Soc Choice Welfare (2008) 31:357–365
DOI 10.1007/s00355-007-0287-3

ORIGINAL PAPER

Games of school choice under the Boston mechanism with general priority structures

Fuhito Kojima

Received: 17 May 2007 / Accepted: 9 November 2007 / Published online: 8 December 2007

Abstract The Boston mechanism is a centralized student assignment mechanism used in many school districts in the US. We investigate strategic behavior of students under the Boston mechanism when schools may have complex priority structures. We show that a stable matching is supported as an outcome of a Nash equilibrium under a general environment. We further show that any outcome of a Nash equilibrium is a stable matching when the school priorities are substitutable.

1 Introduction

The theory of two-sided matching is an area of game theory with many practical applications.[1] It is applied to the analysis and the design of entry-level labor markets, such as those for medical residents and hospitals.[2] More recently, Balinski and Sönmez (1999) and Abdulkadiroğlu and Sönmez (2003) investigate a closely related problem of school choice. Their theories lead to the practical design of school choice programs in New York City and Boston.[3]

[1] Roth and Sotomayor (1990) is a comprehensive survery.

[2] See Roth (1984) and Roth and Peranson (1999).

[3] See Abdulkadiroğlu et al. (2005a), Abdulkadiroğlu et al. (2005b) and Abdulkadiroğlu et al. (2006).

Fuhito Kojima is grateful to Alvin E. Roth, Tayfun Sönmez and anonymous referees for insightful comments.

F. Kojima (✉)
Department of Economics, Harvard University,
Cambridge, MA 02138, USA
e-mail: kojima@fas.harvard.edu

Many school districts use the so-called Boston mechanism.[4] Under this mechanism, students submit their preference lists to the central clearing house. Given the information, the clearinghouse follows an algorithm that tries to match as many students to their stated first choice schools as possible subject to pre-specified priorities of students at each school. Chance that a student is admitted in her preferred school is affected by her stated preferences, and the Boston mechanism induces a nontrivial game among students. In a recent paper, Ergin and Sönmez (2006) analyze the game of school choice induced by the Boston mechanism. Under simple priority structures of schools, they show that the set of Nash equilibrium outcomes coincide with the set of stable matchings. This result has a strong policy implication. The equilibrium assignment is Pareto dominated by the assignment by a student-optimal stable mechanism (Gale and Shapley 1962). Thus the analysis gives some support to using this version of stable mechanism in school choice problems rather than the Boston mechanism.

A natural question is how students behave under the Boston mechanism when school priorities are more complex than assumed in the previous analysis. In many school districts, priorities are quite complex. For instance, the Seattle school district used to utilize a "racial tiebreaker": if the racial balance of the school's student body varied more than fifteen percent from the district average, students are given priority if they make the ratio closer to the average ratio.[5] Similarly, some schools in New York City require balanced distributions of test score. Most existing analyses on school choice mechanisms including Ergin and Sönmez (2006) concentrate on simple priority structures rather than such complex priority structures.[6] Given that school districts need to meet their unique and often quite complex requirements with respect to balance of their student distributions, it is important to know how students behave and if the welfare comparison between the Boston mechanism and alternative algorithms carries over under complex environments.

This paper analyzes the Boston mechanism in such complex environments. We propose a natural extension of the existing Boston mechanism under a general priority structure of schools. We show that, under general priority structures, a stable matching is an outcome of a Nash equilibrium. We then show that, when the school priority is substitutable, any outcome of a Nash equilibrium is a stable matching. Hatfield and Milgrom (2005) show that truthtelling is a dominant strategy for students under the student-optimal stable mechanism when the priority structure satisfies a condition called the law of aggregate demand in addition to substitutability. These results imply that there are welfare gains in changing the assignment mechanism from the Boston mechanism to the student-optimal stable mechanism as long as schools have substitutable priority structures that satisfy the law of aggregate demand. We also give an

[4] The Boston mechanism is recently replaced by a student-optimal stable mechanism in Boston. We nevertheless use this term since it is common in the literature.

