Best-Response Dynamics in the Boston Mechanism

Oğuzhan Çelebi*
MIT

October 21, 2022

Abstract

I analyze a setting where the Boston Mechanism (BM) is applied repeatedly and students form their application strategies by best-responding to the admission cutoffs of the previous period, a process I call the Repeated Boston Mechanism (RBM). If students are truthful in the initial period, the allocation under RBM converges in finite time to the student optimal stable matching (SOSM), which is the Pareto-dominant equilibrium of BM and the outcome of the strategy-proof Deferred Acceptance Mechanism. If some students are sincere and do not strategize, then the allocation under RBM with sincere students converges to the SOSM of a market in which sincere students lose their priorities. When students best-reply to some initial cutoffs in the first period, RBM converges to SOSM if students are initially optimistic about their admissions chances but may cycle between allocations Pareto-dominated by SOSM if they are pessimistic. My results provide a foundation for equilibrium analysis under BM and help explain why students play suboptimal and overcautious strategies.

*MIT Department of Economics, 50 Memorial Drive, Cambridge, MA 02142. Email: ocelebi@mit.edu
I am grateful to Roberto Corrao, Joel Flynn, Parag Pathak and Tayfun Sönmez for helpful comments.

1. Introduction

The Boston Mechanism (BM) is a centralized assignment mechanism that has been widely used in many parts of the world to assign students to schools. The most important feature of BM is that it is not strategy-proof: students lose their priority in schools they rank lower in their preference lists to students who rank those schools higher. Therefore, researchers have studied the equilibria of BM to understand its properties. Starting with the characterization of equilibria of BM (Ergin and Sönmez, 2006), a large body of literature that has analyzed BM, and compared it to its main competitor, the Deferred Acceptance Mechanism (DA).

This paper contributes to the literature on BM by asking the following question: (when) should we expect the equilibrium of BM to be played? Motivated by the fact that BM is used repeatedly across periods in many markets,[1] I set up a multi-period model where each period students submit rankings of the schools to a centralized clearinghouse which uses BM to determine the allocation. In the baseline model, which I call the Repeated Boston Mechanism (RBM), in each period, students play a best-response to the admission cutoffs of the previous period, with the exception of the initial period, where they apply truthfully. This paper focuses on characterizing the (non-)convergence of the best-response dynamics under RBM and its modifications.

My analysis builds on a surprisingly close connection between the steps of DA and the periods of RBM.[2] To illustrate this connection, I define a slightly modified version of DA, the Modified Deferred Acceptance Mechanism (MDA) where students skip schools that will reject them for sure during the implementation of DA. This modification does not change the outcome of the mechanism and MDA is equivalent to DA. Moreover, each step of MDA is analogous to the corresponding period of RBM: for each t, the set students who are tentatively accepted by each school in step t of MDA are identical to the set of students who are accepted by that school in the first step of BM in period t. This reveals a surprising relationship between DA and best-response dynamics under BM. My main result, Theorem 1, shows that the matching implemented in RBM converges in finitely many periods to the Student Optimal Stable Matching (SOSM), which is the dominant-strategy outcome of DA and the Pareto-dominant equilibrium of BM. Theorem 1 provides a foundation to the equilibrium analysis of BM based on best-response dynamics by showing that if BM is repeatedly used in the same market in consecutive periods and students form their strategies by myopically

[1] Examples of settings where BM or variants are used include college admissions in China (Chen and Kesten, 2017), public school systems in Charlotte (Hastings, Kane, and Staiger, 2009), Beijing (He, 2015) and Barcelona (Calsamiglia, Fu, and Güell, 2020). See Agarwal and Somaini (2018) for more examples and details.

[2] In both BM and DA, students submit a ranking of schools to the mechanism, which then proceeds iteratively and returns a matching of students to schools as the outcome after finitely many steps.

best responding to previous period's cutoffs, the implemented matching converges to a stable matching in finitely many periods. The rest of the paper relaxes some assumptions of RBM and to analyzes convergence in more general settings.

Section 4, considers an environment with unsophisticated students, who always report their true preferences, and sophisticated students, who strategize as they do under RBM. Pathak and Sönmez (2008) characterize the equilibria of BM in this setting, show that unsophisticated students lose their priorities to sophisticated ones and sophisticated students prefer the Pareto-dominant Nash equilibrium of BM to DA. I show that RBM with unsophisticated students converge in finitely many periods to the Pareto-dominant equilibrium of BM with unsophisticated students, providing a foundation to their equilibrium characterization and their equilibrium selection for the comparison between BM and DA.

Section 5 relaxes my assumptions on the first period behavior and allows students to best-reply to some initial cutoff scores, instead of applying truthfully. The initial cutoff scores determine the optimism of students about their admission chances: if cutoff scores are low, they believe that their score is enough to obtain a place in most schools, while if cutoffs are high, they believe most schools are not achievable. I show that if students are initially optimistic enough about their admission chances, then RBM converges to the SOSM, while if students are pessimistic (in particular, if they believe their school under SOSM is not achievable), then it is possible that RBM does not converge and implements matchings that are Pareto-dominated by the SOSM in each period. These extensions help reconcile the theory with the observations of non-equilibrium play under BM in real-world markets (He, 2015; Kapor, Neilson, and Zimmerman, 2020; Song, Tomoeda, and Xia, 2020).

Section 6 extends the analysis to large, random markets. First I extend results to a market with continuum students (Abdulkadiroğlu, Che, and Yasuda, 2015; Azevedo and Leshno, 2016). Then, I study the best-response dynamics when a different finite market is sampled from a stationary continuum market and show that for each period, the matching implemented in the finite random market under RBM converges to the matching implemented in the continuum market. This shows that the results are robust to uncertainty in large markets.

Related Literature. This paper contributes to the literature on school choice. Abdulkadiroğlu and Sönmez (2003) formalize and study the Boston Mechanism and compare it with its alternatives, DA and the Top Trading Cycles (TTC) mechanisms. The equilibria of BM is characterized in Ergin and Sönmez (2006) and this characterization is extended to unsophisticated students in Pathak and Sönmez (2008). Chen and Kesten (2017) characterize a parametric family of mechanisms, the application-rejection assignment mechanisms, where BM and DA constitute two limiting cases and empirically demonstrate that authorities move away from BM towards stable mechanisms such as DA. Dur, Pathak, Song, and Sönmez (2022) study the assignment mechanism in Taiwan, which is a hybrid of BM and DA. Akbarpour, Kapor, Neilson, van Dijk, and Zimmerman (2022) show that when students have the same ordinal preferences and some have outside options, students with outside options prefer manipulable BM to DA. Abdulkadiroğlu, Che, and Yasuda (2011) study BM from an ex-ante perspective where students do not know their priorities and have identical ordinal preferences and show that the equilibria of BM generate higher ex-ante welfare for each student than the dominant strategy outcome of DA. In a similar setting, Babaioff, Gonczarowski, and Romm (2018) show that it is possible that non-trivial fraction of students prefer to be sincere rather than strategic under BM.

Dur, Hammond, and Kesten (2021) study a related model where students sequentially submit their preferences to BM and DA mechanisms and compare the efficiency of equilibria. Most relatedly, they also consider a model where students can update their submitted preferences as many times as they want and observe the latest submission of each student while they are resubmitting their preferences. When all students rank the best achievable school first, they show that this process converges to SOSM. My results complement their work by considering a setting where students observe the last period's cutoff scores and submit their preferences once, rather than a setting where students can update their preferences observing what others have done. Thus, this paper manages to uncover the relationship between repeated BM and DA, show that convergence may be attained under repeated application of BM even if students submit preference once and demonstrate how convergence may depend on the beliefs of students and their sophistication.

This paper also builds on the empirical literature that study manipulable mechanisms, and BM in particular. He (2015) examines choice data from Beijing where the Boston mechanism is used, and show that parents are overcautious and play safe strategies too often. Agarwal and Somaini (2018) find evidence that students engage in strategic behavior under BM. Song et al. (2020) analyze college admissions in China and find that equilibrium is not being played. They show that two different types of behavioral students, unsophisticated students who reveal their preferences truthfully and cautious students, who are pessimistic about their admission chances, make up a large portion of student populations. The extensions that show non-convergence to SOSM under unsophisticated or initially pessimistic students are in line with their results. Kapor et al. (2020) demonstrates that students' beliefs about admissions chances differ from rational expectations values and affect their choices. They evaluate the effects of switching to DA, and of improving households' belief accuracy, and find that both would improve welfare.

2. Model

Let $I = \{i_1, ..., i_n\}$ denote the set of students and $C = \{c_1, ..., c_m\}$ denotes the set of schools. Schools' capacities are given by $Q = \{q_{c_1}, ..., q_{c_m}\}$. Each student has a strict preference over the set of schools and being unmatched, denoted by $\succ_I = (\succ_{i_1}, ..., \succ_{i_n})$.[3] Let $\succeq_i$ denote the "at least as good as" relation induced by $\succ$. Σ denotes the set of all strict student preferences. Each student has score $s_c(i) \in [0, 1]$ in school c, where $s_c(i) \neq s_c(j)$ for all i, j and c. The scores of the lowest and highest scoring students in each school are normalized to 0 and 1, respectively. Strict school priorities $\succ_C = (\succ_{c_1}, ..., \succ_{c_m})$ are derived from scores by ranking students with respect to their scores. A market is a tuple $\omega = \{I, C, Q, \succ_I, \succ_C\}$.

A *matching* is a function $\mu : I \cup C \to 2^I \cup C$, where $\mu(i) \in C \cup \{i\}$ for all $i \in I$, $\mu(c) \subseteq I$, $|\mu(c)| \leq q_c$, and $\mu(i) = c$ if and only if $i \in \mu(c)$. $\mathcal{U}$ is the set of all matchings. A matching μ is *blocked* by student i and school c if i prefers c to $\mu(i)$ and either c prefers i to some $i' \in \mu(c)$ or $|\mu(c)| < q_c$. A matching μ is individually rational if $\mu(i) \succeq_i i$ for all i. A matching μ is *stable* if it is individually rational and is not blocked.

A mechanism $\phi : \Sigma \to \mathcal{U}$ produces a matching given preference reports from the students. $\phi(\sigma) = \mu_\phi(\sigma, \cdot)$ for all $\sigma \in \Sigma$ where $\mu_\phi(\sigma, i)$ and $\mu_\phi(\sigma, c)$ denote the school and the set of students i and c are matched to under ϕ. Next, I describe the Deferred Acceptance and Boston Mechanisms.

The Deferred Acceptance Mechanism (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 his rank-order list. If all remaining schools are below the outside option, he applies nowhere. 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 stops when no rejections are issued. Each school is matched to the students it is holding at the end.

The Boston Mechanism (BM)

Step 1: Students apply to their first choice school. Schools reject the lowest-ranking students in excess of their capacity. All other offers are immediately accepted and become permanent matches. School capacities are adjusted accordingly.

