Promoting School Competition Through School Choice: A Market Design Approach*

John William Hatfield
Graduate School of Business
Stanford University

Fuhito Kojima
Department of Economics
Stanford University

Yusuke Narita
Department of Economics
MIT

November 18, 2011

Abstract

We study the effect of different school choice mechanisms on schools' incentives for quality improvement. To do so, we introduce the following criterion: A mechanism respects improvements of school quality if each school becomes weakly better off whenever that school becomes more preferred by students. We first show that no stable mechanism, or mechanism that is Pareto efficient for students (such as the Boston and top trading cycles mechanisms), respects improvements of school quality. Nevertheless, for large school districts, we demonstrate that any stable mechanism approximately respects improvements of school quality; by contrast, the Boston and top trading cycles mechanisms fail to do so. Thus a stable mechanism may provide better incentives for schools to improve themselves than the Boston and top trading cycles mechanisms.

JEL classification: C78, D78, H75, I21

Keywords: Matching; School Choice; School Competition; Stability; Efficiency

*We are grateful to Parag Pathak for continuous discussions and suggestions on related issues. We also thank Daron Acemoglu, Mustafa Oguz Afacan, Itai Ashlagi, Gabriel Carroll, Glenn Ellison, Dan Fragiadakis, Ben Golub, Tadashi Hashimoto, James Heckman, Eun Jeong Heo, Brent Hickman, Scott Duke Kominers, Jonathan Levin, Hitoshi Matsushima, Paul Milgrom, Herve Moulin, Muriel Niederle, Tayfun Sönmez, Jean Tirole, Juuso Toikka, Peter Troyan, Alvin E. Roth, Xin Wei, Muhamet Yildiz and seminar participants at Chicago, Columbia, Cornell, MIT, Northwestern, Rochester, Stanford, Tokyo, Zurich, the Fair Trade Commission of Japan, the NSF/NBER Math Econ/GE Conference, and the NBER Market Design Working Group for useful comments. Michihiro Kandori, Yusuke Kasuya, and Dan Sasaki provided comments that motivated the analysis in Section 5.2.

> If we...implement choice among public schools, we unlock the values of competition in the educational marketplace. Schools that compete for students...will by virtue of their environment make those changes that allow them to succeed.
>
> *Time for Results*,
> 1991 National Governors' Association Report[1]

1 Introduction

School choice has grown rapidly in the United States and many other countries such as Japan, South Korea, and the United Kingdom. In contrast to traditional neighborhood-based placement, school districts with school choice programs allow children and their parents to express preferences over public schools and use these preferences to determine student placement. Many politicians, school reformers, and academics have embraced school choice as a policy that will substantially improve educational outcomes; for instance, in their influential book *Politics, Markets, and America's Schools*, scholars John E. Chubb and Terry M. Moe (1990) argue that school choice is "the most promising and innovative reform" available to improve the quality of public schooling.

Motivated by this interest in school choice, a large body of research in the market design literature now investigates how to assign school seats to students efficiently and fairly, recommending specific school choice mechanisms. In particular, beginning with the seminal paper by Abdulkadiroğlu and Sönmez (2003), it has been demonstrated that an extensively used school choice mechanism called the "Boston mechanism" provides strong incentives for students to misreport their preferences. Given this, two strategy-proof mechanisms have been proposed: the "student-optimal stable mechanism" (or "deferred acceptance algorithm") and the "top trading cycles mechanism". In fact, prompted by this research, the former has been adopted in New York City and Boston, while San Francisco has announced plans to adopt the latter.[2]

However, prior work on school choice in the market design literature has not analyzed the effect of different school choice mechanisms on overall school quality, but rather has always

[1] The National Governors' Association is a bipartisan public policy organization composed of the governors of the U.S. states and territories.

[2] See Abdulkadiroğlu, Pathak, and Roth (2005, 2009) and Abdulkadiroğlu, Pathak, Roth, and Sönmez (2005, 2006) for details of the implementation of these new school choice procedures in New York and Boston, respectively. The announcement by the San Francisco Unified School District of the plan to implement the top trading cycles mechanism (which the school district calls "assignment with transfers") can be found at `http://www.sfusd.edu/en/assets/sfusd-staff/enroll/files/board-of-eduation-student-assignment-policy.pdf`.

assumed that school quality is given and fixed. This is a serious omission, given that the major impetus for the introduction of school choice has been the argument, advanced by both academics and policymakers, that school choice will improve the quality of the public educational system as a whole by introducing competition among schools. For instance, Moe (2008) argues that school choice will induce schools "to educate, to be responsive, to be efficient, and to innovate". Similarly, the 1991 National Governors' Association Report argues that the nation can "increase excellence by increasing choice", that is, school choice will induce schools to improve through competitive pressures, as expressed in the epigraph. Nevertheless, formal analysis of the effects of different school choice mechanisms on schools' incentives to improve has heretofore been absent from the market design literature.

This paper approaches this question by studying how the design of a school choice mechanism affects the competitive pressure on schools to improve. We start by introducing a criterion of whether a mechanism promotes school competition: A mechanism *respects improvements of school quality* if the set of students assigned to a school always becomes weakly better for that school whenever that school becomes more preferred by students.[3] If a school's effort to improve its quality makes it more attractive to students, then requiring that a school choice mechanism assign a (weakly) better set of students to that school is a natural and mild condition in order for that school choice mechanism to incentivize that school to improve.

Despite the mildness of this criterion, we demonstrate that no stable mechanism (such as the student-optimal stable mechanism) or mechanism that is Pareto efficient for students (such as the Boston and top trading cycles mechanisms) respects improvements of school quality. That is, for any such mechanism, there exist preference profiles for the schools and students such that the outcome for a school becomes strictly worse as the school rises in the preference orderings of the students. Given this impossibility result, we consider domain restrictions on school preference profiles to ensure that the school choice mechanisms discussed above respect improvements of school quality. We show that the necessary and sufficient condition is that school preferences are *virtually homogeneous*, that is, all schools have essentially identical rankings over students; this result implies that no standard mechanism is guaranteed to respect improvements of school quality under reasonable school preferences.[4]

Even though our results show that none of the standard school choice mechanisms respects improvements of school quality perfectly, it may be that instances where a school

[3]As discussed below, this concept adapts the concept of *respecting improvements of student quality* introduced by Balinski and Sönmez (1999).

[4]For stable mechanisms, the characterization holds under the presumption that at least one school has a capacity strictly greater than one; when each school has a capacity of one, the school-optimal stable mechanism respects improvements of school quality.

benefits from discouraging student interest are rare for some mechanisms. If so, then that mechanism may provide schools with incentives to improve in practice. To investigate this possibility, we consider "large market" environments, with many schools and students, where we find clear differences in the incentives for quality improvement provided by different school choice mechanisms. We demonstrate that any stable mechanism (such as the student-optimal stable mechanism) approximately respects improvements of school quality; that is, for "almost all" preference profiles, a school is made weakly better off whenever students rank that school more highly. By contrast, surprisingly, the Boston and the top trading cycles mechanisms do not even approximately respect improvements in large markets. These results suggest that the student-optimal stable mechanism is a better school choice mechanism for promoting school competition than other competing mechanisms, particularly the Boston and the top trading cycles mechanisms.

We also study how robust the above results are to changes in the criterion of promoting school competition. It may be socially beneficially for different schools to cater to the needs of different types of students and, if so, it may be enough that a school has incentives to improve for students it finds desirable. To formalize this concept, we say that a mechanism *respects improvements of school quality for desirable students* if the outcome for a school becomes weakly better whenever a set of students, each of whom that school prefers to one of its current students, ranks that school more highly. While no stable mechanism always satisfies this requirement, any stable mechanism satisfies this criterion approximately in large markets; the Boston and top trading cycles mechanisms, however, do not satisfy this criterion even approximately in large markets. Alternatively, each school may not have any intrinsic preferences over individual students and may be concerned solely with its enrollment: A mechanism *respects improvements of school quality in terms of enrollment* if the number of students attending a school weakly increases whenever that school is ranked more highly by students. Any stable mechanism, as well as the Boston mechanism, satisfies this criterion in general markets, while the top trading cycles mechanism does not. These results suggest an additional sense in which the student-optimal stable mechanism provides schools with better incentives to improve quality than the competing top trading cycles mechanism.

Another natural question is whether the mechanisms discussed here *respect improvements of student quality*, that is, whether a student is always weakly better off when schools rank that student more highly. We show that not only the student-optimal stable mechanism, but also the Boston and the top trading cycles mechanisms satisfy this property.

Related Literature

Theoretical analyses such as Abdulkadiroğlu and Sönmez (2003) and Ergin and Sönmez (2006) have advocated for the student-optimal stable mechanism and the top trading cycles mechanism based on their incentive, fairness, and efficiency properties. Their research has lead to several school choice reforms, which were organized and reported by Abdulkadiroğlu, Pathak, and Roth (2005, 2009) and Abdulkadiroğlu, Pathak, Roth, and Sönmez (2005, 2006). This line of studies is extensively surveyed by Roth (2008), Sönmez and Ünver (2009), and Pathak (2011). As we have already emphasized, all of these papers focus on the evaluation of mechanisms in terms of the efficiency and fairness of allocations, assuming (implicitly) that the quality of every school is fixed. While drawing extensively on this literature, we offer a new perspective for distinguishing desirable school choice mechanisms from undesirable ones by analyzing their effect on schools' incentives for improving their educational quality.

The closest work to ours is the pioneering study by Balinski and Sönmez (1999), who introduce the concept of respecting improvements of student quality. According to their definition, a mechanism "respects improvements of student quality" if whenever a student is ranked higher by schools, the student becomes weakly better off. Our definition is a natural adaptation of their notion to the case in which a school improves in students' preference rankings.[5] However, the results of Balinski and Sönmez (1999) cannot be directly applied, as the model of school choice is asymmetric between schools and students since schools have multiple seats while each student can attend only one school. In fact, while Balinski and Sönmez (1999) show that the student-optimal stable mechanism respects improvements of student quality, we show that no stable mechanism, not even the school-optimal stable mechanism, respects improvements of school quality.[6]

From the methodological point of view, the current paper uses two types of analytical methods from the market design literature. First, we show impossibility results on the compatibility of some desirable properties and then find domain restrictions on the class of preferences such that the desirable properties hold simultaneously. In the context of school choice, previous studies such as Ergin (2002), Kesten (2006), and Haeringer and Klijn (2009) find domain restrictions for the student-optimal stable mechanism and the top trading cycles mechanism to satisfy several desirable properties. Similarly to these studies, we find new domain restrictions for a stable or Pareto efficient mechanism to respect improvements; our domain restriction, virtual homogeneity, is more restrictive than any of those identified in

[5] Baïou and Balinski (2000) analyze whether stable mechanisms respect improvements in the many-to-many matching setting, but their results are incorrect (Hatfield, Kojima, and Narita, 2011).

[6] Sönmez and Switzer (2011) build on the work of Balinski and Sönmez (1999), showing that the student-optimal stable mechanism respects improvements of student quality in the more general setting of matching with contracts (Hatfield and Milgrom, 2005).

these previous studies. Second, our paper also uses the large market approach used by, among others, Roth and Peranson (1999), Immorlica and Mahdian (2005), and Kojima and Pathak (2009). As these studies point out, large market analysis can often provide a positive result in cases where more traditional approaches cannot, and thus may help make a clear distinction between good mechanisms and bad ones. The current paper is another example in which the large market approach enables us to make such a distinction and thus to provide a clear policy recommendation.

The remainder of this paper is organized as follows. In Section 2, we present our model and formally define the student-optimal stable mechanism, the Boston mechanism, and the top trading cycles mechanism. In Section 3, we formally define respecting improvements of school quality and present our impossibility results. In Section 4, we present our large market results. Section 5 analyzes alternative criteria of promoting school competition, and Section 6 discusses a number of related topics. We conclude in Section 7. All proofs are in the Appendix unless explicitly noted otherwise.

2 Model

There is a finite set S of students and a finite set C of schools. Each student $s \in S$ has a strict preference relation $\succ_s$ over $C \cup \{\emptyset\}$, where $\emptyset$ denotes the outside option of the student.[7] The weak preference relation associated with $\succ_s$ is denoted by $\succsim_s$ and so we write $c \succsim_s \bar{c}$ (where $c, \bar{c} \in C \cup \{\emptyset\}$) if either $c \succ_s \bar{c}$ or $c = \bar{c}$. A preference profile of all students is denoted $\succ_S \equiv (\succ_s)_{s \in S}$.

Each school $c \in C$ has a strict preference relation $\succ_c$ over the set of subsets of S. We assume that the preference relation of each school is responsive (Roth, 1985): The preferences of school c are **responsive with capacity** q_c if

(1) For any $s, \bar{s} \in S$, if $\{s\} \succ_c \{\bar{s}\}$, then for any $S' \subseteq S \setminus \{s, \bar{s}\}$, $S' \cup \{s\} \succ_c S' \cup \{\bar{s}\}$.

(2) For any $s \in S$, if $\{s\} \succ_c \varnothing$, then for any $S' \subseteq S$ such that $|S'| < q_c$, $S' \cup \{s\} \succ_c S'$, and

(3) $\varnothing \succ_c S'$ for any $S' \subseteq S$ with $|S'| > q_c$.

If a school's preferences are responsive, then that school acts as if it has preferences over students and a quantity constraint, and the school takes the highest-ranking students available

[7] We distinguish $\emptyset$ and $\varnothing$, where $\emptyset$ denotes an outside option while $\varnothing$ is the empty set in the set-theoretic sense.

to that school up to that quantity constraint.[8]

In addition, we assume that every student is acceptable to every school as we are primarily interested in problems such as the assignment of students to public schools.[9] The preference profile of all schools is denoted $\succ_C \equiv (\succ_c)_{c \in C}$. A preference profile of all agents is denoted $\succ \equiv (\succ_S, \succ_C)$.

A **matching** is a vector $\mu = (\mu_s)_{s \in S}$ that assigns each student s a seat at a school (or the outside option) $\mu_s \in C \cup \{\emptyset\}$, and where each school $c \in C$ is assigned at most q_c students. We denote by $\mu_c \equiv \{s \in S : c = \mu_s\}$ the set of students who are assigned to school c.

A matching μ is **Pareto efficient for students** if there exists no matching μ' such that $\mu'_s \succsim_s \mu_s$ for all $s \in S$ and $\mu'_s \succ_s \mu_s$ for at least one $s \in S$.

A matching μ is **individually rational** if $\mu_s \succsim_s \emptyset$ for every $s \in S$. A matching μ is **blocked** by $(s, c) \in S \times C$ if $c \succ_s \mu_s$ and there exists $S' \subseteq \mu_c \cup s$ such that $S' \succ_c \mu_c$.[10] A matching μ is **stable** if it is individually rational and not blocked.

Two remarks are in order. First, in this model, schools are assumed to have preferences over sets of students. Thus, our analysis can be utilized for other applications such as certain entry-level labor markets (Roth, 1984) without modification. Second, in some school districts such as Boston, the preference orderings of schools over students is determined by priorities given by law (Abdulkadiroğlu and Sönmez, 2003).[11] In such cases, it may not be reasonable to assume that the priorities set by law represent real preferences for schools (or school principals). We address this issue in Section 5.1.

2.1 Mechanisms

Given the set of students S and schools C, a **mechanism** is a function φ from the set of preference profiles to the set of matchings. A mechanism φ is **Pareto efficient for students** if $\varphi(\succ)$ is a Pareto efficient matching for students for every preference profile $\succ$.

[8] Responsive preferences rule out the case where a school's ranking of two students depends on the students the school currently holds. However, responsive preferences do allow for certain forms of peer effects: For instance, if a school values high-achieving students, and the presence of one high-achieving student makes other high-achieving students more desirable, then the ordinal ranking of students would be unaffected by the set of students the school currently holds, and hence the school's preferences would be responsive.

[9] This assumption is needed only for our large market result for stable mechanisms (Theorem 3) and our characterization results (Propositions 6 and 7). All of our other results hold even without this assumption.

[10] Throughout the paper, we denote singleton set $\{x\}$ by x when there is no confusion.

[11] However, in many cases, schools do have such preferences. For instance, in New York City, there are many schools, commonly referred to as screened schools, which explicitly rank students based on their academic record, attendance, test scores, and other similar criteria. Another example where schools have explicit preferences is the Turkish higher educational system, the motivating example of Balinski and Sönmez (1999). In this system, each university ranks students by calculating a weighted sum of the students' scores on different parts of a national exam; different universities use different weighting formulas. (We thank Tayfun Sönmez for suggesting this example.)

A mechanism φ is **stable** if $\varphi(\succ)$ is a stable matching for every preference profile $\succ$. We now define three mechanisms of particular interest for school choice problems.

2.1.1 The Student-Optimal Stable Mechanism

Given $\succ$, the (student-proposing) **deferred acceptance (DA) algorithm** of Gale and Shapley (1962) is defined as follows.

- Step 1: Each student $s \in S$ applies to her most preferred acceptable school (if any). Each school tentatively keeps the highest-ranking students up to its capacity, and rejects every other student.

In general, for any step $t \geq 2$,

- Step t: Each student s who was not tentatively matched to any school in Step $(t-1)$ applies to her most preferred acceptable school that has not rejected her (if any). Each school tenatively keeps the highest-ranking students up to its capacity from the set of students previously tenatively matched to this school and the students newly applying, and rejects every other student.

The algorithm terminates at the first step at which no student applies to a school. Each student tentatively kept by a school at that step is allocated a seat in that school, resulting in a matching which we denote by $\varphi^S(\succ)$. The **student-optimal stable mechanism** is a mechanism φ^S that produces $\varphi^S(\succ)$ for every preference profile $\succ$. It is well known that φ^S is a stable mechanism (Gale and Shapley, 1962). Moreover, the outcome of this mechanism is the student-optimal stable matching, that is, the matching that is weakly preferred to any other stable matching by all students. (The name of the student-optimal stable mechanism is due to this property.) In addition, φ^S is known to be strategy-proof for students, that is, for each student it is a weakly dominant strategy to report her true preferences (Dubins and Freedman, 1981; Roth, 1982).[12] Due to these properties, the deferred acceptance algorithm has been implemented in both New York City (Abdulkadiroğlu, Pathak, and Roth, 2005) and Boston (Abdulkadiroğlu, Pathak, Roth, and Sönmez, 2005).

Another canonical stable mechanism is the **school-optimal stable mechanism**. That mechanism is based on the school-proposing version of the deferred acceptance algorithm, in which schools make offers to students and students keep their most preferred offers at each step. We denote the student-optimal stable mechanism by φ^S and the school-optimal stable

[12] In fact, the student-optimal stable mechanism is (weakly) group strategy-proof, in the sense that there is no group deviation which makes all the members of the group strictly better off (Dubins and Freedman, 1981).

mechanism by φ^C; φ^C is also the student-pessimal stable mechanism, i.e., it produces the stable matching that every student weakly disprefers to every other stable matching (See Theorem 2.13 of Roth and Sotomayor (1990)).

2.1.2 The Boston Mechanism

Given $\succ$, the **Boston mechanism** (Abdulkadiroğlu and Sönmez, 2003), denoted φ^B, is defined through the following algorithm.[13]

- Step 1: Each student $s \in S$ applies to her most preferred acceptable school (if any). Each school accepts its most-preferred students up to its capacity and rejects every other student.

In general, for any step $t \geq 2$,

- Step t: Each student who has not been accepted by any school applies to her most preferred acceptable school that has not rejected her (if any). Each school accepts its most-preferred students up to its remaining capacity and rejects every other student.

