Promoting School Competition Through School Choice: A Market Design Approach

John William Hatfield\textsuperscript{1}, Fuhito Kojima\textsuperscript{2}, and Yusuke Narita\textsuperscript{3}

May 10, 2014

\textsuperscript{1}University of Texas at Austin
\textsuperscript{2}Stanford University
\textsuperscript{3}MIT
School Choice: Economics of Education.

The matching market design literature:

- studies how to assign school seats to students efficiently and fairly
- offers specific school choice mechanisms:
  - the so-called **Boston mechanism** (immediate acceptance mechanism) is bad,
  - the **deferred acceptance (DA)** or the **Top Trading Cycles (TTC)** mechanisms are good.

School choice reforms happened/are happening with help of economists: Boston, New York, New Orleans, etc.
School Choice and School Competition

But prior work in matching market design has not analyzed the effect of different school choice mechanisms on overall school quality (rather it has assumed that school quality is given and fixed.)

This is a serious omission, as the major impetus for school choice has been that school choice will improve overall educational quality (c.f. Neilsen 2014).

*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
Research goals:

1. Study how the design of a school choice mechanism affects competitive pressure on schools to improve themselves.
2. Provide a new perspective on what mechanisms should be used (e.g., DA versus TTC versus Boston)
What We Do

Research goals:

1. Study how the design of a school choice mechanism affects competitive pressure on schools to improve themselves.
2. Provide a new perspective on what mechanisms should be used (e.g., DA versus TTC versus Boston)

![Graph showing frequency of violating respecting improvements against the number of schools. The graph compares Boston, TTC, and SOSM mechanisms.]

John William Hatfield, Fuhito Kojima, and Yusuke Narita
Promoting School Competition Through School Choice
In this project, we

1. Introduce several criteria for improvement:
   - “respecting improvements (of school quality)” (cf. Balinski and Sonmez 1999): if students rank a school higher than before, then the outcome of the mechanism should become better for that school.

2. Show that all stable mechanisms (e.g. DA) or Pareto efficient mechanisms (e.g. TTC/Boston) violate this criterion.

These results say that no reasonable mechanism in the literature induces schools to improve themselves. Then, is it impossible to give useful information to policymakers?
A mechanism **approximately respects improvements** if: For “almost all” preference profiles, a school is not made strictly better off when students demote the school in their rankings (in large markets; see Immorlica and Mahdian 2005, Kojima and Pathak 2009).

1. Any stable mechanism approximately respects improvements (we also show a school becomes strictly better off when students promote it).
2. Boston and the TTC mechanisms (two major competitors to stable mechanisms) do not even approximately respect improvements.

Also we introduce several alternative criteria, and show similar conclusions.

→ Stable mechanisms promote school competition for higher quality better than other mechanisms!
Related Literature


**Analytical Methods in Matching Theory:**
Basic Setup

There are a set of students $S$ and a set of schools $C$.

Preferences:

- Each student $s$ has strict preferences $\succ_s$ over schools and the outside option, $\emptyset$.
- Each school $c$ has strict preferences $\succ_c$ over groups of students. Preferences are “responsive” (Roth 1985), with capacity $q_c$, and all students are acceptable (=better than leaving a seat vacant).

A matching $\mu$ specifies which student is matched with which school: The match for student $s$ is denoted by $\mu_s$, and the match for school $c$ is denoted by $\mu_c$. 
The Deferred Acceptance (DA) Algorithm

Gale and Shapley (1962) propose the (student-proposing) deferred acceptance algorithm.

- Step 0: Start from a matching in which no one is matched.
- Step $t \geq 1$: Each student who was not tentatively matched to any school in Step $(t-1)$ applies to her next highest acceptable school (if any). Each school considers these students and students who are temporarily kept from the previous step together, and temporarily keeps the highest-ranking students up to its capacity and rejects everyone else (so students not rejected at this step may be rejected in later steps.)

The algorithm terminates at a step in which no rejection occurs, producing a matching.

→ stability, student-optimality (though not Pareto efficient for students), strategy-proofness.
Our Central Concepts (1): Improvement

Preference $\succ'_s$ of a student $s$ is an **improvement for school** $c$ over another preference $\succ_s$ if

1. For any $c' \in C \cup \{\emptyset\}$, if $c \succ_s c'$ then $c \succ'_s c'$, and
2. For any $c', c'' \neq c$, $c' \succ_s c''$ if and only if $c' \succ'_s c''$.

