PREFERENCE RESPECTING STABLE MATCHINGS IN SCHOOL CHOICE PROBLEMS

by

ALİ ŞİMŞEK

Submitted to the Institute of Social Sciences
in partial fulfillment of the requirements for the degree of
Master of Arts

Sabancı University

JUNE 2017

PREFERENCE RESPECTING STABLE MATCHINGS
IN SCHOOL CHOICE PROBLEMS

APPROVED BY:

Asst. Prof. Mustafa Oğuz Afacan ....
(Thesis Supervisor)

Asst. Prof. Mehmet Barlo ....

Asst. Prof. İpek Gürsel Tapkı ....

DATE OF APPROVAL: 29.06.2017

Abstract

PREFERENCE RESPECTING STABLE MATCHINGS
IN SCHOOL CHOICE PROBLEMS

Ali Şimşek
Economics, Master of Arts Thesis, June 2017
Thesis Supervisor: Asst. Prof. Mustafa Oğuz AFACAN

We introduce a new stability notion called preference respecting stability that incorporates tolerance values for schools and attaches importance to both the preferences of students and the priorities of schools, and study its properties. We find that a preference respecting stable allocation exists in any school choice problem, and it Pareto-dominates the Gale-Shapley stable allocation. We construct a two part mechanism that depends on improvement cycles to reach a constrained efficient preference respecting stable allocation. Our mechanism is a natural generalization of a broad class of mechanisms and admits the student-optimal Stable Mechanism and the Boston Mechanism as special cases. We also study its strategic properties under complete and incomplete information settings and find that truthful reporting of preferences is an ordinal Bayesian Nash equilibrium for the students.

Keywords: Matching Theory, School Choice, Boston Mechanism, Student-Optimal Stable Mechanism, Pareto-Efficiency.

Özet

OKUL SEÇİMİ PROBLEMLERİNDE TERCİHE RİAYETLİ SABİT EŞLEŞMELER

Ali Şimşek
Ekonomi, Yüksek Lisans Tezi, Haziran 2017
Tez Danışmanı: Yrd. Doç. Dr. Mustafa Oğuz AFACAN

Tercihe riayetli sabitlik adında, okullar için tolerans değerleri içeren ve hem okulların önceliklerine hem de öğrencilerin tercihlerine önem veren, yeni bir sabitlik nosyonu sunduk. Her okul seçimi problemi için bir tercihe riayetli sabit eşleşmenin var olduğunu ve Gale-Shapley sabit eşleşmesine Pareto-üstün olduğunu bulduk. Sınırlı-verimli bir tercihe riayetli sabit eşleşmeye ulaşmak için geliştirme çemberlerine dayanan, iki adımlı bir mekanizma geliştirdik. Mekanizmamız geniş bir sınıf mekanizmaların doğal bir genellemesi ve öğrenci-optimal Sabit Mekanizma ve Boston Mekanizması'nı özel vakalar olarak kapsamakta. Ayrıca, mekanizmanın stratejik özelliklerini tam ve eksik bilgi durumları altında inceledik ve tercihleri dürüst bildirmenin, öğrenciler için ordinal Bayes Nash Dengesi olduğunu bulduk.

Anahtar Kelimeler: Eşleşme Teorisi, Okul Seçimi, Boston Mekanizması, Öğrenci-Optimal Sabit Mekanizma, Pareto-Verimlilik.

⟨⟨ ***To my mother Sema Şimşek;***

and my father Mehmet Şimşek. ⟩⟩

Acknowledgements

First and foremost, this thesis would not have come into existence without my advisor, Mustafa Oğuz Afacan. I am eternally indebted to him, both for this thesis, and also for any scientific achievement I might earn in the future. I am also profoundly thankful to Mehmet Barlo, who is also a jury member for this thesis, for introducing me to the beauty of the economic sciences and patiently teaching me the foundations of it, year after year.

I am also immensely grateful to my parents, Sema and Mehmet Şimşek, for their continuous support in every possible dimension.

I would also like to appreciate the support of my friends; Ceren Bengü Çıbık, Candan Kurşungöz and Kaan Erdemli for their academic and mental support, and also Bihter Özturk, Eda Derinel, Dilara İnam, Büşra Karacan and Elif Birced for their mental support, and essentially putting up with me in the process, day in and day out.

I am also grateful to the remaining jury member İpek Gürsel Tapkı, for her helpful comments and remarks.

Contents

1 Introduction

Since their formulation by Abdulkadiroğlu and Sönmez (2003), school choice problems generated a unique interest in market design research, mostly due to the importance of its policy applications[1]. In the process of placing pupils to public schools within a school district with a centralized placing system, each family submits a list of schools ranked by their preference order, and all the schools have a priority ranking over the pupils based on their features, and these priorities are decided by the individuals-in-charge of the schools districts. These features can be the distance of their house to the school, whether they have a sibling at that school, some kind of affirmative action policy, etc. The policies of school districts that have centralized systems affect a lot of families every year, hence the welfare implications of this design problem is sizable.

School choice problems were introduced to the economic literature more than a decade ago, but the tools employed in solving the problem are older. Although they have differences, another tangent design problem to school choice is the college admissions problem, and it has been extensively studied since the seminal paper by Gale and Shapley (1962)[2], which also started the now-voluminous literature of two-sided matching markets[3]. They introduced a stability notion as a desired property of any college admission problem's outcome and constructed an algorithm, namely the Deferred Acceptance Algorithm, that can produce a stable outcome in any college admission problem, and hence, equivalently provided the important existance result, that a stable matching exists in any college admission problem.

[1]Some excellent surveys on the school choice problems literature are written by Abdulkadiroğlu (2011), Pathak (2011), Abdulkadiroğlu and Sönmez (2013) and Sönmez and Ünver (2011).

[2]The origin of the problem comes from the stable marriages problem stated in the same paper, which studies the existence of a monogamic "stable marriage" between a set of men and a set of women in which all individuals have preferences over the agents of the other set.

[3]Although it does not survey the school choice problems, Roth and Sotomayor (1990) is a very thorough introduction to the two-sided matching markets, and also their game theoretic modelling.

But the most important distinction between the college admission problems and the school choice problems is the role of supply side of the market. In college admission problems, the universities have preferences over the set of students, which are, most of the time, determined by the officials of the university itself. In contrast, in the school choice problems, the schools do not (and can not) have a preference ranking over the students, the priorities are given to them exogenously. Hence, the seats at a school in a school choice problem are mere consumption objects and the schools cannot be seen as strategic agents in the market, whereas this not the case in the college admission problem. Moreover, due to this property, only the welfare of the students and their families are of concern [4,5].

Just as with the other market design problems, Gale and Shapley's stability notion and the Deferred Acceptance Algorithm are translated into school choice problems as well, and they are employed in real life in a lot of school districts today. Properties of stable assignments in school choice problems are extensively studied since then. In the school choice context, stability is a desirable property because the parents might be unhappy when a lower priority pupil is placed to a school that they desired but could not place (the formal definition of stability will be introduced later) and they might object to this placement by taking legal action. A well-known result in the school choice literature is that a stable matching exists in any school choice problem, a result that follows Gale and Shapley's (1962) existence results for the marriage market and for the college admission problems. Abdulkadiroğlu and Sönmez (2003)

[4] Another very important difference between the school choice problems and the college admission problems is the nature of the priorities of the supply side of the market. In the college admission problems, the universities generally have strict priority orders without any indifference classes. But in the school choice problems, due to the nature of the problem in the real life as well, the schools assumed to have "coarse priorities", that is, the priority profiles of the schools contain large classes of indifferences. Hence, in most of the school choice mechanisms, a tie-breaking rule is needed. Interested readers should refer to Abdulkadiroğlu and Sönmez (2013) and Sönmez and Ünver (2011).

[5] Another closely related problem is the student placement problem, as introduced by Balinski and Sonmez (1999). The main difference between the college admission problem and the student placement problem (which also constitutes the main difference between the school choice problem and the college admission problem) is the roles of the colleges in the design. In the college admission problem, the colleges are active agents in the market, can behave strategically and their welfare might be the main concern of the designer. Whereas in the student placement problem, the college seats are merely objects to be consumed and the welfare implications of the allocation for the colleges are not considered by the designer. What makes the school choice problem different from the student placement problem is that the schools have exogenously given priorities that might have indifferences, rather than endogenously determined strict preferences of universities over the students in the student placement problem.

showed that stable outcomes produced by the student-optimal Stable Mechanism, which employs student-proposing Deferred Acceptance Algorithm, (although they are Pareto-efficient matchings in the set of stable matchings for a school choice problem as shown by Balinski and Sönmez (1999)) can be Pareto-dominated by other assignments that are not stable, hence they might lead to efficiency losses. Another very prominent school choice mechanism is observed by Abdulkadiroğlu and Sönmez (2003) for the Boston Public Schools system, hence the mechanism aptly named the Boston Mechanism (the mechanism will be formalized in Section 3). They comment that the Boston Mechanism is Pareto-efficient (in the general sense, not restricted to any class of matchings) when the families report their preferences truthfully[6, 7]. The main difference between the student-optimal Stable Mechanism and the Boston Mechanism is the step at which the assignments are made permanent in their respective algorithms. In the Deferred Acceptance Algorithm (which is the underlying algorithm of the student-optimal Stable Mechanism), the assignments are not permanent until the last round of the algorithm finishes. Whereas in the Boston Mechanism, the assignments are made permanent at the end of every round of the algorithm[8]. Beginning with this observation, trade-off between stability and Pareto-efficiency became an important aspect of the research in school choice problems, due to the scale of the problem and the probable efficiency cost mentioned before.

The Boston and student-optimal Stable Mechanisms have been used commonly in a lot of districts to match the students to available seats in the schools they want to attend. The Boston Mechanism enables the schools to favor the the applicants who list the schools at the higher ranks in their preference lists and might increase the ex-ante welfare properties of the allocation, but it does not lead to stable outcomes, as defined by Gale and Shapley (1962), and the procedure can be manipulated by the students, a situation that might create a sizable welfare loss, especially

[6]Notice that, due to the fact that the Boston Mechanism is not strategy-proof, this can hardly be the case in real life, and there is evidence about families misreporting their preferences. When families misreport their preferences in order to behave strategically, Boston Mechanims is not Pareto-efficient.

[7]Abdulkadiroğlu and Sönmez (2003) also show that Top Trading Cycles Mechanism is Pareto-efficient, and also it is strategy proof. In 2012, a mechanism based on Top Trading Cycles Mechanism was used in New Orleans as well. Interested reader should refer to Abdulkadiroğlu et al. (2017)

[8]Due to this property, the algorithm that produces the outcome of the Boston Mechanism is sometimes called the Immediate Acceptance Algorithm (in contrast to the Deferred Acceptance Algorithm).

for the unsophisticated agents. But the stable allocations that is produced by the student-optimal Stable Mechanism might be Pareto-dominated by the outcome of the Boston Mechanism[9]. Hence, there is a trade off between efficiency and stability (and strategy-proofness) in the implementation of the Boston and the student-optimal Stable Mechanisms. This problem can be solved via a hybrid mechanism that tries to reconcile between the nice properties of these two mechanisms. Moreover, as Abdulkadiroğlu (2011) pointed out, the schools have a demand for a hybrid mechanism. Some school might want to increase efficiency and others might prefer stable allocations. One of the main problems is that these schools might be within the same district. For example, within a single school district, some types of schools might be legally bounded to process a specific priority profile, but other types of schools might be more independent and try to admit the students who desire them the most[10].