The algorithm terminates at the first step in which no student applies to a school. Each student accepted by a school during some step of the algorithm is allocated a seat in that school. The Boston algorithm differs from the deferred acceptance algorithm in that when a school accepts a student at a step, in the Boston algorithm, the student is guaranteed a seat at that school, while in the deferred acceptance algorithm, that student may be later displaced by another student whom the school likes better. Note that this mechanism is Pareto efficient for students with respect to any reported preference profile. In Boston, the Boston mechanism has been replaced by the student-optimal stable mechanism, but is still in use in many school districts, such as Denver and Minneapolis (Miralles, 2009).

2.1.3 The Top Trading Cycles Mechanism

The **top trading cycles (TTC) mechanism**, denoted φ^{TTC}, is defined as follows: For any $t \geq 1$,

- Step t: Each student $s \in S$ points to her most preferred school (if any); students who do not point at any school are assigned to $\emptyset$. Each school $c \in C$ points to its most preferred student. As there are a finite number of schools and students, there exists at least one cycle, i.e., a sequence of distinct schools and students $(s_1, c_1, s_2, c_2, \ldots, s_K, c_K)$ such

[13] Alcalde (1996) calls this rule the "now-or-never" mechanism for the special case in which the capacity of each school is one.

that student s_1 points at school c_1, school c_1 points to student s_2, student s_2 points to school c_2, ..., student s_K points to school c_K, and, finally, school c_K points to student s_1. Every student s_k ($k = 1, \ldots, K$) is assigned to the school she is pointing at. Any student who has been assigned a school seat or the outside option as well as any school $c \in C$ which has been assigned students such that the number of them is equal to its capacity q_c is removed. If no student remains, the algorithm terminates; otherwise, it proceeds to the next step.

This algorithm terminates in a finite number of steps as at least one student is matched with a school (or $\emptyset$) at each step and there are only a finite number of students. The TTC mechanism is defined as a rule that, for any preference profile $\succ$, produces $\varphi^{TTC}(\succ)$ through the above algorithm.

The current version of the top trading cycles algorithm was introduced by Abdulkadiroğlu and Sönmez (2003) for the school choice problem.[14] While it does not necessarily produce a stable matching, the mechanism has a number of desirable properties. First, it always produces a Pareto efficient matching, unlike the student-optimal stable mechanism.[15] Second, it is group strategy-proof, that is, no coalition of students can jointly misreport their preferences in such a way that every student in the coalition is made weakly better off with at least one student strictly better off. Based on these advantages, the top trading cycles mechanism has been considered for use in a number of school districts in the United States, such as Boston (which ultimately decided to use the student-optimal stable mechanism) and San Francisco (which recently announced plans to implement a top trading cycles mechanism).

3 Respecting Improvements of School Quality

The main goal of this paper is to analyze how the design of a school choice mechanism affects the incentives of schools to improve themselves. To do this, we now define a criterion for evaluating school choice mechanisms in terms of the incentives they provide for school improvement. We first formally specify the notion of an improvement of school quality in our model.

[14] The original top trading cycles algorithm was defined in the context of the housing market and is attributed to David Gale by Shapley and Scarf (1974).

[15] While the Boston mechanism is Pareto efficient with respect to the stated preferences, it is well-known that it is not, in general, a Nash equilibrium for students to report their preferences truthfully. In fact, the set of Nash equilibrium outcomes under the Boston mechanism is equivalent to the set of stable matchings (Ergin and Sönmez, 2006) and so we would not, in general, expect that the Boston mechanism would result in a Pareto efficient outcome.

Definition 1. A preference relation $\succ'_s$ is an **improvement for school** c over the preference relation $\succ_s$ if

(1) For all $\bar{c} \in C \cup \{\emptyset\}$, if $c \succ_s \bar{c}$, then $c \succ'_s \bar{c}$, and

(2) For all $\bar{c}, \hat{c} \in (C \cup \{\emptyset\}) \setminus \{c\}$, $\bar{c} \succ'_s \hat{c}$ if and only if $\bar{c} \succ_s \hat{c}$.

The student preference profile $\succ'_S$ is an improvement for school c over $\succ_S$ if for every student s, $\succ'_s$ is an improvement for school c over $\succ_s$.

We also say that $\succ'_s$ is a **disimprovement for school** c over $\succ_s$ if $\succ_s$ is an improvement for school c over $\succ'_s$ and that $\succ'_S$ is a disimprovement over $\succ_S$ if $\succ_S$ is an improvement over $\succ'_S$.[16]

Put simply, a student preference profile $\succ'_S$ is an improvement for school c over the preference profile $\succ_S$ when every student ranks c weakly higher under $\succ'_S$ while the ordering of other schools is unchanged between the two preference profiles. When a school improves its quality, it should become more attractive to every student without changing the relative rankings of other schools, and the concept of school improvement is meant to capture this intuition in the standard ordinal setting of the matching literature. With this concept in hand, we now define the property by which we will evaluate school choice mechanisms in this work.

Definition 2. A mechanism φ **respects improvements of school quality** at the school preference profile $\succ_C$ if, for all $c \in C$ and student preference profiles $\succ_S$ and $\succ'_S$, if $\succ'_S$ is an improvement for school c over $\succ_S$, then $\varphi_c(\succ'_S, \succ_C) \succsim_c \varphi_c(\succ_S, \succ_C)$.

Equivalently, a mechanism φ respects improvements of school quality at school preference profile $\succ_C$ if there do not exist a school c and student preference profiles $\succ_S$ and $\succ'_S$ such that $\succ'_S$ is a disimprovement for school c over $\succ_S$ while $\varphi_c(\succ'_S, \succ_C) \succ_c \varphi_c(\succ_S, \succ_C)$.

This definition requires that the outcome of a mechanism be weakly better for a school if that school becomes more preferred by students. If a school's effort to improve its quality makes it more attractive to students, then the concept of respecting improvements of school quality seems to be a natural and mild criterion for schools to have incentives to invest in quality improvement.

The concept of respecting improvements was introduced by Balinski and Sönmez (1999) in the context of centralized college admission. In their work, a mechanism is said to respect improvements of *student* quality if whenever a student improves in colleges' preference rankings, that student is better off. They show that the student-optimal stable mechanism is the

[16] We will also say that $\succ'$ is a (dis)improvement for school c over $\succ$ if $\succ'_S$ is a (dis)improvement for school c over $\succ_S$ and $\succ'_C = \succ_C$.

unique stable mechanism that respects improvements of student quality.[17] Our definition is a natural adaptation of their notion to the case in which a school improves in students' preference rankings. The main difference between our concept and that of Balinski and Sönmez (1999) is that we consider improvements of school quality rather than those of student quality. Because the school choice matching model is asymmetric between schools and students in the sense that schools have multiple seats while each student can attend only one school, the result by Balinski and Sönmez (1999) mentioned above cannot be directly applied. In fact, as we will see in the next section, no stable mechanism respects improvements of school quality, which is in sharp contrast to the result by Balinski and Sönmez (1999).

3.1 Stable Mechanisms

We first investigate whether stable mechanisms such as the student-optimal stable mechanism respect improvements.[18] The following example offers a negative answer to this question.

Example 1. Let $S = \{s, \bar{s}\}$, $C = \{c, \bar{c}\}$. Consider the following preferences:

$$\succ_s : \bar{c}, c, \emptyset, \qquad \succ_c : s, \bar{s},$$
$$\succ_{\bar{s}} : \bar{c}, c, \emptyset, \qquad \succ_{\bar{c}} : \bar{s}, s,$$

where the notational convention for students is that student s prefers $\bar{c}$ most, c second, and $\emptyset$ third, and so forth, and the notational convention for schools is that they have some responsive preferences consistent with preferences over students as described above. (This notation is used throughout.) The capacities of the schools are given by $q_c = 2$ and $q_{\bar{c}} = 1$.[19]

Note that at the first step of the student-proposing deferred acceptance algorithm under the preference profile $\succ \equiv (\succ_s, \succ_{\bar{s}}, \succ_c, \succ_{\bar{c}})$, both students s and $\bar{s}$ apply to $\bar{c}$. Since $q_{\bar{c}} = 1$, $\bar{c}$ rejects s. Then s applies to c, where she is accepted. The algorithm terminates at this step, producing the student-optimal stable matching,

$$\varphi^S(\succ) = \begin{pmatrix} c & \bar{c} \\ s & \bar{s} \end{pmatrix},$$

where this matrix notation represents the matching where c is matched with s while $\bar{c}$

[17] See our discussion in Section 6.2.

[18] For brevity, we will often write "respecting improvements" for the longer phrase "respecting improvements of school quality".

[19] Note that, strictly speaking, the information on school preferences over individual students and the capacity does not uniquely specify that school's preference relation over groups of students. Whenever we specify a school's preferences over individual students and its capacity only, it should be understood to mean an arbitrary responsive preference relation consistent with the given information.

is matched with $\bar{s}$. (Again, this notation is used throughout.) At the first step of the school-proposing deferred acceptance algorithm under preference profile $\succ$, school c proposes to both s and $\bar{s}$ while $\bar{c}$ proposes to $\bar{s}$. Student $\bar{s}$ keeps $\bar{c}$ and rejects c while student s keeps c. Since school c has proposed to all students, the algorithm terminates. Thus the school-optimal stable matching $\varphi^C(\succ)$ is equal to $\varphi^S(\succ)$. Since it is well-known that $\varphi^S_s(\succ) \succsim_s \mu_s \succsim_s \varphi^C_s(\succ)$ for any stable matching μ, it follows that this market has a unique stable matching, $\varphi^S(\succ) = \varphi^C(\succ)$.

Now, consider the preference relation $\succ'_{\bar{s}}$ such that

$$\succ'_{\bar{s}}: c, \bar{c}, \emptyset.$$

Note that $\succ'_{\bar{s}}$ is an improvement for school c over $\succ_{\bar{s}}$. At the first step of the student-proposing deferred acceptance algorithm under the preference profile $(\succ'_{\bar{s}}, \succ_{-\bar{s}})$,[20] student s applies to $\bar{c}$ while student $\bar{s}$ applies to c. The algorithm terminates immediately at this step, producing the student-optimal stable matching

$$\varphi^S(\succ'_{\bar{s}}, \succ_{-\bar{s}}) = \begin{pmatrix} c & \bar{c} \\ \bar{s} & s \end{pmatrix}.$$

On the other hand, at the first step of the school-proposing deferred acceptance algorithm under preference profile $(\succ'_{\bar{s}}, \succ_{-\bar{s}})$, school c proposes to both s and $\bar{s}$ while $\bar{c}$ proposes to $\bar{s}$. Student $\bar{s}$ rejects $\bar{c}$. Rejected from its first choice $\bar{s}$, $\bar{c}$ proposes to s. Now student s rejects c. Because school c has proposed to all students, the algorithm terminates. Thus the school-optimal stable matching $\varphi^C(\succ'_{\bar{s}}, \succ_{-\bar{s}})$ is equal to $\varphi^S(\succ'_{\bar{s}}, \succ_{-\bar{s}})$. This implies that this market has a unique stable matching, $\varphi^S(\succ'_{\bar{s}}, \succ_{-\bar{s}}) = \varphi^C(\succ'_{\bar{s}}, \succ_{-\bar{s}})$.

From the arguments above, we have that, for any stable mechanism φ,

$$\varphi_c(\succ) = s \succ_c \bar{s} = \varphi_c(\succ'_{\bar{s}}, \succ_{-\bar{s}}),$$

even though $\succ'_{\bar{s}}$ is an improvement for c over $\succ_s$; hence, φ does not respect improvements of school quality at the school preference profile $\succ_C$.

The finding from Example 1 can be summarized in the following statement.

Theorem 1. *There exists no stable mechanism that respects improvements of school quality at every school preference profile.*

[20] Subscript $-i$ indicates $C \cup S \setminus \{i\}$, that is, the set of all agents except for i. For instance, $\succ_{-\bar{s}}$ is the profile of preferences of all students and schools except for student $\bar{s}$.

3.2 Pareto Efficient Mechanisms for Students

As in many other resource allocation problems, Pareto efficiency for students is a popular desideratum in school choice because students are considered to be the beneficiaries of public schooling. While the student-optimal stable mechanism is not Pareto efficient for students, there are other mechanisms that are. The popular Boston mechanism (under truth-telling by students) and the theoretically favored TTC mechanism are such examples. Thus it would be of interest to investigate whether these mechanisms or any other Pareto efficient mechanism respects improvements of school quality. As the following example shows, it turns out that there exists no mechanism that is Pareto efficient for students and that respects improvements of school quality.

Example 2. Suppose that there exists a mechanism φ that is Pareto efficient for students and respects improvements of school quality. Let $S = \{s, \bar{s}\}$, $C = \{c, \bar{c}\}$, and the preferences of the schools be given by

$$\succ_c : \bar{s}, s, \emptyset,$$
$$\succ_{\bar{c}} : s, \bar{s}, \emptyset,$$

with capacities of $q_c = q_{\bar{c}} = 1$. First, consider the following preference profile of students:

$$\succ_s : \bar{c}, \emptyset,$$
$$\succ_{\bar{s}} : c, \emptyset.$$

Under $\succ \equiv (\succ_s, \succ_{\bar{s}}, \succ_c, \succ_{\bar{c}})$, the unique Pareto efficient matching is

$$\varphi(\succ) = \begin{pmatrix} c & \bar{c} \\ \bar{s} & s \end{pmatrix}.$$

Thus, in the outcome of the mechanism under $\succ$, school $\bar{c}$ is matched with student s.

Now consider the student preference profile $\succ'_S \equiv (\succ_s, \succ'_{\bar{s}})$ where the preference of $\bar{s}$ has changed to

$$\succ'_{\bar{s}} : \bar{c}, c, \emptyset.$$

Note that $\succ'_{\bar{s}}$ is an improvement for school $\bar{c}$ over $\succ_{\bar{s}}$; hence, $\bar{c}$ must obtain at least as good an outcome under $\succ' \equiv (\succ'_S, \succ_C)$ as under $\succ$, and so $\bar{c}$ must be matched to s. By Pareto efficiency, then, $\bar{s}$ must be matched to c and so $\varphi(\succ') = \varphi(\succ)$.

Finally, consider another student preference profile $\succ''_S \equiv (\succ''_s, \succ'_{\bar{s}})$ where

$$\succ''_s : c, \bar{c}, \emptyset.$$

Note that $\succ''_S$ is an improvement for school c over $\succ'_S$. Under $\succ'' \equiv (\succ''_S, \succ_C)$, the unique Pareto efficient matching for students is

$$\varphi(\succ'') = \begin{pmatrix} c & \bar{c} \\ s & \bar{s} \end{pmatrix},$$

which implies that c is matched with s in the outcome of the mechanism. However, note that $\varphi_c(\succ') = \bar{s} \succ_c s = \varphi_c(\succ'')$, even though $\succ''_S$ is an improvement for school c over $\succ'_S$. This means that this mechanism does *not* respect improvements of school quality, which is a contradiction.

The finding from Example 2 can be summarized in the following statement.

Theorem 2. *There exists no mechanism that is Pareto efficient for students and respects improvements of school quality for every school preference profile.*

Recall that the Boston and the TTC mechanisms are Pareto efficient for students. The above theorem shows that these popular mechanisms do not respect improvements of school quality.

Remark. The above conclusion of Theorem 2 for the Boston mechanism is with respect to the students' reported preferences, but it is well known that truthtelling is not a dominant strategy under the Boston mechanism. However, Theorem 1 in Section 3.1 sheds some light on the Boston mechanism when students behave strategically. Although the Boston mechanism is not stable, the set of Nash equilibrium outcomes under that mechanism is equivalent to the set of stable matchings (Ergin and Sönmez, 2006). Therefore, our Theorem 1 implies that the Boston mechanism does not respect improvements under strategic play if students play a Nash equilibrium.

3.3 Conditions on Preferences for Respecting Improvements

Given that the above representative mechanisms do not respect improvements at every school preference profile, a natural question is what conditions, if any, on the school preference profile $\succ_C$ enable a stable or Pareto efficient mechanism to respects improvements. A school preference profile is **virtually homogeneous** if the rankings of students are identical across all schools except possibly for the "highest-ranked" students, i.e., students that every school

would accept, regardless of the other students available to that school; the precise definition is given in Section 6.1. Clearly this condition is a very strong requirement on school preferences and, in fact, many domain restrictions used in the literature are implied by virtual homogeneity.[21]

We show that virtual homogeneity is the key condition on school preferences for a stable or Pareto efficient mechanism to respect improvements of school quality. In particular, Propositions 6 and 7 in Section 6.1 imply the following fact: When at least one school has capacity larger than one, *there exists a stable mechanism or a Pareto efficient mechanism for students that respects improvements of school quality if and only if the school preference profile is virtually homogeneous.*

Since virtual homogeneity is an extremely strong requirement, this result suggests that the concern that stable or Pareto efficient mechanisms may provide perverse incentives to schools cannot be easily precluded by any mild preference domain restriction. All details including the formal definition of virtual homogeneity and the statements of Propositions 6 and 7 are offered in Section 6.1. This negative result motivates our study in the next section on the properties of mechanisms in large markets.

4 Respecting Improvements in Large Markets

While the results of Section 3 show that no standard school choice mechanism always respects improvements of school quality, violations of this condition may be rare for some school choice mechanisms. In this section, we investigate this possibility by considering random markets, especially in the context of large environments which contain many students and schools; in that environment, we study the probability that a mechanism respects improvements.

4.1 The Large Market Model

We now introduce the following large markets environment, which is (a slight generalization of) the environment studied by Kojima and Pathak (2009). A **random market** is a tuple $\tilde{\Gamma} = (C, S, k, \mathcal{D})$, where k is a positive integer and $\mathcal{D}$ is a pair $(\mathcal{D}_C, \mathcal{D}_S)$ of probability distributions: Each random market induces a market by randomly generating preferences of students and schools. First, $\mathcal{D}_S = (p_c)_{c \in C}$ is a probability distribution on C. Preferences of each student s are drawn as follows (Immorlica and Mahdian, 2005):

[21] We discuss the relationship between virtual homogeneity and existing domain restrictions in the literature in Section 6.1.

- Step 1: Select a school independently from distribution $\mathcal{D}_S$. List this school as the top ranked school of student s.

In general,

- Step $t \leq k$: Select a school independently from distribution $\mathcal{D}_S$ until a school is drawn that has not been drawn in any previous step. List this school as the tth most preferred school of student s.

In other words, each student chooses k schools repeatedly from $\mathcal{D}_S$ without replacement. Student s finds these k schools acceptable, and all other schools unacceptable. For example, if $\mathcal{D}_S$ is the uniform distribution on C, then the preference list is drawn from the uniform distribution over the set of all preference lists of length k.

For schools, preference profile $\succ_C$ is drawn from the given distribution $\mathcal{D}_C$ over school preference profiles. We do not impose any restriction on $\mathcal{D}_C$ at this point. In particular, we allow correlations in school preferences and even the possibility that $\mathcal{D}_C$ is a degenerate distribution, in which case school preferences are deterministic.

A **sequence of random markets** is denoted by $(\tilde{\Gamma}^1, \tilde{\Gamma}^2, \dots) = (\tilde{\Gamma}^n)_{n \in \mathbb{N}}$, where $\tilde{\Gamma}^n = (C^n, S^n, k^n, \mathcal{D}^n)$ is a random market in which $|C^n| = n$ is the number of schools.[22] Consider the following regularity conditions defined by Kojima and Pathak (2009).[23]

Definition 3. A sequence of random markets $(\tilde{\Gamma}^n)_{n \in \mathbb{N}}$ is **regular** if there exist positive integers k, $\tilde{q}$ and $\hat{q}$ such that

(1) $k^n \leq k$ for all n,

(2) $q_c \leq \hat{q}$ for all n and $c \in C^n$,

(3) $|S^n| \leq \tilde{q}n$ for all n, and

(4) for all n and $c \in C^n$, every $s \in S^n$ is acceptable to c at any realization of preferences for c at $\mathcal{D}_{C^n}$.