Preference profile $\succ'$ is an **improvement for school** $c$ over student preference profile $\succ$ if $\succ'_s$ is an improvement for school $c$ over $\succ_s$ for all $s \in S$ while all school preferences are unchanged. We also say that $\succ$ is a **disimprovement** over $\succ'$. An improvement for school $c$ means that students rank $c$ higher while keeping rankings between other schools unchanged.
Our Central Concepts (2): Respecting Improvement

Definition (see Balinski and Sonmez 1999)

A mechanism $\varphi$ respects improvements (of school quality) at school preference profile $\succ_C$ if, for every school $c$ and student preference profiles $\succ_S$ and $\succ'_S$, if $\succ' = (\succ'_S, \succ_C)$ is an improvement for school $c$ over $\succ = (\succ_S, \succ_C)$, then $\varphi_c(\succ') \succeq_c \varphi_c(\succ)$.

Equivalently, a mechanism $\varphi$ respects improvements of school quality if there exist no school $c$ and student preference profiles $\succ_S$ and $\succ'_S$ such that $\succ'$ is a disimprovement for school $c$ over $\succ$ and $\varphi_c(\succ') \succeq_c \varphi_c(\succ)$.

This condition requires that no school has incentives to make itself disliked by students.
DA Does Not Respect Improvements

Students 1, 2, schools $a, b$, and preferences given by

\[
\succ_1: b, a, \emptyset, \quad \succ_a: 1, 2; q_a = 2, \\
\succ_2: b, a, \emptyset, \quad \succ_b: 2, 1; q_b = 1.
\]

DA matches $a$ to 1 and $b$ to 2.

Change student 2’s preferences by $\succ'_2: a, b, \emptyset$ while preferences of everyone else are unchanged. This is an improvement for school $a$.

At this preference profile, DA matches $a$ to 2 and $b$ to 1.

School $a$ is worse off even if its ranking improves in student preferences.
Impossibility

More generally,

**Theorem**

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

Thus no stable mechanism can induce incentives for schools to be attractive to students.

Proof: In the last example, one can verify that there is a unique stable matching under each preference profile.

A standard approach in traditional matching literature is to find a condition on school preferences ("domain restriction"). The necessary and sufficient condition is

- Each school has only one seat, or
- School preferences are **virtually homogeneous**, that is, the ranking of students are almost the same across all schools.
Two other mechanisms are popular.

1. **Boston mechanism**: similar to DA, but all matches in each step of the algorithm are final. Used in many school districts, although recently under attack because of poor incentive and fairness properties.

2. **Top Trading Cycles (TTC) mechanism**: Allow students to “trade priorities at schools.” A favorite competitor to DA in market design research due to its efficiency and incentive properties (used in new school system in San Francisco?).

These mechanisms are **Pareto efficient (for students)**: there is no other matching that is weakly preferred to the outcome of the mechanism by every student.
Impossibility For Pareto Efficient Mechanisms

Theorem

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

Thus no Pareto efficient mechanism can induce incentives for schools to be attractive to students.  

Intuition for Boston mechanism.

Analogously to stable mechanisms, one can show that the necessary and sufficient condition is that school preferences are virtually homogeneous.
School districts usually have many schools and students.

We will seek the “approximate” criterion: For most preference profiles, do schools have incentives to improve?

Formally, consider “random markets” as defined by Immorlica and Mahdian (2005), extended by Kojima and Pathak (2009) to fit the current setting: The market size is indexed by the number of schools, and preferences are drawn from a certain distribution.
Main Result: Improvement in Large Markets

For any school $c$ and mechanism $\varphi$, let $\alpha_c$ be the probability that under the realized preference profile, school $c$ can be made strictly better off when some students demote the school’s ranking.

We say that a mechanism **approximately respects improvements in large markets** if

$$\max_{c \in C^n} \alpha_c \to 0 \quad (\text{as } n \to \infty),$$

**Theorem**

Any stable mechanism approximately respects improvements in large markets. By contrast, neither the Boston mechanism nor the TTC approximately respects improvements even in large markets.
Intuition for stable mechanisms (focus on student-proposing DA):
Recall the first example,

Example
Students 1, 2, schools a, b,

\[\succ_1: b, a, \emptyset,\]  \[\succ_a: 1, 2, q_a = 2,\]

\[\succ'_2: a, b, \emptyset,\]  \[\succ_b: 2, 1, q_b = 1,\]

\[\succ_2: b, a, \emptyset\]

Student 2 changes preferences from \(\succ'_2\) to \(\succ_2\), demoting school a. School a is better off under \(\succ_2\) even though its ranking is worse.