[3] Being unmatched is denoted by i_k for student i_k.

Step t: If a student is rejected in Step $t-1$, he applies to the next school on his rank-order list. If all remaining schools are below the outside option, he applies nowhere. 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 no rejections are issued.

It is well known that DA is strategy proof (Dubins and Freedman, 1981; Roth, 1982), and its outcome is the student optimal stable matching (SOSM), which is individually rational and stable, but is not efficient for students. BM, on the other hand, is neither strategy proof nor stable, but is efficient with respect to the submitted preferences.

To make the connection between DA and RBM clear, I introduce a slightly modified version of the DA, which is equivalent to the original one:

The Modified Deferred Acceptance Mechanism (MDA)

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 highest school on his rank-order list within the schools that either (i) did not fill its capacity in the last round or (ii) temporarily hold the offer of a student who has lower ranking in that school.* If there are no schools in his list that satisfies either (i) or (ii), or that school is ranked below the outside option, he applies nowhere. 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 stops when no rejections are issued. Each school is matched to the students it is holding at the end.

Under MDA, students skip applications to schools that already have enough applicants to fill their capacity with more preferred students. MDA is useful in understanding the relationship between DA, the student optimal stable matching and BM under repeated play. The following lemma shows that this mechanism is equivalent to DA.

Lemma 1. *Both DA and MDA terminate in finite time and yield the student optimal stable matching.*

3. Best-Response Dynamics under Boston Mechanism

I now describe the setting for the repeated application of the Boston Mechanism. Time is discrete and denoted by $t \in \mathbb{N}$. These periods correspond to the consecutive years where students are matched to schools. In each period, each student in I submits a ranking of schools in C and being unmatched. A strategy $\sigma \in \Sigma$ is a best-response to σ_{-i} under mechanism ϕ if

$$\mu_\phi(\sigma, \sigma_{-i}, i) \quad \succeq_i \quad \mu_\phi(\sigma', \sigma_{-i}, i) \text{ for all } \sigma' \in \Sigma \tag{1}$$

While modelling the behavior of the students, it is not very realistic to consider a setting where students best reply to the strategy used by other students in the previous period for a couple of reasons. First, the strategy of all students is a high dimensional and complicated object. Therefore, it is not feasible the school district to report this statistic and students to process it to calculate their best-responses. Second, disclosing the rank-order lists of the students in previous years would be challenging for the school board from a legal perspective. However, the cutoff scores of each school, which correspond to the score of the lowest scoring student who is assigned to a school, is a much simpler object which is reported in various settings.

I use $S^c \in [0, 1]$ to denote the cutoff in school c, and $S = \{S^1, \ldots, S^n\}$. Let $S^c_t(\sigma)$ denote the score of the lowest scoring student who is assigned to school c in the first step of BM in period t if the school fills its capacity in that step (set $S^c_t(\sigma) = 0$ otherwise). Moreover, let $S_t = \{S^{c_1}_t, \ldots, S^{c_m}_t\}$. A strategy σ is *compatible* with a vector of cutoffs S if the cutoff scores in the first step of BM under σ is S, that is $S_t(\sigma) = S$. Let $\mathcal{S}$ denote set of all cutoffs that are compatible with at least one $\sigma \in \Sigma$. These are the cutoffs that might arise in the first step of the implementation of BM. A school c is *achievable for i at cutoffs S* if $s_c(i) \geq S^c$. Moreover, let $FC_i(S)$ denote the set of all strategies that (i) ranks the most preferred school that is achievable at S as the first choice (if there is no achievable school, then students rank their most preferred school first), (ii) ranks all schools preferred to being unmatched above being unmatched and (iii) ranks all schools less preferred to being unmatched below being unmatched. For the rest of the paper, I assume that ϕ is the Boston Mechanism and suppress dependence on ϕ.

Lemma 2. *If $\sigma_i \in FC_i(S)$, then it is a best-response to all σ' such that σ' is compatible with S.*

This lemma shows that students do not need to observe σ_{-i} to determine their best-response. Rather, observing S and determining the most preferred school that is achievable under S, a simple exercise, is sufficient for a student to compute a best-response.[4] I study the following best-response process, which I call *Repeated Boston Mechanism* (RBM):

- In period 1, all students apply truthfully, $\sigma_i^1 = \succ_i$. BM is used to determine the allocation.
- In period t, all students choose a strategy $\sigma_i^t \in FC(S_{t-1})$. BM is used to determine the allocation.

Let T denote the step where MDA terminates. We have the following result.

Proposition 1. *The set of students who are accepted by school c in the first step of the Boston Mechanism in period t of RBM is identical to the set of students who are tentatively accepted by school c in step t of the MDA for all $t \leq T$.*

This proposition shows that there is a close connection between periods of RBM, where students myopically best respond using information on the cutoffs from the previous period, and the steps in MDA. To get intuition about the result, suppose that the result holds for some t. If student i is tentatively accepted to school c at step t of MDA or applies to c in step $t+1$ of MDA, then that student has already either been rejected by or skipped more preferred schools in previous rounds of MDA. Given that the result holds for period t, in both cases, we know (i) the period t first step cutoffs of all schools that i prefers to c are higher than i's score at those schools and (ii) the period t first step cutoff of school c is lower than i's score at c. Therefore, i applies to c in step 1 of period $t+1$ of RBM. Thus, the set of applicants for each school is identical in the first step of BM in period $t+1$ of RBM and step $t+1$ of MDA, which means that the set of students who are accepted in the former and tentatively accepted in the latter are the same. Moreover, Proposition 1 implies that for in all periods after T, the BM terminates at step 1 and the outcome is SOSM:

Theorem 1. *The matching implemented in RBM converges in finitely many periods to the student optimal stable matching.*

As BM is not strategy-proof, to make any predictions about its outcome, one needs to study the equilibria of a preference revalation game where students submit their preferences to the mechanism. Ergin and Sönmez (2006) show that the set of Nash equilibria under BM correspond to the set of stable matchings and interpret their result as evidence in favor of DA. First, the Pareto-dominant equilibrium of BM is SOSM, which is attained under the

[4] Some other strategies that does not rank the most preferred achievable school first might be best-responses. However this is only possible if all schools ranked above the the most preferred achievable school are not achievable and the most preferred achievable school has a cutoff of 0 under $\mathcal{S}$.

DA when students report their preferences truthfully. Therefore, switching from the BM to DA cannot harm any student, but can potentially improve the outcomes of some. Second, as being truthful is a dominant strategy under DA, DA does not require the students to strategic about their applications.

Theorem 1 complements Ergin and Sönmez (2006) by providing a foundation to their analysis based on best-response dynamics by showing that if BM is repeatedly used in the same market, and students form their strategies by myopically best responding previous period's cutoffs, the implemented matching converges to a stable matching in finitely many periods. Therefore, the set of equilibria characterized in Ergin and Sönmez (2006) can be reached under best-response dynamics with minimal information about play in the previous period. Moreover, it selects SOSM as the equilibrium. Even though this result indicates the BM would converge to its Pareto-dominant equilibrium, DA still has a couple of advantages over BM. First, the convergence to SOSM is not immediate, while under DA it is reached in the first period. Moreover, even after the convergence is reached, students still need to be strategic about their applications in each period, while under DA they can rank the schools truthfully.

The rest of the paper studies extensions of the model to analyze convergence of RBM under different conditions.

Unsophisticated Students. Under RBM, all students play a best-response to the previous period's cutoffs. An important question is the following: what happens if some subset of students are not able to play strategically, but report their preferences truthfully? Pathak and Sönmez (2008) develop a framework to allow for unsophisticated students who report their preferences truthfully, extending the analysis of Ergin and Sönmez (2006), while Song et al. (2020) empirically demonstrate that a non-trivial fraction of students behave this way under BM. In Section 4, I extend the analysis in this section by allowing some students to be unsophisticated and show that RBM converges to the the Pareto-dominant equilibrium of BM that Pathak and Sönmez (2008) characterize, providing foundation for their equilibrium characterization as well as their focus on the Pareto-dominant equilibrium when analyzing preferences of sophisticated students over these two mechanisms.

First Period Behavior. Under RBM, all students are truthful in the first period. However, it still possible students are actually strategic in the initial period, depending on their beliefs about their admission chances in different schools. In particular, if students are best-responding to some initial cutoffs, truthful revelation of preferences correspond to a setting where they believe all schools are achievable, in other words, these students are optimistic about their admission chances. In Section 5, I relax the assumption of truthful revelation in the first period and show that if students are optimistic enough, then RBM converges to SOSM, while if students are pessimistic (in particular, if they believe their outcome under SOSM is not achievable), the implemented matching may not converge but instead cycle between matchings that are Pareto-dominated by SOSM.

Large Random Markets. Under RBM, in each period, the preferences and scores of students are the same, which is a reasonable assumption for large student assignment markets. In Section 6, I first study a continuum matching model (Abdulkadiroğlu et al., 2015; Azevedo and Leshno, 2016) and extend the results to that setting. Next, I study a setting where n students are sampled independently from a given continuum market each period and prove the convergence of RBM is obtained asymptotically (as $n \to \infty$), showing that if a market is large and stationary and students best-respond to the previous period's cutoffs, then the implemented matching converges to the SOSM.

4. Sophisticated and Unsophisticated Students

The set of students is given by $I = I_S \cup I_U$. If $i \in I_s$, then in each period, i behaves strategically as they did under RBM, while if $i \in I_U$, then $\sigma_i^t = \succ_i$, that is, students in I_U apply truthfully in every period. Following Pathak and Sönmez (2008), we refer to these students as sophisticated and unsophisticated students, respectively.

Given a market $\omega = \{I_S, I_U, C, Q, \succ_I, \succ_C\}$, construct the augmented preferences $\tilde{\succ}_C$ where each school c ranks students as follows

> Rank all sophisticated students and unsophisticated students who ranks c first according to $\succ_c$
> Rank all unsophisticated students who ranks c second according to $\succ_c$
> $\vdots$
> Rank all unsophisticated students who ranks c last according to $\succ_c$

Under $\tilde{\succ}_C$, unsophisticated students lose their priorities to sophisticated students in all schools apart from the one they rank first. Pathak and Sönmez (2008) define the *augmented market*, $\tilde{\omega} = \{I, C, Q, \succ_I, \tilde{\succ}_C\}$ and show that (i) the Nash equilibria of BM where unsophisticated students mechanically submit their true preferences correspond to the set of stable matchings of the augmented market $\tilde{\omega}$ and (ii) the Pareto-dominant Nash equilibrium of this game corresponds to the SOSM of the augmented market $\tilde{\omega}$ and (iii) unsophisticated students become better off if they become sophisticated. The following proposition shows that RBM with unsophisticated students converges to the Pareto-dominant equilibrium of BM.

Proposition 2. *The matching implemented in RBM with unsophisticated students converges in finitely many periods to the student optimal stable matching under the augmented market* $\tilde{\omega}$.

