The next proposition considers the game induced by the sequential Boston mechanism when schools' only strategic variable is their reported capacity.

Theorem 3. *Holding schools' preferences fixed, the set of SPNE outcomes of the sequential Boston mechanism may have no stable matching when schools are only able to manipulate their capacities. Moreover all SPNE outcomes may be strictly preferred by the schools over all stable allocations.*

Proof. Consider the example in the proof of Proposition 4. The only stable allocation, under true capacities is μ:

$$\mu = \begin{pmatrix} s_1 & s_2 \\ i_2, i_3 & i_1, i_4 \end{pmatrix}$$

Clearly, the profile $(3, 1)$ cannot be part of an SPNE strategy profile, since school s_2 would only obtain student s_4, which is worse than the outcome under $(3, 2)$. If school s_1 reports $(1, 2)$, instead, the equilibrium outcome is unique and given by:

$$\hat{\mu} = \begin{pmatrix} s_1 & s_2 \\ i_1 & i_3, i_4 \end{pmatrix}$$

which is preferred by both schools to μ and for school s_2 is its most preferred set of students. Therefore, $(3, 2)$ is not part of an SPNE strategy profile. Moreover, it follows that $(1, 1)$ is not part of an SPNE strategy profile, as s_2 would prefer to state its capacity truthfully. Consider next what happens when the reported capacities are $(2, 2)$. In that case there are two stable matchings (and thus two possible continuation equilibrium outcomes, given students' strategies). These are μ and $\tilde{\mu}$:

$$\tilde{\mu} = \begin{pmatrix} s_1 & s_2 \\ i_1, i_2, & i_3, i_4 \end{pmatrix}$$

Note that under $\tilde{\mu}$ both schools get their most preferred set of two students, while μ is the unique stable outcome. There are two equilibrium outcomes for the subgame after reported capacities $(2, 2)$. When μ is the outcome, then $(2, 2)$ is not part of an SPNE strategy profile, since school s_1 prefers the outcome of $(1, 2)$ to μ, which it can reach by deviating. If instead the outcome is $\tilde{\mu}$, then $(2, 2)$ is part of a SPNE strategy profile. Lastly, $(2, 1)$ is not part of an SPNE strategy profile, since s_2 only obtains student s_4, which is worse than the outcome under $(2, 2)$, irrespective of the following equilibrium outcome.

Finally, note that the two possible equilibrium outcomes, $\tilde{\mu}$ and $\hat{\mu}$, are strictly preferred by both schools over the unique stable allocation μ. □

The proof of Theorem 3 shows that there are situations in which no stable allocation is supported by a subgame perfect Nash equilibrium when schools can only misrepresent their

Manipulations by Schools		Simultaneous Boston	Sequential Boston
Ranking Changes	Manipulable	No (Ergin and Sönmez, 2006)	Yes (Proposition 3)
	Stable	Yes (Ergin and Sönmez, 2006)	No (Theorem 2)
Acceptability	Manipulable	Yes (Proposition 1)	Yes (Remark 1)
	Stable	No (Theorem 1)	No (Remark 1)
Capacity	Manipulable	No (Proposition 2)	Yes (Proposition 4)
	Stable	Yes (Proposition 2)	No (Theorem 3)

Table 1: Summary of the results for manipulability and stability of equilibrium outcomes for the simultaneous and sequential Boston mechanisms.

capacities, but not their preferences. By misstating their capacities schools can obtain an outcome that all of them strictly prefer to the stable allocation.

5 Concluding Remarks

The contributions of our paper with respect to the previous literature are summarized in table 5.

We point out that all the results shown in this paper also hold under the modified Boston mechanism, proposed by Dur (2015). The first round of the modified Boston mechanism and of the (standard) Boston mechanism are identical, but for any subsequent round we instead have:

- **Round $k \geqslant 2$:** Each student who was rejected in round $k-1$ applies to her next most preferred school that at the end of round $k-1$ still has at least one open seat. Otherwise the student is assigned to the outside option. Each school that has remaining spots accepts students who applied to it in round k in order of their priority according to the preference order reported by the school until there are either no more students who applied or the school has reached its capacity. Students whose applications are unsuccessful are rejected.

The difference is that under the modified Boston mechanism students are somewhat protected from strategic errors as they never apply to a school that has reached full capacity before the student has even applied to it. Note that when students play optimally there is no difference in the outcomes that are obtained under the modified Boston mechanism and the standard Boston mechanism, which implies that the set of Nash equilibria that can be obtained is identical for both mechanisms. Additionally, for the examples we use in our results, the outcome of the Boston Mechanism and the Modified Boston Mechanism are identical.

Remark. Propositions 1 and 3, and theorems 1 and 2 hold for the modified Boston mechanism (Dur, 2015).