A school (school a) is made better off when a student (student 2) demotes its ranking because it increased competition in a different school (school b), thus creating a “rejection chain” that reaches the original school (school a).
Recap from last slide: A school can be made better off when a student demotes the school in her ranking because it increases competition in a different school, thus creating a “rejection chain” that reaches the original school.

In a large market, (we can show that) there is a high probability that there are many schools with vacant seats. So students in the rejection chain are likely to apply to those schools with vacancies and be matched. So the school is unlikely to be made better off.

A formal proof needs to deal with additional issues through a number of steps:

- Does the intuition go through for arbitrary stable mechanisms?
- Is the rejection chain the *only* reason?
Boston and TTC mechanisms

Theorem

Neither the Boston mechanism (very popular in practice) nor the TTC (considered to be appealing in existing studies) approximately respect improvements even in large markets.

Intuition for the Boston mechanism:

1. In Boston mechanism, a student who applies in an earlier step is matched over a more preferred student who applies later.
2. So the school has incentives to make itself disliked by undesirable students.
3. This effect does not necessarily become small even in large markets.
Neither the Boston mechanism (very popular in practice) nor the TTC (considered to be appealing in existing studies) approximately respect improvements even in large markets.

Intuition for TTC:

1. In TTC, even an undesired student may be matched if the student can trade priorities with another student who has a high priority in that school.
2. Thus a bad student pointing to a school may be bad news.
3. This effect can remain even in large economies.
Theorem

Neither the Boston mechanism (very popular in practice) nor the TTC (considered to be appealing in existing studies) approximately respect improvements even in large markets.

Intuition for TTC:

1. In TTC, even an undesired student may be matched if the student can trade priorities with another student who has a high priority in that school.
2. Thus a bad student pointing to a school may be bad news.
3. This effect can remain even in large economies.

→respecting improvements in large markets can be used to distinguish good mechanisms from bad ones!!
Simulation: Method

Simulation motivated by typical school choice setting.

- Two types of schools: With and without a good ESL (English as a second language) teacher.
- Two types of students: need ESL class or not.
- On average, an ESL student is less desirable to schools.
- Each school can fire an ESL teacher → becomes undesirable to ESL students, but not to non-ESL students.
- 300 iterations for each mechanism (DA, TTC, and Boston)

Goals: Study

- Quantitative relevance of our theory.
- If strategic disimprovement works when a school has limited ability to manipulate (i.e., the only choice is keeping or firing a good ESL teacher).
Manipulation possibility exists for each mechanism.

Frequency: Boston > TTC > DA

Bad incentives vanishes for DA as the market size becomes large, but not in Boston or TTC.
Alternative Concept: Enrollment

Schools’ “preferences” may be just priorities set by law.

In such cases, schools may primarily be interested in admitting as many students as possible:

1. School budgets are often determined by enrollment.
2. Schools attended by too few students are often closed.

Definition

A mechanism $\varphi$ respects improvements in terms of enrollment if, for any $c \in C$, if $\succ'$ is an improvement for $c$ over $\succ$, then $|\varphi_c(\succ')| \geq |\varphi_c(\succ)|$.

If a mechanism respects improvements in terms of enrollment, then schools that only want to admit more students have (weak) incentives to become attractive to students.

Remark: There is no logical relationship between respecting improvements and respecting improvements in terms of enrollment.
Results for Enrollment

Theorem

For any school preference profile, any stable mechanism respects improvements in terms of enrollment.

So, there is another sense (in addition to those in large economies) that stable mechanisms provide good incentives for schools to improve.

Proposition

For any school preference profile, the Boston mechanism respects improvements in terms of enrollment.

Proposition

TTC does not necessarily respect improvements in terms of enrollment.

→ enrollment criterion can be used to distinguish good mechanisms from bad ones!!
The paper investigates several additional topics:

1. Similar conclusions are obtained for “respecting improvements for desirable students,” → Detail.

2. What about students’ incentives to improve? DA, Boston, and TTC “respect improvements of student quality.” → Detail.

3. Domain Restriction: Characterization results (briefly mentioned before). → Detail.
Conclusion

We studied how the design of a school choice mechanism affects competitive pressure on schools to improve. We

1. formalized the concept of “respecting improvements of school quality” as a criterion.
2. showed that most mechanisms (all stable or Pareto efficient mechanisms) violate the criterion and, even worse, the condition for guaranteeing it is very stringent (virtual homogeneity).
3. showed that any stable mechanism approximately respects improvements, while Boston and TTC do not.
4. showed that similar conclusions can be obtained for other criteria such as “respecting improvements in terms of enrollment” and “respecting improvements for desirable students.”