Proposition 2 provides a foundation for the equilibria characterized in Pathak and Sönmez (2008) based on best-response dynamics. Moreover, Pathak and Sönmez (2008) also show that sophisticated students prefer the Pareto-dominant equilibrium of BM to the dominant strategy outcome under DA. This indicates that BM favors sophisticated parents if the Pareto-dominant Nash equilibrium is played. Their result explains why some parents were in favor of BM and provides formal support for Boston Public School's position on changing their student assignment system to level the playing field for students do not have resources to be strategic about their applications.[5] Proposition 2 complements their result by showing that the Pareto-dominant equilibrium would be obtained under best-response dynamics, providing further justification for focusing on the Pareto-dominant equilibrium instead of other equilibria, which may not be preferred by the sophisticated students.

5. Initial Conditions and First Period Play

In this section, I study how the outcome of RBM depends on the behavior in the initial period, which turns out to be an important determinant of the convergence of RBM. Let S_0 denote the initial cutoffs to which students best reply in the first period of RBM. If $S_0 = \{0, \ldots, 0\}$, then students are optimistic in the sense that they believe all schools are achievable and in the first period and they rank first their most preferred school, in other words, they are truthful. If initial cutoffs are higher, then some schools become unachievable for some students, and students become more pessimistic about their admission chances in the initial period. In particular, if the initial cutoffs are above the DA cutoffs, then there are students who are pessimistic enough that they believe their match under SOSM is not achievable. The following example shows that if the initial cutoffs are above the DA cutoffs (denoted by S_{DA}), then they may stay above the DA cutoffs in each period, which implies

[5]See Pathak and Sönmez (2008) for a detailed discussion of this topic. In particular, they note that the BPS Strategic Planning Team recommended the implementation of a strategy-proof algorithm by saying:

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

that (i) the cutoffs and the matchings do not converge and (ii) the realized matchings are Pareto dominated by SOSM.

Example 1. *There are three schools with unit capacity and three students. The priorities and preferences of students are as follows.*

Students	s_{c_1}	s_{c_2}	s_{c_3}	$\succ$
i_1	3	1	2	$c_2 \succ c_3 \succ c_1$
i_2	2	3	1	$c_3 \succ c_1 \succ c_2$
i_3	1	2	3	$c_1 \succ c_2 \succ c_3$

In the first step of DA and MDA, all students apply to their most preferred school and are accepted. Thus both mechanisms terminate in the first period and the cutoffs are $S_{DA} = \{1, 1, 1\}$. *However, the matchings where all students obtain their second choice and their third choice are also stable, with cutoffs* $\{2, 2, 2\}$ *and* $\{3, 3, 3\}$ *respectively. Let* $S_0 = \{2, 1, 1\}$. *The following table shows the applications and cutoffs in the first step of first 4 periods of RBM, as well as the implemented matching in each period.*

	Applications			*cutoffs*			*Matching*		
	c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3
P1		$i_1, \underline{i_3}$	i_2	1	2	1	i_1	i_3	i_2
P2	i_3		$\underline{i_1}, i_2$	1	1	2	i_3	i_2	i_1
P3	$\underline{i_2}, i_3$	i_1		2	1	1	i_2	i_1	i_3
P4		$i_1, \underline{i_3}$	i_2	1	2	1	i_1	i_3	i_2

In the first step of the BM in the first period of RBM, i_1 *and* i_3 *apply to* c_2 *and* i_2 *apply to* c_3. *The cutoffs for next period then becomes* $S_1 = \{1, 2, 1\}$ *and* i_3 *applies to* c_1 *while* i_1 *and* i_2 *apply to* c_3. *This results in* $S_2 = \{1, 1, 2\}$ *and in the next period,* i_1 *applies to* c_2 *while* i_2 *and* i_3 *apply to* c_1, *which results in the cutoffs* $S_3 = \{2, 1, 1\}$. *Since* $S_0 = S_3$, *period* 4 *is identical to period* 1, *and the the matchings implemented under RBM cycles between matchings implemented in periods* 1, 2 *and* 3 *without converging to a matching. In each period, one student is assigned to her most preferred school, one student is assigned to her second most preferred school while one student is assigned to her least preferred school. This outcome is dominated by SOSM, which is attained under DA.*

The close connection between MDA and RBM, and the fact that under MDA the cutoffs are increasing and converges to SOSM cutoffs suggests that under RBM, if initial cutoffs are below the DA cutoffs, then the convergence could be attained. The following example shows that convergence may not be attained even in that case.

Example 2. *There are two schools with unit capacity and five students. The preferences and scores of the students are as follows.*

Students	s_{c_1}	s_{c_2}	$\succ$
i_1	5	3	$c_2 \succ_i c_1$
i_2	3	5	$c_1 \succ_i c_2$
i_3	2	4	$c_1 \succ_i c_2$
i_4	0	1	$c_2 \succ_i c_1$
i_5	1	0	$c_2 \succ_i c_1$

The applications and admissions in each step of MDA, as well as the realized cutoffs at the end on each step are given in the following table:

	Applications		*Cutoffs*	
	c_1	c_2	c_1	c_2
S1	$\underline{i_2}, i_3, i_5$	$\underline{i_1}, i_4$	2	2
S2	$\underline{i_2}$	$i_1, \underline{i_3}$	2	3
S3	$\underline{i_1}, i_2$	$\underline{i_3}$	4	2
S4	$\underline{i_1}$	$\underline{i_2}, i_3$	4	4
S5	$\underline{i_1}$	$\underline{i_2}$	4	4

By Proposition 1, if $S_0 = (0,0)$, then the students who are accepted by the schools in the step 1 of first 5 periods of RBM are given by the table above. Moreover, in later periods, the outcome is the same as the outcome of step 5. Suppose that $S_0 = (4,2)$. RBM proceeds as follows:

	Applications		*Cutoffs*	
	c_1	c_2	c_1	c_2
P1S1	$\underline{i_5}$	$i_1, \underline{i_2}, i_3, i_4$	1	4
P2S1	$\underline{i_1}, i_2, i_3, i_5$	$\underline{i_4}$	4	1
P3S1	$\underline{i_5}$	$i_1, \underline{i_2}, i_3, i_4$	1	4
⋮	⋮	⋮	⋮	⋮

Under initial cutoffs $(4,2)$*, most students think that* c_2 *is achievable, while* c_1 *is not. They apply to school* c_2*, increasing its cutoff score, while decreasing the cutoff of* c_1*. However, both schools are unachievable for* i_4 *and* i_5*, who apply to their most preferred school and are matched to that school in even and odd periods, respectively. Thus, the neither the cutoffs nor the matchings under RBM converges to the SOSM, but cycles among matchings that are not stable.*

In Example 2, the cutoffs start below S_{DA} and stay below S_{DA} in all periods but do not converge to S_{DA}. Next proposition shows this result is true in general.

Proposition 3. *If* $S_0 \leq S_{DA}$*, then* $S_t \leq S_{DA}$ *for all* t*. Therefore,* S_t *either converges to* S_{DA} *in finitely many periods, or cycles below* S_{DA}*.*

Thus, if students are optimistic enough that they believe their match under the student optimal stable matching is achievable, then they believe this would be the case in all future periods. Unlike Example 1, the reason of non-convergence is not that students do not apply to the school they will receive in the SOSM, thinking it is unachievable. The convergence does not happen because students may be optimistic in each period and believe that schools more preferred to their match in SOSM are achievable, and never apply to that school.

Given my result on possible non-convergence of RBM, the next question is the efficiency of the matchings that are implemented in a cycle. The matchings that constitute the cycle in Example 2 neither Pareto dominate, nor are Pareto dominated by SOSM. Without any restrictions on strategies, *i.e.*, if students can submit any ranking which is a best-response to previous period's cutoffs, a matching implemented in a cycle of RBM can also be Pareto dominated by the SOSM, as students can be matched to suboptimal schools after the first step in each period. However, if students preserve the relative ranking of schools conditional on best replying, then this cannot happen. Formally, σ_i is an *order-preserving best reply* to cutoffs S if σ_i ranks the most preferred achievable school first and preserves the relative rankings of other schools (as well as the outside option) under $\succ_i$.

Proposition 4. *If* $S_0 \leq S_{DA}$ *and students' strategies are order-preserving best replies, then any matching implemented in an RBM cycle cannot be Pareto dominated by the SOSM.*

The reason behind this result is that when students use order-preserving best replies, if a student is matched to a school worse than their match under SOSM, then there must be at least one student who has become strictly better off compared to SOSM. I now present an example where all matchings in a cycle Pareto dominate SOSM.

Example 3. *There are four schools and five students. The preferences and scores of the students are as follows.*

Students	s_{c_1}	s_{c_2}	s_{c_3}	s_{c_4}	$\succ$	*SOSM*
i_1	1	2	3	0	$c_1 \succ_i c_3 \succ_i c_2$	c_2
i_2	3	1	0	0	$c_2 \succ_i c_1$	c_1
i_3	0	0	2	3	$c_3 \succ_i c_4$	c_4
i_4	0	0	4	2	$c_4 \succ_i c_3$	c_3
i_5	2	0	1	1	$c_3 \succ_i c_4 \succ_i c_1$	i_5

SOSM is inefficient, as i_1 and i_2 as well as i_3 and i_4 can exchange their schools and become better off. However, these exchanges are blocked by i_5 and i_1, respectively. Suppose that $S_0 = (1,1,4,1)$. RBM proceeds as follows:

	Applications				*Cutoffs*				*Matching*			
	c_1	c_2	c_3	c_4	c_1	c_2	c_3	c_4	c_1	c_2	c_3	c_4
P1S1	$\underline{i_1}$	$\underline{i_2}$	$\emptyset$	$\underline{i_3}, i_4, i_5$	1	1	1	4	i_1	i_2	i_4	i_3
P2S1	$\underline{i_1}$	$\underline{i_2}$	$i_3, \underline{i_4}, i_5$	$\emptyset$	1	1	4	1	i_1	i_2	i_4	i_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

As the cutoffs at the end of period 2 are identical to the initial cutoffs, $S_t = (1,1,4,1)$ for odd t and $S_t = (1,1,1,4)$ for even t, while the same matching is implemented in each period. In this matching, students i_1 and i_2 are strictly better off compared to SOSM, while all other students get the same outcome.[6] *Thus, in each period of RBM, the implemented matching Pareto dominates the student optimal stable matching.*

The examples and propositions in this section show how convergence may not be attained and how students may be worse or better off under BM. I now turn to the following question: when does RBM converge to the student optimal stable matching? It turns out that MDA is useful in understanding when convergence happens. Let $\tilde{S}_t^c$ denote the score of the lowest scoring student who is matched to school c in period t of MDA.[7] As at each step of MDA, the students who are tentatively admitted to any school apply to that school, the cutoffs under MDA, $\tilde{S}_t^c$, are increasing in t. Intuitively, as the algorithm proceeds, each school replaces

[6] This more efficient matching is not stable, as i_5 would block it. However, i_5 applies to either c_3 or c_4 in each period, and does not initiate a rejection cycle between i_1, i_2 and i_5.