This observation creates a natural research direction: Can we find a midpoint between these two mechanism while giving the schools the flexibility they desire? A natural way of finding such a midpoint is to loosen the requirements of the stability notion. In this paper, we define a weaker stability notion called preference respecting stability, in which the allocations that satisfy this stability notion might Pareto-dominate the student-optimal Stable Mechanism allocation in any given school choice problem, and it also maintains some of the desiderata of the usual Gale-Shapley stability. We say that a matching is preference respecting stable if (*i*) no student prefers being unassigned to her assigned school, (*ii*) assignment is non-wasteful, and (*iii*) if a student prefers another school c to her assigned school, then cannot be ranked l_c higher than any student that is assigned to school c, and if she is ranked higher at a rank less than l_c, then the envied student ranks school c higher than her in their respective preference profiles, where l_c is the toleration profile of a school and can be interpreted as the school's the answer to the following

[9]One of the most important results is provided by Troyan (2012), who showed that the Boston Mechanism Pareto-dominates any other stable mechanism, from an ex-ante utility perspective.

[10]Abdulkadiroğlu (2011) provides an example from the Boston Public Schools (BPS). In BPS, exam schools and regular schools are both parts of the centralized system. In the exam schools, the priorities are based on an entrance exam score and GPA, whereas in the regular schools, the priorities are based on proximity to the school and siblings' schools. It is commented that violating priorities in exam schools are "politically and legally unfeasible", while the BPS authorities and the public considered violating priorities to increase the student welfare in public schools. A similar account also provided for the New York City high school market.

question: "How much rank difference can you tolerate to get a student who ranks you higher in her preference profile?". Hence in our framework, a student i can justifiably envy another student j at school c if i is ranked l_c higher in the priority profile of the school, or if i fails to do that i is ranked higher in the priority profile of the school and i ranked school c weakly higher in her preference profile compared to student j. Intuitively, we enable schools to favor the students who ranked them higher in their preference lists, and l_c acts as a bound on how much the schools can favor the students while violating the priorities. To the best of our knowledge, this is the first study that incorporates the preferences of the students into the priority profiles in such a way.

One important issue to notice is that any given (Gale-Shapley) stable allocation is also preference respecting stable. Due to this property, a preference respecting stable allocation exists in every school choice problem, stemming from the existence result of Gale and Shapley (1962). We then create a two-step mechanism, aptly called the Preference Respecting Stable Mechanism, that would produce a preference respecting stable allocation in any school choice problem. Our formulation of the mechanism defines a class of mechanisms depending on the tolerance values of the schools in the market, and admits the Boston Mechanism and the student-optimal Stable Mechanism as special cases. The first step of our algorithm is the usual Deferred Acceptance Algorithm, and second step consists of applying efficiency improvement cycles that would not violate the requirements of preference respecting stability.

Our mechanism also taps into a unique design approach. Under our setting and mechanism, schools can decide whether they want to eliminate justified envy or they want to favor the applicants who simply want them more, by adjusting their own tolerance values. Hence, if a school wants (or bounded by law) to eliminate justified envy, then it can directly equate its tolerance value to 0 (and hence do not violate any priorities even for accepting the "more-willing" students), whereas if a school want to favor applicants who ranked it higher in her preference report, then it can adjust its tolerance value and increase it to the point that it is willing to violate the priorities. And most importantly, the needs of all these types of schools can be satisfied with our centralized mechanism. This is not the case in many of the existing mechanisms today.

The rest of this paper is organized as follows: In section 2, we talk about the related literature. In Section 3, we formally introduce the usual school choice problem, the student-optimal Stable Mechanism and the Boston Mechanism and comment about their respective properties. In Section 4, we introduce our model (that incorporates the tolerance vector to the school choice problem) and our stability notion. In Section 5, we construct the Preference Respecting Stable Mechanism and uncover some of its properties. In Section 6 and 7, we analyze the strategic properties of the Preference Respecting Stable Mechanism under complete and incomplete information settings, respectively. All the proofs for our results are relegated to Section 8, which constitutes the Appendix.

2 Related Literature

Trade off between efficiency and stability has been studied extensively in the school choice literature. The most important result that induced this area is that a matching mechanism cannot be Pareto-efficient (in the unconstrained sense) and stable at the same time. Although, at the beginning, the direction of the research pointed towards the student-optimal Stable Mechanism as the ultimate solution to the school choice problems (and many policy-makers followed suit and switched to the student-optimal Stable Mechanism); as some nice efficiency features of the Boston Mechanism were uncovered, researchers became more and more interested in examining the nature of this trade-off. Hence, in line with that research direction, there are other studies that employ hybrid mechanisms and try to improve the efficiency properties of the matching allocations. Just like in this paper, some researchers weakened the stability notion in various ways to increase the welfare gains, but maintain some kind of stability. Moreover, in the vast part of the literature, stability is considered an equivalent notion to fairness, which is achieved via eliminating "justified envy". Hence most of these weak stability notions are researchers interpretation of what fairness and justified envy should (or can) be in school choice problems.

Kesten (2010), motivated by the efficiency loss that can be created by the student-optimal Stable Mechanism, and the supporting empirical evidence of Abdulkadiroğlu et al. (2009), proposed a new mechanism called the Efficiency-Adjusted Deferred-Acceptance Mechanism (EADAM). In the school choice problem setting they consider, the students are asked whether they consent to the central clearinghouse to waive her priority rights when it does not affect her assignment. With this consents, they create a mechanisms that identify "interrupters" that cause rejection cycles and reduce the welfare created by the allocation. EADAM relies on removing these rejection cycles via an iterative algorithm that builds upon the Deferred Acceptance

Algorithm, and eliminate the efficiency loss that would have been created by these rejection cycles under the student-optimal Stable Mechanism.

Another paper that is related to our study and precedes Kesten's study is by Erdil and Ergin (2008). In their setting, schools have "coarse" priorities, and randomly breaking priority ties creates inefficiency problems. In order to solve this problem, they introduce the Stable Improvement Cycles Mechanism, and relies on, as the name suggests, improvement cycles that respects stability of the allocation. In this mechanism, initially the Deferred Acceptance Algorithm is applied with exogenous tie-breaking, and then improvement cycles that preserves stability are identified and implemented. With this study, Erdil and Ergin also initiated a new class of mechanisms that rely on improvement cycles that are implemented on the outcome of the Deferred Acceptance Algorithm, a class that also our mechanism belongs to.

For a recent example, Morrill (2015) created their interpretation of fairness, called "justness", which is a direct weakening of the usual justified envy notion. They say that a student i depends on another student j if j can replace i in her assigned school by reporting another preference profile than her current one. Building on this idea, they say that an assignment is unjust if a student i prefers a school c to her assignment, has a higher priority than a student j at the priority profile of school c and none of the students that have higher priority than i at the priority profile of school c depend on j. Therefore if an objection might harm a higher priority student, than that objection is classified as unjust, but there is no way in which that objection can harm a higher priority student, than that objection is just. They observe that the Top Trading Cycles Mechanism creates an efficient, strategy-proof and just mechanism.

Another intuitive weakening of stability is studied by Afacan, Aliogullari and Barlo (2016). They observe a real life fact: appeals to assignments are costly. They create a framework that incorporates appeal costs, and since parents would appeal if the benefit from appealing and placing at a higher ranked school exceeds the cost of it. In order to achieve this, they let the designer of the mechanism to ask parents for the least rank difference that parents would appeal for. Notice that in their case they confine themselves to the case where appeals arise from priority violations, and hence the appeals always concludes with granted appeals. They call this rank difference minus one the stickiness degree of a student. With these stickiness degrees, they define their own version of justified envy (and the corresponding stability notion called “sticky stability”) which allows priority violations that would not contradict with the stickiness degrees. With their weaker stability notion, they create sticky stable assignments that Pareto-dominate the usual stable assignments.

The main difference between the student-optimal Stable Mechanism and the Boston Mechanism is that in their respective underlying algorithms, the step at which the assignments are finalized are different. In the algorithm to calculate the outcome of the Boston Mechanism, all the assignments at each step are permanent. Whereas in the Deferred Acceptance Algorithm, the assignments are temporary until the algorithm stops. Chen and Kesten (2017) observes this feature and creates a cluster of mechanisms that differs with respect to their periodic steps at which the assignments are finalized (denoted by $e \in \{1, 2, ..., \infty\}$), concordantly called application-rejection mechanisms. Their setting encapsulates all such mechanisms, in particular, the Boston mechanism (the case where $e = 1$) and student-optimal Stable Mechanism (the case where $e = \infty$). Hence for any mechanism with $e \in \mathbb{N}\backslash\{0, 1\}$, we have a mechanism that lies between the spectrum created by the Boston mechanism at one end and the student-optimal Stable Mechanism at the other. They construct this model to create a general framework to analyze mechanism that are employed in various regions that uses centralized mechanism for university admission processes in China.

Another early weakening of the stability notion is done by Abdulkadiroğlu (2011). He emphasize the efficiency differences between the one-sided and two-sided formulations of the school choice problem; more precisely, their handling of the priorities in their respective settings. Most importantly, he argues that the efficiency loss that might be created by the student-optimal Stable Mechanism due to its strict stability notion. Its strictness comes from the fact that it requires respecting priorities at all schools. In order to increase the efficiency, they propose a new matching model with a weaker stability notion. They partition the set of schools into two: “stability constrained“ schools and “stability-unconstrained schools“. They say that a matching is pseudo-stable if it does not violate priorities in stability-constrained schools. The mechanism’s process is starting with a pseudo-stable matching and implementing improvement cycles that do not violate pseudo-stability. The model reduces to one-sided matching if all the schools are stability-unconstrained and reduces to two-sided matching if all the schools are stability-constrained. The model is hybrid when it is neither of those cases.

Dur et al. (2015) proposed a new mechanism that is called Student Exchange under Partial Fairness for the school choice problem that allows priority violations. Their motivation comes from the consideration of schools districts to allow priority violations for certain schools. An example for this setting is given by Abdulkadiroğlu (2011), in which the assignment process for exam schools and regular schools is unified. What makes this case interesting is that there are legal restrictions of the priorities that an exam school can admit whereas regular schools are much more flexible in choosing their priorities. Their mechanism also depends on identifying the outcome of the Deferred Acceptance Algorithm and implementing improvement cycles that do not violate partial fairness.

Papai (2013) also tried to construct a hybrid mechanism, but they tried to reconcile the nice properties of the student-optimal Stable Mechanism and the Top Trading Cycles Mechanism. It is a well known result that both mechanisms are strategy-proof, the student-optimal Stable Mechanism is stable and the Top Trading Cycles Mechanism is efficient, but the reverse of the last two statements are not true. They also introduce a weaker stability notion, induced by an objection rule that is more demanding than the objection rule in the usual stability. The usual objection rule is also what defines a blocking pair: an agent can object to an allocation if there exist a more-preferred object (which is a school seat in the school choice problem setting) that is allocated to another agent with a lower priority than her. Papai's objection rules takes into consideration the possibility that the agent got this allocation via trade. For illustration, take two agents i and j such that i prefers j's assignment to her own assignment, and also has higher priority than j in j's object. In the usual objection rule, i can rightfully object to such an allocation. But Papai takes into account that i might have ended up with that allocation by trading her previous allocation. More specifically, it might be the case that i traded with another agent who has a lower priority than j in the priority profile of j's assignment. The objection rule Papai introduces do not render these objection rightful, and defines a new stability notion that builds onto that rule, called Individual Trade Stability. This notion is obviously weaker than the usual stability notion, and is satisfied by both the student-optimal Stable Mechanism and the Top Trading Cycles Mechanism.

It is important to notice that all the results listed above are obtained under the complete information assumption. This assumption requires that the student not only know the complete priority profiles of all the schools, but also know the complete preference profile of all the students in the market. This is a very demanding assumption, which we can safely assume that is not satisfied in the school choice problems in real life. In order to obtain results for an incomplete information setting, Roth and Rothblum (1999) considers a setting in which the agents have symmetric beliefs[11] about the preferences of the other agents in the market which is a worker-firm matching market in their context. They show that, under the firm-proposing Deferred Acceptance Mechanism, the workers cannot benefit from switching the position of two firms whenever they have symmetric beliefs about the said firms[12]. Building on this, Ehlers (2008) provides the sufficiency conditions for a mechanism to be robust to such misreported preferences under the symmetric beliefs incomplete information setting: Anonymity[13] and Positive Association[14].

