Ex Ante Efficiency in School Choice Mechanisms: An Experimental Investigation

Clayton Featherstone* Muriel Niederle†

December 17, 2008

Abstract

Criteria for evaluating school choice mechanisms are first, whether truth-telling is sometimes punished and second, how efficient the match is. With common knowledge preferences, Deferred Acceptance (DA) dominates the Boston mechanism by the first criterion and is ambiguously ranked by the second. Our laboratory experiments confirm this. A new ex ante perspective, where preferences are private information, introduces new efficiency costs borne by strategy-proof mechanisms, like DA. In a symmetric environment, truth-telling can be an equilibrium under Boston, and Boston can first-order stochastically dominate DA in terms of efficiency, both in theory and in the laboratory.

Keywords: School Choice Mechanisms, Market Design, Experimental Economics.

*Department of Economics, Stanford University

†Department of Economics, Stanford University and NBER, http://www.stanford.edu/~niederle

1 Introduction

School districts that allow students some leeway to choose which school to attend have various mechanisms in place to produce outcomes that match students to schools while taking student preferences into account (see Abdulkadiroğlu and Sönmez 2003). In many districts schools themselves do not have preferences over students, though students may receive differing priorities at different schools, reflecting whether they live in the walk zone or have a sibling at the school. School choice mechanisms are judged by two criteria. The first is that students, in some sense, will not be penalized for truthfully revealing their preferences over schools. The second is the efficiency of the match, that is it should assign as many students as possible to schools they like. In this paper, we examine two such mechanisms: the Boston mechanism and the Deferred Acceptance mechanism. We take them to the laboratory to see if students submit preference reports in a way predicted by equilibrium. Furthermore, we introduce a broader, ex ante view of the school choice problem that casts the efficiency properties of the strategy-proof Deferred Acceptance mechanism and the manipulable Boston mechanism in a new light. The ex ante view also suggests potentially higher efficiency costs that have to be incurred to achieve strategy-proofness than previously thought (see Abdulkadiroğlu, Pathak and Roth, forthcoming).

Some mechanisms explicitly favor students at a school if they have ranked it highly. Note that this immediately implies that truth-telling may not be an optimal strategy. Such mechanisms are known as priority mechanisms and are currently used in many school districts, such as Cambridge, Charlotte-Mecklenberg, Denver, Miami-Dade, Rochester, Tampa-St. Petersburg, and White Plains.[1] A particularly extreme priority mechanism was used by Boston Public Schools, which is where the Boston mechanism got its name (see Abdulkadiroğlu and Sönmez 2003 and Abdulkadiroğlu, Pathak, Roth, and Sönmez 2005).

In 2005, the Boston Public Schools replaced the Boston mechanism with a student-proposing Deferred Acceptance mechanism (DA). They did so for reasons best described by the first criterion that we introduced above, namely that participants should not be harmed by submitting their preferences truthfully. Under DA, truthful revelation of the ordinal preference is a dominant strategy, while under Boston, the equilibrium prediction is often that students will manipulate their preference reports. Multiple published sources of advice discuss strategies that yield better outcomes than truth-telling (such as not ranking unachievable

[1] Mechanisms that explicitly use the rank students assign to possible outcomes are also common in two-sided matching problems, see Roth (1990) and Roth (1991).

schools first). These suggest that students (or parents) manipulate their preferences in the Boston mechanism (Abdulkadiroğlu, Pathak, Roth, and Sönmez 2006).

Empirically, Abdulkadiroğlu, Pathak, Roth, and Sönmez (2006) provide some indirect evidence that some families strategize when they report their preferences over schools, while others seem not to and are potentially harmed by failing to do so. Student-proposing DA, on the other hand, does not have this problem, as regardless of what other students submit, truth-telling remains a best response. This seems to have been a key argument in convincing the board of Boston Public Schools to switch from the Boston mechanism to student-proposing DA.

Unfortunately, since students' true preferences are not observed in the field, it is difficult to directly assess any inefficiency in the allocation produced by Boston. Experiments will not only allow us to do this, but will also provide an immediate comparison to outcomes under different mechanisms. Furthermore, we can study whether some types of manipulations are more easily learned than others. The first part of the paper is most closely related to Chen and Sönmez (2006), which shows that students manipulate submitted preferences under the Boston mechanism (see also Pais and Pintér, 2008). In their experiments, however, players only knew the mechanism, remaining entirely ignorant of how the preferences of the other students were generated. This lack of information made it such that neither participants nor researchers could tell if the observed manipulations were in equilibrium. In the present paper, participants always receive information about the distribution of preferences, but may not know their exact realization.

Everything we have presented thus far seems to confirm the view of the existing literature that the Boston mechanism should simply be thrown out of the market designer's toolbag. This result is based on the point of view that all participants (apart from the school board that allocates school seats) know the preferences of other participants. Below, we will highlight some positive features of the Boston mechanism as we change the perspective under which outcomes are evaluated, specifically, as we move to an ex ante viewpoint.

A second important feature of mechanisms is their efficiency. Since schools are objects to be allocated, only the welfare of students is taken into account. The existing literature on school choice problems focuses on two views of efficiency.[2] The first of these is the ex post view, in which lottery draws and preferences of all students are common knowledge. For example, Ergin and Sönmez (2006) show the outcome of student-proposing Deferred Acceptance is

[2]For a summary of the literature, see e.g. Sönmez and Ünver, forthcoming.

weakly preferred to any Nash equilibrium under the Boston mechanism. A second view which has lately received more attention is the interim view, which holds that lottery draws are not known and should not enter efficiency considerations. Erdil and Ergin (2008) point out that DA, by using explicit tie-breakers, may result in an outcome that is not the student-optimal stable matching given priorities only (see also Kesten and Ünver (2008) and Abdulkadiroğlu, Che and Yasuda (2008)). While they provide a mechanism to improve upon the DA outcome, Abdulkadiroğlu, Pathak and Roth (forthcoming) show that any mechanism that improves upon the DA outcome is not strategy-proof. They also try to estimate the costs of strategy-proofness, albeit still from the view that all student preferences are given.

This paper stresses a different view of school choice problems which holds that students, while aware of their own preferences, do not know the preferences of other students, and instead are only aware of their underlying distribution. In this ex ante viewpoint, additional potential gains can be reaped by allowing trade-offs across different realizations of preferences. Furthermore, incentives to manipulate preferences change as well.[3]

Specifically, we will show that there are some special environments where, from the ex ante viewpoint, truth-telling is not only an equilibrium under DA, but also under the Boston mechanism. What's more, in those environments the Boston mechanism can (ex ante) yield outcomes that first-order stochastically dominate those of DA for every student.

It remains an empirical question whether the theoretical advantages of the Boston mechanism can be realized in practice. One possible hurdle could be that students are more likely to submit preferences truthfully when doing so is not merely a Bayesian equilibrium strategy (as under Boston), but is a dominant strategy (as under DA). We will show that the gains of the Boston mechanism can be realized not only in theory, but also in practice.

2 The Theory

2.1 The School Choice Problem

We start with a set of n students, $\mathcal{I} \equiv \{i_1, \ldots, i_n\}$, and a set of m schools, $\mathcal{S} \equiv \{s_1, \ldots, s_m\}$. Each school s has a corresponding capacity q_s. Additionally, each student i has a strict preference, P_i, over the elements of $\mathcal{S}$ and the outcome of being unmatched. Student i has

[3]The ex ante viewpoint has been used in the two-sided matching literature when discussing strategic incentives to manipulate reported preferences, see e.g. Roth and Rothblum 1999, Erdil and Ergin 2008 and Kojima and Pathak (forthcoming). This viewpoint has, however, been absent in the school choice literature and when comparing efficiency properties of various mechanisms.

an associated vector of von Neumann-Morgenstern utilities, V_i.

Schools do not have preferences in the same way, as they are merely objects to be allocated.[4] Instead, they are endowed with a weak priority ordering over students, denoted as SC_s, which is based on student characteristics. The school "preferences" are a way to encode the privileges that certain students might have due to their characteristics. For example, in Boston, SC_s is based on whether a student is within walking distance of s or has a sibling at s (see Abdulkadiroğlu and Sönmez, 2003). Because SC_s is a weak ordering, sometimes a school will need to decide between two students over whom it is indifferent. For these tie-breaking decisions, each school s has a uniform lottery L_s, which provides a strict ordering over students. For most purposes, L_s is either drawn independently of $L_{s'}$ for all schools s and s' (multiple lotteries) or is drawn once and used as the tie-breaker for all schools (single lottery).

The weak "preferences" of schools can be extended to strict "preferences" via a lexicographic ordering that considers L_s whenever SC_s is indifferent. We denote this "constructed" ordering by $\succ_s \equiv SC_s \circ L_s$. A **realized school choice problem** is defined by $\{\mathcal{I}, \mathcal{S}, \{P_i\}_{i\in\mathcal{I}}, \{q_s\}_{s\in\mathcal{S}}, \{SC_s\}_{s\in\mathcal{S}}\}$ (a cardinal version of this concept replaces $\{P_i\}_{i\in\mathcal{I}}$ with $\{V_i\}_{i\in\mathcal{I}}$).