Condition (1) above assumes that the length of students' preference lists is bounded from above even when the market size grows. Condition (2) requires that the number of seats

[22] Unless specified otherwise, our convention is that superscripts are used for the number of schools present in the market whereas subscripts are used for agents.

[23] A careful reader may notice that our regularity conditions are more general than the ones presented in the main text of Kojima and Pathak (2009). More specifically, the main text of Kojima and Pathak (2009) assumes that $k^n = k$ (rather than $k^n \leq k$), $\hat{q} = \tilde{q}$, and that the distribution of school preference profiles is degenerate (that is, school preferences are deterministic). However, as Kojima and Pathak (2009) point out, all of their results hold under the set of assumptions introduced here.

in any one school is bounded even in large school districts. Condition (3) requires that the number of students does not grow much faster than that of schools (it is allowed, on the contrary, that the number of students does not grow as fast as the number of schools). Condition (4) requires that, at any realized preference profile, each school finds any student acceptable, but preferences are otherwise arbitrary.[24]

We introduce another concept defined by Kojima and Pathak (2009). Let

$$V_T(n) \equiv \left\{ c \in C^n : \frac{\max_{\bar{c} \in C^n}\{p^n_{\bar{c}}\}}{p^n_c} \leq T \text{ and } |\{s \in S^n : c \succ_s \emptyset\}| < q_c \right\}.$$

In words, $V_T(n)$ is a set of schools such that (i) each school c in this set is sufficiently popular ex ante, i.e., the ratio of $p^n_{\bar{c}}$ to p^n_c, where $\bar{c}$ is the most popular school, does not grow without bound as n grows large, while (ii) there are fewer students who find the school acceptable than the capacity of the school ex post. Note that $V_T(n)$ is a random set because student preferences are stochastic.

Definition 4. A sequence of random markets is **sufficiently thick** if there exists $T \in \mathbb{R}$ such that $E[|V_T(n)|]$ approaches infinity as n goes to infinity.

This condition requires that the expected number of schools that are popular enough ex ante, yet have fewer students ex post who find the school acceptable than their numbers of seats, i.e., $V_T(n)$, grows infinitely large as the market becomes large. As we will see later, this condition guarantees that the market is "thick enough" to absorb certain market disruptions. To gain intuition, consider a change in the market in which an additional student needs to be placed at a school. If the market is sufficiently thick, such a student is likely to find a seat at a school that has a seat for her in a stable matching without changing the assignment of many other students. In other words, the sufficient thickness condition implies that a small disruption of the market is likely to be absorbed by vacant seats.

While the condition itself is technically involved, many types of distributions satisfy sufficient thickness (in fact, the concept is not really intended to offer an "intuitive" notion, but rather to subsume as many practical cases as possible). For instance, if all student preferences are drawn from the uniform distribution, the market will be sufficiently thick. To describe another, more general, example, we say that a sequence of random markets satisfies **moderate similarity** if there is a bound T such that $\frac{p_{\bar{c}}}{p_c} \leq T$ for all $c, \bar{c} \in C^n$ for all n. Such a restriction has been employed in studies such as Manea (2009), Kojima, Pathak, and Roth (2011), and Ashlagi, Braverman, and Hassidim (2011).[25] Kojima and Pathak

[24] As mentioned by Kojima and Pathak (2009), it is possible to weaken this condition such that many, but not all, schools find all students to be acceptable.

[25] The term "moderate similarity" follows Manea (2009).

(2009) show that moderate similarity implies sufficient thickness and offer other examples of student preference distributions that satisfy the sufficient thickness condition.

Remark. Condition (1) of regularity requires that the number of schools acceptable to each student is bounded. This assumption is motivated by the fact that in practive observed preference lists in many school districts are quite short: In New York City, almost three quarters of students rank less than 12 schools even though there were over 500 school programs. In Boston, more than 90% of students rank 5 or fewer schools at the elementary school level out of about 30 different schools in each zone. Still, it is of interest to consider alternative assumptions. We say that the sequence of random markets has an **excess supply of school capacities** if there exists $\lambda > 0$ such that $\sum_{c \in C^n} q_c - |S^n| \geq \lambda n$ for all n.[26] This condition requires, as is usually the case in the public school context, there are more than sufficient capacities in schools to accommodate all students in the district. For instance, if the capacity of each school is q and $|S^n| = (q-1)n$ for all n, then this sequence of random markets has an excess supply of school capacities. The conclusion of our main result, Theorem 3, holds even without condition (1) of regularity—so students can find any number of schools acceptable—under the conditions of excess supply of school capacities and moderate similarity.[27] See Appendix A.1.1 for details.

4.2 Main Results

For any random market $\tilde{\Gamma}$, school c, and mechanism φ, let $\alpha_c(\tilde{\Gamma}, \varphi)$ be the probability that the realized preference profile $\succ$ has the property that there exists a student preference profile $\succ'_S$ such that $\succ'_S$ is a disimprovement over $\succ_S$ for c while $\varphi_c(\succ'_S, \succ_C) \succ_c \varphi_c(\succ)$. We say that a mechanism φ **approximately respects improvements of school quality in large markets** if, for any sequence of random markets $(\tilde{\Gamma}^n)_{n \in \mathbb{N}}$ that is regular and sufficiently thick, $\max_{c \in C^n} \alpha_c(\tilde{\Gamma^n}, \varphi) \to 0$ as $n \to \infty$, i.e., for any $\varepsilon > 0$, there exists an integer m such that, for any random market $\tilde{\Gamma}^n$ in the sequence with $n > m$ and any $c \in C^n$, we have that $\alpha_c(\tilde{\Gamma}^n, \varphi) < \varepsilon$. As the name suggests, a mechanism approximately respects improvements in large markets if the probability that a school is made better off by being less preferred by students converges to zero as the size of the markets approaches infinity; in this sense, violations of the respecting improvements criterion are rare for mechanisms that satisfy this condition. With this concept, we are ready to state our main results.

[26] This condition is a slight modification of the "excess number of positions" condition assumed by Ashlagi, Braverman, and Hassidim (2011) in the environment of matching with couples.

[27] In fact, the result can be shown even without excess supply of school capacities, if the number of schools that are acceptable to each student grows without a bound (thus violating Condition (1) of regularity) but sufficiently slowly. More specifically, one of the key results by Kojima and Pathak (2009), which is used in our proof, holds as long as $k^n = o(\log(n))$, and our proof goes through. See their footnote 32 for detail.

Theorem 3. *Any stable mechanism approximately respects improvements of school quality in large markets.*

Proof. See Appendix A.1.1. □

This theorem suggests that while no stable mechanism always respects improvements, such a perverse outcome occurs only very rarely in large markets. More specifically, as the number of participating schools approaches infinity (while the number of students can also grow but does not have to), the probability of such an incident converges to zero.

We defer the formal proof of the theorem to the Appendix and offer an outline of the argument here. For simplicity, we focus our attention on the student-optimal stable mechanism. First, recall Example 1. In that example, school c is better off when student $\bar{s}$ prefers school $\bar{c}$ to c than when student $\bar{s}$ prefers school c to $\bar{c}$. The reason for this is that when student $\bar{s}$ prefers school $\bar{c}$ to c, the student $\bar{s}$ displaces student s from school $\bar{c}$ and then student s applies to school c, which in turn makes school c better off. More generally, a school can be made better off when a student demotes the school in her preference ranking because it increases competition at a different school, thus creating a "rejection chain" in the algorithm that reaches the original school.

Despite this fact, the above theorem says that the probability of such a perverse outcome becomes small in large markets. The intuition behind this result is as follows. If there are a large number of schools in the market, then it can be shown (under the sufficient thickness assumption) that with high probability, there are also a large number of schools with vacant seats. Hence, when the ranking of a school c falls for some student s, the probability that a student involved in a rejection chain will apply to a school with vacant seats is much higher than the probability that the student will apply to c, as there are a large number of schools with vacant seats. Since every student is acceptable to any school by assumption, if such an application happens, then the rejection chain terminates without reaching c. Thus, the probability that the rejection chain reaches and benefits c is small.

The main technical contribution of the proof is to rigorously establish that the above intuition goes through. To do so, we need to overcome at least two difficulties. First, in spite of the plausibility of the above example, it is not clear whether the occurrence of such a rejection chain is the only reason that a stable mechanism does not respect improvements. Second, while the above intuition is directly applicable only to the student-optimal stable mechanism, we must show that the conclusion of the theorem holds not only for the student-optimal stable mechanism but also for an arbitrary stable mechanism. To address these issues, our proof proceeds in three steps. The first step is to establish the following relationship between stable mechanisms that fail to respect improvements of school quality and

stable mechanisms that are subject to strategic preference manipulation by schools.

Lemma 1. *Let φ be a stable mechanism.*

(1) *Suppose that the preference profile $\succ$ and student preference profile $\succ'_S$ are such that $\succ'_S$ is a disimprovement for some $c \in C$ over $\succ_S$ and $\varphi_c(\succ'_S, \succ_C) \succ_c \varphi_c(\succ)$. Then there exists a (reported) preference relation $\succ''_c$ for c such that $\varphi_c(\succ''_c, \succ_{-c}) \succ_c \varphi_c(\succ)$.*

(2) *Suppose that the preference profile $\succ$ and (reported) preference relation $\succ''_c$ for some $c \in C$ are such that $\varphi_c(\succ''_c, \succ_{-c}) \succ_c \varphi_c(\succ)$. Then there exists a student preference profile $\succ'_S$ such that $\succ'_S$ is a disimprovement for c over $\succ_S$ and $\varphi_c(\succ'_S, \succ_C) \succ_c \varphi_c(\succ)$.*[28]

This lemma shows that for stable mechanisms, there is a certain equivalence between the failure of respecting improvements of school quality and the vulnerability to strategic manipulations by schools. In particular, Part 1 of the lemma shows that whenever there exists a school preference profile such that φ does not respect improvements for a school at that school preference profile, then there exists a reported preference profile for that school that makes the school strictly better off than when that school truthfully reports its preferences. Thus, to prove Theorem 3, it is sufficient to show that in any stable mechanism it is approximately optimal for schools to report their true preferences in large markets.

The second step of the proof enables us to focus on the student-optimal stable mechanism. To do so, we invoke the fact that whenever a stable mechanism can be profitably manipulated by a school, the student-optimal stable mechanism can be profitably manipulated by the same school at that preference profile (Pathak and Sönmez, 2011). By this result and the preceding argument, the probability of a school preference profile such that a stable mechanism φ does not respect improvements at that preference profile is bounded from above by the probability that the student-optimal stable mechanism can be profitably manipulated by a school.

The last step of the proof is to bound the probability that the student-optimal stable mechanism can be profitably manipulated by a school. Under our assumptions, Kojima and Pathak (2009) show that this probability converges to zero as the market size approaches infinity. This result and the arguments in the preceding paragraphs complete the proof.

Remark. In Theorem 3, the order of convergence is $O(\frac{1}{E[V_T(n)]})$, which by the sufficient thickness assumption converges to zero. For instance, if the sequence of random markets satisfies moderate similarity (Section 4.1), then the order of convergence is $O(\frac{1}{n})$. See Appendix A.1.1 for details.

[28]Note that Part 2 of Lemma 1 is not needed for showing Theorem 3. Still, it is of independent interest in that, for instance, Corollary 1 in Section 6.2 utilizes not only Part 1 but also Part 2 of Lemma 1. Furthermore, our Theorem 1 is a corollary of Lemma 1, as an impossibility theorem of Roth (1984) shows that there exists no stable mechanism that is strategy-proof for schools while our Lemma 1 shows that strategy-proofness for schools is equivalent to respecting improvements of school quality within the class of stable mechanisms.

In contrast to stable mechanisms, neither the Boston mechanism nor the TTC mechanism approximately respects improvements even in large markets. More precisely, the following results show that, even for arbitrarily large markets, under these mechanisms a school can be made better off if some students demote the school in their preference rankings with a nonnegligible probability.[29]

Theorem 4. *The Boston mechanism does not approximately respect improvements of school quality in large markets.*

Proof. See Appendix A.1.2. □

Theorem 5. *The top trading cycles mechanism does not approximately respect improvements of school quality in large markets.*

Proof. See Appendix A.1.3. □

The negative results of Theorems 4 and 5 provide a sharp contrast to the positive result of Theorem 3. These results indicate that schools may not be incentivized to reduce school quality for any student under the student-optimal stable mechanism, while they will be incentivized to reduce school quality for some students under the Boston or the TTC mechanisms.

The intuition for Theorem 4 is as follows. Recall that in the Boston mechanism, every acceptance is final in each step. Therefore, if a student applies to a school in an earlier step than its more preferred student, the mechanism can match the former to the school at the expense of the latter. Hence, if the less preferred student changes her preferences to like the school better, it can lead to an inferior outcome for the school as it may induce that student to apply earlier. This logic is relatively simple and does not depend on the size of the market: Roughly speaking, a randomly chosen student is less preferred to another randomly chosen student with a fixed probability, whether or not the market is large.[30] The formal proof in the Appendix makes this intuition precise, by presenting a random market in which a less preferred student applies for a position at a school earlier than a more preferred student.[31]

The intuition for Theorem 5 is only slightly more complicated. In the TTC mechanism, even an undesirable student may be matched to a school if the student can trade priorities

[29] In fact, the proofs of these theorems show that the failure of respecting improvements occurs not only for large markets, but for any market with the number of schools $n \geq 2$. (Of course, both the Boston and the TTC mechanisms respect improvements trivially for the case with $n = 1$.)

[30] Of course, one needs to consider the *conditional* probability that one student is more preferred than another given what happens in the mechanism. This issue is considered in the formal proof in the Appendix.

[31] Note that Theorem 4 does not depend on choosing a particular sequence of school preference profiles as the number of schools grows large. In particular, the proof of Theorem 4 shows that the Boston mechanism does not respect improvements in large markets even when schools' preferences over students are uniformly distributed over all possible rankings of students.

with another student who has a high priority for that school. Such a trade can crowd out a student whose priority is higher than the first student but lower than the second. Thus if an undesirable student changes her preferences to like a school better, it may lead to an inferior outcome for the school as such a crowding out may occur. As in the Boston mechanism, this effect can remain even in large economies.[32] The precise argument, again, can be found in the Appendix.

The intuition of the proofs of **??** 3–5 also illustrates the likely strategic behavior of schools under different school choice mechanisms. When the student-optimal stable mechanism is used, it only behooves a school to discourage a student if that student will begin a rejection chain which ends with another student, whom the school likes better, applying to that school; these students are very hard to identify in practice, and so such strategies by schools will be rare. By contrast, when either the Boston or the TTC mechanism is used, the proofs of the theorems show that a school could benefit by ensuring that students the school finds undesirable do not wish to attend the school; these students are likely easy for the school to identify. These "undesirable" students are often members of the most vulnerable parts of society, and the Boston or TTC mechanisms may induce schools to intentionally make themselves less hospitable for these students.[33]

5 Alternative Criteria

5.1 Respecting Improvements of School Quality in Terms of Enrollment

In the preceding discussion on respecting improvements of school quality, whether a mechanism respects improvements is judged in terms of schools' preferences. This means that we implicitly assume that school preferences in the model are the preferences by which schools evaluate matchings. However, in many real-life school choice systems, school preferences do not necessarily reflect schools' true preferences. Rather, they are often priorities set by law, as is the case for schools in Boston. In such cases, a primary objective of schools is likely to be to enroll as many students as possible. Reasons for this include that school budgets

[32] As in the case of Theorem 4 (see footnote 31), Theorem 5 does not depend on choosing a particular sequence of school preference profiles as the number of schools grows large. In particular, the proof of Theorem 5 shows that the TTC mechanism does not respect improvements in large markets even when schools' preferences over students are uniformly distributed over all possible rankings of students.

[33] However, note that we obtain similar results to **??** 3–5 using the more restrictive criterion of **respecting improvements of school quality for desirable students**, i.e., those students whom the school prefers to some current student: see Section 5.2. Hence, the Boston and the TTC mechanisms may induce schools to make themselves less hospitable for some "desirable" students as well.

are often determined based on enrollments and that schools attended by too few students are often closed.[34] If schools desire to increase enrollment as much as possible, the following variant of our criterion, *respecting improvements of school quality in terms of enrollment*, would be a natural requirement for a mechanism to promote school competition.

Definition 5. A mechanism φ **respects improvements of school quality in terms of enrollment** at the school preference profile $\succ_C$ if, for all $c \in C$ and student preference profiles $\succ_S$ and $\succ'_S$, if $\succ_S$ is an improvement for school c over $\succ'_S$, then $|\varphi_c(\succ_S, \succ_C)| \geq |\varphi_c(\succ'_S, \succ_C)|$.

In other words, a mechanism respects improvements of school quality in terms of enrollment if the enrollment of a school weakly increases whenever that school becomes more preferred by students. Note that respecting improvements in terms of enrollment and the original definition of respecting improvements of school quality are logically independent.

As in the case with the original notion of respecting improvements, we first consider whether stable mechanisms, particularly the student-optimal stable mechanism, respect improvements in terms of enrollment. As shown by the following result, in contrast to Theorem 1, it turns out that any stable mechanism respects improvements in terms of enrollment.

Proposition 1. *Any stable mechanism respects improvements of school quality in terms of enrollment at every school preference profile.*

Proof. See Appendix A.1.4. □

In addition, the next result demonstrates that the Boston mechanism also respects improvements in terms of enrollment.

Proposition 2. *The Boston mechanism respects improvements of school quality in terms of enrollment at every school preference profile.*[35]

Proof. See Appendix A.1.5. □

[34] For example, the Chicago Public Schools' School Closing Guidelines, `http://www.cps.edu/SiteCollectionDocuments/SchoolClosingGuidelines.pdf`, cites under-enrollment as a criterion of school closing. In fact, under-enrollment is often used as a criterion for closing. See, for instance, the recent controversy over the closing of the once-venerable Jamaica High School in New York City partly due to declining enrollment (Daily News, 2011).

[35] In Proposition 2, we implicitly assume that students report true preferences. The Boston mechanism is not strategy-proof, so it is of interest to analyze whether the result holds even when students are strategic. As mentioned in the Remark in Section 3.2, Ergin and Sönmez (2006) show that the set of Nash equilibrium outcomes under the Boston mechanism coincides with the set of stable matchings. By this fact and Proposition 1, it follows that the Boston mechanism also respects improvements of school quality in terms of enrollment when students play Nash equilibria.

Given that all stable mechanisms and the Boston mechanism, a Pareto efficient mechanism for students, respect improvements in terms of enrollment, some readers may suspect that the TTC mechanism, which is also Pareto efficient as well as strategy-proof for students, would satisfy the criterion as well. However, as demonstrated by the following result, the TTC mechanism does *not* necessarily respect improvements of school quality in terms of enrollment.

Proposition 3. *The top trading cycles mechanism does not respect improvements of school quality in terms of enrollment at all school preference profiles.*

Proof. Consider the following environment. There are schools c_1, c_2, c_3, and c_4, and students s_1, s_2, s_3, and s_4. School c_1 has a capacity of 2 seats while each of the other schools has a capacity of 1 seat. The preference profile $\succ$ of students and schools is given by:

$$\begin{array}{ll}
\succ_{s_1} : c_3, c_1, \emptyset, & \succ_{c_1} : s_1, s_2, s_3, s_4, \emptyset, \\
\succ_{s_2} : c_2, c_1, \emptyset, & \succ_{c_2} : s_1, s_2, \ldots, \emptyset \\
\succ_{s_3} : c_3, c_1, \emptyset, & \succ_{c_3} : s_4, s_3, s_2, s_1, \emptyset \\
\succ_{s_4} : c_2, c_4, \emptyset, & \succ_{c_4} : s_4, \ldots, \emptyset.
\end{array}$$

Under this preference profile, the TTC outcome is

$$\begin{pmatrix} c_1 & c_2 & c_3 & c_4 \\ \{s_2, s_3\} & s_4 & s_1 & \emptyset \end{pmatrix},$$

where two positions of c_1 are filled.