[5] Seattle Public Schools suspended the use of the racial tiebreaker after the US. Court of Appeals ordered to stop its use in 2002. The issue is yet to be settled down completely as of this writing.

[6] Ergin and Sönmez (2006) analyze the controlled choice problem with type-specific quotas. This class of priorities includes those found in some cities, such as Minneapolis and Boston before 1999 while not others such as the one presented above.

example to show that outcomes of Nash equilibria are not necessarily stable when the priority structure is not substitutable.

2 Model

A school choice problem is tuple $(S, C, (\succeq_i)_{i \in S \cup C})$. S and C are finite and disjoint sets of students and schools. For each student $s \in S$, $\succeq_s$ is a strict preference relation over C and being unmatched (being unmatched is denoted by $\emptyset$). For each school $c \in C$, $\succeq_c$ is a **priority structure**, which is a strict, complete and transitive binary relation over the set of subsets of students. If $j \succ_i \emptyset$, then j is said to be **acceptable** to i.

Given a set of students $S' \subseteq S$, define $Ch_c(S')$ to be the set such that $Ch_c(S') \subseteq S'$ and $Ch_c(S') \succeq_c S''$ for any $S'' \subseteq S'$. In words, $Ch_c(S')$ is the set of students that school c chooses if it can freely choose from S' according to its priority structure.

Given two sets of students S' and S'' with $S'' \subseteq S'$, let $Ch_c(S'|S'')$ be the best group of students for c when c has to accept all the students in S''. Formally, $Ch_c(S'|S'')$ is a set such that $S'' \subseteq Ch_c(S'|S'') \subseteq S'$ and $Ch_c(S'|S'') \succeq_c T$ for any T with $S'' \subseteq T \subseteq S'$.

A **matching** μ is a mapping from $C \cup S$ to $C \cup S$ such that (i) $\mu(s) \subseteq C$ and $|\mu(s)| \leq 1$ for every s, (ii) $\mu(c) \subseteq S$, and (iii) $\mu(s) = c$ if and only if $s \in \mu(c)$.[7] For any matchings μ and μ' and for any $i \in S \cup C$, we write $\mu \succeq_i \mu'$ if and only if $\mu(i) \succeq_i \mu'(i)$.

Given a matching μ, we say that it is **blocked** by $(s, c) \in S \times C$ if $c \succ_s \mu(s)$ and $s \in Ch_c(\mu(c) \cup s)$. A matching μ is **individually rational** if $\mu(s) \succeq_s \emptyset$ for every $s \in S$ and $\mu(c) = Ch_i(\mu(c))$ for every $c \in C$. A matching μ is **stable** if it is individually rational and is not blocked.

For each school $c \in C$, its priority structure $\succeq_c$ is **substitutable** if for any $S', S'' \subseteq S$ with $S'' \subseteq S'$, we have $Ch_c(S') \cap S'' \subseteq Ch_c(S'')$ (Kelso and Crawford 1982). That is, a student who is chosen from a larger set of students is always chosen from a smaller one. If every school has substitutable priorities, then there exists a stable matching (Kelso and Crawford 1982). In the school choice context, stronger conditions are often assumed. A priority structure is **responsive** if it is consistent with priorities over individual students up to its capacity. A priority structure is **responsive with type-specific quotas** if it is consistent with priorities over individual students as long as the number of a specific type does not exceed a given type-specific quota. Gender and test scores are examples of types, and race can also be an example in some school districts. Both of these priorities can be shown to satisfy substitutability (Roth and Sotomayor 1990, Abdulkadiroğlu 2005).[8]

Consider the following mechanism, called the **Boston mechanism**. At the beginning of the mechanism, each student submits a preference list, which is a sequence of schools and the option of being unmatched without repetition. The first element of the

[7] We abuse notation and denote a singleton set $\{x\}$ by x whenever there is no concern for confusion.