A **general school choice problem** is a probability distribution φ over realized school choice problems, where the uncertainty is over $\{P_i\}_{i\in\mathcal{I}}$. That is, the identity of students, and their student characteristics are known, but the particular preferences of each student are drawn according to the distribution φ over possible $P = P_1 \times .. \times P_n$. This is analogously done for cardinal preferences.

A **school choice matching** $\mu : \mathcal{I} \mapsto \mathcal{S} \cup \mathcal{I}$ is a function such that (i) for all $i \in \mathcal{I}$, if $\mu(i) \notin \mathcal{S}$, then $\mu(i) = i$, and (ii) for all $s \in \mathcal{S}$, it must be that $|\{i : \mu(i) = s\}| \leq q_s$.[5] A **pair** $(i, s) \in \mathcal{I} \times \mathcal{S}$ **blocks** matching μ **with respect to** $\boldsymbol{SC_s}$ if $s\, P_i\, \mu(i)$ and one or more of the following is true: (i) There is some j with $\mu(j) = s$ such that $i\, SC_s\, j$, or (ii) $|\{i : \mu(i) = s\}| < q_s$. A student i is said to **block** a matching μ if $i\, P_i\, \mu(i)$. A matching μ is **stable with respect to** $\boldsymbol{SC_s}$ if there is no individual, and no pair that blocks the matching with respect to SC_s. A **direct revelation school choice mechanism** is a function that maps reported student preferences $\{\widetilde{P}_i\}_{i\in\mathcal{I}}$ and "constructed" school preferences $\{\succ_s\}_{s\in\mathcal{S}}$ to a school-choice matching, μ.[6]

[4]Note that schools are constrained to find all students acceptable.

[5]We denote the outcome where student i is unmatched as $\mu(i) = i$.

[6]Note that school preferences are not submitted by the schools, but rather they are codified by administrative rules. Schools have no chance to behave strategically.

We are primarily interested in examining the efficiency and incentive properties of direct revelation school choice mechanisms. This can be done from three perspectives. We describe the scenarios in which students only know ordinal preferences; the case of cardinal preferences is constructed analogously.

Ex post:

- *Information*: Each student i knows the true preferences of all other students (i.e. P_i and P_{-i}) as well as the "constructed preferences" for all schools, (i.e. $\{\succ_s\}_{s\in\mathcal{S}}$). Implicitly then, he knows both $\{SC_s\}_{s\in\mathcal{S}}$ and $\{L_s\}_{s\in\mathcal{S}}$.
- *Efficiency*: For each vector of submitted preferences $\widetilde{P}$ and school preferences $\succ$, the mechanism returns a deterministic matching. As such, efficiency is measured in terms of the true rank of the school to which each student is matched.

Interim:

- *Information*: Each student i knows P_i and P_{-i}, as well as $\{SC_s\}_{s\in\mathcal{S}}$, but remains ignorant of $\{L_s\}_{s\in\mathcal{S}}$. Essentially, the Interim view is just the Ex post view before the lotteries are drawn.
- *Efficiency*: For each vector of submitted preferences $\widetilde{P}$ and school preferences SC, the mechanism returns a matching that depends on what lotteries are drawn. From each student's perspective this induces a distribution over final matches. Efficiency is measured in terms of these distributions.

Ex ante:

- *Information*: Each student knows their own preference, P_i, but now only knows the distribution of the other students preferences' (i.e. φ). Students still know $\{SC_s\}_{s\in\mathcal{S}}$ and remain ignorant of $\{L_s\}_{s\in\mathcal{S}}$.
- *Efficiency*: For each vector of submitted preferences $\widetilde{P}$ and school preferences SC, the mechanism returns a matching that depends on what lotteries are drawn. For each student, uncertainty about both the lottery draws and the true preferences of the other students induces a distribution over final matches. Efficiency is measured in terms of these distributions.

The literature on school choice mechanisms so far has mostly dealt with complete information, and hence the ex post view. In this paper we will also study school choice under incomplete information, especially the ex ante case, which will cast the properties of the Boston mechanism in a new light. We also look at the ex ante view because it seems to provide a more realistic amount of information to students. Additionally, we consider the ex ante conception of efficiency to be a natural fit for a policy maker's objective. Since they must decide on mechanisms that will be in place for several years, they might be more concerned with how the mechanism performs relative to the distribution underlying any given year's realized set of preferences rather than relative to those realized preferences themselves.

2.2 Two school choice mechanisms

We consider two mechanisms which are currently being used for school choice in the U.S. The first is based on the Deferred Acceptance algorithm and has been introduced in Boston and New York City schools by Abdulkadiroğlu et al. (2006, 2008).

The Deferred Acceptance Algorithm (DA)

- **Step 1:** Students apply to their first choice school. Schools reject the lowest-ranking students in excess of their capacity. All other offers are held *temporarily*.
- **Step** t: If a student is rejected in Step $t-1$, he applies to the next school on their rank-order list. Schools consider both new offers and the offers held from previous rounds. They reject the lowest ranked students in excess of their capacity. All other offers are held *temporarily*.
- **STOP:** The algorithm ends when no rejections are issued. Each school is matched to the students it is holding at the end.

Important properties of the Deferred Acceptance algorithm are summarized in the following proposition.

Proposition (Deferred Acceptance) *Suppose students have strict preferences over schools, and schools have either weak (SC_s) or strict ($\succ_s$) priorities over students, which are used as the schools' submitted rank orders. Consider the preference revelation game induced by the Deferred Acceptance algorithm. Then:*

1. *It is a dominant strategy for students to submit their true preferences (Dubins and Freedman 1981, Roth 1982a).*

2. *The outcome achieved by DA is stable. Furthermore, it is the stable matching that all students prefer, known as the student-optimal stable matching (Gale and Shapley 1962).*

The second mechanism we consider is based on a specific priority algorithm, first described as the Boston algorithm by Abdulkadiroğlu and Sönmez (2003).

The Boston Algorithm

- **Step 1:** Students apply to their first choice school. Schools reject the lowest-ranking students in excess of their capacity. All other offers become *permanent* matches. School capacities are adjusted accordingly.
- **Step t :** Is a student is rejected in Step $t-1$, she applies to the next school on her rank-order list. Schools reject the lowest ranked students in excess of their capacity. All other offers become *permanent* matches. School capacities are adjusted accordingly.
- **STOP:** The algorithm stops when all students have been matched or when all schools have been filled.

Note that in, contrast to the Deferred Acceptance algorithm, applications in the Boston algorithm are *permanently accepted* in each step, as opposed to only *tentatively held*. In the Boston algorithm, acceptance fails to be deferred until it is clear that no student with a higher priority will apply at a later step.

This immediately suggests that the Boston algorithm is manipulable, that is, truthful preference revelation is not a dominant strategy (as it is under DA). Even so, in complete information environments, the set of possible equilibrium matchings in a Boston mechanism can be compared to the outcome under DA via the following proposition.

Proposition (Boston - Ergin and Sönmez, 2006) *Consider the preference revelation game where each student i submits a preference report, $\widetilde{P}_i$, to a direct revelation mechanism that assigns a matching through the Boston algorithm. If all students submit their reports from the ex post perspective (that is under complete information), then the set of Nash equilibrium outcomes of this game equals the set of stable matchings under the true student preferences, P, and the strict constructed school preferences, $\succ$.*

Consider the set of matchings stable relative to the schools' "constructed preferences".[7] DA chooses the element of this set that all students prefer most, while Boston is only constrained to choose *some* element of this set (in equilibrium). For any student, the equilibrium outcome under Boston in a complete information, ex post environment can only be, at best, as good as the one obtained under DA. Once more, there is no reason to favor Boston over DA.

The next example shows that with incomplete information, the resultant outcomes of Boston and DA mechanism may cease to be easily comparable (another example of this can be found in Ergin and Sönmez (2006)). This environment will be one of the treatments in our experiment.

The correlated environment There are three schools – Best (1), Second (2) and Third (3). Best has two seats ($q_{best} = 2$), while the other schools only have one. There are two types of students, called "Top", and "Average"; three students are Top, and two are Average. All students have the same von Neumann-Morgenstern payoffs over schools.

TABLE I

School	Best	Second	Third	No School
Seats	2	1	1	–
Payoff	100	67	25	0

Each school prefers a Top over an Average student (this distinction forms SC), and each school has an independently drawn uniform lottery to order students of the same priority class.

The previous proposition shows that in the complete information environment, the Nash equilibrium outcome is a stable matching. Since the Correlated environment only has one stable matching, we then know that this matching must be the unique complete information Nash equilibrium outcome. One set of supporting strategies is the following: The two Top students with the best lottery numbers at Best rank it first, while the third Top student ranks Second first. The Average student who has the better lottery number at Third ranks it first,

[7] Remember that the only meaningful "preferences" of school s are embodied in the weak ordering SC_s, which reprents privileges which lawmakers have extended to students. The "constructed preferences" include the lotteries, and hence include many strict preferences that are merely artifacts of the need to break ties. As such, stability relative to $\succ$ is a concept with many more constraints than stability relative to SC. Hence, the set of matchings that are stable relative to $\succ$ is a subset of the set of matchings that are stable relative to SC.

and the last Average student submits arbitrary preferences, as he will not receive any school. This is also the matching that results under the Deferred Acceptance algorithm.