[7] I set $\tilde{S}_0^c = 0$ as there are no student matched to any school at the start of the algorithm and $\tilde{S}_t^c = 0$ if c has empty seats in step t.

lower scoring tentatively admitted students with higher scoring ones and students move towards less preferred schools. Therefore, $\tilde{S}_t^c$ can be interpreted as an index of optimism regarding a student's admission chances at school c. A cutoff S_0 is *compatible* with MDA if there exists k such that for all c, $S_0^c \in [\tilde{S}_k^c, S_{k+1}^c]$. This means that the initial cutoffs of all schools are between the cutoffs in steps k and $k+1$ of MDA. The following proposition extends the close connection between MDA and RBM to this setting.

Proposition 5. *If S_0 is compatible with MDA, RBM converges to student optimal stable matching in finitely many periods.*

To prove this result, I first show that if S_t is compatible with round k cutoffs of MDA, then S_{t+1} is compatible with round $k+1$ cutoffs of MDA (Lemma 5). To see why, first note that when the cutoffs of other schools are higher, more students (in set inclusion sense) demand school c and next period cutoff of school c increase. Moreover, when other schools cutoffs are exactly $\tilde{S}_k^{-c}$ (regardless of school c's cutoff, as long as it is compatible), next period cutoff of c under RBM is $\tilde{S}_{k+1}^c$. Thus, whenever S_t is between $\tilde{S}_k$ and $\tilde{S}_{k+1}$, then next period cutoffs under RBM is between $\tilde{S}_{k+1}$ and $\tilde{S}_{k+2}$. As the cutoffs under MDA converges to S_{DA} in finitely many rounds, so does the cutoffs under RBM.

The main take-away from Proposition 5 is that, if students' levels of optimism about their admissions chances at different schools are similar in the sense that $S_0^c \in [\tilde{S}_k^c, \tilde{S}_{k+1}^c]$ for all $c \in \mathcal{C}$ for some k, then RBM converges to SOSM. Given S_0, the students are *optimistic enough* if $S_0 \leq \tilde{S}_1$. As this implies compatibility, we have the following theorem:

Theorem 2. *If students are optimistic enough, then RBM converges to the student optimal stable matching in finitely many periods.*

Theorem 2 reinterprets truthfulness in the first period of RBM as optimism regarding admissions chances and further emphasizes the importance of beliefs of students over the behavior of best-response dynamics under BM.

Finally, I analyze a special case of the model to understand the reasons behind non-convergence. Given $\{s_c\}_{c\in\mathcal{C}}$, *scores are common across schools* if $s(i) \equiv s_c(i) = s_{c'}(i)$ for all i, c and c'.[8] It is well known that when scores are common across schools, DA is equivalent to the Serial Dictatorship Mechanism and both mechanisms return the unique stable matching of the market.

Proposition 6. *When scores are common across schools, then the cutoffs and the allocation under RBM converges to S_{DA} and the unique stable allocation in finitely many periods.*

[8] Examples where this property holds include Taiwanese high school admissions (Dur et al., 2022), Chicago Public schools (Dur, Pathak, and Sönmez, 2020) and Turkish high school admissions.

This proposition shows that the non-convergence is due to heterogeneity of scores between schools. This heterogeneity can create cycles where students remain pessimistic and do not apply to the schools that are attainable in the stable matching (Example 1) or remain optimistic and they keep applying to schools that are not attainable in the stable matching (Example 2). The proof of the proposition shows that once a student gets their stable matching in a period, they will be matched to that school in every other period and in each period, at least one additional student will be matched to their school under the stable matching. Therefore, when the scores are common across schools, such cycles cannot exist.

Sections 4 and 5 highlight the limitations of Theorem 1. First, if some students do not behave strategically, then RBM does not converge to SOSM. Second, students' initial beliefs are important and RBM may not converge to SOSM if they are initially pessimistic. Moreover, even if convergence eventually happens, it is not immediate and may take time. These results are in line with the empirical observations of non-equilibrium behavior in BM such as truthful reporting of preferences and overcautious strategies (Song et al., 2020).

6. Large Random Markets

Under RBM, the same market is repeated in every period. In this section, I analyze the robustness of the results to this assumption by building a large market model. First, I consider a continuum matching model based on Abdulkadiroğlu et al. (2015) and Azevedo and Leshno (2016). Second, I consider the setting where in each period, n students from the continuum market are randomly drawn and matched using BM. I show that the results about convergence of RBM to SOSM continue to hold in the first case and hold asymptotically (as $n \to \infty$) in the second case.

6.1. Continuum Markets

There are a finite set of schools, denoted by $\mathcal{C} = \{c_0, c_1, \ldots, c_n\}$ and a unit measure of students, where c_0 is a dummy school which denotes being unmatched. Let $\theta = (\succ_\theta, \{s_c^\theta\}_{c\in\mathcal{C}})$ denote the type of a student whose preferences over the set of schools is $\succ_\theta$ and has score $s_c^\theta \in [0, 1]$ in school c. The set of student types is denoted by Θ, over which there is a probability measure η, which admits a full support density.[9] $Q = (q_0, q_1 \ldots, q_n)$ denotes the capacities of schools. There are no capacity constraints for being unmatched, that is, $q_0 \geq 1$. A matching in this environment is a function $\mu : \mathcal{C} \cup \Theta \to 2^\Theta \cup \mathcal{C}$ where $\mu(\theta) \in \mathcal{C}$ is the school any type θ is assigned and $\mu(c) \subseteq \Theta$ is the set of students assigned to authority c such that

[9]Formally, the score distribution of students with each preference profile $\succ$ has a full support density.

no school is assigned to a measure of students larger than its capacity.[10]

A student-school pair (θ, c) *blocks* a matching μ if the student prefers c to its match and either school c does not fill its quota or school c is matched to another student who has a strictly lower score than θ. Formally, (θ, c) blocks μ if $c \succ_\theta \mu(\theta)$ and either $\eta(\mu(c)) < q_c$ or there exists $\theta' \in \mu(c)$ with $s_c^{\theta'} < s_c^{\theta}$. A matching μ is *stable* if it is not blocked by any student-school pair.

Stable matchings can be represented by cutoffs for n non-dummy schools.[11] For each $S \in [0,1]^n$, $\tilde{D}_\theta(S)$ denotes the *demand* of student θ at S, which is the most preferred school such that θ clears its threshold. $\tilde{D}_c(S)$ denotes the set of students who demand c. Given S, define the assignment $\mathcal{M}(S) = \nu$ induced by S as $\nu(\theta) = \tilde{D}_\theta(S)$. An assignment ν is a matching if $\eta(\tilde{D}_c(S)) \leq q_c$ for all c.

The definitions of the mechanisms are almost identical previous definitions, with the appropriate changes to adapt them to the continuum model. The formal definitions are provided in Appendix B. As η has full support, there is a unique stable matching in this market (Theorem 1 in Azevedo and Leshno (2016)). Moreover, this matching can be represented by a set of cutoffs (S_{DA}) where each student θ is matched to $\tilde{D}_\theta(S_{DA})$. As in Section 2, $\tilde{S}_t^c$ denotes the step t cutoffs in the MDA mechanism. The following proposition shows that MDA implements the stable matching.

Proposition 7. *$\tilde{S}_t$ converges to S_{DA}. The offers each school hold in step t of MDA converge to the unique stable matching as $t \to \infty$.*

The main difference of this result from the convergence results in finite markets is the fact that convergence may not happen in finitely many rounds when there is a continuum of students. Although the measure of students who are rejected in each step of MDA (and also, DA) converges to 0, it is possible that a positive measure of students are rejected in each step.[12]

Let R_0 denote the initial cutoffs that students best-respond in the initial period of RBM, while R_t denote the step 1 cutoffs in the period t of RBM. We can extend Proposition 5 and Theorem 2 to the continuum setting.

Proposition 8. *If R_0 is compatible with MDA, then RBM converges to SOSM.*

[10]The mathematical definition of a matching for the continuum market we study follows Azevedo and Leshno (2016) and requires that μ satisfies the following four properties: (i) for all $\theta \in \Theta$, $\mu(\theta) \in \mathcal{C}$; (ii) for all $c \in \mathcal{C}$, $\mu(c) \subseteq \Theta$ is measurable and $\eta(\mu(c)) \leq q_c$; (iii) $c = \mu(\theta)$ iff $\theta \in \mu(c)$; (iv) (open on the right) for any $c \in \mathcal{C}$, the set $\theta \in \Theta : c \succ_\theta \mu(\theta)$ is open.

[11]The cutoff for c_0 is always 0 as the outside option of being unmatched is always available.

[12]See Azevedo and Leshno (2016) for an example.

Theorem 3. *If students are optimistic enough at R_0, then allocation under RBM converges to the unique stable matching.*

These results serve two purposes. First, they show that the convergence properties of RBM are also true in continuum markets. Second, and more importantly, they are useful in proving the convergence in finite large markets drawn from a stationary distribution, showing the assumption that the repetition of the same market is not necessary in large markets.

6.2. Large Random Markets

To extend the results to large finite markets sampled from a continuum market, I first define finite markets in the continuum setting following Azevedo and Leshno (2016). A finite market $F = [\tilde{\Theta}, \tilde{Q}]$ specifies a finite set of students $\tilde{\Theta} \subset \Theta$ and an integer vector of capacities $\tilde{q}_c > 0$, where $\tilde{q}_{c_0} \geq |\tilde{\Theta}|$. A matching for a finite market is a function $\tilde{\mu} : \mathcal{C} \cup \Theta \to 2^{\Theta} \cup \mathcal{C}$ such that (i) for all $\theta \in \tilde{\Theta}$, $\tilde{\mu}(\theta) \in \mathcal{C}$, (ii) for all $c \in \mathcal{C}$, $\tilde{\mu}(c) \in 2^{\tilde{\Theta}}$ and $|\tilde{\mu}(c)| \geq \tilde{Q}_c$ and (iii) for all $\theta \in \tilde{\Theta}$ and $c \in \mathcal{C}$, $\tilde{\mu}(\theta) = c$ iff $\theta \in \tilde{\mu}(c)$. The definition of blocking pairs, as well as the definition of stability is the same as in Section 6.1. A finite market $F = [\tilde{\Theta}, \tilde{Q}]$ is associated with the following empirical distribution of types

$$\eta = \sum_{\theta \in \tilde{\Theta}} \frac{1}{|\tilde{\Theta}|} \delta_\theta$$

where δ_θ denotes the probability distribution that places probability one on the point θ. The supply of seats per student is given by $Q = \tilde{Q}/|\tilde{\Theta}|$. Either $[\tilde{\Theta}, \tilde{Q}]$ or $[\eta, Q]$ uniquely determine a discrete market F.[13] Fix a continuum market with full support, (η, Q). (η, Q) has a unique stable matching μ with cutoffs $S_{DA} \in [0, 1]^{\mathcal{C}}$. To remove any confusion, I use $\tilde{R}_t$ to denote the first step cutoffs of BM in period t. $\tilde{R}_0$ denotes the initial cutoffs.