Now consider an alternative preference relation of student s_1, $\succ'_{s_1}$: $c_1, c_3, \emptyset$. Note that this is an improvement for school c_1 over $\succ_{s_1}$. However, the TTC outcome under the preference profile $(\succ'_{s_1}, \succ_{-s_1})$ is

$$\begin{pmatrix} c_1 & c_2 & c_3 & c_4 \\ s_1 & s_2 & s_3 & s_4 \end{pmatrix},$$

and so c_1 obtains strictly fewer students. □

The results on respecting improvements in large markets suggest a sense in which stable mechanisms provide better incentives for schools to improve than the TTC mechanism. In addition to that, the results in this section provide another, similar case for stable mechanisms, particularly the student-optimal stable mechanism, in contrast to the TTC mechanism.

5.2 Respecting Improvements of School Quality for Desirable Students

Respecting improvements of school quality requires that a mechanism respects all possible improvements. However, it may be socially beneficial for schools to differentiate and offer different educational experiences to different students; for instance, a school may focus on either math and science, music, or vocational training. If so, then it may be sufficient that a school obtains a (weakly) more preferred set of students when a "desirable" student, i.e., a student with a characteristic that school values, ranks that school more highly.[36] One possible definition of a *desirable* student in this context is simply a student that the school prefers to one of its current students. Hence, we formalize the notion of *respecting improvements of school quality for desirable students* with the following defintion.

Definition 6. A mechanism φ **respects improvements of school quality for desirable students** at the school preference profile $\succ_C$ if the following condition is satisfied: Consider any $c \in C$ and student preference profiles $\succ_S$ and $\succ'_S$ such that

(1) $\succ'_S$ is an improvement for school c over $\succ_S$, and

(2) if $|\varphi_c(\succ_S, \succ_C)| = q_c$, for any s such that $\bar{s} \succ_c s$ for every $\bar{s} \in \varphi_c(\succ_S, \succ_C)$, $\succ'_s$ is the same as $\succ_s$.

Then, $\varphi_c(\succ'_S, \succ_C) \succsim_c \varphi_c(\succ_S, \succ_C)$ holds.

In the above definition, we consider a change of student preferences where a school improves in the ranking of students, each of whom is preferred to a current student, while remaining unchanged in other students' rankings. We say that a mechanism respects improvements of school quality for desirable students if the school always obtains a weakly better set of students as a result of such a change.[37] Clearly, if a mechanism respects improvements of school quality, then it also respects improvements for desirable students. In this sense, respecting improvements for desirable students is a weaker notion than respecting improvements.

Even if we adopt this alternative criterion, the impossibility result for the compatibility of stability and respecting improvements in general markets continues to hold: In Example 1, add another student $\hat{s}$ with $\succ_{\hat{s}}$: $c, \emptyset$ and change school preferences to $\succ_c$: $s, \bar{s}, \hat{s}, \emptyset$ and $\succ_{\bar{c}}$: $\bar{s}, s, \hat{s}, \emptyset$. This modified example shows the desired impossibility.

[36] Here we assume that school preferences truly reflect their intrinsic preferences, as opposed to being priorities set by law.

[37] For schools with unfilled capacity, every student is considered to be a desirable student.

Furthermore, neither the Boston nor the TTC mechanisms respects improvements for desirable students in general markets. For the Boston mechanism, consider the following example:

Example 3. Let $S = \{s_1, s_2, s_3, s_4\}, C = \{c_1, c_2\}$. The capacity of school c_1 is 2 while the capacity of school c_2 is 1. Preferences of students and schools are as follows:

$$\begin{aligned} &\succ_{s_1} : c_2, c_1, \emptyset, && \succ_{c_1} : s_1, s_2, s_3, s_4, \emptyset, \\ &\succ_{s_2} : \emptyset, c_1, c_2, && \succ_{c_2} : s_4, s_1, s_2, s_3, \emptyset, \\ &\succ_{s_3} : c_1, \emptyset, c_2, \\ &\succ_{s_4} : c_2, \emptyset, c_1. \end{aligned}$$

Under this preference profile $\succ$, the Boston mechanism φ^B produces the following matching:

$$\varphi^B(\succ) = \begin{pmatrix} c_1 & c_2 & \emptyset \\ \{s_1, s_3\} & s_4 & s_2 \end{pmatrix}.$$

Now consider an alternative preference relation for student s_2, $\succ'_{s_2}$: $c_1, \emptyset, c_2$. Note that this is an improvement for school c_1 over $\succ_{s_2}$ and $s_2 \succ_{c_1} s_3 \in \varphi^B(\succ)$. However, the Boston mechanism outcome under preference profile $(\succ'_{s_2}, \succ_{-s_2})$ is

$$\varphi^B(\succ'_{s_2}, \succ_{-s_2}) = \begin{pmatrix} c_1 & c_2 & \emptyset \\ \{s_2, s_3\} & s_4 & s_1 \end{pmatrix},$$

which is strictly worse for c_1 than $\varphi^B(\succ)$. Hence, the Boston mechanism does not respect improvements for desirable students.

In order to show that the TTC mechanism does not respect improvements for desirable students, consider the counterexample in the proof of Proposition 3 and suppose that $\{s_2, s_3\} \succ_{c_1} \{s_1\}$. (Note that this does not contradict the assumption that the preferences of c_1 are responsive.) This modfied counterexample provides a school preference profile at which the TTC mechanism does not respect improvements for desirable students.

Given these negative results, a natural question is, as in the case of the original concept of respecting improvements, whether these mechanisms respect improvements for desirable students in large markets. First of all, it is clear that the result for stable mechanisms in large markets remains true since respecting improvements (for all students) implies respecting improvements for desirable students. For the Boston and the TTC mechanisms, as in the case with our original criterion of respecting improvements of school quality, we show that

neither of them respects improvements for desirable students even in large markets. Let $\hat{\alpha}_c(\tilde{\Gamma}, \varphi)$ be the probability that the realized preference profile $\succ$ has the property that there exists a student preference profile $\succ'_S$ such that $\succ'_S$ is a disimprovement for c over $\succ_S$ with the properties (1) and (2) in Definition 6, and $\varphi_c(\succ'_S, \succ_C) \succ_c \varphi_c(\succ_S, \succ_C)$. We say that a mechanism φ **approximately respects improvements of school quality for desirable students in large markets** if, for any sequence of random markets $(\tilde{\Gamma}^n)_{n\in\mathbb{N}}$ that is regular and sufficiently thick, $\max_{c\in C^n} \hat{\alpha}_c(\tilde{\Gamma^n}, \varphi) \to 0$ as $n \to \infty$, i.e., for any $\varepsilon > 0$, there exists an integer m such that, for any random market $\tilde{\Gamma}^n$ in the sequence with $n > m$ and any $c \in C^n$, we have that $\hat{\alpha}_c(\tilde{\Gamma}^n, \varphi) < \varepsilon$.

Proposition 4. *The Boston mechanism does not approximately respect improvements of school quality for desirable students in large markets.*

Proof. See Appendix A.1.6. □

Proposition 5. *The top trading cycles mechanism does not approximately respect improvements of school quality for desirable students in large markets.*

Proof. See Appendix A.1.7. □

Hence, even if we adopt the alternative, weaker criterion of respecting improvements for desirable students, the implications obtained by using our original criterion, respecting improvements of school quality (for all students), are unchanged.[38]

6 Discussion

6.1 Conditions on Preferences for Respecting Improvements

As mentioned in Section 3.3, given that the mechanisms we consider do not respect improvements at every school preference profile, a natural question is under what conditions on the school preference profile $\succ_C$ does a stable or Pareto efficient mechanism respect improvements. Let $r^\ell(\succ_c)$ be the student who is ℓ-th ranked in $\succ_c$.

Definition 7. A school preference profile $\succ_C$ is **virtually homogeneous** if $r^\ell(\succ_c) = r^\ell(\succ_{\bar{c}})$ for all $c, \bar{c} \in C$ and $\ell > \min\{q_{\hat{c}} : \hat{c} \in C\}$.

[38]Similarly to Theorems 4 and 5, the proofs for Propositions 4 and 5 show that the failure of respecting improvements for desirable students occurs not only for large markets, but for any market with the number of schools $n \geq 2$. Furthermore, as in the case of Theorems 4 and 5 (see footnotes 31 and 32), Propositions 4 and 5 do not depend on choosing a particular sequence of school preference profiles as the number of schools grows large.

This condition requires that the same student should be the ℓ-th preferred student for all schools for every ℓ that is larger than the minimum of school capacities. As the name suggests, virtual homogeneity allows for almost no variation in preferences over individual students among different schools. To illustrate this condition, consider a special case in which each school has only one seat, that is, $q_{\hat{c}} = 1$ for all $\hat{c} \in C$. Then $r^{\ell}(\succ_c) = r^{\ell}(\succ_{\bar{c}})$ for all $c, \bar{c} \in C$ and $\ell \geq 2$ and hence for $\ell = 1$ as well. This means that preferences over students are exactly identical between any pair of schools.

When school capacities are larger than one, virtual homogeneity allows for slight variations in school preferences. Still, any allowed variation involves only the top $\min\{q_{\hat{c}} : \hat{c} \in C\}$ students. Such a student is admitted to any school whenever she applies to it in any stable mechanism, so how highly she is ordered within those top students does not affect the allocation as long as a stable mechanism is employed. In other words, the apparent heterogeneity in school preferences involving only the top $\min\{q_{\hat{c}} : \hat{c} \in C\}$ students is irrelevant for the purpose of choosing an allocation from the set of stable allocations.[39] We now characterize the set of school preference profiles under which there exists a stable mechanism that respects improvements of school quality.

Proposition 6. *There exists a stable mechanism that respects improvements of school quality at $\succ_C$ if and only if one of the following conditions is satisfied:*

(1) *The school preference profile $\succ_C$ is virtually homogeneous.*

(2) *For every school $c \in C$, the capacity q_c (associated with $\succ_c$) is one.*

Proof. See Appendix A.1.8. □

While the proposition provides a complete characterization of when a stable mechanism respects improvements of school quality, the main significance of this result is the necessity direction: Virtually homogenous preferences are necessary for a stable mechanism to respect improvements of school quality (when at least one school has a capacity greater than one). Given that virtual homogeneity is an extremely restrictive condition which is rarely satisfied in practice, this result suggests that school preferences in practice are unlikely to exclude the possibility that stable mechanisms may provide perverse incentives for schools to lower their qualities and divert some students' demand for those schools.

The proof of Proposition 6 is quite involved, but the intuition is straightforward: If preferences are not virtually homogenous and at least one school has a capacity greater than one, then with some work one can construct a preference profile of the students such as

[39] On the other hand, however, the variation in school preferences over students may affect school preferences over allocations even when preferences are virtually homogeneous.

that in Example 1. On the other hand, when preferences are virtually homogenous, any stable mechanism is equivalent to a serial dictatorship where students choose school seats in the order that they are preferred by the schools. Such a serial dictatorship clearly respects improvements of school quality. Finally, if the capacity is one for every school, then the school-optimal stable mechanism is the unique stable mechanism that respects improvements of school quality by Theorem 5 in Balinski and Sönmez (1999).

We now show that the set of preference profiles for which a Pareto efficient mechanism respects improvements of school quality is very similar to the set of preference profiles for which a stable mechanism respects improvements of school quality, which is specified in Proposition 6.

Proposition 7. *There exists a mechanism that is Pareto efficient for students and respects improvement of school quality at $\succ_C$ if and only if $\succ_C$ is virtually homogeneous.*

Proof. See Appendix A.1.9. □

As virtual homogeneity is a very strong restriction on school preferences, the significant part of Proposition 7 is that virtually homogenous preferences are required for a Pareto efficient mechanism to respect improvements of school quality. This conclusion implies that Pareto efficient mechanisms may very often provide perverse incentives for schools to lower their qualities and divert some students' demand for those schools.

The proof of Proposition 7 is similar to that of Proposition 6 in spirit, though the technical details differ substantially: if preferences are not virtually homogenous, then it is possible to construct a preference profile for the students such as that in Example 2. On the other hand, when preferences are virtually homogenous, the serial dictatorship where students choose in the order that they are preferred by the schools is both Pareto efficient and respects improvements of school quality.

Remark. When the virtual homogeneity condition in Proposition 7 is satisfied, the TTC mechanism is an example of a mechanism that is Pareto efficient for students and respects improvements of school quality. If a school preference profile is virtually homogeneous, the TTC mechanism coincides with a serial dictatorship using an arbitrary school's preference profile as the priority order. As explained above, such a serial dictatorship respects improvements of school quality. On the other hand, the Boston mechanism does not respect improvements even when $\succ_C$ is virtually homogeneous.[40] An implication of these results is that the student-optimal stable mechanism respects improvements for a wider class of school preference profiles than the TTC mechanism, and the TTC mechanism respects improvements for a wider class of school preference profiles than the Boston mechanism.

[40]For an example showing this point, see Appendix A.2.

Remark. Virtual homogeneity is stronger than *acyclicity* by Ergin (2002) and all of its variants proposed in the literature: *(strong) x-acyclicity* by Haeringer and Klijn (2009), the stronger notions of acyclicity by Kesten (2006), and *essential homogeneity* by Kojima (2011). Note that even these acyclicity-like conditions have been considered to be so restrictive that it seems difficult to find any real-life cases where the conditions are satisfied. This fact demonstrates how restrictive virtual homogeneity is. For a more detailed explanation on the relationship between virtual homogeneity and (the variants of) acyclicity, see Appendix A.3.

6.2 Respecting Improvements of Student Quality

While we have considered competitive pressures on schools to improve, it is also important that a student not have incentives to make schools rank her lower in order to obtain a more preferred school. In addition, it would be natural to suspect that there is a tradeoff between providing incentives for schools to improve and doing so for students. In this section, we consider whether the school choice mechanisms considered in this work respect improvements of student quality.

Definition 8. A mechanism φ **respects improvements of student quality** at the student preference profile $\succ_S$ if, for all $s \in S$ and school preference profiles $\succ_C$ and $\succ'_C$, if $\succ'_C$ is an improvement for student s over $\succ_C$, then $\varphi_s(\succ_S, \succ'_C) \succsim_s \varphi_s(\succ_S, \succ_C)$.[41]

This definition is analogous to that for respecting improvements of school quality. A mechanism respects improvements of student quality if whenever a student's ranking improves in schools' preferences, that student obtains a weakly better placement. We now show that the student optimal stable mechanism, the Boston mechanism, and the TTC mechanism all respect improvements of student quality.

In addition to being a building block for Theorem 3, Lemma 1 allows us to easily prove the following corollary, which was first shown by Balinski and Sönmez (1999).

Corollary 1. *The unique stable mechanism that respects improvements of student quality at every student preference profile is the student-optimal stable mechanism.*

Proof. By Lemma 1, a stable mechanism respects improvements of student quality if and only if it is strategy-proof for students. This fact, and the result by Alcalde and Barberà

[41] Analogously to the definition of an improvement for a school, a preference profile $\succ_C$ is an improvement for student s over preference profile $\succ'_C$ if, for all $c \in C$,

(1) For all $\bar{s} \in S$, if $\{s\} \succ_c \{\bar{s}\}$, then $\{s\} \succ'_c \{\bar{s}\}$, and

(2) For all $\bar{s}, \hat{s} \in S \setminus \{s\}$, $\{\bar{s}\} \succ'_c \{\hat{s}\}$ if and only if $\{\bar{s}\} \succ'_c \{\hat{s}\}$,

and the capacity associated with $\succ'_c$ is equal to that with $\succ_c$.

(1994) that the student-optimal stable mechanism is the only stable mechanism that is strategy-proof for students, complete the proof. □

The Boston mechanism also respects improvements of student quality. Intuitively, when a student improves his ranking, at each step of the algorithm in the Boston mechanism, the student is more likely to be kept by the school. Hence the outcome for the student must become weakly better when the student's ranking improves.

Proposition 8. *The Boston mechanism respects improvements of student quality at every student preference profile.*

Proof. See Appendix A.1.10. □

The TTC mechanism also respects improvements of student quality. At each step of the algorithm in the TTC mechanism, a school is more likely to point at a student if that student is ranked higher. Hence, at each step of the algorithm, a higher-ranked student will have more schools pointing (directly or indirectly) at her, and so she will have a greater set of schools to choose from, and therefore obtain a weakly better outcome.

Proposition 9. *The top trading cycles mechanism respects improvements of student quality at every student preference profile.*

Proof. See Appendix A.1.11. □

7 Conclusion

In this work, we considered how the design of a school choice mechanism affects the incentives of schools to improve their educational quality. We first defined the concept of respecting improvements of school quality, which requires that the outcome of a mechanism becomes weakly better for a school whenever that school becomes more preferred by students.[42] No stable mechanism (such as the student-optimal stable mechanism) or mechanism that is Pareto efficient for students (such as the Boston and the TTC mechanisms) respect improvements of school quality. However, as the size of the school district grows, any stable mechanism approximately respects improvements; in contrast, the Boston and the TTC

[42] Of course, one could also reward schools directly based on students' submitted preferences, such as by providing a monetary reward to a school based on the number of students who rank it first. However, we do not analyze this issue and instead focus on criteria based on the actual student placement. Even were school districts to implement such a direct incentive scheme, our analysis remains applicable as schools would still have preferences over students.

	SOSM	Boston	TTC
RI in General Markets	×	×	×
RI by Desirable Students in General Markets	×	×	×
RI in Large Markets	✓	×	×
RI for Desirable Students in Large Markets	✓	×	×
RI in Terms of Enrollment	✓	✓	×

Table 1: Summary of the Main Results. RI stands for respect improvements.

mechanisms do not even approximately respect improvements in large markets. Similar conclusions were obtained with respect to other criteria: Respecting improvements in terms of enrollment and for desirable students. The main results are summarized in Table 1 (an exhaustive list of our results is in Table 2 in Appendix A.4). These results suggest that the student-optimal stable mechanism may be a better school choice mechanism compared to the Boston and the TTC mechanisms if the goal of public school choice is to "increase excellence by increasing choice" (National Governors' Association, 1991).

We regard this paper as one of the first attempts to use the analytical tools of market design to study the effects of different school choice mechanisms on improving the quality of public schooling. As such, there are a number of promising avenues of future research. First, if data on submitted preferences in real school choice systems is available, it would be possible to analyze how often schools in practice are better off when less preferred by certain students, i.e., how often schools have incentives to discourage student interest. Second, and more ambitiously, empirical work could quantify the effect of different school choice mechanisms on the quality of a public school system and its rate of improvement. We would further suggest that empirical work in this area also study how different school choice mechanisms affect different types of students: As discussed at the end of Section 4, the Boston and the TTC mechanisms provide incentives for schools to make themselves less attractive to "less desirable" students. As these "less desirable" students are likely to be students who are already low-achieving, members of a disadvantaged minority group, or have special needs, the use of the Boston and the TTC mechanisms may further disadvantage these students.

Another important research direction would be to relate the current study, which focuses on public school choice, with other forms of school choice, such as vouchers and charter school systems.[43] Potentially fruitful questions include the following ones: Which system provides the best incentives for schools to improve? How does the form of school competition affect the

[43] Although charter school admissions could be integrated into the public school choice programs in principle, admissions in charter schools are usually operated independently from other public schools in the United States.

quality of the educational experience for different students? What sort of mechanisms should be used to allocate students to charter schools and/or schools accepting vouchers? Answering these questions will require a much more stylized model than the current general matching-theoretic one, but we believe that answering these questions is crucial to the continuing debate over public education.