Under Boston, this outcome ceases to be achievable in equilibrium when lottery numbers are not known; in fact, when participants are risk-neutral, there is a unique pure strategy Bayesian Nash equilibrium:

Proposition 1 (Correlated Environment) *In the Correlated environment, consider the preference revelation game induced by the Boston mechanism where students have to report preferences without knowing the lottery draws. If agents are not too risk-averse, there is a unique Bayesian Nash equilibrium in which the three Top students are matched to the seats at Best and Third, one Average student is matched with school Second and one Average student is unmatched. The outcome is achieved by the following strategies (where x, y, and z are arbitrary):*

1. *Top students submit preferences* $(1, 3, x)$.
2. *Average students submit preferences* $(2, y, z)$.

The proof is in the Appendix. The equilibrium of Proposition 1 entails two types of manipulations. Top students need to misreport their second choice school, submitting school Third instead of school Second. We call this type of manipulation "skipping the middle". Average students need to truncate their preferences from the top, submitting school Second as their most preferred school. We call this type of manipulation "skipping the top". Note that, relative to DA, the Average students are, in expectation, better off, while the Top students are, in expectation, worse off.

2.3 Efficiency properties of Boston and DA

The Correlated environment, which we use in our experiment, shows that it is possible that some students benefit from the Boston algorithm, while others are harmed (see also Ergin and Sönmez (2003)). There is, however, no existing account of a case in which the Boston algorithm can produce uniformly better outcomes than DA in the sense of first-order stochastic dominance.[8] This is not altogether surprising, since, from the ex post perspective, there does not exist a matching, stable or not, that every student strictly prefers to the student-optimal stable matching relative to SC (see Roth 1982 for strict student preferences,

[8]We focus here on ordinal preferences, as students, in general, submit ordinal rankings. When cardinal preferences are considered, and efficiency is measured as the sum of student welfares, then Boston may dominate DA in the Interim case, see Abdulkadiroğlu, Che, Yasuda (2008) and Miralles (2008).

and Abdulkadiroğlu et al. (forthcoming) for weak student preferences). While DA does not always yield the student-optimal matching relative to SC, it does so sometimes when SC is weak and always when SC is strict.

Furthermore, Pathak and Sönmez (2008) show that in a complete information environment, when there are both sophisticated students (who best-respond to others' strategies) and sincere students (who simply submit their true preferences), the sophisticated students may prefer the Boston outcome to the DA outcome, while sincere students may be worse off under Boston.

With these combined findings – the lack of strategy-proofness, the failure to achieve better outcomes, and the possibility of hurting less sophisticated participants – it might seem that market designers would be better off to eliminate the Boston algorithm from their toolbag altogether.

One objective of this paper is to show that the Boston algorithm can have some attractive features and can potentially outperform the Deferred Acceptance algorithm in specific environments.[9] In order to show this, we turn to a case in which there is non-trivial uncertainty about student preferences. In a very symmetric environment, truth-telling can be an equilibrium even under the Boston mechanism (and, of course, it remains a dominant strategy equilibrium under DA). The following example shows how in the incomplete information case, ex ante, every student may strictly prefer Boston to DA. More generally, though, this example points out a potentially more important weakness of strategy-proofness: it is inherently tied to a realized set of student preferences and hence, to a realized school choice problem. In general school choice problems, insisting on strategy-proofness may prevent trade-offs across different preference realizations which, ex ante, could benefit each student. Sometimes these trade-offs may even be realized without distorting straightforward behavior. Truth-telling can still be an equilibrium strategy even if it is not a dominant strategy.

Example Suppose there are three students, i, j and k, and two one-seat schools, an art school and a science school. Each student is equally likely to have the science or art school as his first choice. Note that, as we discuss below, in this environment, it is a Bayesian Nash equilibrium (and in a sense,as we will see, the unique equilibrium) for all students to report preferences truthfully under Boston. Obviously, truth-telling remains

[9] Far from advocating widespread use of the Boston mechanism, we are instead attempting to point out that mechanisms that are not strategy-proof can yield significantly better results in some environments. Showing preference to a student who ranks a school first over a student who ranks it fifth necessarily eliminates strategy-proofness, but it also allows the mechanism to give a more efficient outcome (so long as truth-telling is implemented in equilibrium).

a dominant strategy under Deferred Acceptance.

Suppose schools use a universal lottery, so that we can order students according to their lottery number as Students 1, 2, and 3. The outcome under the Deferred Acceptance algorithm is that Student 1 receives his first choice school, while Student 2 has a $^1/_2$ chance of receiving either her first or second choice school, depending on whether her preferences are different from or the same as Student 1's. Student 3 will always be unmatched. Since all students are equally likely to be either Student 1, 2 or 3, the expected outcome distribution of any student can be calculated by averaging down the columns of the following table.

Deferred Acceptance

Lottery rank	First choice	Second choice	No school
1	1	0	0
2	$^1/_2$	$^1/_2$	0
3	0	0	1
Average	$^1/_2$	$^1/_6$	$^1/_3$

Under the Boston algorithm, Student 1 still receives his first choice school, and Student 2 still receives her first choice school whenever her preferences are different from Student 1's. The Boston algorithm behaves differently through the rest of the table though. Suppose Student 2 has the same preferences as Student 1. If Student 3 has different preferences, then Student 3 receives his first choice school, and Student 2 receives no school. Only if Student 3 also has the same preferences does Student 2 receives his second choice school. Each of these events is equally likely; hence, the expected outcome can be calculated as follows.

Boston

Lottery rank	First choice	Second choice	No school
1	1	0	0
2	$^1/_2$	$^1/_4$	$^1/_4$
3	$^1/_4$	0	$^3/_4$
Average	$^1/_2 + ^1/_{12}$	$^1/_6 - ^1/_{12}$	$^1/_3$

Clearly, the distribution under Boston first-order stochastically dominates the distribution under DA, that is, in this environment, when students play equilibrium strategies, every student is ex ante strictly better off under the Boston mechanism.[10]

The reason for our finding is that the Boston mechanism effectively eliminates "bad" draws from the support of the lottery. Say that we have two artists and one scientist who are drawn by the lottery in just that order. Under DA, one artist will get the art school and the other will get the science school, leaving the scientist unmatched. In a sense, this is a bad outcome, as it gives one student her first choice and the other his second choice. A better outcome is to give two students their first choice, which is accomplished by giving the art school seat to an artist and the science school seat to the scientist. This is just what the Boston mechanism prescribes in our environment. Generally, any lottery draw that is of the form (*artist, artist, scientist*) or (*scientist, scientist, artist*) is suboptimal in the sense we just described. If a planner using DA proposed to remove those orderings from the support of the lottery draw for these preference realizations, any student would agree to the proposal ex ante, as it would make him or her better off. This is because there are just as many (*scientist, scientist, artist*) draws as there are (*artist, artist, scientist*) draws. In our environment, Boston implements this planner's proposal.

Finally, note that this result does not carry through to either the interim or the ex post worlds. In both, which mechanism is preferred depends on the realization of preferences. For instance, suppose students knew that there were two artists and only one scientist. From the interim view, the artists would strictly prefer that DA be used, while the scientist would strictly prefer that Boston be used. If instead we had two scientists and one artist, these conclusions would be reversed. In the ex post world, when the lottery draws an ordering of the form (*artist, artist, scientist*), under Boston, the second artist would report himself as a scientist to avoid being unmatched. The dominance of Boston over DA in our environment is driven by incomplete information about the realized student preferences.

For the experiment, we will use a similar, but slightly more complicated environment. To discuss incentive properties in this new environment, we first introduce a definition and a proposition. A strategy in a preference revelation game is called a **preference permutation** if, for each preference realization, the student submits the *same* permutation of his true preferences to the mechanism.

Proposition 2 (Truthtelling in Boston) *Consider an incomplete information environment, in which all schools are equally large (i.e. $q_s \equiv q$ for each s). Furthermore, assume that*

[10] Note that this is a first-order stochastic dominance; hence students prefer Boston for all cardinal preferences that correspond to their ordinal preferences.

preferences of all students are independently drawn from a uniform distribution over the set of all possible rank orderings over $\mathcal{S}$. Then, truth-telling is a Bayesian Nash equilibrium of the preference revelation game induced by the Boston algorithm when viewed from the ex ante perspective. Furthermore, truth-telling is the unique Bayesian Nash equilibrium where all students use preference permutations.

See the Appendix for a proof. To compare the expected outcomes under Boston and DA, note that it is obvious that the Boston mechanism gives more students their submitted first choice school. Whenever students submit their preferences truthfully, this means that more students will receive their true first choices under Boston than under DA.

The Uncorrelated environment There are five students and four one-seat schools. The preferences of all students are drawn independently from the uniform distribution over all possible orderings of the set of schools, $\mathcal{S}$. All lotteries are drawn from the uniform distribution over all possible orderings of the students. SC is indifferent between all students, i.e. there is no longer any Top/Average distinction.

In this environment, we can compare the theoretical outcome distributions of the students under Boston and DA, assuming a single lottery (DA-SL).[11] Note that using multiple lotteries would affect the outcome of DA but not of Boston. Table II shows each student's expected probability of matching with her first choice, second choice, etc.