→ Help design school choice mechanisms by informing what mechanisms promote school competition for higher quality.
Thank you! All comments and questions are greatly appreciated.
Matching $\mu$ is **blocked** by $(s, c) \in S \times C$ if

1. $c \succ_s \mu_s$ and
2. either (1) $|\mu_c| < q_c$ or (2) $s \succ_c s'$ for some $s' \in \mu_c$.

A matching $\mu$ is **individually rational** if $\mu_s \succ_s \emptyset$ or $\mu_s = \emptyset$ for every $s \in S$.

A matching $\mu$ is **stable** (Gale and Shapley 1962) if it is individually rational and it is not blocked.

Stability can be understood to be the absence of “justified envy.”
What Condition Guarantees Respect?

As standard in matching literature, we will first seek a condition on school preferences such that there exists a stable mechanism that respects improvements.

Let \( r^t(c) \) be the student who is the \( t \)-th ranked student according to \( \succ_c \).

**Definition**

A school preference profile \( \succ_c \) is **virtually homogeneous** if \( r^t(a) = r^t(b) \) for all \( a, b \in C \) and \( t > \min \{ q_c | c \in C \} \).

Virtual homogeneity requires that preferences over students are almost identical across schools.

It implies most restrictions studied in the literature, including “acyclicity” (Ergin 2002), “Kesten-acyclicity” (Kesten 2006), strong \( x \)-acyclicity (Haeringer and Klijn 2009), and “essential homogeneity” (Kojima 2010).
Theorem

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

1. $\succ_C$ is virtually homogeneous, or
2. For every school, its capacity is one.

The main implication is from the “only if” part: the condition that makes stability compatible with respecting improvements is extremely restrictive.

Proof:

1. “Only if” part: With some work, we can “embed” the previous example failing to respect improvements into any economy that violates the conditions.
Theorem

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

Recall that virtual homogeneity requires that school preferences of schools are almost identical to one another.

So the condition that guarantees respecting improvements is extremely restrictive.

Proof:

1. “Only if” part: With some work, we can “embed” the previous example into any economy that violates virtual homogeneity.

Proof of the Impossibility Result

Assume a mechanism is Pareto efficient and respects improvements. Students 1, 2, schools \( a, b \), where \( 2 \succ_a 1 \), \( 1 \succ_b 2 \), and \( q_b = 1 \).

1. For preference profile: \( \succ_1: b, \emptyset \), \( \succ_2: a, \emptyset \), the unique Pareto efficient matching is \( (1 \ 2) \) \( (b \ a) \).

2. Under new preferences \( \succ' := (\succ_1, \succ'_2) \) where 2’s preference has changed: \( \succ_1: b, \emptyset \), \( \succ'_2: b, a, \emptyset \),
   - Pareto efficient matchings are \( (1 \ 2) \) \( (b \ a) \) and \( (1 \ 2) \) \( (\emptyset \ b) \).
   - School \( b \) is matched with 1 under \( \succ \) and \( 1 \succ_b 2 \), so respecting improvements means \( b \) has to be matched with 1, so the resulting matching should be \( (1 \ 2) \) \( (b \ a) \).

3. Under another preference profile \( \succ'': = (\succ'_1, \succ'_2) \) such that \( \succ'_1: a, b, \emptyset \), \( \succ'_2: b, a, \emptyset \), the unique Pareto efficient matching is \( (1 \ 2) \) \( (a \ b) \).

Note that school \( a \) is worse off at \( \succ'' \) than at \( \succ' \) although \( \succ'' \) is an improvement for school \( a \) over \( \succ' \), a contradiction.
First, we show that, for stable mechanisms, there is a formal connection between respecting improvements and manipulability of the mechanism.

**Proposition**

Let $\varphi$ be a stable mechanism and preference profiles $\succ$ and $\succ'$ be such that $\varphi_c(\succ') \succ_c \varphi_c(\succ)$ for a school $c \in C$, where $\succ'$ is a disimprovement for $c$ upon $\succ$. Then there exists a (stated) preference $\succ'_c$ for $c$ such that $\varphi_c(\succ'_c, \succ_{-c}) \succ_c \varphi_c(\succ)$.