[8] Formal definitions of responsiveness and responsiveness with type-specific quotas are omitted. See Roth and Sotomayor (1990) and Abdulkadiroğlu (2005).

list is interpreted as her "first choice" school, the second element as her second choice, and so on. Given stated preference lists of students and exogenously given priorities of schools, the mechanism follows the algorithm below.

- *Step 1*: Each student applies to her first choice school. Let the set of students who apply to school c at this step be denoted by $A^1(c)$. School c admits all the students in $Ch_c(A^1(c))$ and rejects everyone else. Let $B^1(c) = Ch_c(A^1(c))$.

In general,

- *Step t*: Each student who was rejected in Step $(t-1)$ applies to her next highest choice. Let the set of students who apply to school c at this step be denoted by $A^t(c)$. School c admits all the students in $Ch_c(A^t(c) \cup B^{t-1}(c)|B^{t-1}(c))$ and rejects everyone else. Let $B^t(c) = Ch_c(A^t(c) \cup B^{t-1}(c)|B^{t-1}(c))$.

The algorithm terminates either when every student is matched to a school or every unmatched student has been rejected by every school that is acceptable according to her stated preference list. The final assignment is given by $\mu(c) = B^T(c)$ for every $c \in C$, where T is the step in which the algorithm terminates.

Abdulkadiroğlu and Sönmez (2003) and Ergin and Sönmez (2006) define and analyze the Boston mechanism when school priorities are responsive or responsive with type-specific quotas. The current definition of the Boston mechanism is a generalization of theirs for general priority structure.

In real school choice problems, students may have incentives to report false preferences. Hence the above mechanism induces a nontrivial game of school choice. More specifically, we assume that schools are not strategic and only students are players of the game, except in Remark 1 where strategic schools are considered. School priorities and student preferences are common knowledge. Students (strategically) submit their preference lists simultaneously. The outcome is given by the Boston mechanism with respect to the submitted preference lists.

In the rest of the paper, we analyze the set of Nash equilibria of this game. Once one defines the Boston mechanism appropriately, it turns out that the analysis and insight of Ergin and Sönmez (2006) carry over quite naturally to general priority structures.

Proposition 1 *Under a general priority structure, if μ is a stable matching, then there exists a Nash equilibrium with outcome μ under the Boston mechanism.*

Proof Let μ be a stable matching. Consider a strategy profile such that every student lists $\mu(s)$ as her first choice. Under this strategy profile, for each school c, $\mu(c)$ is the set of students who apply to c at the first step of the Boston mechanism. Since $\mu(c) = Ch_c(\mu(c))$ by individual rationality of μ, c accepts all these students. Thus the mechanism terminates at the first step with outcome μ. Now we show that the above strategy profile forms a Nash equilibrium. Suppose that there exists a student s and a strategy such that s is matched to a school $c \neq \mu(s)$. Since any other student s' applies to $\mu(s')$ at the first stage, this means that $s \in Ch_c(\mu(c) \cup s)$. Since μ is stable, this implies that $\mu(s) \succ_s c$, implying that such a deviation of s is not profitable. This shows that the above strategy profile is a Nash equilibrium. □

The above proposition asserts that any stable matching can result as a Nash equilibrium outcome under the Boston mechanism. A natural question is whether other outcomes may arise as an equilibrium. The next proposition answers this question when school priorities are substitutable.