TABLE II

Mechanism	Boston	(Cumul.)	DA-SL	(Cumul.)
First choice	0.610		0.500	
Second choice	0.117	(0.727)	0.167	(0.667)
Third choice	0.055	(0.782)	0.083	(0.750)
Fourth choice	0.018	(0.800)	0.050	(0.800)
No School	0.200	(1.000)	0.200	(1.000)

The cumulative columns make it clear that the ex ante distribution under Boston first-order stochastically dominates the distribution under DA.

While the Boston mechanism may in theory generate outcomes that surpass those of DA, this may not be the case in practice. It could be that truth-telling rates are much lower when

[11] These calculations assume truth-telling, which is a dominant strategy under DA and a Bayesian Nash strategy under Boston.

truth-telling is only a Bayesian Nash equilibrium strategy instead of a dominant strategy. Whether the potential gains of a Boston mechanism can be realized in practice remains, therefore, an empirical question.

3 The Experiment

We have two aims for the experiment. The first is to provide a clean empirical test of whether agents manipulate their preferences under the Boston algorithm and, if so, how. Two other papers have addressed this question. Chen and Sönmez (2006) have subjects play a one-shot preference revelation game under either Boston or DA. Additionally, participants in their experiment know neither the preferences of other students nor the distribution that was used to generate those preferences. As a result, equilibrium strategies cannot be computed. The paper instead focuses on how students respond to a large and complex school choice environment, which has many students vying for seats in schools that are distinguished by being either small or large and either specialized or general interest. They analyze whether students use heuristics, such as applying to small schools or large schools, instead of truth-telling.[12] Pais and Pintér (2008) follow the same approach. In their experiments students know all students' preferences or know only their own preferences, remaining ignorant of the distribution from which the other students' preferences are drawn. Their paper, like Chen and Sönmez (2006), also uses a one-shot game. In contrast we run a repeated game to allow for learning.[13] We also focus on an environment in which we can compute the Bayesian Nash equilibrium so that we can compare submitted strategies to the equilibrium prediction, allowing us to detect whether agents manipulate strategies optimally or whether they make "mistakes". This simpler design also allows us to analyze which deviations from equilibrium may be more common.

Testing whether and how agents manipulate their preferences under the Boston algorithm is difficult to address in the field since the true preferences of students are unobserved. The only empirical paper along these lines thus far is Abdulkadiroğlu et al. (2006), which analyzes submitted preferences under the Boston algorithm. They find some evidence of potentially

[12]Interestingly, in the DA mechanism, they found truthtelling rates of only 55% and 72%. Pais and Pintér (2008) found truthtelling rates of 67% in the DA mechanism.

[13]This allows us to see whether behavior remains stable and whether participants may eventually learn to manipulate in a Boston mechanism and report truthfully in a DA mechanism. While parents in general particpate in school choice mechanisms only once, they often draw from the experiences of other parents, and many districts have school choice mechanisms at several points in a child's education. In using multiple rounds, we also follow the tradition of two-sided matching experiments, see Kagel and Roth (1999), Ünver (2001), McKinney, Niederle and Roth (2005).

sub-optimal behavior, albeit through indirect inference. The advantage of an experiment is that we know the participants' true preferences directly.

The second aim of our experiment is to explore environments in which the Boston mechanism provides arguably better outcomes for students than DA. Here, having many preference realizations is crucial, as the dominance of Boston over DA is driven by allowing for trade-offs across preference realizations. The one caveat from our theoretical analysis was that truth-telling can only be implemented as a Bayesian Nash equilibrium strategy under Boston, while it is implemented as a dominant strategy under DA. It remains, therefore, an empirical question whether the theoretical gains of a Boston mechanism can be realized in practice.

3.1 Correlated environment

Both incomplete information environments that we use in our experiment consist of five students vying for four school seats. Since we have already introduced both environments, we limit ourselves here to a quick summary of the environments and the behavior we expect to observe under Boston and DA.

In the first environment, the Correlated environment, all students have the same preferences, and the only uncertainty is over the draw of the lotteries. There are only three schools: Best, Second, and Third. Best has two seats, while Second and Third are one-seat schools. Participants are paid according to their match as follows (payoffs are in points, each of which is worth 1.5 cents):

Correlated payoffs

School	Best	Second	Third	No School
Seats	2	1	1	–
Payoff	100	67	25	0

Of the five students, three are Top, and two are Average. Each school prefers a Top over an Average student (this is the *SC* ordering). Furthermore, each school independently draws a uniform lottery to order students of the same *SC* class.[14] Equilibrium behavior under our two mechanisms is as follows:

Deferred acceptance: Under DA, it is a dominant strategy to submit preferences truthfully.

The outcome is thus for the two Tops with the best lottery draws to get seats at Best,

[14] For a discussion on using single (universal) versus multiple lotteries (one for each school) see Abdulkadiroğlu et al. (forthcoming).

the other Top to get a seat at Second, and for the Average with the best lottery draw at Third to get a seat there. The other Average remains unmatched.

Boston: The unique pure strategy Bayesian Nash equilibrium strategies are for Top students to submit preferences $(1, 3, x)$, and for Average students to submit preferences $(2, y, z)$, where x, y, and z are arbitrary. Hence Top students must "skip the middle", while Average students must "skip the top". The outcome is for the two Tops with the best lottery draws to get seats at Best, the other Top to get a seat at Third, and for the Average with the best lottery draw at Second to get a seat there. The other Average remains unmatched.

Comparing the strategies and outcomes under Boston and DA can show whether students manipulate preferences in the correlated environment when the mechanism calls for manipulation. Furthermore, the experiment can show whether students will learn the optimal manipulations, and whether one kind of manipulation seems easier to learn than the other.

3.2 Uncorrelated environment

In the Uncorrelated environment, there are five students, and *four* one-seat schools. Each student has independently drawn preferences over the four schools, where all possible preference profiles are equally likely. The payoffs depend on a participant's true ranking of the school to which they are matched; hence, it is possible for two students both to both be paid the highest amount if they have different first choice schools. Also note that, while the specific payoffs are close to the ones used in the correlated environment, the equilibrium predictions for the Uncorrelated environment are not sensitive to their cardinal values. Subject payoffs in the uncorrelated environment are as follows:

Uncorrelated payoffs

School	1^{st} Choice	2^{nd} Choice	3^{rd} Choice	4^{th} Choice	No School
Seats	1	1	1	1	–
Payoff	110	90	67	25	0

For the schools' preferences, we use a different lottery for each school, that is, they too have random preferences over students, uniformly drawn from the set of all possible rankings.[15]

[15] We chose multiple lotteries mostly for simplicity. Note, however, that in a Deferred Acceptance algorithm, using a single lottery versus multiple lotteries may have an effect. While outcomes cannot be directly compared, it seems that on average using a single lottery provides somewhat better outcomes for students than using multiple lotteries. In our final efficiency comparison, we will take this into account.

Expected behavior is as follows:

DA: Truth-telling is a dominant strategy.

Boston: Truth-telling is a Bayesian Nash equilibrium strategy.

As a reminder, in equilibrium, Boston first-order stochastically dominates Deferred Acceptance (see Table II). This environment allows us to test whether truth-telling rates are higher under DA or under Boston and whether the theoretical ex ante gains of Boston over DA in the truth-telling equilibrium can be realized in practice.

3.3 Experimental design

We ran four sessions under the Deferred Acceptance mechanism and seven sessions under the Boston mechanism. In each session, five Stanford undergraduate students played for 15 periods in the Correlated environment, during which players kept their role, as either a Top or an Average student. Then, after a pause to learn about the new environment, they played another 15 periods in the Uncorrelated environment.[16] Within a session, the mechanism was held constant, and each participant participated in only one session. The experiment was conducted on computers, using z-Tree (Fischbacher 2007). At the start of a session, we read instructions concerning the environment and the mechanism, and we checked each player's understanding by having them solve the outcome of a test environment, where participants were given submitted preferences and had to determine the outcome of the relevant mechanism. We repeatedly checked understanding by correcting and explaining outcomes through each subsequent step of the algorithm. Participants earned 1.5 cents for every payoff point and were paid based on their cumulative earnings over all 30 periods of the session.

4 Results from the correlated environment

4.1 Manipulation in the Boston Mechanism

For the duration of this section, we will focus on the last five of the fifteen periods of the Correlated environment. There is, however, not much change between the first five and the last five periods, and most results are qualitatively the same were we to consider all 15 periods.

[16] In our design, participants always see the Correlated environment before they see the Uncorrelated environment. This may lead to less truth-telling in the Boston Uncorrelated environment, and hence seems to work *against* what we are trying to show.

We start our analysis of behavior under the Boston mechanism by focusing on who receives what outcome. Table III shows, for each student type, the fraction of matches at each school. The predicted equilibrium outcome is also included for comparison.

TABLE III: BOSTON

Proportion of seats received by different participants

School	Best	Second	Third	No School
Top	0.67	0.11	0.05	0.17
Top Equil.	$^2/_3$	0	$^1/_3$	0
Average	0.00	0.33	0.43	0.24
Average Equil.	0	$^1/_2$	0	$^1/_2$

The most striking aspect of the outcome distribution is that Top students who do not receive a seat at Best sometimes end up unmatched. Average students, in turn, are too often matched relative to the equilibrium outcome.