As a matter of fact, there is an “if and only if” relationship.
A More Formal Proof Sketch: Step 2

Then we use the following result to focus on the student-proposing DA.

**Result (Pathak and Sonmez 2008)**

_The student-proposing DA is the “most manipulable stable mechanism” for colleges: Given preference profile, if a stable mechanism can be profitably manipulated by a school, then the student-proposing DA is manipulable by the same school at that preference profile._
A More Formal Proof Sketch: Step 3 (Last Step)

Finally, invoke an incentive-compatibility result of student-proposing DA in large markets.

Result (Kojima and Pathak 2009)

*Then for any stable mechanism and* $\varepsilon > 0$, *there exists* $n_0$ *such that, for any random market with more than* $n_0$ *schools and any school* $c$, *the probability that a school can profitably manipulate the student-proposing DA is smaller than* $\varepsilon$.

This finishes the proof.
It may be natural to think that each school tries to improve and make itself more attractive only for desirable students for that school.

Consider a condition that requires a mechanism to respect at least a part of improvements, i.e., those in preferences of desirable students for each school.

Definition

A mechanism $\varphi$ respects improvements of school quality for desirable students at the school preference profile $\succ_C$ if schools are made better off by every improvement such that only students who are more preferred to some student who is currently matched ranks the school higher.
It turns out that stable mechanisms, Boston, and TTC all fail to respect improvements for desirable students.

On the other hand, clearly any stable mechanism approximately respect improvements for desirable students in large markets.

Neither Boston nor TTC satisfies this condition (intuitions are similar to the “approximate” result for the original concept).
Respecting Improvements of Student Quality

It is also important that a student not have incentives to make schools rank her lower in order to obtain a more preferred school.

**Definition**

A mechanism $\varphi$ 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 $\varphi_s(\succ'_C, \succ_s) \succeq_s \varphi_s(\succ_C, \succ_s)$.

This definition is analogous to that for respecting improvements of school quality.
Lemma 1 of the paper allows us to easily prove the following corollary, which was first shown by Balinski and Sonmez (1999).

**Corollary**

_The unique stable mechanism that respects improvements of student quality at every student preference profile is the student-optimal stable mechanism._

It turns out that both Boston and TTC also respect improvements of student quality.
The setup follows Kojima and Pathak (2009):

- **A random market** is a tuple $\tilde{\Gamma} = (C, S, k, D)$, where
  - $k$ is a positive integer and
  - $D$ is a pair $(D_C, D_S)$ of probability distributions.

Each random market induces a market by randomly generating preferences of students.

- $D_S = (p_c)_{c \in C}$ is a probability distribution over $C$.
- Preferences of each student are drawn independently without replacement using probability distribution $D_S$ to form the preference list of students of length $k$.

The preference distribution of schools is completely general: $D_C$ may be any distribution (or even degenerate).
Definition (Regularity)

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 \(D_{C^n}\).

- We also impose the condition that the market is sufficiently thick, i.e. that there are no ‘super-popular’ schools.
- For example, if \(\frac{p_c}{p_{\bar{c}}} \leq T\) for some \(T \in \mathbb{R}\) for all \(c, \bar{c} \in C\), the market is sufficiently thick.
Simulation Method in Detail

1. For each student $s$, her type $t_s$ is either ESL or NESL.
2. Each school $c$ can choose type $t_c \in \{Good, Bad\}$.
3. Utility for student $s$:

$$u_s(c) = a_{t_s}[b\epsilon_{sc} + (1 - b)\zeta_c] + (1 - a_{t_s})1_{t_c=Good},$$

where $\epsilon_{sc}$ and $\zeta_c$ are i.i.d. in $U[0, 1]$, $a_{NESL} = 1$, $a_{ESL} = 0.5$, $b = 0.5$.
4. Utility for school $c$: additive with quota (Poisson with mean 20), utility for an individual student $s$ is

$$v_c(s) = \bar{a}[\bar{b}\eta_{cs} + (1 - \bar{b})\theta_s] + (1 - \bar{a})1_{t_s=NESL},$$

where $\bar{a} = \bar{b} = 0.5$ and $\eta_{cs}, \theta_s$ are random variables i.i.d. in $U[0, 1]$.
5. # of students is from $U[\# \text{ of seats}, 1.5 \times (\# \text{ of seats})]$