References

Abdulkadiroğlu, A., P. A. Pathak, and A. E. Roth (2005): "The New York City High School Match," *American Economic Review Papers and Proceedings*, 95, 364–367.

——— (2009): "Strategy-proofness versus efficiency in matching with indifferences: Redesigning the NYC high school match," *The American Economic Review*, 99(5), 1954–1978.

Abdulkadiroğlu, A., P. A. Pathak, A. E. Roth, and T. Sönmez (2005): "The Boston Public School Match," *American Economic Review Papers and Proceedings*, 95, 368–372.

——— (2006): "Changing the Boston School Choice Mechanism," Harvard University, Unpublished mimeo.

Abdulkadiroğlu, A., and T. Sönmez (1998): "Random Serial Dictatorship and the Core from Random Endowments in House Allocation Problems," *Econometrica*, 66, 649–701.

——— (2003): "School Choice: A Mechanism Design Approach," *American Economic Review*, 93, 729–747.

Alcalde, J. (1996): "Implementation of stable solutions to marriage problems," *Journal of Economic Theory*, 69(1), 240–254.

Alcalde, J., and S. Barberà (1994): "Top dominance and the possibility of strategy-proof stable solutions to matching problems," *Economic theory*, 4(3), 417–435.

Ashlagi, I., M. Braverman, and A. Hassidim (2011): "Matching with Couples in Large Markets Revisited," Unpublished mimeo, MIT Sloan School.

Baïou, M., and M. Balinski (2000): "Many-to-many matching: stable polyandrous polygamy (or polygomous polyandry)," *Discrete Applied Mathematics*, 101, 1–12.

Balinski, M., and T. Sönmez (1999): "A tale of two mechanisms: student placement," *Journal of Economic Theory*, 84, 73–94.

Blum, Y., A. Roth, and U. Rothblum (1997): "Vacancy Chains and Equilibration in Senior-Level Labor Markets," *Journal of Economic Theory*, 76, 362–411.

Chubb, J. E., and T. M. Moe (1990): *Politics, markets, and America's schools.* Brookings Institution Press.

Daily News (2011): "Politicians charge that Jamaica High School's being neglected, vow to fight closure," January 05, 2011.

Dubins, L. E., and D. A. Freedman (1981): "Machiavelli and the Gale-Shapley algorithm," *American Mathematical Monthly*, 88, 485–494.

Ergin, H. (2002): "Efficient Resource Allocation on the Basis of Priorities," *Econometrica*, 88, 485–494.

Ergin, H., and T. Sönmez (2006): "Games of School Choice under the Boston Mechanism," *Journal of Public Economics*, 90, 215–237.

Gale, D., and L. S. Shapley (1962): "College admissions and the stability of marriage," *American Mathematical Monthly*, 69, 9–15.

Gale, D., and M. A. O. Sotomayor (1985a): "Ms. Machiavelli and the Stable Matching Problem," *Discrete Applied Mathematics*, 11, 223–232.

——— (1985b): "Some remarks on the stable matching problem," *American Mathematical Monthly*, 92, 261–268.

Haeringer, G., and F. Klijn (2009): "Constrained school choice," *Journal of Economic Theory*, 144(5), 1921–1947.

Hatfield, J. W., F. Kojima, and Y. Narita (2011): "Many-to-many Matching with Max-Min Preferences," Unpublished mimeo.

Hatfield, J. W., and P. Milgrom (2005): "Matching with Contracts," *American Economic Review*, 95, 913–935.

Immorlica, N., and M. Mahdian (2005): "Marriage, Honesty, and Stability," *SODA 2005*, pp. 53–62.

Kesten, O. (2006): "On Two Competing Mechanisms for Priority-Based Allocation Problems," *Journal of Economic Theory*, 127, 155–171.

Kojima, F. (2011): "Efficient Resource Allocation under Multi-Unit Demand," mimeo.

Kojima, F., and P. Pathak (2009): "Incentives and stability in large two-sided matching markets," *The American Economic Review*, 99(3), 608–627.

Kojima, F., P. Pathak, and A. Roth (2011): "Matching with couples: Stability and incentives in large markets," Discussion paper, National Bureau of Economic Research.

Manea, M. (2009): "Asymptotic ordinal inefficiency of random serial dictatorship," *Theoretical Economics*, 4(2), 165–197.

McVitie, D. G., and L. Wilson (1970): "Stable marriage assignments for unequal sets," *BIT*, 10, 295–309.

Miralles, A. (2009): "School choice: The case for the Boston mechanism," mimeo.

Moe, T. (2008): "Beyond the free market: The structure of school choice," *BYU Law Review*, p. 557.

National Governors' Association (1991): "Time for Results," 1991 National Governors' Association Report.

Pathak, P. (2011): "The mechanism design approach to student assignment," *Annual Review of Economics*, 3(1).

Pathak, P., and T. Sönmez (2011): "School Admissions Reform in Chicago and England: Comparing Mechanisms by their Vulnerability to Manipulation," mimeo.

Roth, A. E. (1982): "The Economics of Matching: Stability and Incentives," *Mathematics of Operations Research*, 7, 617–628.

——— (1984): "The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory," *Journal of Political Economy*, 92, 991–1016.

——— (1985): "The college admission problem is not equivalent to the marriage problem," *Journal of Economic Theory*, 36, 277–288.

——— (1986): "On the allocation of residents to rural hospitals: a general property of two-sided matching markets," *Econometrica*, 54, 425–427.

——— (2008): "Deferred Acceptance Algorithms: History, Theory, Practice, and Open Questions," *International Journal of Game Theory*, 36, 537–569.

Roth, A. E., and E. Peranson (1999): "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design," *American Economic Review*, 89, 748–780.

Roth, A. E., and M. O. Sotomayor (1990): *Two-sided matching: a study in game-theoretic modeling and analysis*. Econometric Society monographs, Cambridge.

Shapley, L., and H. Scarf (1974): "On cores and indivisibility," *Journal of mathematical economics*, 1(1), 23–37.

Sönmez, T., and T. Switzer (2011): "Matching with (Branch-of-Choice) Contracts at United States Military Academy," .

Sönmez, T., and U. Ünver (2009): "Matching, Allocation, and Exchange of Discrete Resources," forthcoming, Handbook of Social Economics, eds. Jess Benhabib, Alberto Bisin, and Matthew Jackson, Elsevier.

A Appendix

A.1 Proofs

A.1.1 Proof of Theorem 3

We begin by proving Lemma 1, which we restate here for convenience. This lemma shows that a school would benefit when some students like it less if and only if that school has a profitable misreporting of preferences.

Lemma 1. *Let φ be a stable mechanism.*

(1) *Suppose that the preference profile $\succ$ and student preference profile $\succ'_S$ are such that $\succ'_S$ is a disimprovement for some $c \in C$ over $\succ_S$ and $\varphi_c(\succ'_S, \succ_C) \succ_c \varphi_c(\succ)$. Then there exists a (reported) preference relation $\succ''_c$ for c such that $\varphi_c(\succ''_c, \succ_{-c}) \succ_c \varphi_c(\succ)$.*

(2) *Suppose that the preference profile $\succ$ and (reported) preference relation $\succ''_c$ for some $c \in C$ are such that $\varphi_c(\succ''_c, \succ_{-c}) \succ_c \varphi_c(\succ)$. Then there exists a student preference profile $\succ'_S$ such that $\succ'_S$ is a disimprovement for c over $\succ_S$ and $\varphi_c(\succ'_S, \succ_C) \succ_c \varphi_c(\succ)$.*

Proof. We prove each part in the order they are listed:

(1) Suppose $\varphi_c(\succ'_S, \succ_C) \succ_c \varphi_c(\succ)$. Consider a preference relation $\succ''_c$ of school $c \in C$ such that $s \succ''_c \emptyset$ if and only if $s \in \varphi_c(\succ'_S, \succ_C)$. Then

Claim 1. $\varphi(\succ'_S, \succ_C)$ is stable under $(\succ''_c, \succ_{-c})$.

Proof. It is obvious that $\varphi(\succ'_S, \succ_C)$ is individually rational at $(\succ'_c, \succ_{-c})$. To show that there is no blocking pair of $\varphi(\succ'_S, \succ_C)$ at $(\succ''_c, \succ_{-c})$, consider the following cases.

(a) There are no blocking pairs of the form (s, c), that is, blocking pairs involving school c, because $\emptyset \succ''_c s$ for any $s \notin \varphi_c(\succ'_S, \succ_C)$ by construction of preference relation $\succ''_c$.

(b) Suppose that there is a blocking pair $(s, \bar{c})$ at $(\succ''_c, \succ_{-c})$ with $\bar{c} \neq c$ and $s \in \varphi_c(\succ'_S, \succ_C)$. Then $\bar{c} \succ_s c$ and, since $(\succ'_S, \succ_C)$ is a disimprovement for school c over $\succ$, it follows that $\bar{c} \succ'_s c$. This and the fact that $(s, \bar{c})$ is a blocking pair of $\varphi(\succ'_S, \succ_C)$ at $(\succ''_c, \succ_{-c})$ implies that $(s, \bar{c})$ is a blocking pair of $\varphi(\succ'_S, \succ_C)$ at $(\succ'_S, \succ_C)$, which is a contradiction to the assumption that φ is a stable mechanism.

(c) Suppose that there is a blocking pair $(s, \bar{c})$ at $(\succ''_c, \succ_{-c})$ with $\bar{c} \neq c$ and $s \notin \varphi_c(\succ'_S, \succ_C)$. Then, $\bar{c} \succ_s \varphi_s(\succ'_S, \succ_C)$ if and only if $\bar{c} \succ'_s \varphi_s(\succ'_S, \succ_C)$ by definition of

a disimprovement, $\bar{c} \neq c$, and $\varphi_s(\succ'_S, \succ_C) \neq c$. Also, the preferences of $\bar{c}$ are identical under $(\succ'_S, \succ_C)$ and $(\succ'_c, \succ_{-c})$. Therefore $(s, \bar{c})$ is a blocking pair of $\varphi(\succ'_S, \succ_C)$ at $(\succ'_S, \succ_C)$, which is a contradiction to the assumption that φ is a stable mechanism.

This completes the proof of Claim 1. □

Now note that by a version of the rural hospital theorem (McVitie and Wilson, 1970; Roth, 1984, 1986; Gale and Sotomayor, 1985a,b) and Claim 1, we have that

$$|\varphi_c(\succ''_c, \succ_{-c})| = |\varphi_c(\succ'_S, \succ_C)|.$$

But since $s \succ''_c \varnothing$ if and only if $s \in \varphi_c(\succ'_S, \succ_C)$ by construction of $\succ'_c$, this equality implies that

$$\varphi_c(\succ''_c, \succ_{-c}) = \varphi_c(\succ'_S, \succ_C).$$

This relation and the hypothesis that $\varphi_c(\succ'_S, \succ_C) \succ_c \varphi_c(\succ)$ completes the proof of Part 1 of Lemma 1.

(2) Suppose $\varphi_c(\succ''_c, \succ_{-c}) \succ_c \varphi_c(\succ)$. Consider a student preference profile $\succ'_S$ such that preferences of students outside $\varphi_c(\succ''_c, \succ_{-c})$ drop school c from their list but preferences of all other students are unchanged: Formally, define $\succ'_S$ by

(a) For any $s \in S \setminus \varphi_c(\succ''_c, \succ_{-c})$, (i) $\emptyset \succ'_s c$ and (ii) $\bar{c} \succ'_s \hat{c} \iff \bar{c} \succ_s \hat{c}$ for any $\bar{c}, \hat{c} \in C \cup \{\emptyset\} \setminus \{c\}$.

(b) $\succ'_s = \succ_s$ for any $s \in \varphi_c(\succ''_c, \succ_{-c})$.

Claim 2. $\varphi(\succ''_c, \succ_{-c})$ is stable under $(\succ'_S, \succ_C)$.

Proof. It is obvious that $\varphi(\succ''_c, \succ_{-c})$ is individually rational at $(\succ'_S, \succ_C)$. To show that there is no blocking pair of $\varphi(\succ''_c, \succ_{-c})$ at $(\succ'_S, \succ_C)$, consider the following cases.

(a) There are no blocking pairs of the form (s, c), that is, blocking pairs involving school c, because $\emptyset \succ'_s c$ for any $s \in S \setminus \varphi_c(\succ''_c, \succ_{-c})$ by construction of the preference relation $(\succ'_S, \succ_C)$.

(b) Suppose that there is a blocking pair $(s, \bar{c})$ at $(\succ'_S, \succ_C)$ with $\bar{c} \neq c$ and $s \in \varphi_c(\succ''_c, \succ_{-c})$. Then $\bar{c} \succ_s c$ and, since $\succ'_s$ is identical to $\succ_s$ by construction, we obtain $\bar{c} \succ'_s c$. This and the fact that $(s, \bar{c})$ is a blocking pair of $\varphi(\succ''_c, \succ_{-c})$ at $(\succ'_S, \succ_C)$ implies that $(s, \bar{c})$ is a blocking pair of $\varphi(\succ''_c, \succ_{-c})$ at $(\succ''_c, \succ_{-c})$, which is a contradiction to the assumption that φ is a stable mechanism.

(c) Suppose that there is a blocking pair $(s, \bar{c})$ at $(\succ'_S, \succ_C)$ with $\bar{c} \neq c$ and $s \notin \varphi_c(\succ''_c, \succ_{-c})$. Then, $\bar{c} \succ_s \varphi_s(\succ''_c, \succ_{-c})$ if and only if $\bar{c} \succ'_s \varphi_s(\succ''_c, \succ_{-c})$ by definition of a disimprovement, $\bar{c} \neq c$, and $\varphi_s(\succ''_c, \succ_{-c}) \neq c$. Also, the preferences of $\bar{c}$ are identical under $(\succ'_S, \succ_C)$ and $(\succ''_c, \succ_{-c})$. Therefore $(s, \bar{c})$ is a blocking pair of $\varphi_c(\succ''_c, \succ_{-c})$ at $(\succ''_c, \succ_{-c})$, which is a contradiction to the assumption that φ is a stable mechanism.

This completes the proof of Claim 2. □

Now note that by a version of the rural hospital theorem (McVitie and Wilson, 1970; Roth, 1984, 1986; Gale and Sotomayor, 1985a,b) and Claim 2, we have

$$|\varphi_c(\succ'_S, \succ_C)| = |\varphi_c(\succ''_c, \succ_{-c})|.$$

But since $c \succ'_s \emptyset$ if and only if $s \in \varphi_c(\succ''_c, \succ_{-c})$ by construction of $(\succ'_S, \succ_C)$, this equality implies that

$$\varphi_c(\succ'_S, \succ_C) = \varphi_c(\succ''_c, \succ_{-c}).$$

This relation and the hypothesis that $\varphi_c(\succ''_c, \succ_{-c}) \succ_c \varphi_c(\succ)$ completes the proof of Part 2 of Lemma 1.

□

We now prove Theorem 3; to do so, we now show that, when some stable mechanism does not respect improvements for a certain preference profile, the student-optimal stable mechanism does not respect improvements for that preference profile.

Claim 3. Let φ be a stable mechanism and suppose that the preference profile $\succ$ has the property that there exists another preference profile $(\succ'_S, \succ_C)$ such that $(\succ'_S, \succ_C)$ is a disimprovement over $\succ$ for c while $\varphi_c(\succ'_S, \succ_C) \succ_c \varphi_c(\succ)$. Then there exists a preference relation $\succ^*_c$ of c such that $\varphi^S_c(\succ^*_c, \succ_{-c}) \succ_c \varphi^S_c(\succ)$.

Proof. By Lemma 1, there exists $\succ''_c$ such that $\varphi_c(\succ''_c, \succ_{-c}) \succ_c \varphi_c(\succ)$. Pathak and Sönmez (2011) show that, if a stable mechanism φ is manipulable by a school at a given preference profile of students and schools, then the student-optimal stable mechanism φ^S is manipulable by the same school at the same preference profile. Thus, there exists $\succ^*_c$ (which may be different from $\succ''_c$) such that $\varphi^S_c(\succ^*_c, \succ_{-c}) \succ_c \varphi^S_c(\succ)$. □

To prove the theorem, suppose that the preference profile $\succ$, realized from random market $\tilde{\Gamma}^n$, has the property that there exists another preference profile $(\succ'_S, \succ_C)$ such that $(\succ'_S, \succ_C)$

is a disimprovement over $\succ$ for c while $\varphi_c(\succ'_S, \succ_C) \succ_c \varphi_c(\succ)$. Then by Claim 3, there exists a preference relation $\succ^*_c$ of c such that $\varphi^S_c(\succ^*_c, \succ_{-c}) \succ_c \varphi^S_c(\succ)$.

Under the assumptions of regularity and sufficient thickness, Lemmata 1, 3, and 10 of Kojima and Pathak (2009) imply that there exists a constant γ and T such that the following property holds: There exists n_0 such that, for any $\tilde{\Gamma}^n$ with any $n > n_0$ and any $c \in C^n$, the probability that, under the realized preference profile $\succ$, there exists a reported preference relation $\succ^*_c$ such that $\varphi^S_c(\succ^*_c, \succ_{-c}) \succ_c \varphi^S_c(\succ)$ is at most $\frac{\gamma}{E[V_T(n)]}$. By the sufficient thickness assumption, $E[V_T(n)] \to \infty$ for any sufficiently large T.[44] This fact and the conclusion from the last paragraph complete the proof.

Remark. From the last part of the proof, it is clear that the order of convergence in the theorem is $O(\frac{1}{E[V_T(n)]})$. For instance, if the sequence of random markets satisfies moderate similarity as defined in Section 4.1, then the order of convergence is $O(\frac{1}{n})$ because $E[V_T(n)] = O(n)$ by Proposition 1 of Kojima and Pathak (2009).[45]

Remark. As mentioned in Section 4.1, the conclusion of the theorem holds even without condition (1) of regularity—so students can find any number of schools acceptable—if the conditions of excess supply of school capacities and moderate similarity are satisfied. To see this point, note first that the proof of Lemma 5 of Kojima and Pathak (2009) shows that, given $V_T(n)$, the conditional probability that a school can profitably manipulate φ^S is $O(\frac{1}{V_T(n)})$. Under the conditions of excess supply of school capacities and moderate similarity, it is clear that $V_T(n) = O(n)$ (for any sufficiently large T) for any realization of preferences because there are at least λn vacant school seats, and hence at least $(\frac{\lambda}{\hat{q}})n$ schools with at least one vacant seat, in any matching. Thus the (unconditional) probability that a school can profitably manipulate φ^S is $O(\frac{1}{n})$. This and the arguments of the above proof establish the conclusion of the theorem.

A.1.2 Proof of Theorem 4

Theorem 4 follows from Proposition 4 as the latter implies the former. However, we present a separate proof, which is simpler than the proof of Proposition 4, for expositional clarity.

Consider a sequence of random markets where there are n schools and $2n$ students, and the capacity of every school is 1. Assume that preferences of all students are generated

[44] Note that condition (i) in the definition of regularity of a sequence of random markets is weaker than that used by Kojima and Pathak (2009) in that they require that $k^n = k$ for all n. It is easy to extend their result to our more general setting, as claimed in footnote 3 in Kojima and Pathak (2009).