To understand these outcomes, we investigate the strategies used by Top and Average students. First, we consider the distribution of submitted first choices (see Table IV).

TABLE IV: BOSTON

First Choices of Participants

School	Best	Second	Third
Top	0.92	0.07	0.01
Average	0.06	0.67	0.27

The vast majority of strategies submitted by Average students rank some school besides Best as their first choice. Analyzing behavior by participant, we find that 13 out of 14 Average students manipulate as a primary strategy (we describe a strategy as **primary** if it is used at least 80% of the time).[17] Not ranking unachievable schools seems to be easy to learn, and Average students do so almost immediately.

The most popular strategy used by Top students is to rank school Best as their first choice. Analyzing behavior by participant, all but one of the Top students ranks Best first as a primary strategy (the remaining Top does so 60% of the time).[18]

When analyzing the second choice of Top students, note that in the last five rounds, 65.7%

[17] Eleven (of 14) Average students always skip school Best (that is 5 out of 5 times).

[18] 14 of 21 Top students always rank Best first. In general, the first choice of a Top student that is not school Best is school Second, though once it is school Third.

of the submitted rankings of Top students are their true rankings.[19] We have 13 of 21 Top participants using truth-telling as their primary strategy, and only 5 of 21 using any sort of misrepresentation as a primary strategy. The difference in truth-telling rates between Top students (65.7%) and Average students (1.5%) provides evidence that students are responding to the environment.

Average students seem to manipulate their preferences quite often, as is predicted by the risk-neutral Nash equilibrium. While 67% of Average student strategies correspond to the risk neutral equilibrium, the same proportion is only 27% for Top students. Furthermore, in all 15 periods and in all seven sessions, we did not observe a single period in which every agent in the group played risk-neutral Bayesian Nash equilibrium strategies.

Even so, this lack of adherence to the risk-neutral equilibrium might be able to be explained by risk-averse agents. Fortunately, there are instances in the data where behavior can be unequivocally classified as sub-optimal. One of these is when Top players rank Third as their first choice. This happens exactly once in our periods of interest. The other instances where sub-optimal behavior is identifiable occur when all Top students rank Best as their first choice and Average students submit rankings such that the second choice of Top students is clearly constrained by best response. We observe two types of cases like this.

Case 1: One Average student ranks Second first, and the other Average student does not rank Third first; hence, first round choices are $\{1, 1, 1\}$ for Top students and $\{2, 1/2\}$ for Average students. Furthermore, neither of the Average students ranks school Best second. In this case the best response of Top students is to rank school Third second: the strategy $(1, 3, x)$ yields expected payoffs of $\frac{2}{3} \cdot 100 + \frac{1}{3} \cdot 25$ as compared to only $\frac{2}{3} \cdot 100$ for a strategy of $(1, 2, x)$.[20]

Case 2: One Average student ranks school Third first, and the other does not rank school Second first, and neither Average student ranks school Best as their second choice; hence, first choices are $\{1, 1, 1\}$, $\{3, 1/3\}$. In this case the best response of Top students is $(1, 2, x)$ which yields $\frac{2}{3} \cdot 100 + \frac{1}{3} \cdot 67$ as compared to $(1, 3, x)$ which yields only $\frac{2}{3} \cdot 100$.

[19]In comparison only 1% of Average student strategies are truthful preference revelations; in fact only one Average student truthfully submitted in the last five periods (and then, only once). Furthermore, there does not seem to be much of a trend in the data: in the first 5 periods, truth-telling rates for Tops and Averages are 69% and 6%, respectively.

[20]Note that in the last five periods, only three strategies of Average students have school Best ranked as their second choice. Two of these are nonetheless included in Case 1. The other is included in Case 2. We include these because this does not affect the best response of Top students, but only how much they gain by doing so.

Case 1 and Case 2 are the only two cases that are present in the data in which Top students have strong incentives to rank either Second or Third as their second choice, given that they ranked Best as their first choice. We now study sessions in which a large fraction (4 out of 5, labeled Frequency 4/5 in the table) or at least the majority (3 out of 5, labeled Frequency 3/5) of periods are either in Case 1 or Case 2 and in which remaining periods do *not* fall into those cases. More succinctly, we focus on sessions in which students weakly best respond by submitting $(1, 2, x)$ or $(1, 3, x)$ in all five periods, and, in 4 (or 3) of those periods, this behavior is a *strict* best response. Table V below shows the number of sessions that fall into these categories, the best responses of Top students in those sessions (T-BR), and whether Top students submit truthfully (i.e. submit $(1, 2)$ (Truth)) or best respond as a primary strategy (BR).

TABLE V

Case	T-BR	Frequency	# sessions	Truth	BR
1	$(1, 3)$	4/5	2	5/6	0/6
1	$(1, 3)$	3/5	2	2/6	2/6
2	$(1, 2)$	4/5	1	3/3	3/3

We find that the proportion of students that use truth-telling $(1, 2, x)$, or manipulation $(1, 3, x)$ as a primary strategy is largely independent of what the best response would have been; in fact, most Top students just truthfully submit their preferences as a primary strategy. As a result, 9 of the 15 Top students analyzed above incur losses in earnings that are higher than 5%. There seems to be little evidence that Top students use truth-telling because it is a best response in their session, but rather it seems that "skipping the middle" (i.e. submitting $(1, 3, x)$) is simply difficult to learn.

When analyzing empirical data concerning submitted strategies under the Boston mechanism, it is a lack of manipulation of this sort ("skipping the middle") that may be the most identifiable, even if true preferences of students are not known. Indeed, Abdulkadiroğlu et al. (2006) find some indirect evidence that students sometimes rank schools below their first choice that are expected to be filled in the first step of the algorithm.[21]

To summarize, the differences in truth-telling rates across type confirm that participants react to the environment, albeit not optimally. There remains the possibility, however, that Average students would not apply to Best (i.e. "skip the top"), even under DA. Furthermore,

[21] If a school is filled in the first step of a Boston mechanism, then no student who has not ranked the school first has a chance of matching to it. If a student believes that a school will fill in the first step of the algorithm, then it is a clear mistake to rank it lower than first.

as we have seen, Top students who should be manipulating their preference reports under Boston in fact use truth-telling as their most common strategy. It is possible that the behavior we observe may not be thoughtful manipulation due to the Boston mechanism but instead is just a reaction to the Correlated environment. These concerns can be addressed by analyzing truth-telling under DA in the same Correlated environment.

4.2 Boston versus DA

The outcome in the last five periods under the DA mechanism is given by Table VI, which shows for each student type the fraction of matches at each school.

Table VI: DA

Proportion of seats received by different participants

School	Best	Second	Third	No School
Top	0.67	0.33	0.00	0.00
Average	0.00	0.00	0.50	0.50

The outcome corresponds exactly to the equilibrium: the three Top students receive seats at Best and Second, while the Average students either receive a seat at Third or remain unmatched. Note that under DA, when compared to the Boston mechanism, Top students are better off and Average students are worse off. We begin our investigation of the strategies used by the subjects by looking at their submitted first choices (Table VII).

Table VII: DA

First Choices of Participants

School	Best	Second	Third
Top	1	0	0
Average	.70	.05	.25

Sufficient conditions for the outcome under DA to yield the stable match constrain Top students to submit preferences $(1, 2, x)$, while only constraining Average students to rank Third above No School. In spite of the fact that truth-telling is not a *strict* best response, we observe that 92% of Top student strategies and 63% of Average student strategies are truth-telling.[22]

To compare truth-telling rates between the Boston and the DA mechanism, we run Mann-

[22] Note that under DA, 11 out of 12 Top and 4 out of 8 Average students use truth-telling as a primary strategy. In comparison, under Boston the numbers were 13 out of 21 for Top students and 0 out of 14 for Average students.

Whitney tests on session averages, which gives us 7 Boston data points and 4 DA data points. Running the test for the two student types, we find that both Top students ($p = 0.07$) and Average students ($p < 0.01$) are significantly more likely to use truth-telling strategies under DA than under Boston. This shows that the manipulations of both Average and Top players under the Boston mechanism are not due to the environment alone, but rather are driven by the combination of environment and mechanism. Nonetheless, as the outcome of the Boston mechanism showed, participants are not best responding even in this simple environment.

5 Results from the uncorrelated environment

Now we move on to the Uncorrelated environment. Recall that, in this environment, under the Boston mechanism, truth-telling is a Bayesian Nash equilibrium strategy, while under DA, truth-telling is of course still a dominant strategy. If participants use truth-telling strategies, then we expect the Boston outcome to first-order stochastically dominate that of DA.

5.1 Boston versus DA: Strategies

To compare strategies between the Boston and Deferred Acceptance mechanisms, first note that basically all submitted strategies rank all schools.[23] The proportion of truth-telling strategies is 58% under the Boston mechanism, compared with 66% under DA. This difference is not significant: a Mann-Whitney test across mechanisms, comparing mean truth-telling rates in each session, yields a p-value of 0.70 ($n = 11$).[24]

To address what manipulations are submitted, we check truth-telling rates at each ranked position, that is, how often a participants's submitted k^{th}-ranked school corresponds to his true k^{th}-ranked school (see Table VIII).