I study the following repeated implementation of BM where k students are drawn from the continuum market (η, Q):

- In period 1, k students are independently and randomly drawn from η and the vector of capacities is $\tilde{Q} = Qk$. All students choose a strategy that is a best-response to the $\tilde{R}_0$. BM is used to determine the allocation and $\tilde{R}_1$ denotes the first step cutoffs in the BM.
- In period t, k students are independently and randomly drawn from η and the vector of capacities is $\tilde{Q} = Qk$. All students choose a strategy that is a best-response to the $\tilde{R}_{t-1}$. The BM is used to determine the allocation and $\tilde{R}_t$ denotes the first step cutoffs in the BM.

[13] To see why, note that $\tilde{\Theta} = \text{support}(\eta)$ and $\tilde{Q} = Q|\tilde{\Theta}|$.

For a given k, the period t cutoffs of RBM, $\tilde{R}_t(k)$ is a random variable distributed in $[0,1]^{\mathcal{C}}$ and its realization depends on the previous period's cutoffs as well as the random market drawn from the distribution. The following proposition shows that the behavior of RBM in random finite markets converges to its behavior in the continuum market the random markets are sampled from.

Proposition 9. *Suppose that $R_0 = \tilde{R}_0$. Then $\lim_{k\to\infty} \tilde{R}_t(k)$ converges in probability to R_t. If $\tilde{R}_0$ is compatible with MDA, then $\lim_{t\to\infty} \lim_{k\to\infty} \tilde{R}_t(k)$ converges in probability to S_{DA}.*

Given Proposition 9 and Theorem 3, we conclude that RBM converges to the unique stable matching in large random markets.

Theorem 4. *If students are optimistic enough, then the matching implemented in period t of RBM converges to the unique SOSM of the continuum market as the market grows large and $t \to \infty$.*

This theorem shows that repetition of the same market is not necessary for the result. In large markets with a stationary distribution of students and scores across periods, the the cutoffs and the allocation in period t of RBM converges to the cutoffs and tentative allocation in step t of MDA, which converges to the unique stable matching of the continuum market. Therefore, it is reasonable to expect that if BM is used in a market and students get can obtain information about previous periods, then the outcome converges to the student optimal stable matching.

7. Conclusion

This paper studies best-response dynamics under repeated application of the Boston Mechanism. When students best-respond to the admission cutoffs in the previous period (with the exception of the initial period, where they are truthful), the implemented matching converges to the student optimal stable matching, which is the Pareto-dominant equilibrium of the Boston Mechanism and dominant strategy outcome of the competing Deferred Acceptance Mechanism. This result provides a foundation for the equilibrium analysis of the Boston Mechanism based on best-response dynamics. Extending the model to include unsophsiticated students who always apply truthfully and allow for student's initial beliefs to determine their first period behavior, I show how the student optimal stable matching may not be reached if some students cannot strategize or are initially pessimistic about their admissions chances.

Appendices

A. Proofs for Discrete Markets

A.1. Proof of Lemma 1

First, a stable allocation exists and is implemented by DA by Gale and Shapley (1962). The proof follows the proof of stability of DA in Gale and Shapley (1962) with a minor modification for the possibility of skipping schools under MDA. A school c is *"possible"* for a student i if there exists a stable matching where i is matched to c. Assume that up to a given step of MDA, no student has been rejected by, or skips a school that is possible for him and let i denote the first student for which this happens. First, assume that i is rejected by or has skipped c. Then there exists set of students I_c who are accepted in that step and note that, all $i' \in I_c$ has higher scores than i. We will show that c is not a possible school for i. For a contradiction, assume that c is possible for i. Then there exists a stable matching μ such that $\mu(i) = c$. However, this means that there exists a student $i' \in I_c$ such that $c \succ_{i'} \mu(i')$, since i' cannot match any school it prefers to c as all schools i' is rejected by or skipped in the previous steps of MDA are impossible schools for i'. However, this contradicts that μ is a stable matching and therefore c is not possible for i. Since there are finitely many students and schools, and students never apply to same school twice after being rejected, both mechanisms terminate in finite time.

A.2. Proof of Lemma 2

If c' is not an achievable school at S, then in all σ that is compatible with S, there exists q_c students who have a higher score than $S^{c'}$ puts c' as their first choice. Therefore, under any σ_i, i is rejected by c' in any step it applied to it in Boston Mechanism. Moreover, since c is achievable, we have that $S^c \leq s(i, c)$, which means that there are at most $q_c - 1$ other students who rank c first and have higher scores than i. Therefore, i is accepted by c in step 1 of the Boston Mechanism. Since c is the most preferred school i can be matched under all σ_{-i} that is compatible with S, and any $\sigma_i \in FC_i(S)$ guarantees that i is matched to c, all such strategies are best-responses.

A.3. Proof of Proposition 1

The proof is by induction. First, observe that in the first step of MDA and first step of BM in the first round of RBM, all students apply to the highest ranked school in their preferences and all schools admit q_c highest scoring applicants. Therefore, for all schools, the set of students who apply and the set of students who are accepted are exactly the same. Now, assume that for all $k \leq \hat{k}$, in the kth step of MDA and in first step of BM in the kth round of RBM, the set of students who apply (for MDA, this also includes the students who are tentatively assigned) and the set of students who are accepted are exactly the same. Let S_k^c denote the score of the lowest scoring student who is tentatively assigned to school c at the end of kth step of MDA. We now show a useful lemma.

Lemma 3. *S_k^c is increasing in k.*

Proof. Note that all students who are tentatively assigned to c at round k applies to c in round $k+1$. Therefore, either all these students are still tentatively assigned to c at round $k+1$ or lowest scoring l students are replaced with higher scoring l students for some $l \geq 1$. In the first case, $S_k^c = S_{k+1}^c$ while in the second case, $S_k^c < S_{k+1}^c$, which proves the result. □

We will show that if student i either (1) is tentatively assigned to school c in step k or (2) applies to school c in step $k+1$ of MDA, then i applies to c in the first step of $k+1$th round of RBM. Assume $c' \succ_i c$. Then in a previous step of MDA (say, step l) i either was rejected by school c or skipped school c. In both cases, we have that $S_l^c > s(i, c')$. Then as $k > l$, by lemma 3, $s(i, c') \geq S_k^{c'}$, which means that any such c' that is preferred to c has a cutoff above $s(i, c')$. Moreover, since i is assign to c in step k of MDA or applies to school c in step $k+1$ of MDA, we have that $s(i, c) \geq S_k^c$, which proves that c is the most preferred achievable school at S_k. Therefore, i applies to c in the first step of $k+1$th round of RBM. Since this is true for all $i \in I$, the set of applicants to each school is same under the kth step of MDA and in first step of BM in the kth round of RBM. As a result, if i is tentatively accepted to c in $k+1$th step of MDA, it is accepted to c in the first step of BM in the kth round of RBM, which finishes the induction argument.

A.4. Proof of Proposition 2

I first define another market, ω', where if $i \in I_U$, then $\succ_i'$ has only the first choice of i as an acceptable school. The rest of the market is the same. I will compare two cases, the outcome of RBM under ω' where all students are sophisticated and the outcome of RBM under ω where students in I_U always apply truthfully.

Claim 1. *The first round outcome of RBM under ω' (where all students are sophisticated) and the first round outcome of RBM (with unsophisticated students) under ω is same in every period.*

Proof. The proof is by induction. First, note that as in both cases first period applications are truthful for all students and first choice of all students are same in both cases, the applications in the first round of first period are the same. This implies that the set of admitted students, and first round cutoffs are the same.

To prove the induction step, suppose that first round cutoffs in period t in both settings are the same. Then all sophisticated students ($i \in I_S$) apply to the same school in both cases. Moreover, all unsophisticated students ($i \in I_U$) apply to the same school in both cases as the top school for i under $\succ_i$ is the only acceptable school for i under $\succ'_i$. Thus, the set of admitted students are the same in each school and the realized first round cutoffs are also same. Therefore, in each period, the outcome of both procedures are the same. □

As by Theorem 1, the first round outcome of RBM under ω' converges in finitely many periods, the first round outcome of RBM under ω with unsophisticated students also converges to the same outcome in the same period. Let $\hat{\mu}$ denote the matching that is realized in the first round after convergence, which is the student optimal stable matching under ω', by Theorem 1. Moreover, given this, in each period after convergence of the first round, the remaining rounds of the Boston Mechanism is the same and yields the same matching, which I denote by μ^*. Let $\hat{I}$ denote the set of all sophisticated students and unsophisticated students who are matched to a school in $\hat{\mu}$. Note that if $i \in \hat{I}$, then, $\hat{\mu}(i) = \mu^*(i)$. Let μ denote the student optimal stable matching under $\tilde{\omega}$.

Claim 2. *All students in $\hat{I}$ are weakly better off under $\hat{m}u$ compared to μ.*

Proof. If $i \in I_S$, then the result follows as under $\tilde{\omega}$, all students in I_U extend preferences. If $i \in I_U$, they are already matched to their top choice in $\hat{\mu}$, and therefore is weakly worse off under μ. □

Thus, if $i \in \hat{I}$, then i is better off under μ^* compared to μ.

Claim 3. *μ^* is stable under the market $\tilde{\omega}$.*

Proof. Assume for a contradiction i blocks μ^* at c under $\tilde{\omega}$. First, this means that c is acceptable to i. Then, c fills its quota as otherwise, i would either apply to c and get admitted or would be admitted to a more preferred school, contradicting that i blocks μ^* at c. Therefore, there exists $i' \in \mu^*(c)$ such that $i \succ_c i'$.

First, suppose that, $i \in \hat{I}$. Then there are two cases, (i) $i' \in \hat{I}$ or (ii) $i' \notin \hat{I}$. Under (i), as if $i \in \hat{I}$, then, $\hat{\mu}(i) = \mu^*(i)$, this blocking pair also blocks $\hat{\mu}$, which is a contradiction. Under (ii), since $i' \notin \hat{I}$, c does not fill its quota in round 1, which would contradict convergence to μ^* as $c \succ_i \mu^*(i)$.

Second, suppose that, $i \notin \hat{I}$. This means that $i' \notin \hat{I}$ and either (i) i' ranks c lower than i or (i) i' ranks c at the same spot as i and has higher score in c. In both cases, i would be accepted to c or a more preferred school if i' is accepted to c, which is a contradiction. $\square$

As μ is the student optimal stable matching under $\tilde{\omega}$, from Claim 3, all students are weakly better off under μ compared to μ^*. Then Claim 2 imply that all students in $\hat{I}$ are matched to exactly same schools in μ and μ^*. The following claim completes the proof.