[45] Kojima and Pathak (2009) show $E[V_T(n)] = O(n)$ for a slightly more general class of distributions, which they call "nonvanishing proportion of popular colleges" in their Appendix A.3. Moderate similarity corresponds to the special case with $a = 1$ in their class of distributions.

according to the procedure described in Section 4.1 associated with the uniform distribution over all schools and $k = 2$. Moreover assume that school preferences over individual students are drawn identically and independently from the uniform distribution over all preferences for students such that all students are acceptable. These assumptions guarantee that the regularity and sufficient thickness conditions are satisfied.

Given n, fix an arbitrary school c and let Event 1 be the event that there is exactly one student who prefers that school c most. The probability of Event 1 is

$$\binom{2n}{1} \times \frac{1}{n} \times \left(1 - \frac{1}{n}\right)^{2n-1} = 2n \times \frac{1}{n} \times \left(1 - \frac{1}{n}\right)^{2n-1}.$$

This expression converges to $\frac{2}{e^2}$ as n approaches infinity, where e is the basis of the natural logarithm.[46] Therefore, for any sufficiently large n, the probability of Event 1 is at least, say, $\frac{1}{e^2}$. Denote by s the unique student who prefers c most.

Since there are $2n$ students and n school seats, there are at least n students who are not matched in the first step of the algorithm of the Boston mechanism. Since $k = 2$, that is, each student finds two schools to be acceptable, each of these students applies to a school in the second step of the algorithm. Therefore, the conditional probability of the event (call this event Event 2) that there is at least one student who lists c as her second choice and hence applies to it in the second step of the mechanism is at least

$$1 - \left(1 - \frac{1}{n-1}\right)^n.$$

As n approaches infinity, this expression converges to $1 - \frac{1}{e}$, so for any sufficiently large n, the conditional probability of Event 2 given Event 1 is at least, say, $1 - \frac{2}{e}$.

Finally, conditional on Events 1 and 2, the probability that at least one of the applicants to school c in the second step of the algorithm is preferred to s by school c (call this event Event 3) is at least one half: To see this point, observe that conditioning on Events 1 and 2 places no restriction on how students are ranked by school c, so for any student $\bar{s}$, the conditional probability that $\bar{s}$ is more preferred to s by c is exactly one half, which provides a lower bound for the conditional probability of Event 3 given Events 1 and 2. Thus the *un*conditional joint

[46] The computation of this limit is as follows:

$$2n \times \frac{1}{n} \times \left(1 - \frac{1}{n}\right)^{2n-1} = 2 \times \left(\left(1 - \frac{1}{n}\right)^n\right)^2 \times \left(1 - \frac{1}{n}\right)^{-1} \to 2 \times (e^{-1})^2 \times 1 = 2/e^2,$$

as $n \to \infty$, where we have used a well-known formula $\lim_{n\to\infty} \left(1 - \frac{1}{n}\right)^n = e^{-1}$. This formula is used in similar calculations of limits in this paper.

probability that Events 1, 2, and 3 happen is at least $\frac{1}{e^2} \times (1 - \frac{2}{e}) \times \frac{1}{2} = (1 - \frac{2}{e})\frac{1}{2e^2}$, which is independent of n and bounded away from below by zero.

Assume that the realization of preferences is such that Events 1, 2, and 3 hold. Then, under this preference profile, school c is matched with student s. Consider the following disimprovement for c: student s declares c to be unacceptable while keeping the relative rankings of all other schools unchanged, and preferences of all other students are unchanged. Under this preference profile, there is no applicant to c in the first step given Event 1, and there is at least one applicant to c in the second step of the algorithm given Event 2. Thus c is matched with the most preferred student among those who apply in the second step. By Event 3, that student is preferred to s by c.

The above arguments imply that there exists n_0 such that the joint probability of Events 1, 2, and 3 is at least $(1 - \frac{2}{e})\frac{1}{2e^2}$ for any $n \geq n_0$. Thus, denoting by p_n the probability that Events 1, 2, and 3 occur in the market with n schools, the probability that Events 1, 2, and 3 occur in the market with any $n \geq 2$ is at least $\min\{p_2, p_3, \ldots, p_{n_0-1}, (1 - \frac{2}{e})\frac{1}{2e^2}\}$, which is bounded away from below by zero, completing the proof.

A.1.3 Proof of Theorem 5

Theorem 5 follows from Proposition 5 as the latter implies the former. However, we present a separate proof, which is simpler than the proof of Proposition 5, for expositional clarity.

Consider a case where there are n schools and n students, and $q_c = 1$ for every school c. Assume that the preferences of all students are generated according to the uniform distribution over all schools with $k = 1$. Moreover, assume that school preferences are drawn identically and independently from the uniform distribution over all preferences over students such that all students are acceptable. These assumptions guarantee that the regularity and sufficient thickness conditions are satisfied.

Let $n \geq 4$. Take an arbitrary school c and let Event 1 be the event that there are exactly two students who prefer c best. The probability of Event 1 is

$$\binom{n}{2} \times \frac{1}{n^2} \times \left(1 - \frac{1}{n}\right)^{n-2} = \frac{n(n-1)}{2} \times \frac{1}{n^2} \times \left(1 - \frac{1}{n}\right)^{n-2}.$$

As n approaches infinity, this expression converges to $\frac{1}{2e}$, so for any sufficiently large n, the probability of Event 1 is at least, say, $\frac{1}{3e}$.

Under Event 1, there are exactly two students who prefer c best. Call these students h and l.[47] Given Event 1, consider the conditional probability of the event (call this event

[47] Later in the proof, we will impose the assumption that h is ranked higher than l by c.

Event 2) that except school c, there is exactly 1 school whthat ranks h first and exactly 1 school that ranks l first. The conditional probability of this event is given by

$$(n-1)(n-2) \times \frac{1}{n^2} \times \left(1 - \frac{2}{n}\right)^{n-3}.$$

As n approaches infinity, this expression converges to $\frac{1}{e^2}$, so for any sufficiently large n, the conditional probability of Event 2 given Event 1 is at least, say, $\frac{1}{2e^2}$.

Denote the schools identified under Event 2 by c_h and c_l, respectively. Given Events 1 and 2, consider the conditional probability of the event (call this event Event 3) that except h and l, there is exactly 1 student who ranks c_h first and exactly one student who ranks c_l first. The conditional probability is given by

$$(n-2)(n-3) \times \frac{1}{(n-1)^2} \times \left(1 - \frac{2}{n-1}\right)^{n-4}.$$

As n approaches infinity, this expression converges to $\frac{1}{e^2}$, so for any sufficiently large n, the conditional probability of Event 3 given Events 1 and 2 is at least, say, $\frac{1}{2e^2}$.

Denote the students identified in Event 3 by s_h and s_l, respectively. Given Events 1, 2, and 3, the conditional probability of the event (call this event Event 4) that either of s_h and s_l has a higher ranking than both h and l in school c's preference relation is $\frac{1}{2}$. (Note that Events 1, 2, and 3 do not impose any restriction on the rankings of h, l, s_h, and s_l in c's preference relation, and hence c is equally likely to rank any of these four students first.) Given the above calculations, the joint probability of Events 1, 2, 3, and 4 is bounded from below by zero (at least $\frac{1}{24e^5}$) for any sufficiently large n.

Given Event 1, school c is matched with h or l with conditional probability 1 by the assumption that $k = 1$. In addition, given Events 1, 2, 3, and 4, the following event occurs with conditional probability 1: c is matched with h or l while being contained in a cycle involving another agent than c, h, and l. Since the above events are symmetric for h and l, this means that c is matched with l with conditional probability $\frac{1}{2}$. Therefore, c is matched with l at least with *un*conditional probability $\frac{1}{48e^5}$, and thus, c is matched with h with a probability that is bounded away from above by 1.

Now additionally assume that h has a higher ranking than l in c's preference relation. (Given the above argument, the joint probability of this event and Events 1-4 is at least $\frac{1}{96e^5}$ for any sufficiently large n.) Then consider the following disimprovement for c in l's preference relation: l declares c (and hence all schools) unacceptable. This change of preferences leads to the situation where h is the only student who ranks c as an acceptable school, and h ranks

c as her most preferred school, which in turn implies that after the disimprovement for c in l's preference relation, c has to be matched with h with probability 1. Note that c is matched with h only with a probability bounded from above by 1 before the disimprovement.

For n with $2 \leq n < 4$, consider the following random market with n schools and $n+1$ students. Each school has one position and school preferences are independently and identically distributed over all preferences over students such that all students are acceptable. Each student's preference relation is uniformly distributed over all preferences over schools with $k = 1$. Then, for each $n \geq 2$, it is a positive probability event that there are two schools c_1 and c_2 and three students s_1, s_2, and s_3 such that

$$\begin{aligned} &\succ_{s_1}: c_2, \emptyset, && \succ_{c_1}: s_1, s_2, s_3, \ldots, \emptyset, \\ &\succ_{s_2}: c_1, \emptyset, && \succ_{c_2}: s_3, s_2, s_1, \ldots, \emptyset, \\ &\succ_{s_3}: c_1, \emptyset. \end{aligned}$$

Under this preference profile $\succ$, clearly c_1 is matched with s_3. Meanwhile, for $(\succ'_{s_3}, \succ_{-s_3})$ such that $\succ'_{s_3}$ prefers $\emptyset$ best, c_1 is matched to s_2.

To finish the proof, let the probability that the respecting improvements criterion is violated be denoted by p_n. Recall that we have already seen that there is an n_0 such that $p_n \geq \frac{1}{96e^5}$ for any $n \geq n_0$. The preceding arguments has shown that, for any $n \geq 2$, p_n is at least $\{p_2, \ldots, p_{n_0-1}, \frac{1}{96e^5}\}$, which is bounded away from below by zero. This completes the proof.

A.1.4 Proof of Proposition 1

Suppose that $\succ'$ is an improvement over $\succ$ for c. Assume by way of contradiction that $|\varphi_c(\succ)| > |\varphi_c(\succ')|$ for some stable mechanism φ. Without loss of generality assume that there exists one student $s \in S$ and a school $\bar{c} \in C \setminus \{c\}$ such that the only difference between $\succ$ and $\succ'$ is that under $\succ$, $\bar{c}$ is preferred by s to c, while under $\succ'$, c is preferrey by s to $\bar{c}$. Formally, assume that $\bar{c} \succ_s c$, $c \succ'_s \bar{c}$, $c \succ_s \hat{c}$ if and only if $c \succ'_s \hat{c}$ for any $\hat{c} \in C \setminus \{c, \bar{c}\} \cup \{\emptyset\}$, $\bar{c} \succ_s \hat{c}$ if and only if $\bar{c} \succ'_s \hat{c}$ for any $\hat{c} \in C \setminus \{c, \bar{c}\} \cup \{\emptyset\}$, and $\succ'_{-s} = \succ_{-s}$.[48]

Since $|\varphi_c(\succ)| > |\varphi_c(\succ')|$ by assumption, by the rural hospital theorem it follows that $\varphi(\succ')$ is not stable under preference profile $\succ$. Thus there is a blocking pair of $\varphi(\succ')$ under $\succ$. First, note that s is part of the blocking pair because she is the only agent whose preferences are different between $\succ$ and $\succ'$. Moreover, it should be the case that $s \in \varphi_c(\succ')$

[48] This assumption is without loss of generality because, for any $\succ$ and $\succ'$ such that $\succ'$ is an improvement for c over $\succ$, there exists a finite sequence of preference profiles $(\succ^1, \succ^2, \ldots, \succ^T)$ such that $\succ^1 = \succ$, $\succ^T = \succ'$, and for any $t = 1, \ldots, T-1$, the only difference between $\succ^t$ and $\succ^{t+1}$ is that the ranking between c and some other $\bar{c}$ next to it is exchanged for one student s.

and the only blocking pair is $(s, \bar{c})$ because the only change from $\succ'_s$ to $\succ_s$ is that $\bar{c}$ is more preferred to c at $\succ_s$ while c is more preferred to $\bar{c}$ at $\succ'_s$. Now satisfy this blocking pair to obtain a new matching. If $|\varphi_{\bar{c}}(\succ')| < q_{\bar{c}}$, then the resulting matching is stable at $\succ$. If $|\varphi_{\bar{c}}(\succ')| = q_{\bar{c}}$, then reject the least preferred student by $\bar{c}$ in $\varphi_{\bar{c}}(\succ')$, and let him block with his most preferred school that can form a blocking pair with him, and so on.[49] This procedure terminates in a finite number of steps, leading to a matching μ that is stable under $\succ$. Moreover, if c is part of the blocking pair in this algorithm, then the algorithm stops at that step, because no student is rejected by c as there is a vacancy. But the resulting matching μ has the property that $|\mu_c| \leq |\varphi_c(\succ')| < |\varphi_c(\succ)|$ (note that $s \in \varphi_c(\succ')$ by the above discussion), which is a contradiction to the rural hospital theorem. This completes the proof.

A.1.5 Proof of Proposition 2

Let $c \in C$, $\succ$ be a preference profile, and $\succ'_s$ be an improvement over $\succ_s$ for c. Recall that φ^B denotes the Boston mechanism. We will show that $|\varphi^B_c(\succ'_s, \succ_{-s})| \geq |\varphi^B_c(\succ)|$. If $|\varphi^B_c(\succ'_s, \succ_{-s})| = q_c$, then the conclusion trivially holds. Thus we assume $|\varphi^B_c(\succ'_s, \succ_{-s})| < q_c$. This assumption implies that unless

$$\varphi^B_s(\succ'_s, \succ_{-s}) = c, \ \varphi^B_s(\succ) \neq c, \tag{1}$$

by definition of φ^B it follows that $\varphi^B(\succ'_s, \succ_{-s}) = \varphi^B(\succ)$, so there is nothing to prove.[50] Thus we assume relation (1) in the rest of the proof.

Since $\varphi^B_s(\succ'_s, \succ_{-s}) = c$ and $|\varphi^B_c(\succ'_s, \succ_{-s})| < q_c$, we have that

$$\varphi^B_{\bar{s}}(\succ^{\emptyset}_s, \succ_{-s}) = \varphi^B_{\bar{s}}(\succ'_s, \succ_{-s}), \tag{2}$$

for every $\bar{s} \neq s$, where $\succ^{\emptyset}_s$ is a preference relation of s that ranks $\emptyset$ as the most preferred outcome.

Now we compare $\varphi^B(\succ)$ and $\varphi^B(\succ^{\emptyset}_s, \succ_{-s})$. It is clear by the definition of the algorithm that $\varphi^B_{\bar{s}}(\succ^{\emptyset}_s, \succ_{-s}) \succsim_{\bar{s}} \varphi^B_{\bar{s}}(\succ)$ for every $\bar{s} \neq s$. This fact and the property that the matching

[49] This procedure is a variant of the "vacancy chain dynamics" studied by Blum, Roth, and Rothblum (1997).

[50] The reason is as follows. If $\varphi^B_s(\succ) \neq c$, then because $|\varphi^B_c(\succ'_s, \succ_{-s})| < q_c$, it follows from the definition of the Boston mechanism that $\varphi^B_s(\succ'_s, \succ_{-s}) \succ_s c$. Since $\succ'_s$ is an improvement for c over $\succ_s$, this implies that the algorithm for the Boston mechanism terminates without using any part of $\succ'_s$ or $\succ_s$ at or below c in either of the preference profiles, thus the algorithm proceeds in an exactly identical manner under both preference profiles, resulting in the same outcome. If $\varphi^B_s(\succ) = c$, then $\varphi^B(\succ'_s, \succ_{-s}) = \varphi^B(\succ)$ by inspection of the steps of the algorithm.

under the Boston mechanism is individually rational imply that

$$|\{\bar{s} \in S \setminus \{s\} : \varphi^B_{\bar{s}}(\succ^{\emptyset}_s, \succ_{-s}) \in C\}| \geq |\{\bar{s} \in S \setminus \{s\} : \varphi^B_{\bar{s}}(\succ) \in C\}|.$$

Since

$$|\{\bar{s} \in S : \varphi^B_{\bar{s}}(\succ^{\emptyset}_s, \succ_{-s}) \in C\}| = |\{\bar{s} \in S \setminus \{s\} : \varphi^B_{\bar{s}}(\succ^{\emptyset}_s, \succ_{-s}) \in C\}|$$

as s is clearly unmatched at $\varphi^B_{\bar{s}}(\succ^{\emptyset}_s, \succ_{-s})$), and clearly

$$|\{\bar{s} \in S \setminus \{s\} : \varphi^B_{\bar{s}}(\succ) \in C\}| \geq |\{\bar{s} \in S : \varphi^B_{\bar{s}}(\succ) \in C\}| - 1,$$

as s is the only feasible element of $\{\bar{s} \in S : \varphi^B_{\bar{s}}(\succ) \in C\}$ who is not also in $\{\bar{s} \in S \setminus \{s\} : \varphi^B_{\bar{s}}(\succ) \in C\}$. Hence, we conclude that

$$|\{\bar{s} \in S : \varphi^B_{\bar{s}}(\succ^{\emptyset}_s, \succ_{-s}) \in C\}| \geq |\{\bar{s} \in S : \varphi^B_{\bar{s}}(\succ) \in C\}| - 1. \tag{3}$$

On the other hand, it is clear by definition of the Boston mechanism that $|\varphi^B_{\bar{c}}(\succ)| \geq |\varphi^B_{\bar{c}}(\succ^{\emptyset}_s, \succ_{-s})|$ for all $\bar{c} \in C$. Because the matching is bilateral, i.e., $\mu_{\bar{s}} = \bar{c} \iff \bar{s} \in \mu_{\bar{c}}$ for any matching μ, this and relation (3) imply that there is at most one school $\bar{c} \in C$ such that $|\varphi^B_{\bar{c}}(\succ)| > |\varphi^B_{\bar{c}}(\succ^{\emptyset}_s, \succ_{-s})|$, and for such a school, $|\varphi^B_{\bar{c}}(\succ^{\emptyset}_s, \succ_{-s})| \geq |\varphi^B_{\bar{c}}(\succ)| - 1$. In particular, we obtain that

$$|\varphi^B_c(\succ^{\emptyset}_s, \succ_{-s})| \geq |\varphi^B_c(\succ)| - 1. \tag{4}$$

Note that relations (1) and (2) imply that $|\varphi^B_c(\succ'_s, \succ_{-s})| = |\varphi^B_c(\succ^{\emptyset}_s, \succ_{-s})| + 1$.[51] This and relation (4) imply $|\varphi^B_c(\succ'_s, \succ_{-s})| = |\varphi^B_c(\succ^{\emptyset}_s, \succ_{-s})| + 1 \geq |\varphi^B_c(\succ)|$, completing the proof.

A.1.6 Proof of Proposition 4

Consider a sequence of random markets where there are n schools and $3n$ students, and the capacity of every school is 2. Assume that preferences of all students are generated according to the procedure described in Section 4.1 associated with the uniform distribution over all schools and $k = 2$. Moreover assume that school preferences over individual students are drawn identically and independently from the uniform distribution over all preferences for students such that all students are acceptable. These assumptions guarantee that the

[51] To see why this equality holds, note that relation (2) means that all students except for s receive identical assignments between $\varphi^B(\succ'_s, \succ_{-s})$ and $\varphi^B(\succ^{\emptyset}_s, \succ_{-s})$, and the last term accounts for the implication of relation (1) that student s is matched to c at $\varphi^B(\succ'_s, \succ_{-s})$ while she is unmatched at $\varphi^B(\succ^{\emptyset}_s, \succ_{-s})$ by definition of $\succ^{\emptyset}_s$.

regularity and sufficient thickness conditions are satisfied.