Erdil and Ergin (2008) imported the symmetric belief setting to the school choice problems, and showed that if students have symmetric beliefs, the students cannot be better of by switching the schools in their preferences under the stable improvement cycles algorithm.

Kesten (2010) also showed that, for their mechanism EADAM and under the symmetric belief incomplete information setting, truth-telling stochastically dominates any other strategy when other students are truthful, which implies that truth-telling

[11]Symmetric beliefs can be illustrated by the perfect example of Roth and Rothblum (1999): "A new assistant professor candidate in economics might have {Harvard,MIT}-symmetric beliefs, if, despite knowing which of the two she preferred, she couldn't say which of the two was more likely to rank her highly compared to other top candidates, or which of the two would likely to be preferred by other candidates".

[12]They are unable to provide such a result for a truncation strategy. The agents can be better off by misreporting their preferences via truncating their true preferences.

[13]Anonymity requires that, when we switch the position of two firms in the preference profiles of the workers and also switching their respective preferences with each other, the resulting matching should be exactly the same as before, except the assignments of the two said firms are exchanged. This implies that a mechanism that satisfies anonymity should treat all the firms equally.

[14]Positive Association requires that, if a worker switches the ranking of the firm she is placed to with another firm that is ranked higher in her preference profile, then her assignment should remain the same as before.

constitutes an ordinal Bayesian Nash equilibrium in their respective setting.

3 The School Choice Problem, the Student-Optimal Stable Mechanism and the Boston Mechanism

3.1 The School Choice Problem

The school choice problem is introduced to the economic literature by Abdulkadiroğlu and Sönmez (2003). It consists of a set of students S and a set of schools C that represent the two sides of the market. Both the number of students and the number of schools are finite. Each student s should be matched to one school. Each school c has a quota q_c which indicates the maximum number of students that can be matched to that school. It is assumed that there is no shortage of seats in the market (i.e. $\sum_{c=1}^{\#|C|} q_c \geq \#|S|$[15]). Each student s has a preference over the set of schools, P_s, and each school has a priority ranking over the set of students, $\succeq_c$. Let R_s be the weak preference relation induced by P_s[16]. Hence a school choice problem $(S, C, Q, P, \succeq)$ identifies:

- S: the finite set of students in the market
- C: the finite set of schools in the market
- Q: a quota vector for schools $Q = (q_c)_{c \in C}$ such that $q_c \in \mathbb{Z}_{++}$
- P: a preference profile of students $P = (P_s)_{s \in S}$ such that P_s is a strict preference relation ordering of student s over the schools in the market and the empty set (which signifies remaining unmatched[17]), i.e. $P_s \in (C \cup \emptyset) \times (C \cup \emptyset)$

[15] $\#|.|$ denotes the cardinality of a set, a notation we will employ throughout the paper.
[16] Formally, xR_sy if either xP_sy or $x = y$.
[17] Remaining unmatched, which is the outside option for the student, can be interpreted as attending a private school or homeschooling

- $\succeq$: a priority profile of schools $\succeq = (\succeq_c)_{c \in C}$ such that $\succeq_c$ is a binary relation over the set of students that is complete, reflexive and transitive.

Next, we define the notion of matching.

Definition 1. A **matching** of students to schools is a function $\mu : S \to (C \cup \emptyset)$ such that

- $\mu(s) \subset C$ such that $\#|\mu(s)| \leq 1 \ \forall s \in S$;
- $\mu^{-1}(c) \subset S$ with $\#|\mu^{-1}(c)| \leq q_c \ \forall c \in C$;
- $s \in \mu^{-1}(c)$ if and only if $c \in \mu(s)$, $\forall s \in S$ and $\forall c \in C$.

There are some desired properties that are constructed for a matching. A matching is *individually rational* if $\mu(s) P_s \emptyset \ \forall s \in S$. A matching is *non-wasteful* if for any school c and any student s, $c P_s \mu(s) \Rightarrow \#|\mu^{-1}(c)| = q_c$ where $c \neq \mu(s)$. The definition of stability follows these desired properties.

Definition 2. A matching μ is *stable* if ;
(i) it is individually rational,
(ii) it is non-wasteful,
(iii) there exists no student school pair $(s, c) \in S \times C$ such that $c P_s \mu(s)$ and $s \succ_c j$ for some student $j \in \mu^{-1}(c)$.

The pair (s, c) in the third requirement is called a *blocking pair.* The third requirement is also known as *justifiable envy.*

We say that a matching μ *Pareto-dominates* another matching ϕ if $\mu(s) R_s \phi(s) \ \forall s \in S$ and $\mu(j) P_j \phi(j)$ for at least one $j \in S$. A matching μ is *Pareto-efficient* if it is not Pareto-dominated by another matching.

In this paper, we analyze school choice mechanisms. A *mechanism* φ is a function which takes any school choice problem $(S, C, Q, P, \succeq)$ and assigns it to a matching, where the matching outcome is denoted by $\varphi(S, C, Q, P, \succeq)$, the match for a student s is denoted by $\varphi_s(S, C, Q, P, \succeq)$ and a match for a school c is denoted by $\varphi_c(S, C, Q, P, \succeq)$. A mechanism is called a *stable* if it produces a stable matching for all school choice problems. A mechanism φ *Pareto-dominates* another matching φ' if $\varphi(S, C, Q, P, \succeq) R_s \varphi'(S, C, Q, P, \succeq)$ $\forall s \in S$ and $\varphi(S, C, Q, P, \succeq) P_j \varphi'(S, C, Q, P, \succeq)$ for at least one $j \in S$. A matching is *Pareto-efficient* if it produces a Pareto-efficient matching for all school choice problems. A mechanism is *strategy-proof* if it makes the truthful reporting of preferences a best response for all of the students; formally, $\varphi_s(S, C, Q, P, \succeq) R_s \varphi_s(S, C, Q, (P'_s, P_{-s}), \succeq)$, where P'_s denotes any possible misreported preference by student s (i.e. $P'_s \in [(C \cup \emptyset) \times (C \cup \emptyset)] \backslash \{P_s\}$).

Next, we formally define the student-optimal Stable Mechanism, the Boston Mechanism and their corresponding algorithms.

3.2 The Student-Optimal Stable Mechanism

The student-optimal Stable Mechanism builds upon the Deferred Acceptance Algorithm by Gale and Shapley (1962), which they introduced for the college admission problem. The following algorithm is a version of the student-proposing Deferred Acceptance Algorithm translated to the school choice problem setting.

Step 1: Each student proposes to her first choice. Each school tentatively assigns its seats to the students who proposed based on its priority order, up to its quota. Any remaining proposers (if they exist) are rejected.

In general, at

Step t: Each student who was rejected at step $t-1$ proposes to her next choice according to her preference order. Each school consider both the students it tentatively accepted at step $t-1$ and the students proposed in this step, and again tentatively assign students to its seats from this set according to its priority order, up to its quota. Any remaining proposers (if they exist) are rejected.

The algorithm terminates at the step in which no students are rejected or there are no seats left in any of the schools in the market. Each student is assigned permanently to her tentative assignment in the last step.

Since the cardinalities of the sets of schools and students are finite, the algorithm terminates in finite time (Gale and Shapley (1962), Dubins and Freedman (1981)). The direct mechanism induced by the student-proposing Deferred Acceptance Algorithm is called the student-optimal Stable Mechanism. The most important aspect of this mechanism is that it is a stable mechanism. Moreover, it has a number of other nice properties.

Remark 1 (*Gale and Shapley, 1962*). Student-optimal Stable Mechanism Pareto-dominates any other mechanism that eliminates justified envy.

Hence, from this remark, we can conclude that the student-optimal Stable Mechanism is the Pareto-efficient mechanism in the class of stable mechanisms[18].

Remark 2 (*Dubins and Freedman, 1981 and Roth, 1982*). The student-optimal Stable Mechanism is strategy-proof.

A strategy-proof mechanism is desired since it eliminates the incidences in which sophisticated agents can take advantage of the system and reduce the welfare of the naive agents.

Although it seems like the student-optimal Stable Mechanism has all the best properties that a mechanism can possess, its problems come to light when one's concern is Pareto-efficiency, as shown by Roth (1982). An example for the efficiency shortcomings of the student-optimal Stable Mechanism can be seen in Example 1 of Abdulkadiroğlu and Sönmez (2003).

3.3 The Boston Mechanism

Another very popular mechanism is the Boston Mechanism and it was first pinned down by Abdulkadiroğlu and Sönmez (2003). Below, we identify the generalized

[18] The class of stable mechanisms is defined as the set of school choice mechanisms that produces a stable outcome for any school choice problem $(S, C, Q, P, \succeq)$.

version of Boston Mechanism[19,20]:

Step 1: Each student proposes to her first choice. Each school permanently assigns its seats to the students who proposed based on its priority order, up to its quota. Any remaining proposers (if they exist) are rejected.

In general, at

Step t: Consider the remaining students who are not placed in any seats in the previous steps. For each school c with available seats, consider the students who have listed it as their t^{th} choice. Permanently assign the remaining seats of the school to the students who proposed in this step according to its priority order, up to its quota. Any remaining proposers (if they exist) are rejected.

The algorithm terminates at the step in which no students are rejected or there are no seats left in any of the schools in the market.

The appeal of Boston Mechanism may be in part due to its welfare properties. The outcome of the Boston Mechanism might Pareto-dominate the outcome of the student-optimal Stable Mechanism, as shown by Miralles (2009), Abdulkadiroğlu et al. (2011) and Troyan (2012). But the most important shortcoming of the Boston Mechanism is the fact that it is not strategy-proof.

Building upon the manipulability measure constructed by Pathak and Sönmez (2013), Chen and Kesten (2013) find that the Boston Mechanism (employing a definition for a class of mechanisms that encompasses the Boston Mechanism and the student-optimal Stable Mechanism) is extremely vulnerable to manipulability. Moreover, this theoretical results are reinforced with various laboratory experiments, such as Chen and Sönmez (2006) and Chen and Kesten (2013). Both of the studies find a proportion of the participants misreporting their preferences under the Boston Mechanism.

[19]In Abdulkadiroğlu and Sönmez (2003), the Boston Mechanism is identified with the specific priority order that the Boston school system used (namely, first priority is given to the students who have siblings in a specific school and also live within the walk zone, second to the students who have siblings in a specific school, third to the student who live within the walk zone and forth to the rest of the students), but we identify a version of the original Boston Mechanism that can be formulated with any possible priority order.

[20]Some alternative axiomatic characterizations of the Boston Mechanism can be found in Afacan (2013) and Kojima and Ünver (2014).

4 A School Choice Model with Tolerance

Two disjoint sides of the market are denoted by S for students and C for schools. Each student $s \in S$ has a *preference relation* P_s, which is a complete, reflexive and transitive binary relation over $C \cup \{\emptyset\}$, where the empty set represents not being matched to any school. Let R_s be the weak preference relation induced by P_s such that cR_sc' if either cP_sc' or $c = c'$. Let $P = (P_s)_{s \in S}$ be the *preference profile of all the students.* We say that a school $c \in C$ is *acceptable* to student s if $cP_s\emptyset$. Let $Rank(c|P_s)$ be the ranking of the school c in the preference profile of the student s. Let $Rank(s| \succ_c)$ be the ranking of the student s in the priority order ranking of school c. Each school $c \in C$ admits a *strict priority order* $\succ_c$ over S, and the priority order profile of schools is denoted by $\succ = (\succ_c)_{c \in C}$ [21]. Naturally, $Rank(i| \succ_c) > Rank(j| \succ_c) \iff i \succ_c j$. Additionally, every school has a *quota* which is denoted by q_c and the quota profile of the schools is denoted by $Q = (q_c)_{c \in C}$. Up until this point, the model is the same with the other conventional school choice problems. To this setting, we add the *tolerance profile of the schools*, which is denoted by the vector $L = (l_c)_{c \in C}$. For any school c, l_c corresponds to the minimum rank difference for a student s (compared to the students c has already accepted) such that c is willing to break ties with an already accepted student and accept s instead. Naturally, $L \in \mathbb{N}^{\#|C|}$, i.e. $l_c \geqslant 0 \ \forall c \in C$. Hence, a school choice problem in our setting consists of the tuple $(S, C, Q, P, \succ, L)$.