Claim 4. *If $i \notin \hat{I}$, then $\mu^*(i) = \mu(i)$.*

Proof. Let $\tilde{I}$ denote the set of students such that $\mu^*(i) \neq \mu(i)$. Note that any $i \in \tilde{I}$ applies to $\mu(i)$ and is rejected during the implementation of the Boston Mechanism in periods after convergence is achieved.

Let i denote a student that is rejected from $\mu(i)$ in the earliest round k during the implementation of the Boston Mechanism in periods after convergence is achieved (if there are multiple students rejected in this earliest round, i can be any of those students).

Since i is rejected, there must be another student, i' such that $\mu^*(i') \neq \mu(i')$ and i' is accepted to $\mu(i)$ before or at round k. Also note that $\mu^*(i') \neq \mu(i')$ implies that $i \notin \hat{I}$, and therefore is an unsophisticated student. Moreover, as $\mu^*(i') \neq \mu(i')$, and i' has not been rejected from $\mu(i')$ yet, which means that i' prefers $\mu^*(i')$ to $\mu(i')$, which contradicts that μ is the student optimal stable matching and completes the proof of the claim. $\square$

A.5. Preliminaries for Discrete Markets

First, I will define some notation. Let $\tilde{S}_t$ denote the step t cutoffs under MDA. Let $B_i(S)$ denote the budget set of student i under cutoffs i. Formally, $B_i(S) = \{c \in \mathcal{C} : s_i(c) \geq S^c\}$. $D_i(S)$ denotes the demanded school of student i, which is the P_i-maximal school in $B_i(S)$.[14] $D_c(S)$ denotes the demand set of a school, which is the set of students who demand that school under S. Formally, $D_c(S) = \{i : D_i(S) = c\}$. Let $\mathcal{U}_c(\hat{s})$ denote the set of students who score higher than $\hat{s}$ in school c, $\mathcal{U}_c(\hat{s}) = \{i : s_i(c) \geq \hat{s}\}$. Moreover, let $T_c(S)$ denote the first step cutoffs under Boston Mechanism when students apply according to S. Formally,

$$T_c(S) = \begin{cases} 0 & \text{if} \quad |D_c(S)| < q_c \\ \min_{\{s : \exists i \in D_c(S) \text{ s.t. } s_c(i) = s\}} |D_c(S) \cap \mathcal{U}_c(s)| = q_c & \text{if} \quad |D_c(S)| \geq q_c \end{cases} \tag{2}$$

Moreover, define $T = \prod_{j=1}^{n} T_j$, where $T : [0, 1] \to [0, 1]$.

Claim 5. *$D_c(S^c, S^{-c})$ is increasing (in set inclusion sense) in S^{-c} and decreasing in S^c.*

[14] If $B_i(S) = \emptyset$, then $D_i(S)$ is the most preferred school of i.

Proof. To prove the first part, take $i \in D_c(S^c, S^{-c})$. Let $\hat{S}^{-c} \geq S^{-c}$. Then $B_i(S^c, \hat{S}^{-c}) \subseteq B_i(S^c, S^{-c})$. As $i \in D_c(S)$, c is maximal in $B_i(S)$, which implies that c is maximal in $B_i(S^c, \hat{S}^{-c})$. Thus, $i \in D_c(S^c, \hat{S}^{-c})$, proving the first part.

To prove the second part, take $i \in D_c(S^c, S^{-c})$. Let $\hat{S}^c \leq S^c$. $B_i(\hat{S}^c, S^{-c}) \setminus \{c\} = B_i(S^c, S^{-c}) \setminus \{c\}$ as the cutoffs of all other schools are the same. As $i \in D_c(S^c, S^{-c})$ and $\hat{S}^c \leq S^c$, we have that $c \in B_i(\hat{S}^c, S^{-c})$, proving the second part. □

Claim 6. *$T_c(S)$ is increasing in S^{-c}, decreasing in S^c.*

Proof. Suppose that $\hat{S}^{-c} \geq S^{-c}$. From Claim 5, $D_c(S^c, S^{-c}) \subseteq D_c(S^c, \hat{S}^{-c})$, which implies that $D_c(S^c, S^{-c}) \cap \mathcal{U}(S^c) \subseteq D_c(S^c, \hat{S}^{-c}) \cap \mathcal{U}(S^c)$. The first part of the result then follows from the definition of T_c. Similarly, suppose that $\hat{S}^c \leq S^c$. From Claim 5, $D_c(S^c, S^{-c}) \subseteq D_c(\hat{S}^c, S^{-c})$, which implies that $D_c(\hat{S}^c, S^{-c}) \cap \mathcal{U}(\hat{S}^c) \subseteq D_c(S^c, S^{-c}) \cap \mathcal{U}(S^c)$. The second part of the result then follows from the definition of T_c. □

Lemma 4. *S_{DA} is a fixed point of T.*

Proof. Let μ denote the student optimal stable matching. Note that as S_{DA} is the stable matching cutoff, if $\mu(i) = c$, then $i \in D_c(S_{DA})$. To see why, suppose that, for a contradiction, $i \in D_{c'}(S_{DA})$. If $s_{c'}(i) \geq S_{DA}^{c'}$, then i and c' block μ, which is a contradiction. If $s_{c'}(i) < S_{DA}^{c'}$, this means that c is not acceptable to i, which contradicts the stability of μ.

Next, suppose that c fills its quota in the student optimal stable matching and let j denote the lowest scoring student who is matched to c at μ. Note that $D_c(S_{DA} \cap \mathcal{U}_c(S_{DA})) = \mu(c)$. If $T_c(S_{DA}) < S_{DA}^c$, then from definition of T_c, there exists a student $j' \in D_c(S_{DA})$ such that $s_c(j) = T_c(S_{DA})$. However, this is a contradiction as $|D_{S_{DA} \cap \mathcal{U}_c(T_c(S_{DA})}| > q_c$, as it includes j' and all q_c students who are matched to c at μ. Conversely, if $T_c(S_{DA}) > S_{DA}^c$, then $|D_c(S_{DA} \cap \mathcal{U}_c(T_c(S_{DA}))| < q_c$ as $j \notin U_c(S_{DA})$, which is a contradiction.

Finally suppose that c does not fill its quota in the student optimal stable matching, which means that $S_{DA}^c = 0$. Moreover, from stability of μ, $|D_c(S_{DA})| < q_c$, which means that $T_c(S_{DA}) = 0$, proving the result. □

A.6. *Proof of Proposition 3*

Fix an $\hat{S}_t$ such that $\hat{S}_t \leq S_{DA}$. Define S_t by setting $S_t^c = S_{DA}^c$ and $S_t^{-c} = \hat{S}_t^{-c}$. Let $\mathcal{D}(S) = \{i : i \in D_c(S_t)\, s_c(i) \geq S_{DA}^c\}$. From Lemma 4, $T_c(S_{DA}) = S_{DA}^c$. Then by Claim 6, we have $T_c(S) \leq S_{DA}^c$. Therefore, there are at most q_c students who has scores above S_{DA}^c demand c, that is, $|\mathcal{D}(S_t)| \leq q_c$. Moreover, if $i \in \mathcal{D}(\hat{S}_t)$ but $i \notin D(S_t)$, then $s_c(i) \leq S_t^c \leq S_{DA}^c$. Thus, $|\mathcal{D}(\hat{S}_t)| \leq q_c$, which implies $T_c(\hat{S}_t) \leq S_{DA}$, proving the result.

A.7. *Proof of Proposition 4*

Suppose that μ is in a cycle and is Pareto dominated by the SOSM, denoted by μ^*. Suppose that μ appears in round $t+1$, that is, following the cutoffs S_t. First, note that $S_t \leq S_{DA}$. Therefore, in the first step of BM in period t of RBM, all students apply to a school that they weakly prefer to their match in the μ^*. Thus, at the end of first step, all students who are matched to a school are weakly better off compared to μ^*. As μ^* Pareto dominates μ, all such students are matched to the school they match at μ^*.

Suppose that k is the first step in the period t of RBM such that a student is matched to a school which is less preferred than her match under μ^*, or does not have any school to apply even though she is matched to a school at μ^*. Denote this student by i. Note that this is only possible if $\mu^*(i)$ has exhausted its capacity in the previous step. This means that there exists a student i' such that $\mu^*(i') \neq \mu^*(i)$, but i' is matched to $\mu^*(i)$ in a previous step. However, as no student has received a match that is less preferred to their match under μ^* at any previous step, i' must be strictly better off under μ compare to μ^*, which is a contradiction.

A.8. *Proof of Proposition 5*

I first prove a few useful results.

Claim 7. $T_c(\tilde{S}_t) = \tilde{S}^c_{t+1}$

Proof. Immediate from the fact that the set of students who apply to c in step $t+1$ of MDA is $D_c(S)$. □

Claim 8. *If $T_c(S^c, S^{-c}) = \hat{S}^c$ and $\hat{S}^c \geq S^c$, then $T_c(s, S^{-c}) = \hat{S}^c$ for all $s \leq \hat{S}^c$.*

Proof. If $\hat{S}^c = 0$, then the result is immediate as $s = \hat{S}^c$. Suppose that $\hat{S}^c > 0$. Then $|D_c(S^c, S^{-c}) \cap \mathcal{U}(\hat{S}^c)| = q_c$. Then, $T_c(\hat{S}^c, S^{-c}) = \hat{S}^c$. Thus, from Claim 6, $T_c(s, S^{-c}) \leq \hat{S}^c$ for all $s \leq \hat{S}^c$. Moreover, observe that for all $s \leq \hat{S}^c$, $D_c(\hat{S}^c, S^{-c}) \cap \mathcal{U}(\hat{S}^c) = D_c(s, S^{-c}) \cap \mathcal{U}(\hat{S}^c)$, which means that there are q_c students who demand c at cutoffs s, S^{-c} with scores over $\hat{S}^c$. Therefore, $T_c(s, S^{-c}) = T_c(\hat{s}, S^{-c})$. □

Now, assume that $S_k \in [\tilde{S}_t, \tilde{S}_{t+1}]$. We have the following claim.

Claim 9. $T_c(S_k) \leq \tilde{S}^c_{t+2}$.

Proof. From Claim 7, $T_c(\tilde{S}^c_{t+1}, \tilde{S}^{-c}_{t+1}) = \tilde{S}^c_{t+2}$. From Claim 8, as $\tilde{S}^c_{t+2} \geq \tilde{S}^c_{t+1}$, we have that $T_c(S^c_k, S^{-c}_k) = T_c(\tilde{S}^c_{t+1}, \tilde{S}^{-c}_{t+1})$. Moreover, as $S^{-c}_k \leq \tilde{S}^{-c}_{t+1}$, from Claim 6, we have that

$$T_c(S^c_k, S^{-c}_k) \leq T_c(S^c_k, \tilde{S}^{-c}_{t+1}) = T_c(\tilde{S}_{t+1}) = \tilde{S}^c_{t+2} \quad (3)$$

□

Claim 10. $T_c(S_k) \geq \tilde{S}^c_{t+1}$.