Given any $n \geq 2$, fix an arbitrary school c and let Event 1 be the event that there are exactly two students who prefer that school c most. The probability of Event 1 is

$$\binom{3n}{2} \times \left(\frac{1}{n}\right)^2 \times \left(1 - \frac{1}{n}\right)^{3n-2} = \frac{3n(3n-1)}{2} \times \left(\frac{1}{n}\right)^2 \times \left(1 - \frac{1}{n}\right)^{3n-2}.$$

This expression converges to $\frac{9}{2e^3}$ as n approaches infinity, where e is the basis of the natural logarithm. Therefore, for any sufficiently large n, the probability of Event 1 is at least, say, $\frac{1}{e^3}$. Denote by s and $\hat{s}$ the students who prefer c most.

Since there are $3n$ students and $2n$ school seats, there are at least n students who are not matched in the first step of the algorithm of the Boston mechanism. Since $k = 2$, that is, each student finds two schools to be acceptable, each of these students applies for a school in the second step of the algorithm. Therefore, given Event 1, the conditional probability of the event (call this event Event 2) that there is at least one student who lists c as her second choice and hence applies for it in the second step of the mechanism is at least

$$1 - \left(1 - \frac{1}{n-1}\right)^n.$$

As n approaches infinity, this expression converges to $1 - \frac{1}{e}$, so for any sufficiently large n, the conditional probability of Event 2 given Event 1 is at least, say, $1 - \frac{2}{e}$.

Finally, conditional on Events 1 and 2, the probability that at least one of the applicants to school c in the second step of the algorithm is preferred to both s and $\hat{s}$ by school c (call this event Event 3) is at least $\frac{1}{3}$: To see this point, observe that conditioning on Events 1 and 2 places no restriction on how students are ranked by school c. So, for any student $\bar{s}$, the conditional probability that $\bar{s}$ is more preferred to s and $\hat{s}$ by c is exactly 1/3, which provides a lower bound for the conditional probability of Event 3 given Events 1 and 2. Thus the *un*conditional joint probability that Events 1, 2, and 3 happen is at least $\frac{1}{e^3} \times (1 - \frac{2}{e}) \times \frac{1}{3} = \frac{1-\frac{2}{e}}{3e^3}$, which is independent of n and bounded away from below by zero.

Assume that the realization of preferences is such that Events 1, 2, and 3 occur. Then, under this preference profile, school c is matched with students s and $\hat{s}$. Without loss of generality assume that $s \succ_c \hat{s}$ and consider the following disimprovement for c: Student s declares c to be unacceptable while keeping the relative rankings of all other schools unchanged, and preferences of all other students are unchanged. Under this preference profile, the only applicant to c in the first step is $\hat{s}$ by Event 1. Also, there is at least one applicant to c in the second step of the algorithm by Event 2. Thus c is matched with the most preferred student, say $\bar{s}$, among those who apply in the second step. By Event 3,

that student is preferred to s by c. Therefore, the set of students matched with c after the disimprovement is $\{\bar{s}, \hat{s}\}$, which is preferred to $\{s, \hat{s}\}$, the set of students matched with c before the disimprovement. Moreover, the improvement in the preference relation of s (from the disimproved preferences, where s finds c unacceptable, to the improved preferences, where s prefers c most) has properties (1) and (2) in Definition 6 of respecting improvements for desirable students: That is, s is more preferred to $\hat{s}$ by c, while $\hat{s}$ is matched to c under the preference profile after the disimprovement.

The above arguments imply that there exists n_0 such that the joint probability of Events 1, 2, and 3 is at least $\frac{1-\frac{2}{e}}{3e^3}$ for any $n \geq n_0$. Thus, denoting by p_n the probability that Events 1, 2, and 3 occur in the market with n, the probability that Events 1, 2, and 3 occur in the market with any $n \geq 2$ is at least $\min\{p_2, p_3, \ldots, p_{n_0-1}, \frac{1-\frac{2}{e}}{3e^3}\}$, which is bounded away from below by zero, completing the proof.

A.1.7 Proof of Proposition 5

Consider a case where there are n schools and n students, and the capacity of each school is 2. Assume that preferences of all students are generated according to the uniform distribution over all schools with $k = 1$. Moreover, assume that school preferences are drawn identically and independently from the uniform distribution over all preferences over students such that all students are acceptable. These assumptions guarantee that the regularity and sufficient thickness conditions are satisfied.

Let $n \geq 6$. Take an arbitrary school c and let Event 1 be the event that there are exactly 3 students who prefer c best. The probability of Event 1 is

$$\binom{n}{3} \times \frac{1}{n^3} \times \left(1 - \frac{1}{n}\right)^{n-3} = \frac{n(n-1)(n-2)}{3 \times 2} \times \frac{1}{n^3} \times \left(1 - \frac{1}{n}\right)^{n-3}.$$

As n approaches infinity, this expression converges to $\frac{1}{6e}$. Thus, for any sufficiently large n, the probability of Event 1 is at least, say, $\frac{1}{7e}$.

Under Event 1, there are exactly 3 students who prefer c best. Call these students h, m and l.[52] Given Event 1, consider the conditional probability of the event (call this event Event 2) that except school c, there is exactly 1 school that ranks h first, exactly 1 school that ranks m first, and exactly 1 school that ranks l first. This conditional probability is

[52] Later in the proof, we will impose the assumption that h is the highest-ranked, m is the middle-ranked, and l is the lowest-ranked by c among h, m, and l.

given by

$$(n-1)(n-2)(n-3) \times \frac{1}{n^3} \times \left(1 - \frac{3}{n}\right)^{n-4}.$$

As n approaches infinity, this expression converges to $\frac{1}{e^3}$. Thus, for any sufficiently large n, the conditional probability of Event 2 given Event 1 is at least, say, $\frac{1}{2e^3}$.

Given Event 2, denote the schools that rank h, m, and l first by c_h, c_m, and c_l, respectively. Given Events 1 and 2, consider the conditional probability of the event (call this event Event 3) that except h, m and l, there is exactly 1 student who ranks c_h first, exactly 1 student who ranks c_m first, and exactly one student who ranks c_l first. The conditional probability is given by

$$(n-3)(n-4)(n-5) \times \frac{1}{(n-1)^3} \times \left(1 - \frac{3}{n-1}\right)^{n-6}.$$

As n approaches infinity, this expression converges to $\frac{1}{e^3}$, so for any sufficiently large n, the conditional probability of Event 3 given Events 1 and 2 is at least, say, $\frac{1}{2e^3}$.

Given Event 3, denote the students who rank c_h, c_m, and c_l first by s_h, s_m and s_l, respectively. Given Events 1, 2, and 3, the conditional probability of the event (call this event Event 4) that at least two out of s_h, s_m, and s_l have higher rankings than all of h, m, and l in school c's preference relation is $\frac{1}{5}$.[53] (Note that Events 1, 2, and 3 do not impose any restriction on the rankings of h, m, l, s_h, s_m, and s_l in c's preference relation.) Given the above calculations, the joint probability of Events 1, 2, 3, and 4 is bounded from below by zero (at least $\frac{1}{140e^7}$) for any sufficiently large n.

Given Event 1, school c is matched with two students out of h, m and l with conditional probability 1 by the assumption that $k = 1$. In addition, given Events 1, 2, 3, and 4, the following event occurs with conditional probability 1: c is matched with two students out of h, m, and l while being contained in a cycle involving another agent than c, h, m, and l. Since the above events are symmetric for h, m and l, this means that c is matched with $\{m, l\}$ with conditional probability $\frac{1}{3}$. Therefore, c is matched with $\{m, l\}$ with *un*conditional probability of at least $\frac{1}{420e^7}$.

Now additionally assume Event 5, that c prefers h to m to l, which happens with probability $\frac{1}{6}$ (conditional on Events 1-4. Given the above argument, the joint probability of Events 1-5 is at least $\frac{1}{2520e^7}$ for any sufficiently large n. Then consider the following disim-

[53] Note that the above event is equivalent to the event that the two highest-ranked students by c among the six students h, m, l, s_h, s_m, and s_l are from s_h, s_m and s_l. The probability of the latter event is given by $3 \times 2 \times \frac{4!}{6!} = \frac{1}{5}$.

provement for c in m's preference relation: m declares all schools unacceptable. This change of m's preference relation leads to the situation where h and l are the only students who rank c as an acceptable school, which in turn implies that after the disimprovement for c in m's preference relation, c has to be matched with $\{h, l\}$ with probability 1. Since c is matched with $\{m, l\}$ with a probability bounded away from below by zero before the disimprovement, the disimprovement for c makes c strictly better off with a probability that is bounded from below by zero. Moreover, the improvement for c in the preferences of m (from the disimproved preferences, where m finds c unacceptable, to the improved preferences, where m prefers c most) has properties (1) and (2) in Definition 6 of respecting improvements for desirable students: That is, m is more preferred to l by c, while l is matched under the preference profile after the disimprovement.

For n with $2 \leq n < 6$, consider the following random market with n schools and $n + 3$ students. School preferences are independently and identically distributed over all preferences over students such that all students are acceptable. Each student's preference relation is uniformly distributed over all preferences over schools with $k = 1$. Then, for each $n \geq 2$, it is a positive probability event that there are two schools c_1 and c_2, and c_3 and five students $s_1, ..., s_5$ such that

$$\begin{array}{ll}
\succ_{s_1}: c_2, \emptyset, & \succ_{c_1}: s_5, s_1, s_2, s_3, s_4, \ldots, \emptyset, \\
\succ_{s_2}: c_1, \emptyset, & \succ_{c_2}: s_4, s_3, s_2, s_1, \ldots, \emptyset, \\
\succ_{s_3}: c_1, \emptyset. & \\
\succ_{s_4}: c_1, \emptyset. & \\
\succ_{s_5}: c_2, \emptyset. &
\end{array}$$

where both schools have a capacity of two. Under this preference profile $\succ$, clearly c_1 is matched with $\{s_3, s_4\}$. Meanwhile, for $(\succ'_{s_3}, \succ_{-s_3})$ such that $\succ'_{s_3}$ prefers $\emptyset$ best, c_1 is matched to $\{s_2, s_4\}$, which is strictly preferred to $\{s_3, s_4\}$ by c_1. Note that s_3 is a desirable student for c_1 at $(\succ'_{s_3}, \succ_{-s_3})$.

To finish the proof, let the probability that the respecting improvements criterion is violated be denoted by p_n. Recall that we have already seen that there is an n_0 such that the joint probability of Events 1, 2, 3, and 4 is at least $\frac{1}{420e^7}$. The preceding arguments has shown that, for any $n \geq 2$, p_n is at least $\{p_2, \ldots, p_{n_0-1}, \frac{1}{2520e^7}\}$, which is bounded away from below by zero. This completes the proof.

A.1.8 Proof of Proposition 6

It is useful to start with the following result, presenting an equivalent representation of virtual homogeneity. This result is similar in spirit to Theorem 2 of Ergin (2002), which provides an equivalent condition to his acyclicity condition.

Lemma 2. *A school preference profile $\succ_C$ is virtually homogeneous if and only if there exist no $a, b \in C$ and $i, j \in S$ such that*

- *$i \succ_a j$ and $j \succ_b i$, and*
- *There exists a set of students $S_b \subseteq S \setminus \{i, j\}$ such that $|S_b| = q_b - 1$ and $s \succ_b i$ for every $s \in S_b$.*

Proof. **The "only if" direction.** Let $\rho_{\succ_c}(s)$ be the ranking of student s in $\succ_c$. That is, $\rho_{\succ_c}(s) = \ell$ if and only if $r^\ell(\succ_c) = s$.

Suppose that $\succ_C$ is virtually homogeneous and $a, b \in C$ and $i, j \in S$ satisfy

$$i \succ_a j, \quad j \succ_b i. \tag{5}$$

Consider a student $s^* \in \{i, j\}$ whose worst ranking by a or b is the worst among i and j's rankings by a or b. That is, s^* is a student who satisfies $\max\{\rho_{\succ_a}(s^*), \rho_{\succ_b}(s^*)\} = \max\{\rho_{\succ_a}(i), \rho_{\succ_b}(i), \rho_{\succ_a}(j), \rho_{\succ_b}(j)\}$. (If both i and j satisfy this condition, let s^* be one of them arbitrarily). Consider the following cases.

(1) Suppose $s^* = i$. Then, since $\rho_{\succ_a}(i) < \rho_{\succ_a}(j) \leq \rho_{\succ_b}(i)$ by assumption (5), virtual homogeneity implies that $\rho_{\succ_b}(i) \leq \bar{q} \equiv \min\{q_{\hat{c}} : \hat{c} \in C\}$. Therefore there does not exist $S_b \subseteq S \setminus \{i, j\}$ such that $|S_b| = q_b - 1$ and $s \succ_b i$ for all $s \in S_b$.

(2) Suppose $s^* = j$. Then, since $\rho_{\succ_b}(j) < \rho_{\succ_b}(i) \leq \rho_{\succ_a}(j)$ by assumption (5), virtual homogeneity implies that $\rho_{\succ_a}(j) \leq \bar{q}$. Thus we obtain $\rho_{\succ_b}(i) \leq \rho_{\succ_a}(j) \leq \bar{q}$. Therefore there does not exist $S_b \subseteq S \setminus \{i, j\}$ such that $|S_b| = q_b - 1$ and $s \succ_b i$ for all $s \in S_b$.

The "if" direction. We shall prove the contraposition. Thus assume that $\succ_C$ is not virtually homogeneous. Let

$$\lambda = \max\{\ell \in \mathbb{N} : \text{ there exist two schools } c, \bar{c} \in C \text{ such that } r^\ell(\succ_c) \neq r^\ell(\succ_{\bar{c}})\}.$$

Then the assumption that $\succ_C$ is not virtually homogeneous implies $\lambda > \bar{q} \equiv \min\{q_{\hat{c}} : \hat{c} \in C\}$.

Let schools $a, b \in C$ satisfy $r^\lambda(\succ_a) \neq r^\lambda(\succ_b)$ and, without loss of generality, $q_b = \bar{q}$.[54] Denote $i = r^\lambda(\succ_b)$ and $j = r^\lambda(\succ_a)$. By maximality of λ, it follows that $i \succ_a j$ and $j \succ_b i$. Moreover, since $\lambda > \bar{q} = q_b$, there exists $S_b \subseteq S \setminus \{i, j\}$ such that $|S_b| = q_b - 1$ and $s \succ_b i$ for every $s \in S_b$, finishing the proof. □

Next, the following lemma offers an equivalent condition to the requirements in the statement of Proposition 6.

Lemma 3. *The condition that either*

(1) *The school preference profile $\succ_C$ is virtually homogeneous, or*

(2) *For every school $c \in C$, the capacity associated with $\succ_c$ is one,*

is satisfied if and only if the following condition is satisfied: There exist no $a, b \in C$ and $i, j \in S$ such that

- *$q_a \geq 2$,*
- *$i \succ_a j$ and $j \succ_b i$, and*
- *There exists a set of students $S_b \subseteq S \setminus \{i, j\}$ such that $|S_b| = q_b - 1$ and $s \succ_b i$ for every $s \in S_b$.*

Proof. The "only if" direction follows immediately by inspection of the conditions and Lemma 2. To show the "if" direction, assume that $\succ_C$ is not virtually homogeneous and there is at least one school $c \in C$ with $q_c \geq 2$, and we shall show that there exist a, b, i, and j that satisfy the three conditions in the statement of this claim. By Lemma 2, there exist $a, b \in C$ and $i, j \in S$ such that

- $i \succ_a j$ and $j \succ_b i$, and
- There exists a set of students $S_b \subseteq S \setminus \{i, j\}$ such that $|S_b| = q_b - 1$ and $s \succ_b i$ for every $s \in S_b$.

Consider the following cases.

(1) Assume $q_a \geq 2$. Then the three conditions in the statement of this claim immediately follow.

[54] The reason that it is without loss of generality to assume $q_b = \bar{q}$ is as follows. Define b to be a school with $q_b = \bar{q}$. By assumption there exist two schools $\bar{a}$ and $\hat{a}$ such that $r^\lambda(\succ_{\bar{a}}) \neq r^\lambda(\succ_{\hat{a}})$. Then it is clear that at least one of the relations $r^\lambda(\succ_{\bar{a}}) \neq r^\lambda(\succ_b)$ and $r^\lambda(\succ_{\hat{a}}) \neq r^\lambda(\succ_b)$ should hold. Let $a \in \{\bar{a}, \hat{a}\}$ be a school such that the relation holds, which shows the claim.

(2) Assume $q_a = 1$ and $q_b \geq 2$. Then the desired conclusion holds by relabeling (a, b, i, j) to (b, a, j, i).

(3) Assume $q_a = q_b = 1$. Then, by assumption there exists $c \neq a, b$ such that $q_c \geq 2$. If $i \succ_c j$, then the desired conclusion holds by relabeling c to a. If $j \succ_c i$, then the desired conclusion holds by relabeling (c, a, j, i) to (a, b, i, j).

□

With these lemmas at hand, we are now ready to prove Proposition 6.

The "only if" direction. We shall show the contraposition. Assume that the condition in Proposition 6 is not satisfied. By Lemma 3, there exist $a, b \in C$ and $i, j \in S$ such that

- $q_a \geq 2$,
- $i \succ_a j$ and $j \succ_b i$, and
- There exists a set of students $S_b \subseteq S \setminus \{i, j\}$ such that $|S_b| = q_b - 1$ and $s \succ_b i$ for every $s \in S_b$.

Consider a preference profile $\succ_S$ such that

$$\begin{aligned}
&\succ_i : b, a, \emptyset, \\
&\succ_j : b, a, \emptyset, \\
&\succ_k : b, \emptyset, \forall k \in S_b, \\
&\succ_l : \emptyset, \forall l \in S \setminus (\{i, j\} \cup S_b).
\end{aligned}$$

Then the unique stable matching at this preference profile matches i to a and $S_b \cup \{j\}$ to b while leaving every other student unmatched. Now consider an alternative preference profile $\succ' = (\succ'_j, \succ_{-j})$ where $\succ'_j$: $a, b, \emptyset$. Note that $\succ'$ is an improvement for a over $\succ$. The unique stable matching at preference profile $\succ'$ matches j to a and $S_b \cup \{i\}$ to b. Thus a is made worse off at $\succ'$ than at $\succ$ although $\succ'$ is an improvement for a over $\succ$, showing the claim.

The "if" direction. First consider the case (2) of the conditions in the statement of the proposition in which $q_c = 1$ for all $c \in C$. In this case, Theorem 5 of Balinski and Sönmez (1999) shows that the student-optimal stable mechanism respects improvements.

Second, consider the case (1) of the conditions in the statement of the proposition in which $\succ_C$ is virtually homogeneous. We will show the claim by presenting a specific mechanism that is stable and respects improvements. Fix a school $c \in C$ arbitrarily and consider the following **serial dictatorship** with respect to $\succ_c$:

- Step t: Choose student $r^t(\succ_c)$. Let her be matched with a school (or the outside option) that she prefers most among all the schools whose entire capacity has not been exhausted by the end of Step $(t-1)$.

If $\succ_C$ is virtually homogeneous, then clearly the serial dictatorship with respect to $\succ_c$ is identical to the serial dictatorship with respect to $\succ_{\bar{c}}$ for any $c, \bar{c} \in C$ because the top $\bar{q}$ students in every school's preferences are always matched with their most preferred schools regardless of which school's preferences are used. Thus, when convenient, we refer to the mechanism simply as the serial dictatorship.

Claim 4. If $\succ_C$ is virtually homogeneous, then the serial dictatorship with respect to $\succ_c$ is stable for any $c \in C$.