A *matching* μ is a group of many-to-one assignments such that a student $s \in S$ is matched to at most one school and a school $c \in C$ is matched to at most q_c students. Let $\mu(s)$ denote the student s's assignment under the matching μ and let $\mu^{-1}(c)$

[21] Although it is hardly the case in real life school choice problems, one can assume that a prior tie-breaking rule strictly ordered the students within the indifference classes. Moreover, our results are robust to changing the priority order of schools to weak binary relations, after adjusting the tolerance values accordingly.

denote the set of students that assigned to school c under matching μ. Formally, a matching is a function $\mu : S \to C$ such that $\mu(s)$ is always singleton valued and $\#|\mu^{-1}(c)| \leqslant q_c$. The definitions of *individual rationality*, *non-wastefulness*, *Pareto-dominance* and *Pareto-efficiency* for matching allocations follow from Section 3.1. Moreover, the definitions of *Pareto-dominance*, *Pareto-efficiency* and *strategy-proofness* for direct matching mechanisms follow from Section 3.1 as well.

In the majority of school choice problems, there is a trade-off between stability, efficiency and strategy-proofness. Efficiency of a matching is an obvious desirable property. On the other hand, stability enables us to design "robust" mechanisms and avoid legal actions. Also strategy-proofness is an important notion of fairness, and it helps us protect the naive agents in the market from a possible predatory behavior of sophisticated agents. The constrained efficient stable matching (produced by the student-optimal Stable Mechanism) is the most efficient within the class of stable matchings, but it might be Pareto-dominated by another matching that is not stable.

In this paper, we introduce a new notion of stability which is weaker than the conventional stability. It enables us to find a midpoint between efficiency and stability for a matching. Employing our notion, we can improve the efficiency status of an allocation while maintaining some form of stability. In this light, we define the following notion of stability, in which we weaken the *justifiable envy* notion:

Definition 3. Given a school choice problem $(S, C, Q, P, \succ, L)$ and a matching μ for this problem, μ is *preference respecting stable* if
(i) it is individually rational,
(ii) it is non-wasteful,
(iii) $\nexists(s, c) \in S \times C$ s.t. $cP_s\mu(s)$ such that either

$(iiia)$ $\exists j \in \mu^{-1}(c)$ such that $Rank(j|\succ_c) - Rank(s|\succ_c) > l_c$

or

$(iiib)$ $\exists k \in \mu^{-1}(c)$ such that $s \succ_c k$ and $Rank(c|P_k) \geqslant Rank(c|P_s)$

or both

The last part of the definition exposes what *justified envy* is in our setting. This notion can be explained verbally as follows: Take any two students i and j such i envies the seat of student j at school c. Student i can justifiably envy student j if either (a) she is ranked l_c higher than j in the priority profile of c, or (b) she has higher priority than j and ranks c higher in her preference profile, compared to j. This rather involved justifiable envy definition enables us to weaken the conventional notion of stability introduced by Gale and Shapley (1962).

One of the most evident features of this formulation is that, when $l_c = 0\ \forall c \in C$, preference respecting stability and Gale-Shapley stability coincide, which constitutes the following remark.

Remark 3. Preference respecting stability is equivalent to Gale-Shapley stability whenever $l_c = 0\ \forall c \in C$.

The proof of this remark is provided in the Appendix. Moreover, notice that as toleration profile vector L increases, the justifiable envy condition (and hence, our stability condition) gets weaker and weaker. The following remark is a direct consequence of this observation.

Remark 4. Take any preference respecting stable matching μ for a school choice problem $(S, C, Q, P, \succ, L)$. Matching μ is also preference respecting stable for the school choice problem $(S, C, Q, P, \succ, L')$ where $L' \geqq L$ (i.e. $l'_c \geqslant l_c\ \forall c \in C$).

The proof of this remark can be found in the Appendix. Verbally, a preference respecting stable matching continues to be preference respecting stable if we weakly increase the tolerance profile vector. Moreover, employing this remark, we can see that the set of preference respecting stable allocations in a school choice problem is always weakly greater than the set of Gale-Shapley stable allocations for the same problem, i.e. preference respecting stability is a weaker condition than stability. This observation gives way to our existence result.

Corollary 1. *There always exists a preference respecting stable allocation for any given school choice problem.*

Our existence result follows from Gale and Shapley (1962) who showed that there always exists a stable matching in any two-sided matching market, and this matching can be found by the Deferred Acceptance Algorithm. Hence, for any school choice problem, the Deferred Acceptance Algorithm will produce a preference respecting stable allocation. Also, since the problem in which $l_c = 0\,\forall c \in C$ coincides with the usual school choice problem and since a preference respecting stable allocation stays preference respecting stable if we increase l_c for any school c, the outcome of the Deferred Acceptance Algorithm is stable for any school choice problem regardless of the values of l_c's.

Moreover, this result shows that the set of preference respecting stable allocations for any given school choice problem with tolerance values is weakly greater than the set of stable allocations for that school choice problem.

5 The Preference Respecting Stable Mechanism

A natural research direction is to construct a mechanism that produces the constrained efficient preference respecting stable matching for a given school choice problem. Most straightforward way of constructing such a mechanism is to incorporate tolerance values of the schools into their respective priority profiles, i.e. updating the existing priority orders to a new artificial priority order, and finding an algorithm that would directly produce a preference respecting stable matching. However, notice that the artificial priority order profile $\rhd = (\rhd_c)_{c \in C}$ induced by $\succ$ is not transitive, hence there may be cycles in the priority order for a school. In order to see that, the artificial priority order $\rhd$ is defined as follows:

For a school $c \in C$ and any two students $i, j \in S$;

$$
\begin{aligned}
i \rhd_c j \Longleftrightarrow\ & Rank(c|P_j) > Rank(c|P_i) \text{ and } Rank(i|\succ_c) \geqslant Rank(j|\succ_c) - l_c \\
& \text{or} \\
& Rank(c|P_j) = Rank(c|P_i) \text{ and } Rank(i|\succ_c) > Rank(j|\succ_c) \\
j \rhd_c i \Longleftrightarrow\ & \text{otherwise}
\end{aligned}
$$

For a counter-example, take a school c and students i, j, k, and assume that $l_c = 3$, $Rank(c|R_i) = 1$, $Rank(c|R_j) = 2$, $Rank(c|R_k) = 3$, $Rank(i|\succ_c) = 1$, $Rank(j|\succ_c) = 3$ and $Rank(k \succ_c) = 5$. Given $\rhd_c$ induced by $\succ_c$, it is easy to see that $i \rhd_c j \rhd_c k$, but $k \rhd_c i$. Due to this result, it is not possible to construct a direct algorithm that would give a preference respecting stable outcome, since there might be no maximal element in the set of students with respect to the artificial priority that is induced by the priority order and the tolerance vector of schools.

After this negative result, we focus on finding an iterative mechanism that can produce a preference respecting stable outcome from a known allocation. Since we know that the set of stable allocations is a subset of the set of preference respecting stable allocations (as a direct consequence of Corollary 1), we try to find a preference respecting stable allocation that Pareto-dominates (if such an allocation exists) the outcome of the Deferred Acceptance Algorithm, and furthermore, the constrained efficient one. We say that a preference respecting stable matching μ is *constrained efficient* if there exists no other preference respecting stable matching that Pareto-dominates μ. Since out main aim is to increase the efficiency of an allocation, a mechanism that produces *any* preference respecting stable allocation is not of great use, hence the mechanism should identify the most efficient allocation within the set of preference respecting stable allocations.

In accordance with this research motivation, we introduce a two-step mechanism that will produce the preference respecting stable allocation that is Pareto-superior to the outcome of the Deferred Acceptance Algorithm and other preference respecting stable allocations, if there is any; and if there does not exist such an allocation, it will produce the same outcome as the Deferred Acceptance Algorithm. The mechanism is very similar to the "stable improvement cycles algorithm" of Erdil and Ergin (2008). This algorithm is a two-stage algorithm that consists of the conventional Deferred Acceptance Algorithm in the first stage, and implementing the trading cycles in the second stage, which leads to an increase in welfare for the students. Building on top of this idea, we now introduce our algorithm that will produce the constrained efficient preference respecting stable matchings. In order to construct our algorithm, we initially define the *preference respecting stable improvement cycle* as follows:

Definition 4. Given a school choice problem $(S, C, Q, P, \succ, L)$ and a preference respecting stable matching μ, we say that μ admits a preference respecting stability compatible improvement cycle if there are distinct students $\{i_1, ..., i_n\} \in S$ with $n \geqslant 2$ such that for any $k \in \{1, ..., n\}$:

- $\mu(i_k) \neq \emptyset$,
- $\mu(i_{k+1}) P_{i_k} \mu(i_k)$ with $\mu(i_{n+1}) = \mu(i_1)$,
- for every $j \in S \backslash \{i_1, ..., i_n\}$,
  if $Rank(i_k| \succ_{\mu(i_{k+1})}) - Rank(j| \succ_{\mu(i_{k+1})}) > l_{\mu(i_{k+1})}$ then
  $\mu(j) R_j \mu(i_{k+1}) \; \forall i_k \in \{i_1, ..., i_n\}$,

or

if $\mu(i_{k+1})P_j\mu(j)$ for any $i_k \in \{i_1, ..., i_n\}$, then either $i_k \succ_{\mu(i_{k+1})} j$ or $j \succ_{\mu(i_{k+1})} i_k$ and $Rank(\mu(i_{k+1})|R_{i_k}) \leqslant Rank(\mu(i_{k+1})|R_j)$,

- for every $i_k, i_{k'} \in \{i_1, ..., i_n\}$,
  if $Rank(i_{k'}| \succ_{\mu(i_{k'+1})}) - Rank(i_k| \succ_{\mu(i_{k'+1})}) > l_{\mu(i_{k'+1})}$ then $\mu(i_{k+1})P_{i_k}\mu(i_{k'+1})$
  or
  if $\mu(i_{k'+1})P_{i_k}\mu(i_{k+1})$, then either $i_{k'} \succ_{\mu(i_{k'+1})} i_k$ or $i_k \succ_{\mu(i_{k'+1})} i_{k'}$ and $Rank(\mu(i_{k+1})|R_{i_k}) > Rank(\mu(i_{k+1})|R_{i_{k'}})$.

Accordingly, we define the improved matching μ' which is obtained by implementing this preference respecting stable improvement cycle above as follows:

$$\mu' = \begin{cases} \mu'(i_j) = \mu(i_j) & if\ j \in S \backslash \{i_1, ..., i_n\} \\ \mu'(i_j) = \mu(i_{k+1}) & if\ j = i_k\ for\ k \in \{1, ..., n\} \end{cases}$$

Hence, our mechanism's procedure can be verbally explained as follows: It first pins down the outcome of the student-proposing Deferred Acceptance Algorithm. Then, it identifies all the possible improvement cycles that would make the students participating in the cycle strictly better of by assigning them to a school they prefer, and make the students that are not participating in the cycle indifferent, where the cycles do not violate the requirements of preference respecting stability when they are implemented. If there exist multiple of such cycles, then it picks one randomly, or if there exists a unique such cycle, it picks that one directly, and implements the cycle.