Proof. From Claim 7, $T_c(\tilde{S}^c_t, \tilde{S}^{-c}_t) = \tilde{S}^c_{t+1}$. From Claim 8, as $\tilde{S}^c_{t+1} \geq S^c_k$, $T_c(S^c_k, \tilde{S}^{-c}_t) = \tilde{S}^c_{t+1}$. Moreover, as $S^{-c}_k \geq \tilde{S}^{-c}_t$, from Claim 6, we have that

$$T_c(S^c_k, S^{-c}_k) \geq T_c(S^c_k, \tilde{S}^{-c}_t) = \tilde{S}^c_{t+1} \tag{4}$$

The following Lemma is immediate from Claims 9 and 10.

Lemma 5. *If S_k is compatible with step t cutoffs of MDA, then S_{k+1} is compatible with step $t+1$ cutoffs of MDA*

Given the lemma, the proposition follows from the fact that MDA converges to the student optimal stable matching in finitely many steps (Proposition 1 and Theorem 1). □

A.9. Proof of Proposition 6

The proof is by induction. First, note that in the first period, regardless of S_0, the highest scoring student applies to her most preferred school and is accepted. Note that this school is the student's assignment under the serial dictatorship (and therefore, the deferred acceptance) mechanism. Let μ_S denote the serial dictatorship outcome.

Let I_t denote highest scoring t students. Now suppose that at period t, all students in I_t are assigned to their match under the serial dictatorship mechanism at the first round of the Boston Mechanism. I will show that in period $t+1$, all students in I_{t+1} are assigned to their match under the serial dictatorship mechanism at the first round of the Boston Mechanism. First, note that in the first round of period $t+1$, if $i \in I_t$, then i applies to $\mu_S(i)$ as $\mu_S(i)$ is attainable under S_t and all schools more preferred to $\mu_S(i)$ are not attainable under S_t. Moreover, all such i are admitted.

Next, let j denote the student with $t+1$th highest score. As $c \equiv \mu_S(j)$ is the school of j under serial dictatorship and all higher scoring students are assigned to their serial dictatorship outcome at round t, $S^c_t \leq s(i)$. Moreover, for any school c' that j prefers to c, $S^{c'}_t > s(i)$ as otherwise, i would be matched to c' under serial dictatorship. Therefore, j applies to c. Moreover, as all students in I_t are assigned to their match under serial dictatorship in the first round of period $t+1$, j is admitted to c in the first round of period $t+1$ as otherwise, j must be rejected in favor of a lower scoring student, which is a contradiction. This proves the result as the set of students are finite.

B. Proofs for Large Markets

B.1. Preliminaries for Continuum Markets

The Boston Mechanism (BM) - Continuum Market
Step 1: Students apply to their first choice school. Each school c admits all students over some cutoff S_c, where S_c is the infimum over all cutoffs where the school c does not exceed its capacity. All other offers are immediately accepted and become permanent matches. School capacities are adjusted accordingly.
Step t: If a student is rejected in Step $t-1$, he applies to the next school on his rank-order list. If he has no more schools on his list, he applies nowhere. Each school c admits all students over some cutoff S_c, where S_c is the infimum over all cutoffs where the school c does not exceed its capacity. All other offers become permanent matches. School capacities are adjusted accordingly. **Stop**: The algorithm stops when no rejections are issued.

Note that Boston Mechanism terminates in at most $|\mathcal{C}|$ rounds, as all students are either permanently matched or have run out of schools to apply.

The Modified Deferred Acceptance Mechanism (MDA) - Continuum Market
Step 1: Students apply to their first choice school. Each school c tentatively admits all students over some cutoff S_c, where S_c is the infimum over all cutoffs where the school c does not exceed its capacity. Schools reject all students who are not tentatively accepted.
Step t: If a student is rejected in Step $t-1$, he *applies to the highest school on his rank-order list within the schools that either (i) did not fill its capacity in the last round or (ii) temporarily hold the offer of a student who has lower ranking in that school.* If there are no schools in his list that satisfies either (i) or (ii), or that school is ranked below the outside option, he applies nowhere. Schools consider both new offers and the offers held from previous rounds and tentatively admits all students over some cutoff S_c, where S_c is the infimum over all cutoffs where the school c does not exceed its capacity. Schools reject all students who are not tentatively accepted.
Stop: The algorithm stops when no rejections are issued. Each school is matched to the students it is holding at the end.

Let $\tilde{S}_n$ denote the cutoffs in step n of the MDA mechanism. I now define a useful function that maps a distribution of types, capacities and cutoffs to a new set of cutoffs after one round of admissions. As in the finite market model, $B_\theta(S)$ denotes the budget set of student θ while $D_\theta(S)$ denote the demand of student θ. $D_c(S) = \{\theta : D_\theta(S) = c\}$ is the demand set of school c and $\mathcal{U}_c(s)$ return the students with scores higher than s at school c.

$$H_c(S) = \min_{\hat{s}\in[0,1]} |D_c(S) \cap \mathcal{U}_c(\hat{s})| \leq q_c \tag{5}$$

where the minimum exists as η has full support. Let $H(S) = \{H_{c_1}(S), \ldots, H_{c_n}(S)\}$. The following claim is immediate from the definition of MDA.

Claim 11. *$H^n(0, \ldots, 0) = \tilde{S}_n$. That is, starting with zero cutoffs in all schools and applying T n times gives the cutoffs in step n of MDA.*

Let $\tilde{\Theta}^\epsilon_S = \{\theta : s_c(\theta) \in \mathcal{B}_\epsilon(S_c)$ for some $c\}$, where $\mathcal{B}_\epsilon(S_c)$ is the ϵ neighborhood around S_c. Let d denote the euclidean distance. The following lemma is useful in showing the continuity of T.

Claim 12. *For each $\theta \notin \tilde{\Theta}^0_S$, if $d(S, \hat{S}) < \min_c |s_c(\theta) - S_c|$, then $D_\theta(\hat{S}) = D_\theta(S)$.*

Proof. As $\theta \notin \tilde{\Theta}^0_S$, $\hat{\delta} \equiv \min_c |s_c(\theta) - S_c| > 0$. Moreover, for all $\hat{S}$ such that $d(S, \hat{S}) < \hat{\delta}$, $B_\theta(\hat{S}) = B_\theta(S)$. As students demand the same school when their budget set is the the same, $D_\theta(\hat{S}) = D_\theta(S)$, proving the claim. □

Claim 13. *For each $\epsilon > 0$, there exists α_ϵ such that $\eta(\tilde{\Theta}^\alpha_S) < \epsilon$ for all $\alpha \leq \alpha_\epsilon$.*

Proof. Immediate from the fact that η has no mass points and admits a density. □

Lemma 6. *$H_c(S)$ is continuous in S.*

Proof. Let $DD(S, \hat{S}) = \{\theta : D_\theta(S) \neq D_\theta(\hat{S})\}$ denote the set of students whose demanded school is different under S and $\hat{S}$. Note that by Claim 12, $DD(S, \hat{S}) \subseteq \tilde{\Theta}^\alpha_S$ for $\alpha = d(S, \hat{S})$.

Claim 14. *Fix S and let $\bar{\eta} > 0$. There exists $\delta_{\bar{\eta}}$ such that if $d(S, \hat{S}) < \delta_{\bar{\eta}}$, then $\eta(DD(S, \hat{S})) < \bar{\eta}$.*

Proof. From Claim 13, there exists $\alpha_{\bar{\eta}}$ such that $\eta(\tilde{\Theta}^{\alpha_{\bar{\eta}}}_s) < \bar{\eta}$ for all $\alpha < \alpha_{\bar{\eta}}$. From Claim 12, for all $\theta \notin \tilde{\Theta}^{\alpha_{\bar{\eta}}}_S$, $D_\theta(S) = D_\theta(\hat{S})$ for all $\hat{S}$ with $d(S, \hat{S}) < \alpha_{\bar{\eta}}$. As $DD(S, \hat{S}) \subseteq \tilde{\Theta}^{\alpha_{\bar{\eta}}}_S$ for all $\hat{S}$ with $d(S, \hat{S}) < \alpha_{\bar{\eta}}$ and $\eta(\tilde{\Theta}^{\alpha_{\bar{\eta}}}_S) < \bar{\eta}$, we have that $\eta(DD(S, \hat{S})) < \bar{\eta}$ for all $\hat{S}$ with $d(S, \hat{S}) < \alpha_{\bar{\eta}}$, proving the claim. □

To prove the lemma, let $\epsilon > 0$ be given and fix S. Define

$$\Theta^{\epsilon\downarrow} = \{\theta : s_c(\theta) \in (H_c(S) - \epsilon, H_c(S)), c \succ_\theta c_0 \succ_\theta c' \text{ for all } c' \notin \{c, c_0\}\} \tag{6}$$

$$\Theta^{\epsilon\uparrow} = \{\theta : s_c(\theta) \in (H_c(S) + \epsilon/2, H_c(S) + \epsilon), c \succ_\theta c' \text{ for all } c' \neq c\} \tag{7}$$

Let $\bar{\eta} = \min\{\eta(\Theta^{\epsilon\uparrow}), \eta(\Theta^{\epsilon\downarrow})\}$. By Claim 14, there exists $\delta_{\bar{\eta}}$ such that if $d(S, \hat{S}) < \delta_{\bar{\eta}}$, then $\eta(DD(S, \hat{S})) < \bar{\eta}$. Moreover, if $d(S, \hat{S}) < \min\{\epsilon/2, \delta_{\bar{\eta}}\}$, for $\theta \in \Theta^{\epsilon\uparrow} \cup \Theta^{\epsilon\downarrow}$, $D_\theta(S) = D_\theta(\hat{S}) = c$.

I will now show that if $d(S, \hat{S}) < \min\{\epsilon/2, \delta_{\bar{n}}\}$, then $|H_c(\hat{S}) - H_c(S)| < \epsilon$. There are three cases, $H_c(\hat{S}) = H_c(S)$, $H_c(\hat{S}) > H_c(S)$ and $H_c(\hat{S}) < H_c(S)$. If $H_c(\hat{S}) = H_c(S)$, then we are done.

Suppose that $H_c(\hat{S}) > H_c(S)$. Let $\Theta_N = \{\theta : D_\theta(\hat{S}) = c, D_\theta(S) \neq c\}$ denote the set of students who demand c under $\hat{S}$ but not under S. As $d(S, \hat{S}) < \delta_{\bar{\eta}}$, $DD(S, \hat{S}) < \bar{\eta}$. Therefore, $\eta(\Theta_N) < \bar{\eta}$. But this means that, there can be at most measure $\bar{\eta}$ new students who demand c under $\hat{S}$. As all $\theta \in \Theta^{\epsilon\uparrow}$ still demand c under $\hat{S}$ and $\eta(\Theta^{\epsilon\uparrow}) = \bar{\eta}$, $H_c(\hat{S}) - H_c(S) < \epsilon$.