[23]Only one student, for three rounds at the beginning of the Uncorrelated environment, failed to rank all schools.

[24]Truth-telling rates declined somewhat from the first five periods to the last. The rates in the first five periods were 74% under DA and 61% under Boston.

Table VIII: Uncorrelated

	Boston				DA			
	1st.	2nd.	3rd.	4th.	1st.	2nd.	3rd.	4th.
Rank 1	0.76	0.22	0.01	0.01	0.74	0.19	0.07	0.00
Rank 2	0.16	0.61	0.13	0.10	0.14	0.69	0.15	0.02
Rank 3	0.06	0.07	0.80	0.07	0.07	0.11	0.77	0.05
Rank 4	0.02	0.10	0.06	0.82	0.05	0.01	0.01	0.93

Note that there is little discernable difference across mechanisms, as predicted by equilibrium. Furthermore, note that manipulations tend to move a school up one or down one from its true ranking. While participants are not exactly truth-telling, they're not manipulating in an extreme way either.

5.2 Boston versus DA: Outcomes

We want to compare the outcomes in terms of how often a participant receives her k^{th}-ranked school. We have seen that, over all possible preference realizations, the Boston mechanism should first-order stochastically dominate DA. Even so, there are a total of 24^4 possible preference realizations for students, while we draw only 7×15 realizations under Boston and 4×15 under DA. If we only analyze the last five periods instead of all fifteen, we further cut the number of observations over which we average by a factor of three. We will therefore expand our analysis to include all fifteen periods. The results when restricting ourselves to the last five periods are similar, though sometimes only significant in a one-sided test.

The DA mechanism which we ran in the lab uses multiple lotteries; however, recent work by Abdulkadiroğlu et al. (forthcoming) provides simulations that indicate that DA under a universal lottery can in some environments produce a better ex ante outcome distribution than DA under multiple lotteries. Note that in both cases, truth-telling remains a dominant strategy. In our uncorrelated environment, we know that Boston theoretically dominates DA using a single lottery. We therefore compute the outcome had we used a single lottery in the laboratory. We do this by taking the participants' submitted preferences and using the lottery draw for one of the schools as the universal lottery. Since we have four of these lotteries (one per school) for each session, we use each of them as a universal lottery and then average the resultant outcome distributions.[25]

Earlier, we mentioned that we take only a small sample of the possible preference realiza-

[25] We cannot take each of the lotteries as a data point, as they are not independent of each other.

tions in our experiment. As a result, it may be that the dominance results established for the entire distribution over preferences might not hold for the particular preferences that were drawn in the lab. As such, we calculate the counterfactual outcome that would have occurred had truth-telling rates been 100%. In a sense, this is a measure of how much advantage Boston theoretically could have had over DA. The figures always show not the proportion at which participants receive the school they merely ranked first, second, etc., but the proportion at which participants receive a school that corresponds to their *actual* first choice school, second choice school, etc.

Figure 1 shows, for each true school rank, the probability with which a student receives either a school of that rank or a more preferred school. We plot the mean outcomes relative to participants' actual preferences for Boston and DA, and also for the single-lottery DA counterfactual, which we denote by DA-SL. This is the outcome we would have observed had the participants been 100% truthful. Since this was not the case, Figure 2 shows the outcomes using the participants' submitted preferences.

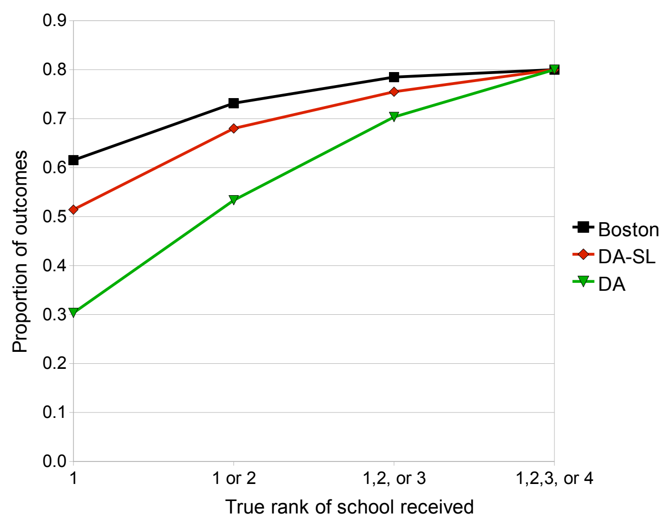

Figure 1: CDF of Outcomes with 100% truth-telling

When looking at the 100% truth-telling outcomes (Figure 1), we find that the Boston mechanism yields significantly more true first choice outcomes than either the DA mechanism ($p < 0.01$) or the DA-SL mechanism ($p < 0.01$).[26] When we compare the proportion of

[26]We use Mann-Whithney test, where the session mean is a data point, that is, we have 7 data points for

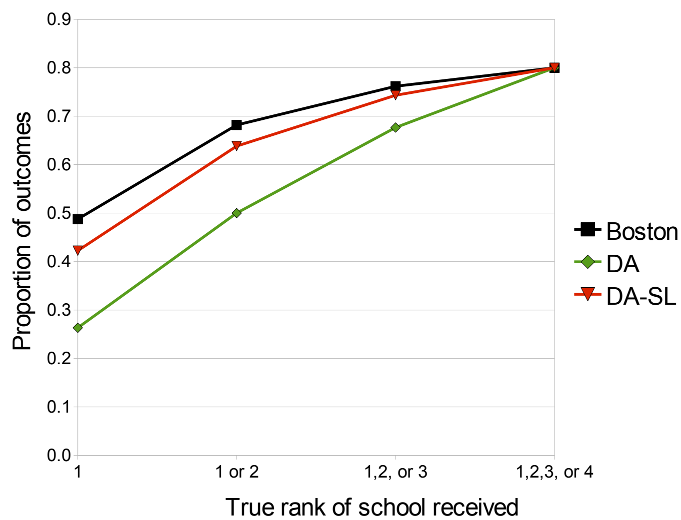

Figure 2: CDF of Outcomes with submitted preferences

students who receive either their true first or second choice, once more Boston has significantly higher proportions than either the DA mechanism or the DA-SL mechanism ($p < 0.01$ in both cases).[27]

Figure 2 shows the actual outcome distribution observed in the lab. As we change from 100% truth-telling to submitted preferences, the cumulative outcomes seem to shift downwards, but do not squeeze closer together that much. This implies that, while deviation from 100% truth-telling decreases efficiency, it does not erase the gains to be made from switching to Boston from either version of DA.

A significantly larger fraction of students receive their true first choice even with submitted preferences when the Boston rather than DA ($p < 0.01$), or DA-SL ($p = 0.06$), is used.[28] When we compared the proportion of students who receive either their first or second choice, Boston once more significantly outperforms both the DA ($p < 0.01$) and DA-SL ($p = 0.09$).[29]

To summarize we found that in the uncorrelated environment truth-telling rates between

Boston, and 4 for DA and DA-SL. When we only consider the last five periods, the p-values are 0.01 and less than 0.01, respectively.

[27]When we look at the last five rounds only, the difference between Boston and DA is still significant at $p < 0.01$, while the one between Boston and DA-SL is not ($p = 0.12$).

[28]When we only consider the last five periods, Boston still gives a higher fraction of participants their first choice. The difference is significant when compared to DA ($p < 0.01$), though not when compared to DA-SL, ($p = 0.18$). Note that a one-sided test would yield significance in all the two-sided tests that we have discussed.

[29]When we consider only the last five preiods, the p values are < 0.01 and 0.03 respectively.

Boston and DA were very similar, in spite of the fact that truth-telling is a dominant strategy under DA and only a Bayesian Nash equilibrium strategy under Boston. A plausible explanation is that truth-telling holds special sway as a focal point.[30] Regardless, it seems that the theoretical advantages of Boston over DA (in our specialized environment) can indeed be realized in practice.

6 Efficiency: The Ex-Ante View

To compare efficiency properties of mechanisms, the matching literature initially focused on the ex post point of view, taking not only preferences but lottery outcomes as given. When students not only know each others' preferences and priorities, but also the lotteries of all the schools, Ergin and Sönmez (2006) show that the set of Nash equilibria under the Boston mechanism is equal to the set of stable matchings. However, the student- proposing DA mechanism selects the student-optimal stable matching (when using both student characteristics and lotteries as real constraints, that is using $\succ_s \equiv SC_s \circ L_s$ for each school s); hence, the efficiency properties of the Boston mechanism are inferior those of the DA mechanism.

The school choice literature recognized that the outcome of the student-proposing DA mechanism may be the student-optimal stable outcome given $SC \circ L$, but that, when computing efficiency, constraints induced by the lottery outcomes L should maybe not be seen as "real" constraints. Once the outcome of a DA mechanism is evaluated in terms of student characteristics (SC) only, it may *not* be a student-optimal stable match. The main issue is that the lotteries induce additional constraints that might make the student-optimal stable match relative to SC unstable relative to $\succ$.