References

Abdulkadiroğlu, Atila and Tayfun Sönmez, "School Choice: A Mechanism Design Approach," *American Economic Review*, 2003, pp. 729–747. 1

____ **, Parag A Pathak, Alvin E Roth, and Tayfun Sönmez**, "The Boston public school match," *American Economic Review*, 2005, pp. 368–371. 2

____ **,** ____ **, and** ____ , "The New York City High School Match," *American Economic Review*, 2005, pp. 364–367. 3

____ **, Yeon-Koo Che, and Yosuke Yasuda**, "Resolving Conflicting Preferences in School Choice: The "Boston Mechanism" Reconsidered," *American Economic Review*, 2011, *101* (1), 399–410. 1

Basteck, Christian, Katharina Huesmann, and Heinrich Nax, "Matching Practices for secondary schools – Germany," 03 2016. 1

Chen, Li, "Matching Practices for Elementary Schools – Ireland," 03 2016. 1

Dur, Umut, "The Modified Boston Mechanism," *Working Paper*, April 2015. 1, 5

____ **and Onur Kesten**, "Sequential versus simultaneous assignment systems and two applications," Technical Report, Boston College, mimeo 2014. 3

Ehlers, Lars, "Truncation strategies in matching markets," *Mathematics of Operations Research*, 2008, *33* (2), 327–335. 1

____ , "Manipulation via capacities revisited," *Games and Economic Behavior*, 2010, *69* (2), 302–311. 1

Ergin, Haluk and Tayfun Sönmez, "Games of school choice under the Boston mechanism," *Journal of public Economics*, 2006, *90* (1), 215–237. 1, 2, 3, 4, 4, 4, 4, 5

Featherstone, Clayton R. and Eric Mayefsky, "Why Do Some Clearinghouses Yield Stable Outcomes? Experimental Evidence on Out-of-Equilibrium Truth-Telling," *Working Paper*, 2011. 1

Gale, David and Lloyd S Shapley, "College Admissions and the Stability of Marriage," *The American Mathematical Monthly*, 1962, *69* (1), 9–15. 1, 2, 3, 4

Haan, Monique De, Pieter A Gautier, Hessel Oosterbeek, and Bas Van der Klaauw, "The performance of school assignment mechanisms in practice," 2015. 1

Kojima, Fuhito, "Games of school choice under the Boston mechanism with general priority structures," *Social Choice and Welfare*, 2008, *31* (3), 357–365. 1

____ **and M Utku Ünver**, "The "Boston" School-Choice Mechanism: An Axiomatic Approach," *Economic Theory*, 2014, *55* (3), 515–544. 1

Konishi, Hideo and M Utku Ünver, "Games of capacity manipulation in hospital-intern markets," *Social choice and Welfare*, 2006, *27* (1), 3–24. 1

Mennle, Timo and Sven Seuken, "The Naïve versus the Adaptive Boston Mechanism," *Working Paper*, 2014. 1

Miralles, Antonio, "School choice: The Case for the Boston Mechanism," in "Auctions, Market Mechanisms and Their Applications," Springer, 2009, pp. 58–60. 1

Pais, Joana and Ágnes Pintér, "School choice and information: An experimental study on matching mechanisms," *Games and Economic Behavior*, 2008, *64* (1), 303–328. 1

Pathak, Parag A and Tayfun Sönmez, "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism," *The American Economic Review*, 2008, *98* (4), 1636–1652. 1

Pathak, Parag A. and Tayfun Sönmez, "School Admissions Reform in Chicago and England: Comparing Mechanisms by Their Vulnerability to Manipulation," *American Economic Review*, 2013, *103* (1), 80–106. 1

Romero-Medina, Antonio and Matteo Triossi, "Games with capacity manipulation: incentives and Nash equilibria," *Social Choice and Welfare*, 2013, *41* (3), 701–720. 1

Roth, Alvin E, "The college admissions problem is not equivalent to the marriage problem," *Journal of economic Theory*, 1985, *36* (2), 277–288. 4

_____ , "A natural experiment in the organization of entry-level labor markets: regional markets for new physicians and surgeons in the United Kingdom," *The American economic review*, 1991, pp. 415–440. 1

_____ **and Andrew Postlewaite**, "Weak versus strong domination in a market with indivisible goods," *Journal of Mathematical Economics*, 1977, *4* (2), 131–137. 1

_____ **and Marilda A Oliveira Sotomayor**, *Two-sided matching: A study in Game-theoretic Modeling and Analysis* number 18, Cambridge University Press, 1990. 4

Discussion Papers of the Research Area Markets and Choice 2016

Research Unit: **Market Behavior**

David Danz, Steffen Huck, Philippe Jehiel SP II 2016-201
Public statistics and private experience:
Varying feedback information in a take-or-pass game

Jana Friedrichsen SPII 2016-202
Signals sell: Designing a product line when consumers have
social image concerns

Uri Gneezy, Silvia Saccardo, Roel van Veldhuizen SPII 2016-203
Bribery: Greed versus reciprocity

Inácio Bó, C.-Philipp Heller SPII 2016-204
Strategic schools under the Boston mechanism revisited

Research Unit: **Economics of Change**

Armin Falk, Nora Szech SP II 2016-301
Pleasures of skill and moral conduct

Thomas Deckers, Armin Falk, Fabian Kosse, Nora Szech SP II 2016-302
Homo moralis: Personal characteristics, institutions, and
moral decision-making

Jenny Simon, Justian Mattias Valasek SP II 2016-303
The political economy of multilateral aid funds

Vittorio Bassi, Steffen Huck, Imran Rasul SP II 2016-304
A note on charitable giving by corporates and aristocrats:
Evidence from a field experiment

Ludwig Ensthaler, Steffen Huck, Johannes Leutgeb SP II 2016-305
Games played through agents in the laboratory – A test of
Prat & Rustichini's model

Maja Adena, Steffen Huck SP II 2016-306
Online fundraising, self-deception, and the long-term impact
of ask avoidance

Research Professorship: **Market Design: Theory and Pragmatics**

Kevin McLaughlin, Daniel Friedman SP II 2016-501
Online ad auctions: An experiment

All discussion papers are downloadable:
http://www.wzb.eu/en/publications/discussion-papers/markets-and-choice