For a clear illustration of the working of our mechanism, we provide the following simple example:

Example 1 Take a school choice problem where $S = \{s_1, s_2, s_3, s_4\}$, $C = \{c_1, c_2, c_3, c_4\}$,

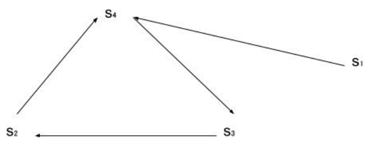
$q_c = 1$ and $l_c = 2$ $\forall c \in C$. P and $\succ$ are defined as below:

P_{s_1}	P_{s_2}	P_{s_3}	P_{s_4}		$\succ_{c_1}$	$\succ_{c_2}$	$\succ_{c_3}$	$\succ_{c_4}$
c_1	c_1	c_2	c_3		s_4	s_2	s_3	s_4
c_4	c_2	c_3	c_1		s_1	s_3	s_4	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$		s_2	$\vdots$	$\vdots$	
					$\vdots$			

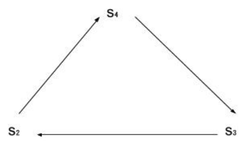
As Step 1 of our mechanism, we apply the student-proposing Deferred Acceptance Algorithm, and the outcome is as follows:

$$\mu^1 = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 \\ c_4 & c_2 & c_3 & c_1 \end{pmatrix}$$

After identifying the outcome of Step 1, we look for possible improvement cycles that would not violate preference respecting stability, as defined in Definition 4. If all the students point towards the students whom they want their respective seats, we have:



From this, we can clearly see that we have the following improvement cycle:



After implementing the cycle above, we have the following matching as a result of Step 2:

$$\mu^2 = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 \\ c_4 & c_1 & c_2 & c_3 \end{pmatrix}$$

It is clear that there exists no improvement cycle to implement after Step 2, hence the resulting matching of the Preference Respecting Stable Mechanism is μ^2

Notice that $\mu^2 \sim_1 \mu^1$ and $\mu^2 P_2 \mu^1$, $\mu^2 P_3 \mu^1$ and $\mu^2 P_4 \mu^1$. This straightforwardly implies that the outcome of the Preference Respecting Stable Matching Pareto-dominates the outcome of the Deferred Acceptance Algorithm.

Notice that, as also explained in Example 1, the resulting matching μ' is still preference respecting stable, and it Pareto-dominates the initial matching μ. Formally, $\mu'(i) R_i \mu(i)\ \forall i \in S$ and $\mu'(j) P_j \mu(j)$ for at least one $j \in S$. More specifically, $\mu'(j) P_j \mu(j)\ \forall j \in \{i_1, ..., i_n\}$ where $\{i_1, ..., i_n\}$ is the set of students who constitute the improvement cycle.

Theorem 1. *Given a school choice problem $(S, C, R, \succ, Q, L)$, a preference respecting stable matching μ is constrained efficient if and only if it does not admit a preference respecting stability compatible improvement cycle.*

The proof is provided in the Appendix.

We now formally construct our two-stage mechanism. Given a school choice problem $(S, C, Q, P, \succ, L)$;

Step 1: Run the Deferred Acceptance Algorithm and obtain the resulting matching μ^1.

Step t$\geqslant$ 2: Take the resulting matching μ^{t-1} and look for any possible preference respecting stability compatible improvement cycles. If there exists any such cycle, then pick one (can be picked randomly) and implement it.

The algorithm terminates when there does not remain any preference respecting stability compatible improvement cycle to implement. Since these cycles does not violate preference respecting stability and also our initial assignment is preference respecting stable, the following corollary is straightforward.

Corollary 2. *Outcome of the preference respective stable mechanism is preference respecting stable.*

Notice that when $l_c = 0 \ \forall c \in C$, the Preference Respecting Stable Mechanism is the same as the student-optimal Stable Mechanism, and when $l_c = \#|S| \ \forall c \in C$, the Preference Respecting Stable Mechanism is the same as the Boston Mechanism. Hence, our mechanism lies somewhere in the spectrum created by the student-optimal Stable Mechanism and the Boston Mechanism as two endpoints, depending on the value of the tolerance profile vector L. Since it improves upon the outcome of the Deferred Acceptance Algorithm by implementing the trading cycles, the outcome of the Preference Respecting Stable Mechanism Pareto-dominates the outcome of the student-optimal Stable Mechanism. Furthermore, due to Theorem 1, the outcome of the Preference Respecting Stable Mechanism is constrained efficient in the class of mechanisms that produce a preference respecting stable allocation for a given tolerance vector.

6 Strategic Properties under Complete Information

From a mechanism design perspective, an implementer must ascertain the the true preferences of the students (under the assumption that the schools are not strategic agents, an assumption we hold throughout this paper). A natural interest for researchers is that whether truthful reporting is an undominated strategy for the students. In order to satisfy this interest, we will first answer this question under complete information. The incomplete information case is studied in the next section.

As a reminder, we say that a mechanism is *strategy-proof* if it makes truthful report of the preferences a best response for the students; formally, $\varphi_s(S, C, Q, P, \succ, L) R_s \varphi_s(S, C, Q, (P_i', P_{-s}), \succ, L)$ $\forall s \in S$ where P_s' denotes any possible misreported preference by student i.

In the general setting of matching markets that incorporate the notion of tolerance values for the schools, this question is answered straightforwardly. As we observed previously, the Preference Respecting Stable Mechanism collapses into the student-optimal Stable Mechanism when the tolerance value for each school equals to zero, i.e. $l_c = 0$ $\forall c \in C$. So, as previous results in the literature has shown, the Preference Respecting Stable Mechanism is strategy-proof when $l_c = 0$ $\forall c \in C$.

Remark 5. The Preference Respecting Stable Mechanism is strategy-proof when the tolerance values for all schools are equal to 0, i.e. $l_c = 0$ $\forall c \in C$.

This result is provided without a proof, and follows Theorem 9 of Dubins and Freedman (1981) and also Roth (1982). Another trivial result is the case in which there is only one school with a tolerance value higher than 0. Since we need at least two students (and hence two schools) in order to create a preference respecting stability improvement cycle and the priorities of the schools are strict, the Preference Respecting Stable Mechanism is strategy-proof whenever there is only one school with a tolerance value higher than 0. This result constitutes the following remark, which is provided without a proof.

Remark 6. The Preference Respecting Stable Mechanism is strategy-proof when there is only one school with a tolerance value higher than 0, i.e. $\exists! t$ such that $l_t > 0$ and $l_c = 0\ \forall c \in C \backslash \{t\}$.

Since our novel contribution to the literature is the case when the tolerance values do not equal to zero, the special interest is on those cases. The relation between Pareto-dominance and stability is studied in the literature previously. Kesten and Kurino (2016) found that[22], in a very general setting that encompasses our setting as well, there is no strategy-proof mechanism that Pareto-dominates the student-optimal Stable Mechanism. Since the Preference Respecting Stable Mechanism Pareto-dominates the student-optimal Stable Mechanism whenever there exists at least two schools that have a tolerance value higher than 0, it cannot be strategy proof.

Theorem 2. *The Preference Respecting Stable Mechanism is not strategy-proof when there is at least two schools with tolerance values greater than 0.*

Proof provided by Kurino and Kesten (2016) directly applies to our setting and proves our result[23]. Below, for illustration, we provide an example that also constitutes a proof via counter-example.

Example 2 The argument is a counter-example. Consider the school choice problem $(S, C, Q, P, \succ, L)$ where $S = \{s_1, s_2, s_3, s_4, s_5, s_6\}$, $C = \{c_1, c_2, c_3, c_4, c_5, c_6\}$, $q_c = 1$ and $l_c = 2\ \forall c \in C$, with $\succ$ and P are defined as below:

[22] Their result is a direct generalization of Kesten (2010), which also directly applies to our setting as well.

[23] Assign $l_{t_1} = l_{t_2} = 1$, and the result follows.

P_{s_1}	P_{s_2}	P_{s_3}	P_{s_4}	P_{s_5}	P_{s_6}	$\succ_{c_1}$	$\succ_{c_2}$	$\succ_{c_3}$	$\succ_{c_4}$	$\succ_{c_5}$	$\succ_{c_6}$
c_5	c_6	c_5	c_3	c_5	c_6	$\vdots$	s_3	s_2	s_4	s_5	s_6
c_2	c_5	c_4	c_1	$\vdots$	$\vdots$		s_1	s_4	s_3	$\vdots$	$\vdots$
c_1	c_2	c_2	c_4				s_2	$\vdots$	$\vdots$		
$\vdots$	c_3	$\vdots$	$\vdots$								

In such a school choice problem, the following matching is the outcome of the Preference Respecting Stable Mechanism:

$$\mu = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 \\ c_2 & c_3 & c_4 & c_1 & c_5 & c_6 \end{pmatrix}$$

which is also the outcome of the Deferred Acceptance Algorithm, i.e., there is no improvement cycle to implement. Now consider the following problem $(S, C, Q, (P'_{s_2}, P_{-s_2}), \succ, L)$ with s_2's misreported preferences:

P_{s_1}	P'_{s_2}	P_{s_3}	P_{s_4}	P_{s_5}	P_{s_6}	$\succ_{c_1}$	$\succ_{c_2}$	$\succ_{c_3}$	$\succ_{c_4}$	$\succ_{c_5}$	$\succ_{c_6}$
c_5	c_5	c_5	c_3	c_5	c_6	$\vdots$	s_3	s_2	s_4	s_5	s_6
c_2	c_2	c_4	c_1	$\vdots$	$\vdots$		s_1	s_4	s_3	$\vdots$	$\vdots$
c_1	c_3	c_2	c_4				s_2	$\vdots$	$\vdots$		
$\vdots$	$\vdots$	$\vdots$	$\vdots$								

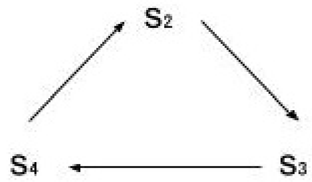
In such a case, s_2 creates a application-rejection cycle that would lead to a possibility of an improvement cycle implementation, and the following matching is the outcome of the Preference Respecting Stable Mechanism,

$$\mu' = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 \\ c_1 & c_2 & c_4 & c_3 & s_5 & s_6 \end{pmatrix}$$

where the assignment in Step 1 of the mechanism is:

$$\mu^1 = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 \\ c_1 & c_3 & c_2 & c_4 & s_5 & s_6 \end{pmatrix}$$

and the improvement cycle is:



7 Strategic Properties under Incomplete Information

In the previous section, we studied the strategic properties of the Preference Respecting Stable Mechanism under complete information, which assumes that the students have complete information concerning the preference relations of all the other students, and the priority orders and the tolerance values of all the schools in the market. This is the main component that enables the students to manipulate the system, rendering the strategy-proofness invalid. But a mechanism being not strategy-proof does not necessarily imply that the mechanism is easy to manipulate. While there is nothing problematic about assuming that the priority orders and the tolerance values of the schools are public knowledge (in real life, the criteria that form the priority orders of schools, such as proximity to the school or the siblings' enrollment, are publicly known, and it is plausible to assume that the tolerance vectors will be announced during a possible implementation), the assumption that the students have complete information about their fellow students' preference relations poses a problem regarding the plausibility of this framework. In reality, the students (or their families) may not possess such a vast knowledge or the mental capacity to process and employ that knowledge. In fact, it is much more possible that they have severely less information about the preferences than the assumption suggests.