Proposition 2 *Suppose that schools have substitutable priorities. If μ is an outcome of a Nash equilibrium under the Boston mechanism, then μ is stable.*

Proof Suppose that μ is not stable. Then there exists a student s and school c such that $c \succ_s \mu(s)$ and $s \in Ch_c(\mu(c) \cup s)$. Consider the first step of the Boston mechanism. By the construction of the Boston mechanism, $B^1(c) = Ch_c(A^1(c))$, the set of students who are accepted by c at the first step, is a subset of $\mu(c)$. Suppose that s submits a profile in which c is put as her first choice. Then, now the set of students who apply to c at the first step is $A^1(c) \cup s$. Since the priority structure of c is substitutable and $B^1(c) \subseteq \mu(c)$, we have that $s \in Ch_c(B^1(c) \cup s) = Ch_c(Ch_c(A^1(c)) \cup s) = Ch_c(A^1(c) \cup s)$.[9] Therefore s is accepted by c at the first step. Since a school never rejects a student who was accepted in an earlier step under the Boston mechanism, s is matched to c in the final matching. Since $c \succ_s \mu(s)$, this implies that the original strategy profile is not a Nash equilibrium. □

Theorems 1 and 3 of Ergin and Sönmez (2006) show that the set of Nash equilibrium outcomes and the set of stable matchings are equivalent when school priorities are responsive and responsive with type-specific quotas, respectively. Since these priorities are special cases of substitutable priorities, Propositions 1 and 2 generalize their results.

The next example shows that the conclusion of Proposition 2 does not necessarily hold when the priority structure is not substitutable.

Example 1 $C = \{c_1, c_2\}$, $S = \{s_1, s_2, s_3, s_4\}$. School priorities are given by

$$c_1 : \{s_1, s_2\}, \{s_1, s_3\}, s_3,$$
$$c_2 : s_4, s_1.$$

Student preferences are given by

$$s_1 : c_2, c_1,$$
$$s_2 : c_1,$$
$$s_3 : c_1,$$
$$s_4 : c_2.$$

Truthtelling of student preferences results in the matching μ given by

$$\mu(c_1) = s_1, s_3,$$
$$\mu(c_2) = s_4,$$
$$\mu(s_2) = \emptyset.$$

Matching μ is not stable since s_2 and c_1 block μ. However, it can be shown that truthtelling is a Nash equilibrium in this market under the Boston mechanism. In

[9] Note that the choice function is **path independent**, i.e. $Ch_c\left(Ch_c\left(S'\right) \cup S''\right) = Ch_c\left(S' \cup S''\right)$ for any $S', S'' \subseteq S$, when $\succeq_c$ is substitutable.

particular, even though s_2 can form a blocking pair with c_1, it can not deviate profitably under the rule of the Boston mechanism: s_2 applied to c_1 at the first step of the algorithm, but c_1 rejected s_2 given since only s_2 and s_3, and not s_1, were applying. On the other hand, if s_2 applies to c_1 at the second step, s_3 is alreay accepted by c_1 and c_1 will never accept s_2 in the presence of s_3.

Propositons 1 and 2 have an implication for welfare comparison of the Boston mechanism and an alternative algorithm. Our propositions show that Nash equilibrium outcomes of the Boston mechanism is unanimously less preferred by all the students to the student-optimal stable matching. Hatfield and Milgrom (2005) show that truth-telling is a dominant strategy for students under the student-optimal stable mechanism when the priority structure satisfies a condition called the law of aggregate demand in addition to substitutability.[10] Therefore there are unambiguous welfare gains in switching the assignment mechanism from the Boston mechanism to the student-optimal stable mechanism if school priorities satisfy substitutability and the law of aggregate demand.

On the other hand, truthtelling may not be a dominant strategy for students under the student-optimal stable mechanism even with substitutable priority, if the law of aggregate demand is violated. Given this fact, analyzing Nash equilibrium outcomes under the student-optimal stable mechanism may be of interest.[11] To pursue this issue, we define the student-optimal stable mechanism formally. At the beginning of the mechanism, each student submits a preference list. Given stated preference lists of students and priorities of schools, the mechanism follows the algorithm below.