Proof. Let μ be the matching resulting from the serial dictatorship. It is obvious that μ is individually rational. To show that there is no blocking pair of μ, assume that $\bar{c} \succ_s \mu_s$ for a student $s \in S$. Then, by the definition of the serial dictatorship with respect to $\succ_c$, it follows that

$$|\mu_{\bar{c}}| = q_{\bar{c}}, \tag{6}$$

$$\bar{s} \succ_c s \text{ for every } \bar{s} \in \mu_{\bar{c}}. \tag{7}$$

Also note that $\rho_c(s) > \bar{q}$ because otherwise s should receive her most preferred school in the serial dictatorship with respect to $\succ_c$. Property (7) and the assumption that $\succ_C$ is virtually homogeneous imply

$$\bar{s} \succ_{\bar{c}} s \text{ for every } \bar{s} \in \mu_{\bar{c}}. \tag{8}$$

Properties (6) and (8) show that $(s, \bar{c})$ does not block μ, showing that the serial dictatorship is a stable mechanism. □

Claim 5. If $\succ_C$ is virtually homogeneous, then the serial dictatorship respects improvement of school quality for any $c \in C$.

Proof. Let φ be the serial dictatorship. Consider two preference profiles $\succ$ and $\succ' = (\succ'_s, \succ_{-s})$ such that $\succ'$ is an improvement for c over $\succ$, where $s \in S$ and $c \in C$.[55]

[55] Focusing on the case where only one students' preferences change is without loss of generality because, for any $\succ$ and $\succ'$ such that $\succ'$ is an improvement for c over $\succ$, there exists a finite sequence of preference profiles $(\succ^1, \succ^2, \ldots, \succ^T)$ and students $(s^1, s^2, \ldots, s^{T-1})$ such that $\succ^1 = \succ$, $\succ^T = \succ'$, and for any $t = 1, \ldots, T-1$, $\succ^{t+1}_{s^t}$ is an improvement for c over $\succ^t_{s^t}$ and $\succ^{t+1}_{-s^t} = \succ^t_{-s^t}$.

(1) Suppose that $\varphi_s(\succ) = c$. Then, $\varphi(\succ') = \varphi(\succ)$ by inspection of the steps of the serial dictatorship.[56]

(2) Suppose that $\varphi_s(\succ) \neq c$ and $\varphi_s(\succ') \neq c$. Then, again $\varphi(\succ) = \varphi(\succ')$ by inspection of the steps of the serial dictatorship.

(3) Suppose that $\varphi_s(\succ) \neq c$ while $\varphi_s(\succ') = c$.

(a) Suppose $\varphi_c(\succ) \setminus \varphi_c(\succ') = \emptyset$. Then $\varphi_c(\succ') \succsim_c \varphi_c(\succ)$ by responsiveness of school preferences as well as the assumption that every student is acceptable to c under $\succ_c$.

(b) Suppose $\varphi_c(\succ) \setminus \varphi_c(\succ') \neq \emptyset$. We show the following claim.

Claim 6. Suppose $\varphi_c(\succ) \setminus \varphi_c(\succ') \neq \emptyset$. Then there exists $\bar{s} \in S$ such that $\varphi_c(\succ') = \varphi_c(\succ) \cup \{s\} \setminus \{\bar{s}\}$ and $s \succ_c \bar{s}$.

Proof. First note that $\varphi_{\hat{s}}(\succ) = \varphi_{\hat{s}}(\succ')$ for every student $\hat{s}$ with $\hat{s} \succ_c s$ because of the definition of the serial dictatorship. Thus every student in $\varphi_c(\succ) \setminus \varphi_c(\succ')$ is less preferred to s by c. Let $\bar{s}$ be the most preferred student according to $\succ_c$ in $\varphi_c(\succ) \setminus \varphi_c(\succ')$. Suppose that $\bar{s}$ is the last student who receives c in the serial dictatorship at preference profile $\succ$. Then, since no student receives c in subsequent steps either at $\succ$ or $\succ'$, clearly $\varphi_c(\succ') = \varphi_c(\succ) \cup \{s\} \setminus \{\bar{s}\}$. Suppose that $\bar{s}$ is not the last student who receives c in the serial dictatorship at preference profile $\succ$. This implies that a seat in c is still available to be received by $\bar{s}$ at that step in both preference profiles $\succ$ and $\succ'$. Therefore, the school that student $\bar{s}$ is assigned to at $\succ'$ is the unique school that has a vacant seat in that step at $\succ'$ but not at $\succ$. This implies that at the end of that step, the numbers of seats available in each school in the serial dictatorships are identical between $\succ$ and $\succ'$. Therefore, assignments for every student who is less preferred to $\bar{s}$ are identical between $\succ$ and $\succ'$, implying that $\varphi_c(\succ') = \varphi_c(\succ) \cup \{s\} \setminus \{\bar{s}\}$. □

Since $\succ_c$ is responsive, Claim 6 implies that $\varphi_c(\succ') \succ_c \varphi_c(\succ)$.

Hence, when $\varphi_c(\succ) \setminus \varphi_c(\succ') \neq \emptyset$, we have that $\varphi_c(\succ') \succsim_c \varphi_c(\succ)$.

As these three cases are exhaustive, this shows Claim 5. □

Claims 4 and 5 complete the proof of Proposition 6.

[56] Technically speaking, this is a consequence of Maskin monotonicity. Note that it is well-known than the serial dictatorship satisfies Maskin monotonicity.

A.1.9 Proof of Proposition 7

The "only if" direction. Assume for contradiction that $\succ_C$ is not virtually homogeneous, but there exists a mechanism that is Pareto efficient for students and respects improvements of school quality. By Lemma 2, there exist $a, b \in C$ and $i, j \in S$ such that

- $i \succ_a j$ and $j \succ_b i$, and
- There exists a set of students $S_b \subseteq S \setminus \{i, j\}$ such that $|S_b| = q_b - 1$ and $s \succ_b i$ for every $s \in S_b$.

First, consider the following preference profile $\succ_S$ of students:

$$\begin{aligned}
&\succ_i : a, \emptyset, \\
&\succ_k : b, \emptyset, \forall k \in S_b \cup \{j\}, \\
&\succ_l : \emptyset, \forall l \in S \setminus (\{i, j\} \cup S_b)
\end{aligned}$$

Under $\succ \equiv (\succ_S, \succ_C)$, the unique Pareto efficient matching matches i to a, $S_b \cup \{j\}$ to b, and leaves all other students unmatched.

Next, consider students' new preferences $\succ'_S \equiv (\succ'_i, \succ_{-i})$ where i's preference is $\succ'_i$: $b, a, \emptyset$. Note that $\succ'$ is an improvement for school b over $\succ$. Since $j \succ_b i, s \succ_b i$ for every $s \in S_b$, and the mechanism is Pareto efficient for students and respects improvement, the outcome of the mechanism under $\succ'$, b has to be matched with $S_b \cup \{j\}$. This in turn means that a must be matched with i under $\succ'$.

Finally, consider another preference profile $\succ'' \equiv (\succ'_i, \succ'_j, \succ_{-\{i,j\}})$ where $\succ'_j$: $a, b, \emptyset$. Note that $\succ''$ is an improvement for school a over $\succ'$. Under $\succ''$, the unique matching that is Pareto efficient for students matches j to a and $S_b \cup \{i\}$ to b, which implies that a is matched with j in the outcome of the mechanism. However, note that $i \succ_a j$ although $\succ''$ is an improvement for school a over $\succ'$. This means that this mechanism does not respect improvements of school quality, which is a contradiction.

The "if" direction. Fix $c \in C$ arbitrarily and consider the serial dictatorship with respect to $\succ_c$. It is well-known that the serial dictatorship is Pareto efficient for students (see Abdulkadiroğlu and Sönmez (1998)). This fact and Claim 5 in the proof of Proposition 6 complete the proof.

A.1.10 Proof of Proposition 8

Consider a student s, a student preference profile $\succ_S$, and two school preference profiles $\succ_C$ and $\succ'_C$, where $\succ'_C$ is an improvement for student s over $\succ_C$. Consider the first step t

at which the Boston algorithm using $\succ_C$ matches a student to a different school than the Boston algorithm using $\succ'_C$. (If no such step occurs, then s must get the same school under both preference profiles, and we are done.) Since all other students besides s are ranked the same relative to each other under preference profiles $\succ_C$ and $\succ'_C$, this step must involve student s applying to some school c. However, since student s is ranked (weakly) higher by all schools, this means that the difference in the outcome of the algorithm at t using the two different inputs must be that student s is assigned to the school c under preference profile $\succ'_C$, but is not assigned to c under preference profile $\succ_C$. Therefore, student s is better off under $\succ'_C$ as she can only recieve a worse outcome in the later steps of the Boston algorithm under preferences $\succ_C$, and so we are done.

A.1.11 Proof of Proposition 9

Consider a student s, a student preference profile $\succ_S$, and two school preference profiles $\succ_C$ and $\succ'_C$, where $\succ'_C$ is an improvement for student s over $\succ_C$. Consider, without loss of generality, the s-avoiding TTC algorithm, where in each step t, we remove one cycle $(s_1, c_1, s_2, \ldots, s_K, c_K)$; if there are multiple cycles, we clear a cycle that does not involve student s. Since the order of cycle removal does not affect the outcome, this is equivalent to the original TTC mechanism.

At each step of the s-avoiding TTC algorithm before s is removed under preferences $(\succ_S, \succ'_C)$, the same cycle is also removed under preferences $\succ_C$, as a school at each step under $\succ'_C$ is pointing either at the same student as under $\succ_C$ or at s.

Now note that, when agent s is removed under $(\succ_S, \succ'_C)$, every school is directly or indirectly pointing at agent s, and so agent s receives his favorite school from those remaining at that step. Hence, as the set of schools left at the step where s is removed under preferences $\succ'_C$ is a superset of the schools left at the step s is removed under preferences $\succ_C$, s is weakly better off.

A.2 The Boston Mechanism Does Not Respect Improvements Even When a School Preference Profile is Virtually Homogeneous: An Example

Let $S = \{s, \bar{s}, \hat{s}\}$ and $C = \{c, \bar{c}\}$. Consider the following preferences:

$$\succ_s : c, \bar{c}, \emptyset, \qquad \succ_c : s, \bar{s}, \hat{s}, \emptyset,$$
$$\succ_{\bar{s}} : c, \bar{c}, \emptyset, \qquad \succ_{\bar{c}} : s, \bar{s}, \hat{s}, \emptyset,$$
$$\succ_{\hat{s}} : c, \bar{c}, \emptyset.$$

The capacities of the schools are given by $q_c = q_{\bar{c}} = 1$. Note that the two schools' preferences are exactly the same and thus this school preference profile is virtually homogeneous.

Under $\succ \equiv (\succ_s, \succ_{\bar{s}}, \succ_{\hat{s}}, \succ_c, \succ_{\bar{c}})$, the Boston mechanism φ^B produces the following matching:

$$\varphi^B(\succ) = \begin{pmatrix} c & \bar{c} & \emptyset \\ s & \bar{s} & \hat{s} \end{pmatrix}.$$

Now, consider student $\hat{s}$'s new preference relation $\succ'_{\hat{s}}$: $\bar{c}, c, \emptyset$. Note that $\succ'_{\hat{s}}$ is an improvement for school $\bar{c}$ over $\succ_{\hat{s}}$. Under $(\succ'_{\hat{s}}, \succ_{-\hat{s}})$, the Boston mechanism produces the following matching:

$$\varphi^B(\succ'_{\hat{s}}, \succ_{-\hat{s}}) = \begin{pmatrix} c & \bar{c} & \emptyset \\ s & \hat{s} & \bar{s} \end{pmatrix}.$$

Hence,

$$\varphi^B_{\bar{c}}(\succ) = \bar{s} \succ_{\bar{c}} \hat{s} = \varphi^B_{\bar{c}}(\succ'_{\hat{s}}, \succ_{-\hat{s}}),$$

even though $\succ'_{\hat{s}}$ is an improvement for $\bar{c}$ over $\succ_{\hat{s}}$. Therefore, the Boston mechanism does not respect improvements of school quality at school preference profile $(\succ_c, \succ_{\bar{c}})$ even though $(\succ_c, \succ_{\bar{c}})$ is virtually homogeneous.

A.3 The Relationship between Virtual Homogeneity and Acyclicity (and Its Variants)

As referenced in the Remark at the end of Section 6.1, virtual homogeneity is stronger than *acyclicity* by Ergin (2002) and all of its variants proposed in the literature: *strong x-acyclicity* by Haeringer and Klijn (2009), the stronger notions of acyclicity by Kesten (2006),

and *essential homogeneity* by Kojima (2011). In this section, we prove this statement. We first introduce the definitions of the above properties.

Definition 9. A school preference profile $\succ_C$ is **Ergin acyclic** if there exist no $a, b \in C$ and $i, j, k \in S$ such that

- $i \succ_a j \succ_a k \succ_b i$ and
- there exist (possibly empty) disjoint sets of students $S_a, S_b \subseteq S \setminus \{i, j, k\}$ such that $|S_a| = q_a - 1, |S_b| = q_b - 1$, $s \succ_a j$ for every $s \in S_a$ and $s \succ_b i$ for every $s \in S_b$.

Definition 10. A school preference profile $\succ_C$ is **essentially homogeneous** if there exist no $a, b \in C$ and $i, j \in S$ such that

- $i \succ_a j$ and $j \succ_b i$, and
- there exist (possibly empty) sets of students $S_a, S_b \subseteq S \setminus \{i, j\}$ such that $|S_a| = q_a - 1, |S_b| = q_b - 1$, $s \succ_a j$ for every $s \in S_a$ and $s \succ_b i$ for every $s \in S_b$.

Definition 11. A school preference profile $\succ_C$ is **strongly x-acyclic** if there exist no $a, b \in C$ and $i, j \in S$ such that

- $i \succ_a j$ and $j \succ_b i$ and
- there exist (possibly empty) disjoint sets of students $S_a, S_b \subseteq S \setminus \{i, j\}$ such that $|S_a| = q_a - 1, |S_b| = q_b - 1$, $s \succ_a j$ for every $s \in S_a$ and $s \succ_b i$ for every $s \in S_b$.

Definition 12. A school preference profile $\succ_C$ is **Kesten acyclic** if there exist no $a, b \in C$ and $i, j, k \in S$ such that

- $i \succ_a j \succ_a k$, $k \succ_b i$, and $k \succ_b j$
- there exists a (possibly empty) set of students $S_a \subseteq S \setminus \{i, j, k\}$ such that $|S_a| = q_a - 1$ and for every $s \in S_a$, either (1) $s \succ_a i$ or (2) both $s \succ_a j$ and $k \succsim_b s$.

Definition 13. A school preference profile $\succ_C$ is **strongly Kesten acyclic** if there exist no $a, b \in C$ and $i, j, k \in S$ such that

- $i \succ_a j \succ_a k$, $k \succ_b i$, and $k \succ_b j$
- there exists a (possibly empty) set of students $S_a \subseteq S \setminus \{i, j, k\}$ such that $|S_a| = q_a - 1$ and $s \succ_a k$ for every $s \in S_a$.

It is easy to see that if a school preference profile is virtually homogeneous, then it is both Ergin acyclic and essentially homogeneous.[57] Also, given that essential homogeneity implies strong x-acyclicity by definition, any virtually homogeneous preference profile is also strongly x-acyclic. Also it is clear that strong Kesten acyclicity implies Kesten acyclicity. Thus, the only thing we have to show is that virtual homogeneity implies strong Kesten acyclicity.

Proposition 10. *If a school preference profile is virtually homogeneous, then it is strongly Kesten acyclic.*

Proof. Suppose that a school preference profile is virtually homogeneous and is not strongly Kesten acyclic, i.e., there exist $a, b \in C$ and $i, j, k \in S$ such that

- $i \succ_a j \succ_a k$, $k \succ_b i$, and $k \succ_b j$
- there exists a (possibly empty) set of students $S_a \subseteq S \setminus \{i, j, k\}$ such that $|S_a| = q_a - 1$ and $s \succ_a k$ for every $s \in S_a$.

This implies that there exist $a, b \in C$ and $i, j, k \in S$ such that

- $i \succ_a k$ and $k \succ_b i$
- there exists a (possibly empty) set of students $S_a \subseteq S \setminus \{i, k\}$ such that $|S_a| = q_a - 1$ and $s \succ_a k$ for every $s \in S_a$.

However, such a, b, i, j, and k cannot exist by the assumption that the school preference profile is virtually homogeneous. To see this point, observe that if such a, b, i, j, and k exist, then b, a, k, and i satisfy the condition in Lemma 2. (It can be verified by simply substituting (b, a, k, i) into (a, b, i, j) into Lemma 2.) However, according to the lemma, such schools and students cannot exist when a school preference profile is virtually homogeneous, a contradiction. □

In summary, the above discussions show that virtual homogeneity is stronger than acyclicity and its variants in the literature. A more detailed description of the relationships among these properties is provided in the following Venn diagram in Figure 1, which combines the results of this section with the Venn diagram on p. 1934 in Haeringer and Klijn (2009).

[57] These conclusions follow clearly by comparing conditions of Ergin acyclicity and virtual homogeneity with those in Lemma 2.

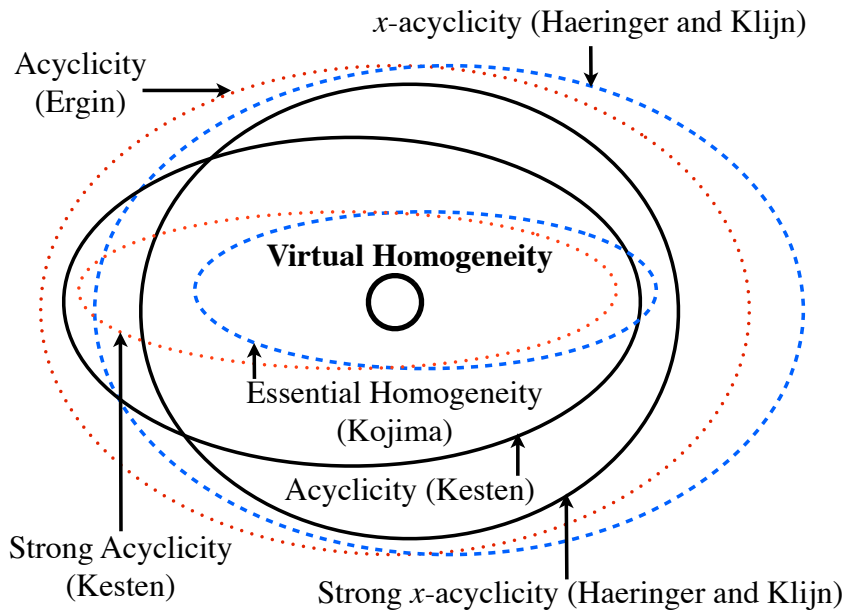

Figure 1: Relationship Between Virtual Homogeneity and Other Properties.

A.4 An Exhaustive List of the Results

The following table provides an exhaustive list of the results in this paper. In this table, "RI" is an abbreviation of respecting improvements. "✓" in a cell means that the corresponding mechanism satisfies the corresponding property (under the assumption that students truthfully report their preferences) while "×" means that the corresponding mechanism does not satisfy the corresponding property. In addition, for the Boston mechanism, which is not strategy-proof, marks in parentheses indicate results under the assumption that students play a Nash equilibrium. Specifically, "(✓)" ("✓" in parentheses) means that for any selection of a Nash equilibrium at each preference profile, the corresponding mechanism satisfies the corresponding property. On the other hand, "(×)" means that there exists a selection of a Nash equilibrium at each preference profile such that the corresponding mechanism does *not* satisfy the corresponding property.

	SOSM	Boston	TTC
RI in General Markets	×	×	×
RI for Desirable Students in General Markets	×	×	×
RI in Large Markets	✓	×(✓)	×
RI for Desirable Students in Large Markets	✓	×(✓)	×
RI in Terms of Enrollment	✓	✓	×
RI of Student Quality	✓	✓(×)	✓

Table 2: An Exhaustive List of the Results.