There are various setting with which incomplete information setting can be studied. The initial setting was constructed by Roth and Rothblum (1999), and furthered by Ehlers (2008) and Kesten (2010). While allowing incomplete information, they add the assumption of "symmetric belief". In their limited information setting, a student's belief is a probability distribution over the profiles of the other agents (both other students and all the schools) in the market. The formalization follows Roth and Rothblum (1999), Ehlers (2008) and Kesten (2010). Let $\mathcal{P}_s$ be the class of all strict preferences for student s, and B_c be the class of all strict priority orders for school c. Let $X_{-s} = (B_c)_{c\in C} \times (\mathcal{P}_{s'})_{s'\in S\setminus\{s\}}$. A random school choice problem is a probability distribution $\widetilde{P}_{-s}$ over X_{-s}. $\widetilde{P}_{-s}$ signifies the belief of student s about the stated preferences of the other students and priority orders of all schools. Let M be the set of all matchings. A random matching $\widetilde{m}$ is a probability distribution over M. Let $\widetilde{m}(s)$ be the distribution that $\widetilde{m}$ induces on C. Let φ be a mechanism. For the ease of notation, we denote the problem $(S, C, Q, P_s, P_{-s}, \succ, L)$ as (P_s, P_{-s}). Given a problem (P_s, P_{-s}), where $P_{-s} \in X_{-s}$, let $\varphi(P_s, P_{-s})$ be the matching selected by φ for this problem. Let $\varphi(P_s, P_{-s})(s)$ denote the placement of student s at this matching. Given a mechanism φ and a student s with preferences P_s, each random preference profile $\widetilde{P}_{-s}$ induces a random matching $\varphi(P_s, \widetilde{P}_{-s})$ in the following way: For all $m \in M$, $Pr(\varphi(P_s, \widetilde{P}_{-s}) = m) = Pr(\widetilde{P}_{-s} = P_{-s}\, and\, \varphi(P_s, P_{-s}) = m)$. Let $\varphi(P_s, \widetilde{P}_{-s})(s)$ be the distribution that $\varphi(P_s, \widetilde{P}_{-s})$ induces over the placements of student s. Given $s \in S$ and P_s, P'_s, $P''_s \in \mathcal{P}_s$ and a random preference profile $\widetilde{P}_{-s}$, we say that strategy P'_s *stochastically* P_s*-dominates* strategy P''_s if for all $c \in C$, $Pr(\varphi(P'_s, \widetilde{P}_{-s})(s) R_s c) \geq Pr(\varphi(P''_s, \widetilde{P}_{-s})(s) R_s c)$.

A student's belief said to be symmetric for two schools c and c', which is called $\{c, c'\} - symmetric$, if given that his or her own preferences are fixed, his or her information assigns the same probability to any problem and to its symmetric problem in which the positions of c and c' are exchanged (both in terms of priorities and quotas), that is, the student is unable to deduce any difference between the two schools from her information. Formally, given $s \in S$ and $P_s \in \mathcal{P}_s$ and $c, c' \in C$, let $P_s^{c\leftrightarrow c'}$ denote the preferences in which the positions of c and c' are exchanged and the other positions in $P_s$are unchanged. Let $P_{-s}^{c\leftrightarrow c'}$ be the profile such that each student $s' \in S\setminus\{s\}$ exchanges the positions of s and s' in her preferences, schools c and c' exchange their priority orders, capacities and tolerance vectors, and the priority orders, capacities and the tolerance vectors of all other schools remain unchanged. Given $s \in S$ and $c, c' \in C$, the belief of student s for schools c and c' is *symmetric* if P_{-s} and $P_{-s}^{c\leftrightarrow c'}$ are equally probable, i.e., $Pr(\widetilde{P}_{-s} = P_{-s}) = Pr(\widetilde{P}_{-s} = P_{-s}^{c\leftrightarrow c'})$.

Due to Ehlers (2008), if a student has symmetric beliefs for two schools, there are two conditions such that under a mechanism that satisfies these two conditions, a student cannot be better off by switching the rankings of the schools. These two conditions are *anonymity* and *positive association.* Anonymity requires that, for all $s \in S$, all $P_s \in \mathcal{P}_s$, all $P_{-s} \in X_{-s}$, and all $c, c' \in C$, if $\varphi(P_s, P_{-s}) = m$, then $\varphi(P_s^{c\leftrightarrow c'}, P_{-s}^{c\leftrightarrow c'}) = m^{c\leftrightarrow c'}$, where given $m \in M$ and $c, c' \in C$, $m^{c\leftrightarrow c'}$ denotes the matching such that for all $s \in S$, if $m(s) \notin \{c, c'\}$ then $m^{c\leftrightarrow c'}(s) = m(s)$, or if $m(s) = c$ then $m^{c\leftrightarrow c'}(s) = c'$, or if $m(s) = c'$ then $m^{c\leftrightarrow c'}(s) = c$. This condition implies that if we exchange the roles of two schools c and c', then the resulting matching must only switch the assignments of c and c', and keep all other assignments same as before. Positive association requires that for all $s \in S$, all $P_s \in \mathcal{P}_s$, all $P_{-s} \in X_{-s}$, and all $c, c' \in C$, if $\varphi(P_s, P_{-s})(s) = c$ and $c' P_s c$, then $\varphi(P_s^{c\leftrightarrow c'}, P_{-s})(s) = c$. This condition implies that, given a student s and her match $m(s)$, if the roles of $m(s)$ and another school c are switched such that s prefers c to $m(s)$, then the student's placement should not change.

Kesten (2010) improves upon this characterization. In their setting, the set of schools are partitioned to different sets based on their quality. This partitioning and the quality signal conveyed by this is common knowledge for students and their families. Hence any student in the market prefers a school from a higher quality class to a school from a lower quality class. But within these classes, information of all the students are symmetric. In our investigation, we adopt this framework by Kesten (2010), while incorporating the tolerance vectors into the system. The main conclusion we acquire is the following theorem.

Theorem 3. *Suppose the following is common knowledge among the students. The set of schools is partitioned into quality classes as follows: Let* $\{C_1, C_2, ..., C_n\}$ *be a partition of* C*. Given any* $k, l \in \{1, ..., n\}$ *such that* $k < l$*, each student prefers any school in* C_k *to any school in* C_l*. Moreover, each student's information is symmetric for any two schools* c *and* c' *such that* $c, c' \in C_r$ *for some* $r \in \{1, ..., n\}$*. Then, truth telling is an ordinal Bayesian Nash equilibrium of the preference revelation game induced by the Preference Respecting Stable Mechanism.*

This results suggest that, although strategy-proofness is an important component of any matching mechanism, the concern for lack thereof might be overstated. The setting that we employ, which assumes that the students only care about the "class of the schools", and that they are not able to distinguish between the schools in the same class, is reflecting the behavior of the parents of the pupils in the matching markets in real life.

8 Conclusion

The student-optimal Stable Mechanism and the Boston Mechanism are two of the most widely used mechanisms in real life school choice markets. While student-optimal Stable Mechanism provides stability and strategy proofness, the Boston Mechanism outperforms the student-optimal Stable Mechanism in terms of welfare in some of the school choice settings. A contemporary direction in the matching literature is to understand the components of these mechanisms that lead to these properties, so that we can create new mechanisms that would pick a compromise point with respect to the designer's concern in terms of stability, efficiency and strategy-proofness; since it is a wide known result that a mechanism cannot be all of them at the same time.

In this study, we constructed a new school choice problem that incorporates tolerance values for schools and is the same with the traditional school choice problems apart from this feature. Tolerance values can be interpreted as follows: the number of priority differences that a school is willing to forego in order to accept the students who want to be in that school more compared to the higher priority-ranked students. We introduced a new notion of stability, called preference respecting stability, for the school choice problems. We formulated a new mechanism, aptly called the Preference Respecting Stable Mechanism, that depends on improvement cycles that would make the students that participate in the cycle better off, while maintaining to satisfy our stability notion. Hence, the most important feature of this mechanism is that it produces higher welfare for the students in the market compared to the student-optimal Stable Mechanism. Moreover, we showed that our mechanism is constrained efficient in the class of mechanisms that produces preference respecting allocations for a given school choice problem.

Although this mechanism is not strategy-proof, we showed that it can perform quite well in environments with incomplete information. In the incomplete information setting that we studied, the families of the students can differentiate between the "classes of schools" (and be able to rank schools according to their classes), but cannot differentiate between the schools when they are in the same class. Classes of schools can be thought as the reputations of the schools in real life. We found that, in such a setting, truthful reporting of the preferences constitutes an ordinal Bayesian Nash equilibrium in the game induced by the school choice problem in hand. This implies that the families of the students cannot be better off by misreporting their preferences to the clearing house when all the other students are reporting their preferences truthfully.

Moreover, our setting and our mechanism might be useful in real life in school choice markets with heterogenous schools, in the sense that some of them want to eliminate justified envy (in the Gale-Shapley sense) but the rest want to favor wishful applicants. Abdulkadioğlu's (2011)'s example from the Boston Public Schools system is a nice example for such a demand from schools. Our setting and mechanism gives the schools individual freedom about their decision between favoring the elimination of justified envy and favoring the acceptance of the students who want those schools more.

This setting and its corresponding mechanism might be complicated to implement in real life, but our motivation was to understand the nature of the differences between two competing mechanisms, namely the student-optimal Stable mechanism and the Boston Mechanism, and try to create another mechanism that would maintain nice properties of both of the mechanism to some extent. This questions possess a vital importance for the future of the research on the school choice problems, as well as the matching theory in general.

Appendix

Proof of Remark 3. The first and the second requirements of both of the stability notions are the same. We will show that the third requirements of the respective stability notions coincide when $l_c = 0$. Notice that, when $l_c = 0$, the requirement $(iiia)$ becomes: $\exists j \in \mu^{-1}(c)$ such that $Rank(j| \succ_c) - Rank(s| \succ_c) > 0$. Since $\succ_c$ is defined discrete and does not admit indifference classes, the condition is equivalent to the following: $\exists j \in \mu^{-1}(c)$ such that $Rank(s| \succ_c) < Rank(j| \succ_c)$. If we convert ranking terms to the priority ordering, the condition becomes: $\exists j \in \mu^{-1}(c)$ such that $s \succ_c j$. Notice that this condition is identical to the (iii) condition in the definition of Gale-Shapley stability, since we also have $cP_s\mu(s)$ in the definition of preference respecting stability. Hence, we conclude that preference respecting stability and Gale-Shapley stability coincide whenever $l_c = 0 \ \forall c \in C$. □

Proof of Remark 4. Fix $(S, C, Q, P, \succ)$ and consider two tolerance vectors L and L' such that $l_c \leq l'_c \ \forall c \in C$. Take a preference respecting stable matching μ for the problem $(S, C, Q, P, \succ, L)$. We need to show that μ is also a preference respecting stable matching for the problem $(S, C, Q, P, \succ, L')$. Since μ is preference respecting stable, there does not exist any pair (s, c) that would contradict with the third requirement of the definition of preference respecting stability. More specifically, since it is the only component that features and might be vulnerable to change, $(iiia)$ holds. This implies that there does not exist a pair (s, c) such that $cP_s\mu(s)$ and $Rank(j| \succ_c) - Rank(s| \succ_c) > l_c$ holds at the same time for any $j \in \mu^{(-1)}(s)$. Now consider the problem $(S, C, Q, P, \succ, L')$.

Case 1: $cP_s\mu(s)$ holds but $Rank(j| \succ_c) - Rank(s| \succ_c) > l_c$ does not hold in $(S, C, Q, P, \succ, L)$. This implies that $Rank(j| \succ_c) - Rank(s| \succ_c) \leq l_c$ holds in $(S, C, Q, P, \succ, L)$, which renders the blocking pair not rightful. Consider $(S, C, Q, P, \succ, L')$ where $l_c \leq l'_c \ \forall c \in C$. This directly implies that

$Rank(j|\succ_c) - Rank(s|\succ_c) \leq l'_c$, which renders the blocking pair not rightful in $(S, C, Q, P, \succ, L')$ as well. Hence the matching continues to be preference respecting stable.