- *Step 1*: Each student applies to her first choice school. Let the set of students who apply to school c at this step be denoted by $A^1(c)$. School c admits all the students in $Ch_c(A^1(c))$ and rejects everyone else. Let $B^1(c) = Ch_c(A^1(c))$.

In general,

- *Step t*: Each student who was rejected in Step (t−1) applies to her next highest choice. Let the set of students who apply to school c at this step be denoted by $A^t(c)$. School c admits all the students in $Ch_c(A^t(c) \cup B^{t-1}(c))$ and rejects everyone else. Let $B^t(c) = Ch_c(A^t(c) \cup B^{t-1}(c))$.

The algorithm terminates either when every student is matched to a school or every unmatched student has been rejected by every school that is acceptable according to her stated preference list. The final assignment is given by $\mu(c) = B^T(c)$ for every $c \in C$, where T is the step in which the algorithm terminates. Note that the only difference between the Boston mechanism and the student-optimal stable mechanism is that a student who is admitted at a school in a step is never rejected in subsequent steps in

[10] Hatfield and Kojima (2007) strengthen their result, showing that the student-optimal stable mechanism is group strategy-proof for students when the priority structure satisfies substitutability and the law of aggregate demand.

[11] I am grateful to an anonymous referee for pointing out that the law of aggregate demand is needed for the welfare comparison in the previous paragraph and suggesting studying Nash equilibria under the student-optimal stable mechanism.

the Boston mechanism while it may happen in the student-optimal stable mechanism. An analogue of Proposition 1 holds for the student-optimal stable mechanism.[12]

Proposition 3 *Under a general priority structure, if μ is a stable matching, then there exists a Nash equilibrium with outcome μ under the student-optimal stable mechanism.*

Proof Let μ be a stable matching. Consider a strategy profile such that every student lists $\mu(s)$ as her first choice, and declare every other school unacceptable. Under this strategy profile, for each school c, $\mu(c)$ is the set of students who apply to c at the first step of the student-optimal stable mechanism. Since $\mu(c) = Ch_c(\mu(c))$ by individual rationality of μ, c accepts all these students. Thus the mechanism terminates at the first step with outcome μ. Now we show that the above strategy profile forms a Nash equilibrium. Suppose that there exists a student s and a strategy such that s is matched to a school $c \neq \mu(s)$. Since any other student s' applies to $\mu(s')$ at the first stage, this means that $s \in Ch_c(\mu(c) \cup s)$. Since μ is stable, this implies that $\mu(s) \succ_s c$, implying that such a deviation of s is not profitable. This shows that the above strategy profile is a Nash equilibrium. □

A natural question is whether there are Nash equilibria whose outcomes are unstable. It turns out that even with responsive preferences, which is a stronger assumption than substitutability, an unstable matching may be realized in equilibrium. Indeed, Haeringer and Klijn (2007) show that the set of Nash equilibrium outcomes of games under the student-optimal stable mechanism coincides with the set of stable matchings for all possible preferences of students if and only if the priority structure satisfies a stringent condition called acyclicity (Ergin 2002).

Remark 1 In many school districts such as Boston, school priorities are set by the district and there is no room for schools to strategize (Abdulkadiroğlu et al. 2005b). In some school districts, by contrast, schools may behave strategically. For instance, schools can set priorities themselves in New York City and they seem to set priorities strategically (Abdulkadiroğlu et al. 2005a). While the main part of Ergin and Sönmez (2006) analyzes the Boston mechanism when schools are not strategic, their Sect. 5 analyzes schools' incentives when schools are allowed to behave strategically. Their Theorem 2 shows that, when it is common knowledge that the priority structure is responsive and every student is acceptable to every school, truthtelling is a dominant strategy for schools.

We present an example which shows that truthtelling may not be a dominant strategy for schools if schools are allowed to report substitutable preferences that are not responsive.[13] Let $C = \{c_1, c_2\}$, $S = \{s_1, s_2, s_3, s_4\}$,

[12] The proof of Proposition 3 is a minor modification of the one for Propositoin 1. We include the proof for completeness.