A new wave of papers takes this *interim view*, that is, continues to take student preferences as given, but evaluates efficiency properties before lotteries (L) are drawn. The main focus of that literature is to improve upon the DA outcome, since it may not always be a student-optimal stable match (relative to SC).

Erdil and Ergin (2008) show that there is no strategy-proof mechanism that guarantees a student-optimal stable matching in the school choice problem in which schools have weak

[30] This may be why Pais and Pintér (2008) found that in environments in which participants have no information about how preferences are generated, a large fraction of participants use truth-telling as a strategy. They found similar truth-telling rates when either a Boston or a DA mechanism was used. One difference is that in our environment, participants have information about the underlying preference distribution, which allows for game-theoretic equilibrium analysis. In a theoretical investigation, Pathak and Sönmez (forthcoming) use truth-telling as *the* non-equilibrium strategy.

priorities. Furthermore, they provide an algorithm of stable improvement cycles that improve on a DA outcome, while maintaining stability relative to SC (though not necessarily L). Kesten and Ünver (2008), in similar vain, try for a mechanism that is less constrained by L than the DA mechanism. Abdulkadiroğlu, Che and Yasuda (2008) and Miralles (2008) move a step further, looking at cardinal, rather than ordinal, preferences of students.[31]

None of these proposed mechanisms is strategy-proof; indeed, Abdulkadiroğlu et al. (forthcoming) show that no mechanisms that Pareto dominate DA are strategy-proof. Furthermore, they try to empirically estimate the costs of strategy-proofness. Their data are the submitted preferences in New York and Boston, both of which recently started to use a DA mechanism (see Abdulkadiroğlu, Pathak, Roth, Sönmez (2005) and Abdulkadiroğlu, Pathak, and Roth (forthcoming). Starting with an outcome given by the DA mechanism, they use stable improvement cycles à la Erdil and Ergin and top trading cycles, which allow for improvements that break stability. They find that the costs of strategy-proofness, while present to some extent in New York, are not very large in Boston.

In this paper we introduce the *ex ante* viewpoint. Efficiency properties are not assessed for a given set of preferences, but rather from a perspective where only the distribution of preferences is common knowledge, and the realized preferences of each student are private information. This viewpoint may be relevant not only for participating students in an actual school choice problem, but also for a school board that plans to use a mechanism over many years. Such a board may be more concerned with achieving consistently good outcomes from year to year, rather than for a specific realized preference profile. The main effect of this ex ante viewpoint, is that it allows for trade-offs across specific preference realizations.

The canonical example of how this might be advantageous is our highly artificial and very symmetric art and science schools example. In this example, truth-telling is an equilibrium under the Boston mechanism. Furthermore, the Boston mechanism achieves, for each realization, a student-optimal stable match that gives as many students as possible their first choice. This outcome first-order stochastically dominates the DA outcome for each student. Note that in this example, for any preference realization, DA produces a student-optimal stable

[31]While DA may yield desired ordinal outcomes, this need not be the case when outcomes are cardinal. In this case, for specific cardinal preferences they construct an example in which the equilibrium outcome of a Boston mechanism is superior to a DA outocme, in an environment in which preferences P and student characetrisctics SC are known, but L is not. The reason is that the equilibrium in a Boston mechanism which uses strategies that are not truth-telling allows some freedom to students in deciding how to break ties. They use that approach to put forward their CADA (Choice Augmented DA) mechanism, which basically gives students the possibility to decide to some extent how to break ties. Miralles (2008) is very similar in spirit.

outcome, just not always the same as Boston. The reason there is room to improve upon DA is that the student- optimal stable match is not unique. Furthermore, DA does not select among these student-optimal stable matching in a way that takes advantage of tradeoffs across different preference realizations, hence maximizing expected utility ex ante.

The advantages of the Boston mechanism, are most easily seen in the context of the art and science schools example. If only one student likes the science school the most, that student should go to the science school, while the two artists should have an equal chance of receiving the one seat at the art school. Similarly, if there is only one artist, the Boston mechanism makes sure that the artist can go to the art school, while the two scientists receive a lottery over the seat in the science school. Since both preference realizations are equally likely, each student strictly benefits ex ante over a mechanism such as DA that treats art and science students the same when considering to whom it should give the remaining school, be it the science or the art school. DA may be preferable to the two artists once it is known there is only one scientist, since they would not want the scientist to receive preferential treatment over them. However, this is not the case ex ante, when both artists and scientists are equally likely to be in the minority. Note that the gains of the Boston mechanism are realized when students report their preferences truthfully, and in an ordinal setting.

The ex ante view and example above also highlight an alternative way to calculate the cost of strategy-proofness. In the example the DA outcome is always a student-optimal stable matching (though not necessarily the one that provides as many students as possible with their first choice school). Nonetheless, neither Erdil-Ergin nor top trading cycles improvements have any impact on any allocation reached by a DA mechanism. As such, the method devised by Abdulkadiroğlu et al. (forthcoming) to measure the costs of strategy-proofness would conclude that in our art and science schools example that there are no costs of strategy-proofness. This is obviously not the case from the ex ante view, as a student halves his chance to get his second choice school, and instead increases his chance to receive his first choice school by the same amount.[32]

While the ex ante view has not been used to determine efficiency properties before, it has been used when discussing incentives to manipulate under various matching mechanisms, see Roth and Rothblum (1999), Ehlers (2008) and Kojima and Pathak (forthcoming). What we

[32]Note that some ordinal efficient mechanisms such as cake eating at equal speed (see Bogolmania and Moulin 2001) are also not able to reach the ex ante efficiency levels of the Boston mechanism. The reason is that equal speed cake eating also gives science students a positive probability to attend the art school, even in the case in which there are two scientists and only one artist. Furthermore, using top trading cycles would yield the same outcome as DA, and both are equivalent to random serial dictatorship in this example.

show in this paper is that the ex ante view may be relevant to efficiency, and that it changes what has been the conventional wisdom concerning the Boston mechanism .

7 Conclusions

This paper provides experimental evidence that optimal manipulations in school choice problems may be hard to achieve even in very simple environments and even when participants receive a lot of feedback. While parents in general participate in matching mechanisms only once, or a few times, they may receive information about strategizing from other sources (see Abdulkadiroğlu et al. 2006 and Abdulkadiroğlu et al. forthcoming). It also seems that in actual mechanisms where agents can benefit by manipulating, participants can take a few years to learn to do so (Roth 1990, 1991). In the laboratory we run the experiment for several periods, showing that behavior seems to stabilize fairly quickly and providing participants with a chance to learn how to play in such mechanisms. Furthermore, experiments with multiple periods allow us to compare the outcome under a Boston and DA mechanism in the Uncorrelated environment, as we can take averages over 15 periods instead of just having only a single round as a data point in each session.

While other experiments on mechanisms in two-sided matching use that same methodology (see e.g. Kagel and Roth, 1999, McKinney, Niederle and Roth 2005, Ünver 2001), many school choice mechanism experiments are played exactly once (Chen and Sönmez, 2006, Pais and Pintér, 2008). The second difference between this paper and other school choice mechanism experiments is that participants are always informed about the environment, specifically about how student preferences are generated. This allows us to compute equilibrium behavior and to compare outcomes to equilibrium predictions.

We find that participants' strategies react both to the mechanism, and also to the environment (Correlated and Uncorrelated). Nonetheless, many participants fail to submit optimal strategies in the Correlated environment under the Boston mechanism. We provide clear evidence that manipulations that call for "skipping the middle" (i.e. leaving out the true second choice school) are hard to learn. This provides additional support for the suggestive empirical evidence of Abdulkadiroğlu et al. (2006).

Finally, we showed that a Boston mechanism which provides as many students as possible with their first choice, may have some very good properties. The problem of the strategy-proof DA mechanism is actually very well described by an objection raised by education officials in Boston and New York when they were discussing whether to adopt DA (private communication from Alvin E. Roth and Parag A. Pathak). In a DA mechanism it can be that at some point in the algorithm two students with the same student characteristics, apply to the same school, one of whom ranks the school as his first choice, while the other ranks it late in his preference list. Under DA it is the lottery number that will decide which of the two students receives a seat at the school; the ranking is not taken into account at all. This is the cost of strategy-proofness and is exactly what a Boston mechanism avoids (and why, in general, truth-telling is not an equilibrium under Boston).

In future work we hope to find new hybrid mechanisms that can extend some of the desirable features of the Boston mechanism to more general environments, while keeping truth-telling as an equilibrium. This paper shows that such a research agenda could have a large impact on student welfare.

References

[1] Abdulkadiroğlu, Atila, Yeon-Koo Che, and Yosuke Yasuda (2008), "Expanding 'Choice' in School Choice," working paper.

[2] Abdulkadiroğlu, Atila, Parag A. Pathak, Alvin E. Roth, and Tayfun Sönmez (2005), "The Boston Public School Match," *American Economic Review, Papers and Proceedings*, 368-371.

[3] Abdulkadiroğlu, Atila, Parag A. Pathak, Alvin E. Roth, and Tayfun Sönmez (2006), "Changing the Boston School Choice Mechanism," working paper.

[4] Abdulkadiroğlu, Atila, Parag A. Pathak, and Alvin E. Roth (forthcoming), "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match," *American Economic Review.*

[5] Abdulkadiroğlu, Atila, and Tayfun Sönmez (2003), "School Choice: A Mechanism Design Approach," *American Economic Review*, 93: 729-747.