Next, suppose that $H_c(\hat{S}) < H_c(S)$. Note that $\Theta^{\epsilon\downarrow}$ has positive measure. Let $\Theta_O = \{\theta : D_\theta(S) = c, D_\theta(\hat{S}) \neq c\}$ denote the set of students who demand c under S but not under $\hat{S}$. As $d(S, \hat{S}) < \delta_{\bar{\eta}}$, $DD(S, \hat{S}) < \bar{\eta}$. Therefore, $\eta(\Theta_O) < \bar{\eta}$. But this means that, there can be at most measure $\bar{\eta}$ students who demand c under S but not under $\hat{S}$. As all $\theta \in \Theta^{\epsilon\downarrow}$ still demand c under $\hat{S}$ and $\eta(\Theta^{\epsilon\uparrow}) = \bar{\eta}$, $H_c(S) - H_c(\hat{S}) < \epsilon$.

This shows that for each ϵ, we can find $\alpha = \min\{\bar{\eta}, \epsilon/2\}$ such that $|H_c(S) - H_c(\hat{S})| < \epsilon$. Thus, for each ϵ, we can find a δ such that $d(T(S), T(\hat{S})) < \epsilon$ whenever $d(S, \hat{S}) < \delta$.

□

Lemma 7. *S_{DA} is the unique fixed point of H.*

Proof. I start by showing that, if S is a fixed point of H, then $\nu = \mathcal{M}(S)$ is a stable matching. First, I show that ν is a matching. To see why, from definition of H_c, for each c there is a measure of q_c students with $D_\theta(S) = c$ and $s_c(\theta) \geq S_c$. Therefore, there is a measure of q_c students with $\tilde{D}_\theta(S) = c$ and $\eta(\nu(c)) \leq q_c$ for all c. Assume for a contradiction (θ, c) blocks ν. From definition of ν, $\tilde{D}_\theta(S) = \nu(\theta)$. Moreover as (θ, c) blocks ν, we have $c \succ_\theta \nu(\theta)$ and $s_c(\theta) \geq S_c$. However, this implies that $\tilde{D}_\theta(S) \neq \nu(\theta)$, which is a contradiction.

As full support assumption implies that the market has a unique stable matching, there is exactly one fixed point of H, which corresponds to the stable matching. □

B.2. Proof of Proposition 7

First, note that $\tilde{S}_1 \geq \tilde{S}_0 = (0, \ldots, 0)$.

Lemma 8. *$\tilde{S}_n$ is increasing. That is, $\tilde{S}^c_{n+1} \geq \tilde{S}^c_n$ for all c, n.*

Proof. Proof is by induction. The base case holds as $\tilde{S}_1 \geq \tilde{S}_0$. Suppose that $\tilde{S}_n \geq \tilde{S}_{n-1}$. Fix a $c \in \mathcal{C}$. If $\tilde{S}^c_n = 0$, then $\tilde{S}^c_{n+1} \geq \tilde{S}^c_n = 0$ and we are done. Therefore, suppose that $\tilde{S}^c_n > 0$. Define

$$\Theta^c_n = \{\theta : D_\theta(\tilde{S}_{n-1}) = c, s_c(\theta) \geq \tilde{S}^c_n\} \tag{8}$$

As $\tilde{S}^c_n > 0$, $\eta(\Theta^c_n) = q_c$. Moreover, as for all c $\tilde{S}^c_n \geq \tilde{S}^c_{n-1}$ and $s_c(\theta) \geq \tilde{S}^c_n$, for all $\theta \in \Theta^c_n$, we have that $D_\theta(\tilde{S}_n) = c$ for all $\theta \in \Theta^c_n$. As $\eta(\Theta^c_n) = q_c$, $\tilde{S}^c_{n+1} = H_c(\tilde{S}_n) \geq \tilde{S}_n$. Repeating this for all c proves the lemma. □

Given Lemma 8, as $\tilde{S}_n$ is bounded, $\lim_{n\to\infty} \tilde{S}_n = S^*$ for some S^*. From continuity of H (Lemma 6) and the fact that $H^n(0, \ldots, 0) = \tilde{S}_n$, S^* is a fixed point of H. As H has a unique fixed point, which is S_{DA} (Lemma 7), the result follows.

B.3. Proof of Proposition 8

First, Claims 5, 6, 8, 9 and 10 as well as Lemma 5 hold in the continuum model, with essentially the same proofs, replacing the function T with H. The result then follows from Proposition 7.

B.4. Preliminaries for Finite Markets Sampled From a Continuum Market

A sequence of finite markets $\{F^k\}_{k\in\mathbb{N}}$ where $F^k = [\eta^k, Q^k]$ *converges* to a continuum market $F = [\eta, Q]$ if the empirical distribution of types η^k converges to η in the weak sense and if capacity per student Q^k converges to Q.

B.5. Proof of Proposition 9

The proof is by induction. Suppose that $\lim_{k\to\infty} \tilde{R}_t(k) \xrightarrow{p} R_t$. I will show that $\lim_{k\to\infty} \tilde{R}_{t+1}(k) \xrightarrow{p} R_{t+1}$, which amounts to showing, for all ϵ,

$$\lim_{k\to\infty} Pr\left(|\tilde{R}_{t+1}(k) - R_{t+1}| > \epsilon\right) = 0 \tag{9}$$

Let F^k_t denote the distribution of $\tilde{R}_t(k)$, while F_t denotes the (degenerate) distribution of R_t. Then the following must be shown:

$$\lim_{k\to\infty} Pr\left(\left|\int T(r,k)dF^k_t(r) - H(R_t)\right| > \epsilon\right) = 0 \tag{10}$$

Claim 15. *The following is true*

$$\lim_{k\to\infty} Pr\left(|H(\tilde{R}_t(k)) - H(R_t)| > \epsilon/2\right) = 0 \tag{11}$$

Proof. Follows from continous mapping theorem given the continuity of H and the assumption that $\lim_{k\to\infty} \tilde{R}_t(k) \xrightarrow{p} R_t$. □

Claim 16. *The following is true*

$$\lim_{k\to\infty} Pr\left(\left|\int T(r,k)dF_t^k(r) - H(\tilde{R}_t(k))\right| > \epsilon/2\right) = 0 \tag{12}$$

Proof. We can rewrite $H(\tilde{R}_t(k))$ as

$$H(\tilde{R}_t(k)) = \int H(r)dF_t^k(r) \tag{13}$$

Thus, Equation 12 becomes

$$\lim_{k\to\infty} Pr\left(\left|\int T(r,k) - H(r)dF_t^k(r)\right| > \epsilon/2\right) = 0 \tag{14}$$

As $T(r,k)$ converges to $H(r)$ pointwise, F_t^k converges to F_t and both $T(r,k)$ and F_t^k are bounded for all k, by dominated convergence theorem,

$$\lim_{k\to\infty} \int |T(r,k) - H(r)|\, dF_t^k(r) = 0 \tag{15}$$

which implies equation 14 and proves the claim. □

Taken together, claims 15 and 16 imply equation 10 and finishes the proof of the inductive step. To prove the base case of induction, we prove the following claim, which finishes the proof of the proposition.

Claim 17. $\lim_{k\to\infty} \tilde{R}_1(k) \xrightarrow{p} R_1$.

Proof. Since $\tilde{R}_0(k) = R_0$, we have that $\lim_{k\to\infty} \tilde{R}_0(k) \xrightarrow{p} R_0$. The result then follows from the proof we had for the inductive step. □

References

Abdulkadiroğlu, A., Che, Y.-K., Yasuda, Y., 2011. Resolving conflicting preferences in school choice: The" boston mechanism" reconsidered. American Economic Review 101, 399–410.

Abdulkadiroğlu, A., Che, Y.-K., Yasuda, Y., 2015. Expanding" choice" in school choice. American Economic Journal: Microeconomics 7, 1–42.

Abdulkadiroğlu, A., Sönmez, T., 2003. School choice: A mechanism design approach. American economic review 93, 729–747.

Agarwal, N., Somaini, P., 2018. Demand analysis using strategic reports: An application to a school choice mechanism. Econometrica 86, 391–444.

Akbarpour, M., Kapor, A., Neilson, C., van Dijk, W., Zimmerman, S., 2022. Centralized school choice with unequal outside options. Journal of Public Economics 210, 104644.

Azevedo, E. M., Leshno, J. D., 2016. A supply and demand framework for two-sided matching markets. Journal of Political Economy 124, 1235–1268.

Babaioff, M., Gonczarowski, Y. A., Romm, A., 2018. Playing on a level field: sincere and sophisticated players in the boston mechanism with a coarse priority structure. arXiv preprint arXiv:1810.06229 .

Calsamiglia, C., Fu, C., Güell, M., 2020. Structural estimation of a model of school choices: The boston mechanism versus its alternatives. Journal of Political Economy 128, 642–680.

Chen, Y., Kesten, O., 2017. Chinese college admissions and school choice reforms: A theoretical analysis. Journal of Political Economy 125, 99–139.

Dubins, L. E., Freedman, D. A., 1981. Machiavelli and the gale-shapley algorithm. The American Mathematical Monthly 88, 485–494.

Dur, U., Hammond, R. G., Kesten, O., 2021. Sequential school choice: Theory and evidence from the field and lab. Journal of Economic Theory 198, 105344.

Dur, U., Pathak, P. A., Song, F., Sönmez, T., 2022. Deduction dilemmas: The taiwan assignment mechanism. American Economic Journal: Microeconomics 14, 164–85.

Dur, U., Pathak, P. A., Sönmez, T., 2020. Explicit vs. statistical targeting in affirmative action: Theory and evidence from chicago's exam schools. Journal of Economic Theory 187, 104996.

Ergin, H., Sönmez, T., 2006. Games of school choice under the boston mechanism. Journal of public Economics 90, 215–237.

Gale, D., Shapley, L. S., 1962. College admissions and the stability of marriage. The American Mathematical Monthly 69, 9–15.

Hastings, J., Kane, T. J., Staiger, D. O., 2009. Heterogeneous preferences and the efficacy of public school choice. NBER Working Paper 2145, 1–46.

He, Y., 2015. Gaming the boston school choice mechanism in beijing .

Kapor, A. J., Neilson, C. A., Zimmerman, S. D., 2020. Heterogeneous beliefs and school choice mechanisms. American Economic Review 110, 1274–1315.

Pathak, P. A., Sönmez, T., 2008. Leveling the playing field: Sincere and sophisticated players in the boston mechanism. American Economic Review 98, 1636–52.

Roth, A. E., 1982. The economics of matching: Stability and incentives. Mathematics of operations research 7, 617–628.

Song, Y., Tomoeda, K., Xia, X., 2020. Sophistication and cautiousness in college applications. Available at SSRN 3711118 .