Case 2: $Rank(j|\succ_c) - Rank(s|\succ_c) > l_c$ holds but $cP_s\mu(s)$ does not hold in $(S, C, Q, P, \succ, L)$. Then the student s is not interested in blocking the matching μ in $(S, C, Q, P, \succ, L)$. Since her preferences do not change in $(S, C, Q, P, \succ, L')$, she is not interested in blocking the matching μ in $(S, C, Q, P, \succ, L')$ as well.

Case 3: Neither $cP_s\mu(s)$ nor $Rank(j|\succ_c) - Rank(s|\succ_c) > l_c$ holds in $(S, C, Q, P, \succ, L)$. Any one of the arguments in Case 1 or Case 2 applies.

Since requirements (i), (ii) or $(iiib)$ does not incorporate tolerance vectors, all of them will still be satisfied if we weakly increase the tolerance vector. Hence we can conclude that, if a matching μ is preference respecting stable for the problem $(S, C, Q, P, \succ, L)$, then it is also preference respecting stable for the problem $(S, C, Q, P, \succ, L)$ where $l_c \leq l'_c$ $\forall c \in C$. This intuitively implies that if I cannot object to an allocation when the tolerance values are low. □

For the proof of Theorem 1, we initially provide a lemma that will be instrumental during the proof of the main theorem of the paper. The lemma is largely due to Erdil and Ergin (2008).

Lemma. *Suppose μ is a preference respecting stable matching that is Pareto-dominated by a (not necessarily preference respecting stable) matching v. Let S' denote the set of students who are strictly better off under v and let $C' = \mu(S')$ be the set of schools to which students in S' are assigned to under μ. Then we have:*

(i) *Students who are not in S' have the same match in both μ and v.*

(ii) *The number of students in S' who are assigned to a school c are the same in μ and v, in particular, $C' = v(S')$.*

(iii) *Each student in S' is assigned to a school both in μ and v.*

Proof. Part (i) follows from the fact that $s \in S \backslash S'$ is indifferent between $\mu(s)$ and $v(s)$ and her preferences are strict.

For part (ii), let us first show that $|S' \cap \mu^{-1}(c)| \geqslant |S' \cap v^{-1}(c)|$ for any school c. For a contradiction, suppose $|S' \cap \mu^{-1}(c)| < |S' \cap v^{-1}(c)|$ for some school c. Together with part (i), this implies that the number of students in S who are assigned to c in μ is less than the number of students who are assigned to c under v. Hence, c must have empty seats under μ. However, for any $s \in S' \cap v^{-1}(c)$, $c = v(s)P_s\mu(s)$, that is, s desires c which has empty seats under μ, which is a contradiction to the preference respecting stability of μ (due to non-wastefulness condition).

Now suppose that the inequality $|S' \cap \mu^{-1}(c)| \geqslant |S' \cap v^{-1}(c)|$ holds strictly for some school c. Summing across all schools, we have:

$$\sum_{c \in C} |S' \cap \mu^{-1}(c)| > \sum_{c \in C} |S' \cap v^{-1}(c)|$$

In other words, the number of number of students in S' who are assigned to some school in μ is more than the number of students in S' who are assigned to some school in v. Hence, there exists a student $s \in S'$ who is assigned to a school in μ but not in v. Since $s = v(s)P_s\mu(s)$, this contradicts with the preference respecting stability of μ. Hence, we know,

$$|S'| > \sum_{c \in C} |S' \cap \mu^{-1}(c)| = \sum_{c \in C} |S' \cap v^{-1}(c)|$$

Hence, there exists a student $s \in S'$, who is unmatched in v. Note that s has to be matched in μ, otherwise she would be indifferent between μ and v, a contradiction to her being in S'. But then, $s = v(s)P_s\mu(s)$, a contradiction to the preference respecting stability of μ. Hence, we must have

$$|S'| = \sum_{c \in C} |S' \cap \mu^{-1}(c)| = \sum_{c \in C} |S' \cap v^{-1}(c)|$$

□

Proof of Theorem 1. ($\Longrightarrow$) μ is preference respecting stable and constrained efficient. We need to show that it does not admit a preference respecting stability compatible improvement cycle (hereafter, improvement cycle). For a contradiction, assume not. Assume that μ is preference respecting stable and constrained efficient, and it admits an improvement cycle. Apply the improvement cycle to μ, and obtain the new matching v. Due to the earlier remark, v is preference respecting stable as well, and it Pareto dominates μ. This is a contradiction to μ being constrained efficient.

($\Longleftarrow$) μ is preference respecting stable and it does not admit an improvement cycle. We need to show that μ is constrained efficient. Assume not. Assume there exists another preference respecting stable matching v that Pareto dominates μ. We will show that, as a contradiction, μ admits an improvement cycle by building an improvement cycle.

Let S' be the set of students that are better off under v compared to μ, and let $C' = \mu(S')$. We know that $\mu(k) = v(k)$ if $k \notin S'$. We also know that $C' = v(S')$. Moreover, we know that all the students in S' are assigned to a school under both μ and v. Define $S'' = S \backslash S'$ and $C'' = C \backslash C'$. Due to the lemma above, for any student $t \in S''$, $\mu(t) = v(t)$, and $\mu(S'') = v(S'') = C''$. Hence, no student in S' is assigned to a school in C'' and no student in S'' is assigned to a school in C' under v.

Now consider any school $c \in C'$. Take the students in S' who prefer school c to their matching under μ, and pick the highest ranked one among them with respect to $\succ_c$ and call this student D_c. Formally, $D_c = \underset{\succ_c}{max}\{s \in S' : cP_s\mu(s)\}$. Repeat this process for every school in C' and find all such D_c's. Notice that there always exists such a student and it is always singleton (due to the properties of $\succ_c$). Call all such students D. Formally, $D = \underset{c \in C'}{\cup} D_c$.

Now take any student in D and call it s_1, and denote $\mu(s_1) = c_1$. Let $s_2 = D_{c_1}$ and $c_2 = \mu(D_{c_1})$. Let $s_3 = D_{c_2}$ and $c_3 = \mu(D_{c_2})$ and so on. In the end, since the set D is finite, this chain will constitute a cycle. We claim that this cycle is an improvement cycle. The first two properties of the definition of an improvement cycle are satisfied trivially.

For the third condition, we need to show that for any student $j \in S \backslash D$, if $Rank(s_k|c_{k+1}) - Rank(j|c_{k+1}) > l_{c_{k+1}}$ then $\mu(j) R_j c_{k+1}$ for any $s_k \in D$, or if $c_{k+1} P_j \mu(j)$ then either $s_k \succ_{c_{k+1}} j$ or $j \succ_{c_{k+1}} s_k$ and $Rank(c_{k+1}|R_{s_k}) \leqslant Rank(c_{k+1}|R_j)$. First, assume that $Rank(s_k|c_{k+1}) - Rank(j|c_{k+1}) > l_{c_{k+1}}$ and $c_{k+1} P_j \mu(j)$. But this is a contradiction to μ being preference respecting stable since (j, c_{k+1}) would have constitute a blocking pair. Now assume that $c_{k+1} P_j \mu(j)$. Due to μ being preference respecting stable, it must be either $s_k \succ_{c_{k+1}} j$ or if $j \succ_{c_{k+1}} s_k$ then $Rank(c_{k+1}|R_j) \leqslant Rank(c_{k+1}|R_j)$ since otherwise (j, c_{k+1}) would have constitute a blocking pair. Hence the third condition is satisfied as well.

For the forth condition, we need to show that for every $i_k, i_{k'} \in D$, if $Rank(s_{k'}| \succ_{c_{k'+1}}) - Rank(s_k| \succ_{c_{k'+1}}) > l_{c_{k'+1}}$ then $c_{k+1} P_{s_k} c_{k'+1}$ or if $c_{k'+1} P_{s_k} c_{k+1}$ then either $s_{k'} \succ_{c_{k'+1}} s_k$ or $s_k \succ_{c_{k'+1}} s_{k'}$ and $Rank(c_{k+1}|R_{s_k}) > Rank(c_{k+1}|R_{s_{k'}})$. Assume $Rank(s_{k'}| \succ_{c_{k'+1}}) - Rank(s_k| \succ_{c_{k'+1}}) > l_{c_{k'+1}}$ and $c_{k'+1} P_{s_k} c_{k+1}$. But this cannot be the case due to construction, since $s_{k'}$ is the highest priority among those who prefer $c_{k'+1}$, i.e. $s_{k'} = D_{c_{k'+1}}$, which makes the first assumption contradictory. Now assume that $c_{k'+1} P_{s_k} c_{k+1}$. Again, due to construction, $s_{k'} \succ_{c_{k'+1}} s_k$ holds for every $s_k \in D$. Hence, the forth condition is satisfied as well, which concludes the proof. □

For the proof of Theorem 3, we will employ some lemmas. Suppose all the conditions stated in Theorem 3 holds. The following lemma from Kesten (2010), which is a direct consequence of Theorem 3 of Ehlers (2008), directly applies to our setting, hence provided without a proof.

Lemma 1. *Consider a student s with true preferences P_s and information $\widetilde{P}_{-s}$ that satisfies the conditions stated in Theorem 3. Under any mechanism satisfying anonymity and positive association, the strategy P_s stochastically P_s-dominates any other strategy P'_s that ranks every school in C_r above every school in C_k for all $r < k$.*

Lemma 2. *The school at which student s is placed under the student-optimal Stable Mechanism and the school she placed under Preference Respecting Stable Mechanism are in the same quality class.*

Proof. Assume not. Assume that a student $s \in S$, who is matched to a school c under the student-optimal Stable Mechanism (which constitutes the first step of the Preference Respecting Stable Mechanism), is matched to school c' under Preference Respecting Stable Mechanism such that $c' \in C_r$ and $c \in C_k$ where $r < k$. This implies that, student s was in the preference respecting stability improvement cycle after Deferred Acceptance Algorithm is applied. Notice that the case where

$r \geqslant k$ does not constitute a viable case due to individual rationality condition which requires that every student that participates in the preference respecting stability compatible improvement cycle must be strictly better off from participation. If student s is matched to c', which is universally preferred to school c by the students, after the implementation of the cycle, some other student s' in the cycle must be matched to a school that is in a lower quality class compared to her assignment under Deferred Acceptance Algorithm, not necessarily c. In order to see this, assume the cycle is constituted by the set of students $\{s_1, s_2, ..., s_t\}$ and the set of schools $\{c_1, c_2, ..., c_t\}$ where $DA(s_i) = c_i$ for all $i \in \{1, 2, ..., t\}$, and $PR(s_i) = c_{i+1}$ if $i \in \{1, 2, ..., t-1\}$ and $PR(s_t) = c_1$; where $DA(.)$ denotes the students' match under student-optimal Stable Mechanism and $PR(.)$ denotes the students match under Preference Respecting Stable Mechanism (i.e., after the implementation of the preference respecting stability compatible improvement cycle following the Deferred Acceptance Algorithm). Notice that this is the usual construction of a preference respecting stability compatible improvement cycle. Assume that, for simplicity, all the students in $\{s_2, s_3, ..., s_t\}$ is matched to a higher quality class school or an equal quality class school under Preference Respecting Stable Mechanism compared to their matches under student-optimal Stable Mechanism while only one of them is matched to a school that belongs to a higher quality class school, and assume for simplicity that this student is s_t, and consider the case for s_1. The assumptions implies that $c_1 P c_t R c_{t-1} R ... R c_3 R c_2$ by all the students in the improvement cycle, where xPy implies that x belongs to a higher quality class than y (which implies that $xP_s y$ for all $s \in S$) and xRy implies that x belongs to a higher quality class than or equal quality class to y (which implies that $xR_s y$ for all $s \in S$). Due to transitivity of P (by construction), it is implied that $c_1 P c_2$ by all the students, which implies that $c_2$belongs to a lower quality class than c_1. This specifically implies that $c_1 = DA(s_1) P_{s_1} PR(s_1) = c_2$. This is a contradiction to the second requirement (individual rationality) of the preference respecting stability compatible improvement cycle definition. □

Lemma 3. *Given a student s with preferences P_s, let P_{-s} be a realization of $\widetilde{P}_{-s}$. Suppose that student s is placed at some school $x \in C_r$ in the problem $P = (P_s, P_{-s})$ under Preference Respecting Stable Mechanism. Suppose that student s considers* *submitting the preference list* $P_s^{c \leftrightarrow c'}$*, in which the positions of the schools* c *and* c' *are switched. Suppose that student* s *is placed at some school* y *in the problem* $(P_s^{c \leftrightarrow c'}, P_{-s})$ *under Preference Respecting Stable Mechanism. If the school* x *is not in the same quality class with neither* c *nor* c'*, then we have* xR_sy*.*

Proof. Notice that, we need to show that the outcomes of the round 1 of the Preference Respecting Stable Mechanisms in both problems are in the same quality class. This suffices due to the result of Lemma 2. Without loss of generality, we assume that cP_sc'. There are three possible cases to consider:

Case 1: If xP_sc, then the Preference Respecting Stable Mechanism terminates with the same match for both problems P and $(P_s^{c \leftrightarrow c'}, P_{-s})$, since position switch of c and c' does not affect the outcome of Round 1. Hence we can conclude that xR_sy.