[13] I am grateful to an anonymous referee for considering the issue of strategic schools with general substitutable priorities.

$$c_1 : \{s_1, s_2\}, \{s_1, s_3\}, s_1, s_2, s_3, s_4,$$
$$c_2 : s_4, s_3, s_2, s_1,$$
$$s_1 : c_1,$$
$$s_2 : c_2, c_1,$$
$$s_3 : c_1,$$
$$s_4 : c_2, c_1.$$

If all students and schools are truthful, the resulting matching μ is given by

$$\mu(c_1) = s_1, s_3,$$
$$\mu(c_2) = s_4,$$
$$\mu(s_2) = \emptyset.$$

If c_1 reports the priority given by $c_1 : \{s_1, s_2\}, s_1, s_2, s_3, s_4,$ while others are truthful, then the resulting matching μ' is given by

$$\mu'(c_1) = s_1, s_2,$$
$$\mu'(c_2) = s_4,$$
$$\mu'(s_3) = \emptyset.$$

As college c_1 strictly prefers $\mu'(c_1)$ to $\mu(c_1)$ under its true priority, truthtelling is not a dominant strategy for c_1. Note that every priority we used in this example is substitutable. Note also that each school finds all students acceptable as in Ergin and Sönmez (2006), that is, $Ch_c(s) = s$ for every college c and every student s.

3 Conclusion

This paper investigated the Boston mechanism when schools have general priority structures. It is shown that the Boston mechanism implements the set of stable matchings in Nash equilibrium when the priority structure is substitutable. A stable matching is still an outcome of a Nash equilibrium with a general priority structure, while some Nash equilibrium outcomes may not be stable.

We assumed complete information, which is a strong requirement. In many real school choice problems, participants may have little information regarding preferences of other participants, or sometimes even school priorities. Analysis under incomplete information is an interesting future research direction.

References

Abdulkadiroğlu A (2005) College admission with affirmative action. Int J Game Theory 33:535–549

Abdulkadiroğlu A, Pathak PA, Roth AE (2005a) The New York City high school match. Am Econ Rev Papers Proc 95:364–367

Abdulkadiroğlu A, Pathak PA, Roth AE, Sönmez T (2005b) The Boston public school match. Am Econ Rev Papers Proc 95:368–372

Abdulkadiroğlu A, Pathak PA, Roth AE, Sönmez T (2006) Changing the Boston school choice mechanism. Harvard University, Unpublished mimeo

Abdulkadiroğlu A, Sönmez T (2003) School choice: a mechanism design approach. Am Econ Rev 93: 729–747
Balinski M, Sönmez T (1999) A tale of two mechanisms: student placement. J Econ Theory 84:73–94
Ergin H (2002) Efficient resource allocation on the basis of priorities. Econometrica 70:2489–2498
Ergin H, Sönmez T (2006) Games of school choice under the Boston mechanism. J Public Econ 215–237
Gale D, Shapley LS (1962) College admissions and the stability of marriage. Am Math Month 69:9–15
Haeringer G, Klijn F (2007) Constrained school choice. mimeo
Hatfield JW, Kojima F (2007) Group incentive compatibility for matching with contracts. mimeo
Hatfield JW, Milgrom P (2005) Matching with contracts. Am Econ Rev 95:913–935
Kelso A, Crawford V (1982) Job matching, coalition formation, and gross substitutes. Econometrica 50:1483–1504
Roth AE (1984) The evolution of the labor market for medical interns and residents: a case study in game theory. J Polit Econ 92:991–1016
Roth AE, Peranson E (1999) The redesign of the matching market for American physicians: some engineering aspects of economic design. Am Econ Rev 89:748–780
Roth AE, Sotomayor MO (1990) Two-sided matching: a study in game-theoretic modeling and analysis. Econometric Society Monographs, Cambridge