[6] Bogomolnaia, A., and H. Moulin (2001), "A New Solution to the Random Assignment Problem," *Journal of Economic Theory*, 100, 295-328.

[7] Chen, Yan, and Tayfun Sönmez, (2006), "School Choice: An Experimental Study," *Journal of Economic Theory*, 127: 2002-231.

[8] Dubins, Lester E. and David A. Freedman, (1981), "Machiavelli and the Gale-Shapley Algorithm," *American Mathematical Monthly*, 88, 485 - 494.

[9] Ehlers, Lars, (2008), "Truncation Strategies in Matching Markets", *Mathematics of Operations Research*, 33(2), 327-335.

[10] Erdil, Aytek, and Haluk Ergin, (2008), "What's the Matter with Tie-Braking? Improving Efficiency in School Choice," *American Economic Review*, 98: 669-689.

[11] Ergin, Haluk and Tayfun Sönmez (2006), "Games of School Choice Under the Boston Mechanism," *Journal of Public Economics*, 90: 215-237.

[12] Fischbacher, Urs (2007), "z-Tree: Zurich Toolbox for Ready-made Economic experiments," *Experimental Economics*, 10(2), 171-178.

[13] Gale, David and Lloyd Shapley (1962), "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69: 9-15.

[14] Kagel, John and Alvin E. Roth (1999), "The Dynamics of Reorganization in Matching Markets: A Laboratory Experiment Motivated by a Natural Experiment," *Quarterly Journal of Economics*, 201-235.

[15] Kesten, Onur, (2008), "School Choice with Consent," working paper.

[16] Kesten, Onur and M. Utku Ünver, (2008) "A Theory of School Choice Lotteries: Why Ties Should Not Be Broken Randomly," working paper.

[17] Kojima, Fuhito and Parag A. Pathak, "Incentives and Stability in Large Two-Sided Matching Markets," forthcoming *American Economic Review.*

[18] McKinney, C. Nicholas, Muriel Niederle, and Alvin E. Roth, (2005) "The collapse of a medical labor clearinghouse (and why such failures are rare)," *American Economic Review,* vol. 95, No.3, June 2005, 878-889.

[19] Miralles, Antonio, "School Choice: The Case for the Boston Mechanism," working paper.

[20] Pais, Joana and Ágnes Pintér, (2008), "School Choice and Information: An experimental study on matching mechanisms", *Games and Economics Behavior*, 64, 303-328.

[21] Pathak, Parag, and Tayfun Sönmez, "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism," forthcoming, *American Economic Review.*

[22] Pathak, Parag, and Tayfun Sönmez, (2008), Comparing Mechanisms by their Vulnerability to Manipulation, working paper, June 2008.

[23] Roth, Alvin E. (1982a), "The Economics of Matching: Stability and Incentives," *Mathematics of Operations Research*, 7: 617-628.

[24] Roth, Alvin E. (1982b), "Incentive Compatibility in a Market with Indivisible Goods," *Economic Letters*, 9, 127-132.

[25] Roth, Alvin E. (1990), "New Physicians: A Natural Experiment in Market Organization." *Science,* 250: 1524-1528.

[26] Roth, Alvin E. (1991), "A Natural Experiment in the Organization of Entry Level Labor Markets: Regional Markets for New Physicians and Surgeons in the U.K." *American Economic Review*, 81: 415-440.

[27] Roth, Alvin E, and Uriel G. Rothblum (1999), "Truncation Strategies in Matching Markets - In Search of Advice for Participants," *Econometrica*, 67, 21-43.

[28] Sönmez, Tayfun and M. Utku Ünver (forthcoming), "Allocation and Exchange of Discrete Resources: Matching Market Design for House Allocation, Kidney Exchange, and School Choice," Jess Benhabib, Alberto Bisin, and Matthew Jackson (eds.) Handbook of Social Economics, Elsevier.

[29] Ünver, M. Utku, (2001), "Backward Unraveling over time: The evolution of strategies in the entry level British medical labor market", *Journal of Economic Dynamics and Control*, 25, 1039-1080.

8 Appendix

Proof of Proposition 1. Since Top students have highest priority, there can be no pure strategy Nash equilibrium where fewer than two Tops rank Best first. Now, we consider what the other Top (call her i) could rank first in equilibrium. Clearly, ranking Third first is dominated by ranking Second first. If she ranks Second first, the best response of Averages is to rank Third first. But if they do this, then i can profitably deviate by ranking Best first and Second second, since it is clear that Second won't be taken in the first round of the Boston algorithm. Thus, it must be that any PSNE has all Tops ranking Best first. This implies that Averages' best response is to rank Second first, and Tops, in turn, best respond by ranking Third second, as Second will be taken in the first round of the Boston algorithm. These strategies form a Nash equilibrium, and since we have ruled out all other PSNE, this one must be unique. □

To prove Proposition 2, we will need a few definitions and lemmas.

Definition 1. *Consider a submitted rank order vector where school s is ranked k^{th}. Now consider a new rank order, where, for $j < k$, the vector is the same from the 1^{st} to $(j-1)^{st}$ positions but substitutes school s in the j^{th} spot. Mathematically,*

$$\begin{aligned} r &= (r(1), \dots, r(j-1), \dots, r(k-1), s, \dots) \\ r' &= (r(1), \dots, r(j-1), s, r'(j+1), \dots) \end{aligned} \tag{1}$$

*A mechanism is called **rank monotonic** if the probability of being matched to school s is weakly higher under r' than under r, regardless of how preference, priority, and lottery draws are distributed.*

Lemma 2. *The Boston mechanism is rank monotonic.*

Proof. Consider r and r' from the previous definition. Also consider any state of the world (i.e. a vector of lottery draws, priorities, and submitted rank order lists for all students) where student i submits r and is matched to school s (his k^{th} ranked school). In such a state, school s is not filled by other students prior to Round k of the Boston algorithm, and student i is not matched to any school he ranked higher than k^{th}. Hence, if student i had instead submitted rank-order r', in the same state of the world, he also would have been matched to s. This implies that the probability of student i being matching to school s is weakly greater when he submits r' instead of r. □

Definition 3. *If the probability of student i being matched to his x^{th} ranked school is independent of his submitted rank order list, we say that his preference revelation problem exhibits* ***school anonymity****.*

Lemma 4. *Say there are k schools of quota q and n students. If the submitted preferences of $\mathcal{I}\backslash i$, as well as all lottery draws are uniformly distributed, then student i's preference revelation problem exhibits school anonymity.*

Proof. Let student i submit a rank order list r_i. Let the other students' rank order lists be given by r_{-i}, and the lottery ordering over students by L. Let A be the set of rank order lists and lottery draws, (r_{-i}, L), such that student i is matched to his j^{th}-ranked school, $r_i(j)$. Now let student i submit a different rank-order list, r_i'. This induces a permutation mapping π, such that $\pi(r_i) = r_i'$. Now, consider the set of rank-order lists and lottery draws such that student s is matched to $r_i'(j)$. Call it A'. This set is easily constructed from A. For each $(r_{-i}, L) \in A$, by symmetry, we know that $(\pi(r_{-i}), L) \in A'$. π is a one-to-one mapping, and as such, has a unique one-to-one inverse. Any distinct element in A must have a corresponding distinct element in A', so we conclude that $|A| = |A'|$, and since our uniformity assumptions tell us that each element of these sets carries the same probability weight, we conclude $\Pr(A') = \Pr(A)$, i.e. the probability of i being matched with his j^{th} choice school is independent of the rank-order list that he submits. □

Lemma 5. *If a student's preference revelation problem exhibits school anonymity, the matching mechanism begin used exhibits rank monotonicity, and $x < y$, then the probability of student i being matched to his x^{th} ranked school is weakly greater than the probability of his being matched to his y^{th} ranked school.*

Proof. School anonymity means that the probability of being matched to a school depends only on what it was ranked. If the probability of being matched to the y^{th} ranked school were strictly larger than the probability of being matched to the x^{th} ranked school, the rank monotonicity condition would clearly be violated. Thus, the theorem is proven. □

Lemma 6. *If all schools are of the same size, and submitted preferences over schools and lotteries over students are all uniformly drawn, then truth-telling is a best-response.*

Proof. Lemma 4 shows that the environment will exhibit school anonymity. The rank monotonicity of the Boston mechanism and school anonymity then tell us that probability of match to the j^{th} ranked school is independent of the submitted rank-order and decreasing in j. The best response in this case is clearly to put the favorite school in the first slot, the second most favorite in the second slot, etc. So, truth-telling is a best response. □

Definition 7. *A strategy is called a **preference permutation** if it calls for a student to submit the same permutation of his true preference ordering regardless of what that true ordering might be.*

Now, we are prepared to prove Proposition 2.

Proof of Proposition 2. If all students but one (call her i) play preference permutations, then the submitted preferences will be uniformly distributed. Lemma 6 tells us that the best response to such an environment is truthful revelation. Truth-telling is a preference permutation, hence the only unrestricted equilibrium where all agents play preference permutations must be the equilibrium where all agents truthfully reveal their preferences. □