Case 2: If $c'P_sx$, then during the Deferred Acceptance Algorithm in the initial round, student s applies to the same schools in the problem $(P_s^{c \leftrightarrow c'}, P_{-s})$ as she would in problem P, and her match in Round 1 must be the same for both problems. Hence we can conclude that xR_sy.

Case 3: If cP_sxP_sc', then the Round 1 of the both problems are identical until she applies to school c in problem P and school c' in problem $(P_s^{c \leftrightarrow c'}, P_{-s})$. If student s matches with school c' at the end of Round 1, due to Lemma 2, the outcome of the Preference Respecting Stable Mechanism is either c' or another school in the quality class of c' (which must be lower than x). Hence we can conclude that xR_sy. If student s gets rejected from school c', then the next schools she applies to in Round 1 are identical to the schools she applies to in Round 1 in the problem $(P_s^{c \leftrightarrow c'}, P_{-s})$ after school c. Thus her placement cannot be different. Hence we can conclude that xR_sy. □

Lemma 4. *The Preference Respecting Stable Mechanism satisfies anonymity and positive association.*

Proof. It is clear that the Preference Respecting Stable Mechanism satisfies anonymity. We will show that it satisfies positive association as well. Round 1 of the Preference Respecting Stable Mechanism directly satisfies positive association due to the strategy-proofness of the student-optimal Stable Mechanism.

Suppose that a student s is placed at school c in the Preference Respecting Stable Mechanism outcome and consider another school c' such that $c' P_s c$. This implies that she must have applied to school c in Round 1 of the mechanism (and also to school c'). Hence there are two cases to consider, in one she is matched to c in Round 1 (and was not a part of a possible improvement cycle implementation), and in the other one she is rejected by c in the process of the algorithm in Round 1 (and later matched to c after implementation of the improvement cycle). Notice that, in the second case, the school that student s is placed in Round 1 must be worse than c due to the second requirement (individual rationality) of the preference respecting stability compatible improvement cycle.

Case 1: Student s is matched to school c in Round 1 of problem P. Then in problem $(P_s^{c \leftrightarrow c'}, P_{-s})$, she must be matched to school c at the end of Round 1 as well, due to the positive association property of the student-optimal Stable Mechanism which constitutes Round 1. Moreover, in problem P, it is evident that student s applied to school c' in Round 1 and she is rejected by it, and also she failed to participate in an improvement cycle implementation. Since the preferences of all the other students are the same under both problems, she must also fail to participate in an improvement cycle implementation in the Preference Respecting Stable Mechanism process of the problem $(P_s^{c \leftrightarrow c'}, P_{-s})$ as well (since all the Round 1 matches are the same and all the preferences are the same except that of student s). Hence, the outcome of the Preference Respecting Stable Mechanism for the problem $(P_s^{c \leftrightarrow c'}, P_{-s})$ must be the same with problem P for student s (namely, school c).

Case 2: Student s is matched to a school that is worse for her than school c in Round 1 of problem P. This implies that student s participated in an improvement cycle implementation in the problem P and matched to school c as a result. This in turn implies that there does not exist an improvement cycle that would place student s at any school that is ranked higher than school c in the preferences of student s in the problem P. Again, due to strategy-proofness of the student-optimal Stable Mechanism, the outcomes for Round 1 are same for both problems, for all the students. Again, since all the preferences are the same for all the students except s, there does not exist a preference respecting stability compatible improvement cycle that would match student s to school c' in the problem $(P_s^{c \leftrightarrow c'}, P_{-s})$, since now c' is ranked lower than c and there exists an improvement cycle that would place student s to school c. Hence, the outcome of the Preference Respecting Stable Mechanism for the problem $(P_s^{c\leftrightarrow c'}, P_{-s})$ must be the same with problem P for student s (namely, school c). □

Proof of Theorem 3. Let s be a student with true preferences P_s, and let P_{-s} be a realization of $\widetilde{P}_{-s}$. Suppose that student s is placed at school $x \in C_r$ in the problem P under Preference Respecting Stable Mechanism. Take two schools c and c' and consider the strategy $P_s^{c\leftrightarrow c'}$. Suppose that student s is placed at some school y in the problem $(P_s^{c\leftrightarrow c'}, P_{-s})$ under Preference Respecting Stable Mechanism. If $C_r \cap \{c, c'\} = \{c\}$, then by Lemma 3, we have xR_sy. If $\{c'\} \subset C_r \cap \{c, c'\}$, then this strategy is equivalent to some other strategy that ranks every school in C_r above every school in C_k for all $r < k$.Then by Lemma 1, Lemma 2 and Lemma 4, for any $c, c' \in C$, strategy P_s stochastically P_s-dominates strategy $P_s^{c\leftrightarrow c'}$. Employing an induction argument (similar to the proof of part (b) of Theorem 3.1 of Ehlers (2008)), we conclude that strategy P_s stochastically P_s-dominates any other strategy $P'_s \in \mathcal{P}_s$. □

Bibliography

Abdulkadiroğlu, Atila (2011). "Generalized Matching for School Choice". Working Paper.

— (2013). "School Choice". In: *The Handbook of Market Design*. Ed. by Nir Vulkan, Alvin Roth, and Zvika Neeman. Oxford University Press, pp. 138–169.

Abdulkadiroğlu, Atila, Yeon Koo Che, and Yosuke Yasuda (2011). "Resolving Conflicting Preferences in School Choice: The "Boston Mechanism" Reconsidered". In: *The American Economic Review*.

Abdulkadiroğlu, Atila, Parag Pathak, and Alvin Roth (2009). "Strategy-Proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match". In: *The American Economic Review*.

Abdulkadiroğlu, Atila and Tayfun Sönmez (2003). "School Choice: A Mechanism Design Approach". In: *The American Economic Review* 93.3, pp. 729–747.

— (2013). "Matching Markets: Theory and Practice". In: *Advances in Economics and Econometrics*. Ed. by Daron Acemoğlu, Manuel Arellano, and Eddie Dekel. Cambridge University Press, pp. 3–47.

Abdulkadiroğlu, Atila, Parag Pathak, Alvin Roth, and Tayfun Sönmez (2005). "The Boston Public School Match". In: *The American Economic Review*.

Abdulkadiroğlu, Atila, Yeon Koo Che, Parag Pathak, Alvin Roth, and Olivier Tercieux (2017). "Minimizing Justified Envy in School Choice: The Design of New Orleans' Oneapp". NBER Working Paper.

Afacan, Mustafa Oğuz (2013). "Alternative Characterizations of Boston Mechanism". In: *Mathematical Social Sciences*.

Afacan, Mustafa Oğuz, Zeynel Harun Alioğulları, and Mehmet Barlo (2016). “Sticky Matching in School Choice”. In: *Economic Theory*.

Balinski, Michel and Tayfun Sönmez (1999). “A Tale of Two Mechanisms: Student Placement”. In: *Journal of Economic Theory*.

Chen, Yan and Onur Kesten (2013). “From Boston to Chinese Parallel to Deferred Acceptance: Theory and Experiments on a Family of School Choice Mechanisms”. WZB Discussion Paper.

— (2017). “Chinese College Admissions and School Choice Reforms: A Theoretical Analysis”. In: *Journal of Political Economy*.

Chen, Yan and Tayfun Sönmez (2006). “School Choice: An Experimental Study”. In: *Journal of Economic Theory*.

Dubins, L. E. and D. A. Freedman (1981). “Machiavelli and the Gale-Shapley Algorithm”. In: *The American Mathematical Monthly*.

Dur, Umut, Arda Gitmez, and Özgür Yılmaz (2015). “School Choice under Partial Fairness”. Working Paper.

Ehlers, Lars (2004). “In Search of Advice for Participants in Matching Markets Which Use the Deferred-Acceptance Algorithm”. In: *Games and Economic Behavior*.

— (2008). “Truncation Strategies in Matching Markets”. In: *Mathematics of Operations Research*.

Erdil, Aytek and Haluk Ergin (2008). “What’s the Matter with Tie-Breaking? Improving Efficiency in School Choice”. In: *The American Economic Review*.

Gale, David and Lloyd Shapley (1962). “College Admission and the Stability of Marriage”. In: *The American Mathematical Monthly*.

Kesten, Onur (2010). “School Choice with Consent”. In: *The Quarterly Journal of Economics*.

Kesten, Onur and Morimitsu Kurino (2016). “Do Outside Options Matter in Matching?: A New Perspective on the Trade-offs in Student Assignment”. Discussion Paper.

Kojima, Fuhito and Utku Ünver (2014). “The “Boston“ School Choice Mechanism: An Axiomatic Approach”. In: *Economic Theory*.

Miralles, Antonio (2009). “School Choice: The Case for the Boston Mechanism”. In: *Auctions, Market Mechanisms and Their Applications*. Ed. by Sanmay Das, Michael Ostrovsky, David Pennock, and Boleslaw Szymanski. Vol. 14. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Springer.

Morrill, Thayer (2015). “Making Just School Assignments”. In: *Games and Economic Behavior*.

Papai, Szilvia (2013). “Matching with Minimal Priority Rights”. Working Paper.

Pathak, Parag (2011). “The Mechanism Design Approach to Student Assignment”. In: *Annual Review of Economics*.

Pathak, Parag and Tayfun Sönmez (2013). “School Admissions Reform in Chicago and England: Comparing Mechanisms by Their Vulnerability to Manipulation”. In: *The American Economic Review*.

Roth, Alvin (1982). “The Economics of Matching: Stability and Incentives”. In: *Mathematics of Operations Research*.

Roth, Alvin and Uriel Rothblum (1999). “Truncation Strategies in Matching Markets: In Seach of Advice for Participants”. In: *Econometrica*.

Roth, Alvin and Marilda Oliveira Sotomayor (1990). *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Econometric Society Monographs No.18. Cambridge University Press.

Sönmez, Tayfun and Utku Ünver (2011). “Matching, Allocation and Exchange of Discrete Resources”. In: *The Handbook of Social Economics*. Ed. by Jess Benhabib, Alberto Bisin, and Matthew O. Jackson. Vol. 1, pp. 781–852.

Troyan, Peter (2012). “Comparing School Choice Mechanisms by Interim and Ex-Ante Welfare”. In: *Games and Economic Behavior*.